MIT Libraries | DSpace@MIT

MIT Open Access Articles

Channel coding for high-speed links: A systematic look at code performance and system simulation

Massachusetts Institute of Technology

Channel Coding For High-Speed Links: A Systematic Look at Code Performance and System Simulation

Natasa Blitvic, *Student Member, IEEE*, Maxine Lee, and Vladimir Stojanovic, *Member, IEEE*

Abstract—While channel coding is a standard method of improving a system's energy efficiency in digital communications, its practice does not extend to high-speed links. Increasing demands in network speeds are placing a large burden on the energy efficiency of high-speed links and render the benefit of channel coding for these systems a timely subject. The low error rates of interest and the presence of residual intersymbol interference (ISI) caused by hardware constraints impede the analysis and simulation of coded high-speed links. Focusing on the residual ISI and combined noise as the dominant error mechanisms, this paper analyzes error correlation through concepts of error region, channel signature, and correlation distance. This framework provides a deeper insight into joint error behaviors in high-speed links, extends the range of statistical simulation for coded high-speed links, and provides a case against the use of biased Monte Carlo methods in this setting. Finally, based on a hardware test bed, the performance of standard binary forward error correction and error detection schemes is evaluated, from which recommendations on coding for high-speed links are derived.

Index Terms—Communication systems, integrated circuit interconnections, intersymbol interference (ISI).

I. INTRODUCTION

THE practice of constraining the data stream in order to mitigate the effects of the communication channel on the received signal, commonly referred to as channel coding, is a fundamental technique in digital communications that is responsible for some of the most dramatic improvements in the modern communication standards. While channel coding is employed in most of today's communication systems, both wireless and wireline, in order to improve on the speed/reliability/energy efficiency of the system, the technique remains unexploited in a ubiquitous class of communication systems, namely the high-speed backplane and chip-to-chip interconnects. More than just a question of unharvested potential, the increasing network speeds place a large burden on high-speed links, which fail to keep up with the scaling trends. The under-

lying problem is the bandwidth-limited nature of the backplane communication channel, exacerbated by severe complexity and power constraints.

Despite several recent efforts [1], [2], the topic of channel coding for high-speed links remains largely unexplored due to a lack of suitable analysis and simulation frameworks. The residual intersymbol interference (ISI), coupled with noise and other circuit impairments, significantly obscures the performance picture and renders both the theoretical and computational approaches more arduous. Specifically, the channel memory introduces error correlation in the received symbol stream, regardless of whether the latter is constrained or unconstrained. The code performance is determined by the joint symbol error statistics and, as the task of accurately accounting for the error correlation due to channel memory is combinatorial in nature, exact expressions are computationally intractable. The problem of estimating the performance of a coded high-speed link is further exacerbated by the low error rates of interest ($\sim 10^{-15}$), which render direct Monte Carlo simulation prohibitive and strain the accuracy of common approximations.

This paper provides a more systematic look at the potential of bringing energy-efficient channel coding to high-speed links. Modeling the high-speed link as a system with additive white noise and ISI, as described in Section II, makes it possible to describe the error correlation in terms of two fundamental quantities: the systems's error region and the channel's sign signature. The error region corresponds to the set of values in the ISI distribution that are responsible for the majority of errors. While the error region is determined by the combined noise and the magnitude of the coefficients forming the channel's pulse response, the channel signature is specified by the signs of those coefficients. The analysis of Section III shows how these quantities conceptually decouple the complex problem of accounting for the effect of a real-valued channel on error behavior and provide a missing insight into error correlation in a high-speed link.

While current statistical simulation techniques for high-speed links ignore error correlation between symbols, Section IV shows that a direct extension of the independent-errors approximation improves the estimate's accuracy by up to five orders of magnitude for the error rates of interest. The approach exploits the physical properties of high-speed links, particularly the nature of the ISI, which limits the range of the error correlation. It relies on accurately capturing the short-term error correlation within nonoverlapping blocks of symbols and assumes independence in the error behavior across distinct blocks, rather than individual symbols. The computational mechanics are

N. Blitvic and V. Stojanovic are with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: blitvic@mit.edu; vlada@mit.edu).

M. Lee is with the Kramer Levin Naftalis & Frankel LLP, New York, New York 10036 USA (e-mail: mlee@kramerlevin.com).
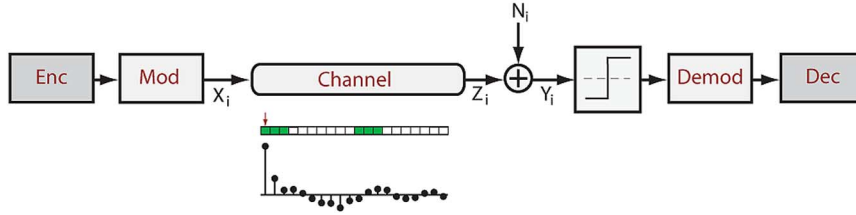
Fig. 1. Simplified model of a high-speed link. Transmit/receive equalization is reflected on the symbol-spaced pulse response.

those of integer partitions and the approach is computationally efficient for high-speed link channels. The proposed analysis and simulation frameworks also present a realistic case cautioning against the use of biased Monte Carlo methods in the performance estimation of coded high-speed links.

Finally, the case studies of Section V evaluate the performance of several binary forward error correction and error detection codes in a high-speed link. The results are based on an experimental setup with a Virtex-II Pro X FPGA and a 20-in Rogers link. They further illustrate the effect of error correlation on the code performance and provide more insight into the tradeoff between error-correcting capability and coding overhead over the bandwidth-limited high-speed link channel.

II. SYSTEM MODEL

A simplified model of a high-speed link is shown in Fig. 1. The bitstream, which can be coded or uncoded (unconstrained), is modulated to produce the equivalent symbol stream and transmitted over a communication channel. The system employs PAM2 modulation with detection performed on a symbol-by-symbol basis with the decision threshold at the origin. The transmitter and receiver may contain equalizers, in which case the channel's impulse response may contain residual ISI. The two main mechanisms that account for the most significant portion of the residual ISI in high-speed links are dispersion and reflection. In addition, residual interference may also include co-channel interference, caused, for instance, by electro-magnetic coupling (crosstalk) [3], [4]. As accounting for co-channel interference involves the same set of mathematical tools as accounting for the ISI, the remainder of the paper focuses on the effects of the ISI.

The quantity of interest is the received signal at the input to the decision circuit at time i, denoted Y_i and expressed as

$$Y_i = Z_i + N_i \qquad (1)$$

where Z_i denotes the received signal in the absence of noise and N_i is the noise term. Specifically, denoting the channel's pulse response by $h_{-k}, \ldots, h_{-1}, h_0, \ldots, h_m$, where $l = k + m + 1$ represents the length of the pulse response and h_0 is associated with the principal signal component, and letting $\{X_i\}$ denote a sequence of transmitted symbols, then

$$Z_i = \sum_{j=-k}^{m} X_{i-j} h_j. \qquad (2)$$

The noise term, representing the combined thermal noise and timing jitter, is assumed to be Gaussian with the standard deviation of $\sigma = 3\,\mathrm{mV}$ relative to the X_i peak values of ± 1 V [5].

III. ERROR BEHAVIOR IN SYSTEMS WITH NOISE AND ISI

In the system of Fig. 1, an error at the receiver occurs if the noise and ISI couple to bring the received signal over the decision threshold. It follows that the marginal symbol error probability for the ith symbol is given by

$$p_i = \mathrm{P}\left(\{Y_i < 0 | X_i = 1\} \cup \{Y_i > 0 | X_i = -1\}\right). \qquad (3)$$

Assuming an unconstrained symbol stream, the marginal symbol error probabilities are equal, that is $p_i = p_j = p$ for all symbols i, j. The quantity p, which becomes the relevant figure of merit, is entirely determined by two factors, namely the channel pulse response $h_{-k}, \ldots, h_{-1}, h_0, \ldots, h_m$, and the probability distribution of the noise. Efficient methods of computing the marginal error probability p are described in [6], [7] among others.

In a coded system with ISI, this picture changes in two important ways. Due to constraints on the symbol stream, the marginal error probabilities are no longer equal across different symbols. An efficient method of computing p_i for different symbol locations in a codeword is described in [8], which focuses on systematic binary linear block codes. However, the performance of a coded system cannot be expressed through marginal error statistics alone, but is instead dependent on the joint error behavior. For instance, the performance of a t-error correcting linear block code is typically expressed through the word error rate (WER), given by the probability of observing at least $t + 1$ errors in a codeword.

The following development shows that the complex relation between the ISI and the joint error behavior can be greatly elucidated by decoupling the effects of the magnitude and the signs of the channel's pulse response. Understanding the effect of system's *error region* and channel's *sign signature* on error correlation lends a deeper insight into the behavior of codes and the shortcomings of common simulation techniques in high-speed links. Further, an analysis of *correlation distance* in high-speed links paves the way for a more reliable simulation approach.

A. Error Region, Channel Signature and Error Correlation

The *error region* for a particular system should be thought of as the set $\mathcal{E}_f$ of ISI values that are responsible for the majority of errors at the receiver. More formally, letting $0 < f \leq 1$ denote a lower bound on the proportion of symbol errors ascribed to

the error region (e.g., 99% of all errors), the error region is the smallest possible interval of the form $\mathcal{E}_f = (-\infty, v]$ such that

$$P(Z_i \in \mathcal{E}_f | E_i, X_i = 1) \geq f \qquad (4)$$

where E_i denotes the error event on the ith symbol. The concept of error region is primarily useful when it is possible to pick f large enough so that the ISI events in $\mathcal{E}_f$ are responsible for the majority of the errors, and at the same time small enough so that every event in $\mathcal{E}_f$ is likely to cause an error.[1] For the system of Fig. 1 and some given f, the error region is entirely determined by the decision threshold, the noise standard deviation σ and the channel pulse response $h_{-k}, \ldots, h_{-1}, h_0, \ldots, h_m$, where $k + m + 1 = l$. Note that for an unconstrained symbol stream, only the *magnitudes* of the pulse response coefficients are needed. The signs of the coefficients are captured by the *channel signature* $\mathbf{s} = (s_{-k}, \ldots, s_0, \ldots, s_m)$, which takes values from the set $\{-1, 0, 1\}^l$ and is defined as

$$\mathbf{s} = (\text{sign}(h_{-k}), \ldots, \text{sign}(h_{-1}), \text{sign}(h_0), \ldots, \text{sign}(h_m)). \qquad (5)$$

To elucidate the link between the error region, channel signature and error correlation, first consider the case where the error region is limited to the worst-case ISI only, that is, where other ISI events cause error with negligible probability. The corresponding regime is referred to as *worst-case-dominant* in [9]. Symbol $X_i = 1$ is affected by worst-case ISI when $(X_{i+k}, \ldots, X_i, \ldots, X_{i-m}) = \mathbf{p}$, where

$$\mathbf{p} = (-s_{-k}, \ldots, -s_{-1}, 1, -s_1, \ldots, -s_m) \qquad (6)$$

is referred to as the channel's *worst-case pattern*.[2] It follows that the joint error statistics are entirely determined by the "nesting" properties of the channel's worst-case patterns, which are in turn determined by the channel's signature. The pattern $\mathbf{p}$ nests if $\mathbf{p}$ and some shifted versions of $\pm\mathbf{p}$, each shifted by at most $l - 1$ symbols relative to the first pattern, align exactly on the overlapping symbols. For instance, the pattern $\mathbf{p} = 1\ 1\text{-}1\text{-}1\ 1\ 1$ nests simultaneously in three locations, as shown below

$$\begin{aligned} 1\ 1\text{-}1\text{-}1\ 1\ 1 \quad &= \mathbf{p} \\ \text{-}1\text{-}1\ 1\ 1\ 1\text{-}1\text{-}1 \quad &= -\mathbf{p} \\ 1\ 1\text{-}1\text{-}1\ 1\ 1 \quad &= \mathbf{p} \end{aligned}$$

Suppose that the channel is such that the worst-case sequences $\pm\mathbf{p}$ nest simultaneously at several locations. Then, among the symbol sequences that are likely to cause an error on at least one symbol, that is, sequences containing $\pm\mathbf{p}$, a relatively large proportion will also contain additional nestings of $\pm\mathbf{p}$. It follows that the error-prone sequences likely cause errors on more than one symbol and the result is generally an increase in the frequency of certain higher-order error events[3]

---

[1]Trivially, for any system, $\mathcal{E}_f = (-\infty, \infty)$ with $f = 1$, but this is not an insightful choice.

[2]When $\mathcal{E}_f$ reduces to the worst-case ISI, the corresponding parameter f becomes the a posteriori probability, conditioned on $X_i = 1$ and on the error event E_i, of symbols $X_i, \ldots, X_{i-l+1}$ forming the worst-case pattern $\mathbf{p}$, i.e. $f = P((X_{i+k}, \ldots, X_i, \ldots, X_{i-m}) = \mathbf{p}|E_i, X_i = 1)$. For noise of infinite support, the regime of interest occurs in the limit $f \to 1$.

[3]In the context of hard-decision decoding, the events of interest correspond to observing k errors in a block of n symbols, for $k = 0, 1, \ldots, n$.
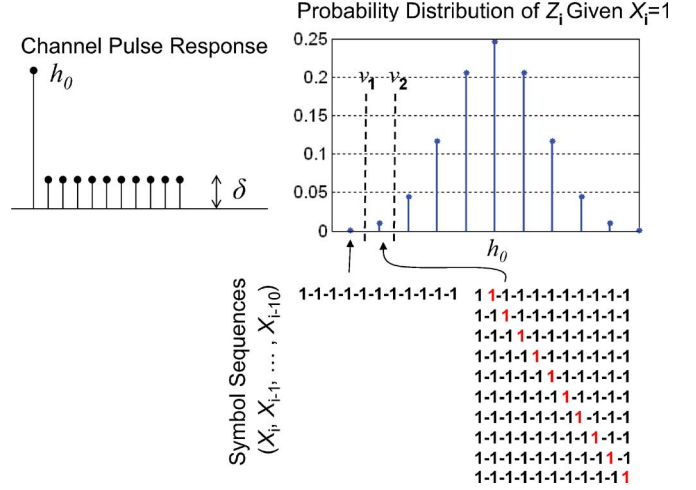

Fig. 2.   Simplified pulse response [left] and the corresponding ISI distribution [right].

compared to the assumption that errors occur independently across distinct symbols. The maximal degree of error correlation is incurred by channels whose worst-case patterns nest fully, that is, simultaneously on every symbol. One example is $\mathbf{p} = 1\ 1\ 1\ 1 \ldots 1$, which has the property that any symbol sequence with a runlength of $2l$ causes worst-case ISI to occur on at least $l + 1$ symbols. Another maximally-nested worst-case pattern is the alternating pattern $\mathbf{p} = 1\text{-}1\ 1\text{-}1 \ldots 1\text{-}1$. These examples become particularly relevant in realistic channels containing only a handful of strong interference coefficients, as further discussed in Section III-B.

To illustrate the effect of extending the error region beyond the worst-case ISI, consider the equal-magnitude, all-positive channel of Fig. 2. The corresponding worst-case sequence, given by $\mathbf{p} = 1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1$, cannot be nested, implying that it is impossible for two symbols to both be affected by the worst-case ISI unless separated by at least $l - 1$ symbols. Suppose that for some sufficiently large f, $\mathcal{E}_f = (-\infty, v_1]$, where v_1 happens to be such that only the worst-case ISI causes significant error, as illustrated. Compared to the assumption that the ISI affects distinct symbols independently, the higher-order error events become significantly less likely. Consider next the case where the system parameters (e.g. noise, threshold margin) are changed so that $\mathcal{E}_f = (-\infty, v_2]$, where $v_2 > v_1$ as shown in the figure. By allowing the error-prone sequences to deviate from $\mathbf{p}$ by one symbol, it becomes possible to nest two error-prone sequences, as illustrated below

$$\begin{aligned} &1\ 1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1 \\ &1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1 \end{aligned}$$

In particular, focusing on symbol X_i, let A_i be the event where the surrounding symbols are given by the error-prone pattern $X_i, \ldots, X_{i-l+1} = 1\ 1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1\text{-}1$. As shown above, if A_i occurs, then not only does Z_i belong to $\mathcal{E}_f$, but so does Z_{i-1} as well, regardless of the remaining symbol X_{i-l}. As a result, if $P(E_i | A_i)$ is sufficiently large compared to the marginal error probability $P(E_i)$, this observation alone is sufficient to guarantee that the proportion of codewords with two errors is

increased compared to the independent-errors approximation.[4] In a similar manner, widening the error region to include deviations from the worst-case pattern by more than two symbols increases the probabilities of the corresponding higher-order error events.

The above example provides a first illustration of the shortcomings of common simulation methods when applied to coded high-speed links. While widening the error region by increasing the noise variance or introducing a mean-shift inflates the error probability, thus making it possible to capture the behavior of the system by Monte Carlo simulation, the behavior at high error probabilities is not necessarily indicative of the system's behavior at low error probabilities. As changes in the error region can alter the error correlation, such methods are unsuitable for the performance estimation of coded systems, whose performance is highly dependent on the joint error statistics. This is further discussed in Section IV-B.

For any channel, the joint error behavior is fully determined by the noise and the nesting properties of the error-prone sequences. For a channel with interference coefficients of equal magnitude, such as that of the previous example, the error-prone sequences take on a particularly simple form. Specifically, all deviations by k symbols from the worst-case patterns $\pm\mathbf{p}$ cause the same ISI. The problem then decouples into considering the error region $\mathcal{E}_f$, which specifies the maximal deviation k that may be responsible for an error, and the channel signature $\mathbf{s}$, which determines the patterns $\pm\mathbf{p}$ and thus controls the underlying nesting properties. For realistic channels whose interference coefficients do not have equal magnitudes, the problem does not decouple fully and numerical simulation is needed to accurately predict joint error behaviors. However, the previous discussion provides insight into the error behavior of high-speed links.

B. Error Correlation in High-Speed Links

In practical channels, the direct link between the channel signature, error region and joint error behavior holds only asymptotically. For instance, in the worst-case-dominant conditions, the problem reduces to the nesting properties of $\mathbf{p}$, while in the limit of large noise, the effect of any channel correlation vanishes as errors become independent. However, both the channel signature and the error region play an important role in determining the joint error statistics. An illustration of the effect of

[4]As a numerical example, following the notation of Fig. 2, let $n = 11$, $h_0 = 1$, and $\sigma = 10^{-3}$ (noise rms). For a decision threshold at the origin, let $\mathcal{E}_f = (-\infty, 0]$ with $f \geq 1 - 10^{-13}$. Two choices of magnitudes for the interference coefficients, namely $\delta_1 = 0.120$ and $\delta_2 = 0.125$, induce two different error behaviors. For δ_1, the error region is limited to the worst-case ISI. Based on the marginal error probability $p_1 = 9.8 \times 10^{-4}$, the proportion of 10-symbol blocks with exactly two errors is given by $\binom{10}{2}p_1^2(1 - p_1)^8 = 4.3 \times 10^{-5}$. However, accurately calculating the joint error statistics, by enumerating all possible symbol sequences affecting a 10-symbol block, shows that the probability of double-error occurrences is *zero*, calculated to the limits of available numerical precision. In contrast, letting $\delta_2 = 0.125$, the error region includes second-to-worst-case ISI. Based on the marginal error probability of $p_2 = 5.9 \times 10^{-3}$ and assuming independent errors, the desired event's probability is given by $\binom{10}{2}p_2^2(1 - p_2)^8 = 1.5 \times 10^{-3}$. Yet, the proportion of 10-symbol blocks with two errors calculated accurately is now five times *higher*, namely 7.3×10^{-3}. Thus, in effectively extending the error region to include deviations from the worst-case ISI, the proportion of symbol blocks with two errors transitions from being significantly below the prediction to exceeding it.
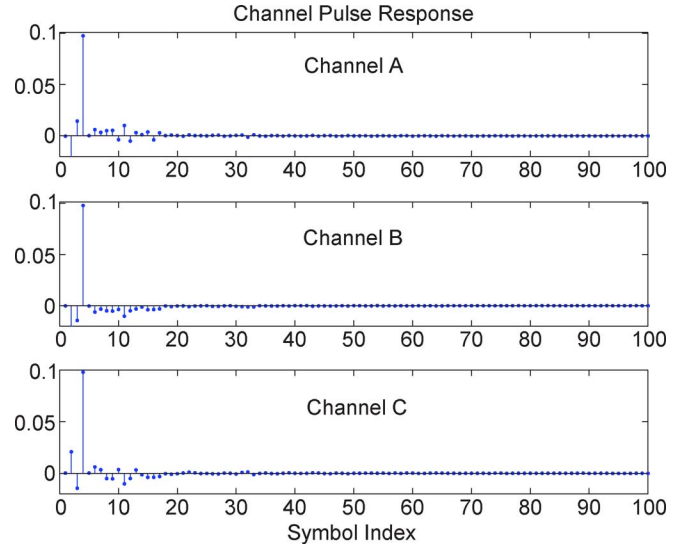

Fig. 3. Equalized pulse response of the standard 802.3ap B32 [10] channel operating at 10 Gb/s. Channel A: Actual channel. Channel B: All-positive signature. Channel C: Randomly generated signature.
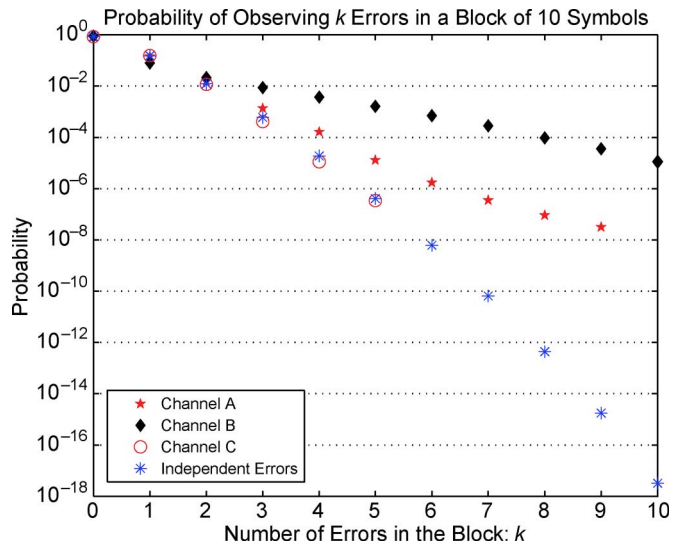

Fig. 4. Joint error statistics for the channels of Fig. 3. Margining by 36 mV increases the error rate to 1.78×10^{-2}, rendering the higher-order error events observable through Monte Carlo simulation (2.5×10^7 symbol blocks). The figure illustrates the extent to which channel signature affects the joint error behavior of the system.

channel signature on error correlation in a realistic high-speed link is shown in Figs. 3 and 4. The communication channel of Fig. 3(a) is a standard 802.3ap B32 [10] channel operating at 10 Gb/s. Single tap decision-feedback and three-tap zero-forcing equalizers are used, yielding an error rate of 2.5×10^{-6}. The channels of Fig. 3(b) and (c) are obtained by altering the signature of the original channel, which preserves the marginal error behavior, but alters the joint error statistics. Margining by 36 mV widens the error region, thus increasing the error rate to 1.78×10^{-2} and rendering the higher-order error events observable through Monte Carlo simulation. The resulting joint error statistics are shown in Fig. 4 for a block of ten consecutive symbols.

Unlike the maximally-correlated channel (Channel B), the original channel (Channel A) does not nest in the sense described in the previous section, neither considering the worst-case patterns formed by the entire channel pulse response, nor those formed by the dominant interference coefficients only. However, both channels show a significant increase in the frequency of the higher-order error events compared to the independent-errors assumption.[5] This behavior is analogous to the example of Fig. 2 and is due to the size of the error region, which is made sufficiently large to generate Monte Carlo estimates. Due to the limitations of the Monte Carlo method, it is difficult to infer the degree and type of error correlation associated with the randomly generated signature (Channel C).[6]

The presence of a handful of strong interference coefficients in the previous example is due to the dispersive nature of the high-speed link channel and the presence of signal reflections. In general, the pulse response of a typical high-speed link can contain several clusters of strong interference coefficients, separated by coefficients of significantly weaker magnitudes. This suggests that a viable method of deriving intuition about the error behavior from the channel pulse response consists of considering the error correlation caused by these dominant interference coefficients separately from that caused by the rest of the channel.

In particular, it is interesting to consider the benefit of focusing on the *largest* interference coefficients so that the error hinges on the occurrence of the resulting worst-case ISI. The relevant error region becomes associated with the worst-case ISI computed relative to the largest interference coefficients, and includes any amount of deviation caused by the remainder of the channel's pulse response. Identifying the worst-case-dominant conditions in a high-speed link is primarily useful from the standpoint of pattern elimination, that is, the use of code constraints to prohibit error-causing symbol patterns in a high-speed link. This idea is further explored in [9]. The worst-case-dominant conditions are, however, of limited use in improving the performance estimation of high-speed links. In particular, while accurately accounting for the effect of the worst-case ISI on joint error behavior is computationally tractable, the error correlation due to secondary interference coefficients can still be large. This is illustrated in Fig. 5 for the high-speed link channel of Fig. 3(a). For a noise level of $\sigma = 3$ mV, considering as dominant all interference coefficients greater than σ, 2σ, 3σ, and 4σ yields the corresponding error-region parameters of $f = 7.7 \times 10^{-3}$, $f = 0.82$, $f = 0.97$, and $f = 1 - 2.3 \times 10^{-3}$, respectively. The most accurate simulation results were obtained for the cutoffs of 3σ and 4σ. The former case yields the error statistics displayed in the figure, which shows the proportion of codewords with a given number of errors in a block of ten symbols. The results are compared to
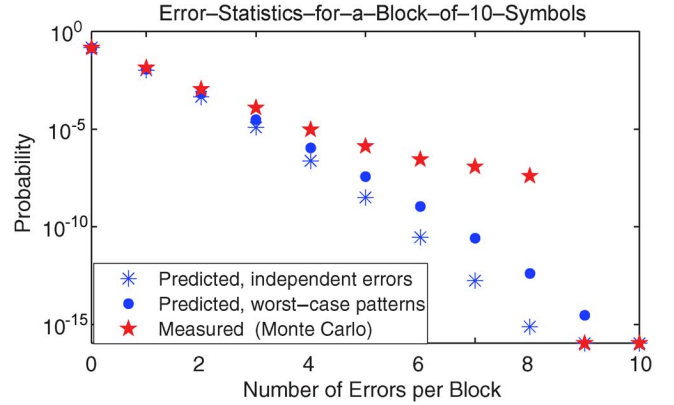

Fig. 5. Error statistics for the channel of Fig. 3(a) margined by 36 mV. The dominant interference coefficients are chosen based on the cutoff of 3σ, yielding $f = 0.97$. Taking into account error correlation due to worst-case patterns improves on the joint statistics computed through the independent-errors assumption. However, by allowing the tail of the pulse response to contain large interference coefficients in order to achieve sufficiently high f, the approach fails in capturing an adequate level of error correlation, as evidenced by the deviation from the corresponding Monte Carlo estimate (2.5×10^7 symbol blocks).

a commonly-employed estimate based on the assumption that errors on distinct symbols are independent, as well as the actual Monte Carlo measurement. While accounting for the error correlation caused by the worst-case patterns improves the error estimate by an order of magnitude relative to the independent-errors approximation, the approach still falls short of accurately predicting the error statistics for this channel.

C. Correlation Distance in High-Speed Links

While the previous sections focus on the effect of the channel signature and error region on the joint error behaviors in a high-speed link, more can be said by considering also the spatial properties of the error correlation embedded in the shape of the channel's pulse response. In particular, it becomes relevant to ask at what time separation can any two symbols effectively become independent. Due to the finite length of the channel's pulse response, the dependence between received signals Y_i and Y_j (and similarly, the error events E_i and E_j) in an unconstrained symbol stream decreases with increasing distance $|i - j|$, to vanish when $|i - j| \geq l$ for received signals (and $|i - j| \geq l - 1$ for errors). More significant, however, is the fact that the principal sources of the ISI in a high-speed link are signal reflections, which attenuate quickly with time and are therefore less pronounced further in the tail of channel's pulse response. It follows that the error correlation typically decays much more rapidly than the previous bound predicts. Thus, despite the pulse response generally spanning hundreds of symbols, the error correlation in a high-speed link is of short-term nature. Continuing with the case study of the high-speed link of Fig. 3(a), the error correlation is quantified based on the measurement of error lengths, that is the maximum observed distances between any two errors in a codeword. The error lengths for a block of 40 symbols containing exactly two errors are shown in Fig. 6.

Taking into account statistical significance, the largest degree of error correlation, measured by the amount of deviation from the independent-errors behavior, is contained within a length of approximately eight to ten symbols. The shape of the error-

---

[5]In that case, the corresponding error statistics are given by the binomial distribution, that is, $\mathrm{P}(k \text{ errors out of } n \text{ symbols}) = \binom{n}{k} p^k (1-p)^{n-k}$, where p is the marginal symbol error probability

[6]Note that the link between the channel signature and error correlation points to nonstandard approaches to energy-efficient equalization in high-speed links. Specifically, understanding the effects of common equalization methods on the channel's sign signature provides means of decreasing the frequency of higher-order error events, thus enhancing the performance of standard error-control codes. This approach requires further exploration.
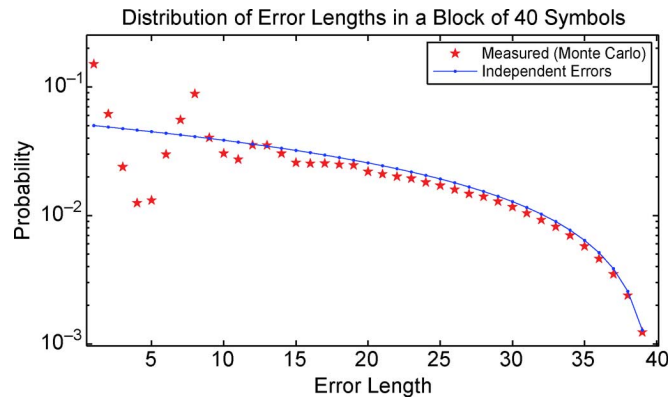
Fig. 6. Distribution of error lengths in a block of 40 symbols with two symbol errors, for the channel of Fig. 3(a) with a threshold bias of 36 mV. The distribution of error lengths is obtained through Monte Carlo simulation (2.5×10^7 symbol blocks) and normalized relative to the total proportion of blocks with two errors. A comparison with the error-length distribution computed assuming error independence suggests that most of the error correlation takes place within relatively short distances.

length distribution is determined by the nesting properties of the error-prone sequences.[7] The short-term nature of the error correlation in high-speed links is at the core of the simulation approach proposed in the following section.

IV. STATISTICAL SIMULATION FOR CODED HIGH-SPEED LINKS

For an arbitrary system with ISI, fully specifying the joint error probabilities for all symbols in a codeword is computationally intractable due to the size of the resulting state space. The same holds even if the code constraints on the symbol stream are ignored and the codeword is replaced by a block of consecutive, independently transmitted symbols. For this reason, it is common practice to account for the ISI only through marginal error statistics while discounting its effect on error correlation. However, considering error events on distinct symbols to be independent frequently yields large inaccuracies in the performance estimate, as demonstrated in the previous sections.

Rather than ignoring the error correlation altogether, accounting for short-term correlation within a codeword is both computationally tractable and yields superior performance estimates. A simple statistical method of estimating the effect of short term error correlation is described below. The method also provides direct means of trading off computational requirements for accuracy and enables computationally tractable code-space explorations. A set of numerical examples illustrates the proposed simulation approach and completes the previous discussion of the effect of error region and channel signature on error correlation. The results point to the inadequacy of biased Monte Carlo techniques in accurate high-speed link simulation. In particular, it is shown that the joint error behavior of a specific high-speed link channel at high error rates need not be indicative of its behavior at low error rates. Thus, without an adequate method of "unbiasing" the performance estimate, biased Monte Carlo techniques should not be used for the accurate simulation of coded high-speed links.

A. Proposed Simulation Method

Based on the physical properties of high-speed links, the previous section develops the motivation for focusing on short-term error correlation in simulation of coded high-speed links. While the independent-errors assumption is by default incapable of capturing any error correlation, the following simple extension provides means of capturing varying degrees of short-term error correlation and thus drastically improves the accuracy of the joint error estimates. The approach consists of subdividing a codeword into non-overlapping blocks of consecutive symbols, accurately computing the error statistics for each block, and combining the results assuming the errors across distinct blocks to be independent. Effectively, this replaces the "independent errors" approximation by the "independent blocks" approximation. Although transmitted symbols in separate blocks need not be independent in a coded symbol stream, as the blocks form parts of a larger codeword, [9] shows that it is relatively difficult for a code to achieve consistent pattern-eliminating properties. It follows that the underlying symbol constraints in a coded system likely have little direct effect on the marginal and joint error statistics prior to decoding.[8] However, more significant inaccuracies may arise from the error behavior at the boundaries between blocks, as at least one symbol in each block is affected by the ISI from the preceding blocks. The quality of this approximation for a given codeword length improves with a decreasing number of blocks, which yields a direct method of trading off computing speed for accuracy.

To accurately compute the joint error statistics in each block, it is convenient to shorten the channel pulse response by removing the portion of the response tail that creates negligible error correlation. The effect of the tail ISI is treated as mean-distortion and added to the noise term.

Based on the above approximations, the performance of a coded high-speed link is estimated as follows. Subdivide the codeword of length n into k blocks of lengths $\tilde{n}_1, \tilde{n}_2, \ldots, \tilde{n}_k$, where the number of blocks and the corresponding block lengths are chosen based on implementation convenience. For each of the blocks, the error statistics can be accurately computed by considering the possible symbol patterns that affect the corresponding received symbols. Specifically, for the jth block, it suffices to enumerate all possible symbol patterns of length $\tilde{l} + \tilde{n}_j - 1$, where $\tilde{l}$ denotes the length of the shortened channel's pulse response, and compute the ISI affecting each symbol in the block. For an uncoded symbol stream, the number of underlying symbol patterns equals $2^{\tilde{l}+\tilde{n}_j-1}$. Then, the probability of observing i errors in the jth block, denoted by $\tilde{p}^{(j)}(i)$, is computed considering the possible error patterns and taking into account the noise.

Given the partial error statistics $\tilde{p}^{(j)}(i)$ for all $i = 0, 1, \ldots, n_j$ and $j = 1, \ldots, k$, it remains to compute p_m, that is, the total probability of observing m errors in a codeword of n symbols, where $m = 0, 1, \ldots, n$. This is achieved by considering all compositions of m into k parts, that is, the possible vectors $(m_1, \ldots, m_k)$, where $m = m_1 + \ldots + m_k$ and $0 \leq m_1, m_2, \ldots, m_k \leq k$. The number of possible compositions

---

[7]Specifically, if no two error-prone sequences can nest at some given distance d, the probability of observing an error length of d for a codeword with two errors is drastically reduced compared to the independent-errors case.

[8]The power of typical error-control codes resides in their postprocessing capabilities, in the form of error detection/correction.
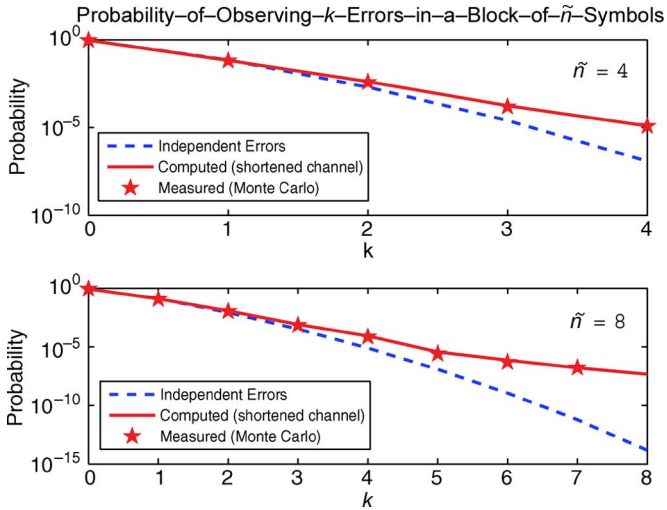
Fig. 7.   Error statistics for blocks of 4 and 8 symbols for the channel of Fig. 3(a) with a threshold bias of 36 mV. The results computed analytically based on a shortened channel with $\tilde{l} = 16$ by considering all $2^{\tilde{n}+\tilde{l}-1}$ symbol patterns match closely the Monte Carlo measurements (2.5×10^7 symbol blocks) for the actual (unshortened) channel. The corresponding result based on the independent-errors assumption is also included.
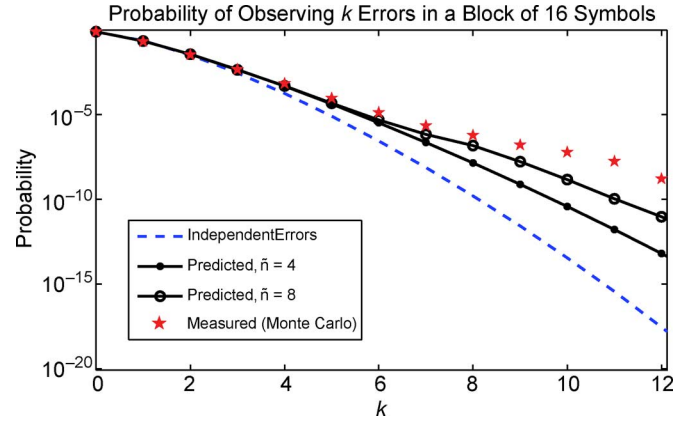

Fig. 8.   Error statistics for a codeword of 16 symbols for the channel of Fig. 3(a) with a threshold bias of 36 mV. The results computed analytically based on the independent-blocks assumption (block lengths $\tilde{n}$ of 4 and 8 symbols) are compared to the corresponding results based on the independent-errors assumption, as well as the actual Monte Carlo measurements (2.5×10^9 symbol blocks).

is given by $\binom{m+k-1}{m}$ and the corresponding probability p_m is given by

$$p_m = \sum_{\substack{\text{all compositions} \\ (m_1,\dots,m_k)}} \tilde{p}^{(1)}(m_1)\tilde{p}^{(2)}(m_2)\dots\tilde{p}^{(k)}(m_k). \quad (7)$$

In addition, if the error statistics are equal across the k blocks, which occurs if $\tilde{n}_1 = \tilde{n}_2 = \dots = \tilde{n}_k$ and the symbol stream is considered to be unconstrained, the ordering among the k blocks does not need to be considered. The problem therefore reduces to dealing with a partition of m into k integers.[9] Finally, given the codeword error statistics $p_0, p_1, \dots, p_n$, the performance of a t-error-correcting code, for example, is given by $\text{WER}_t = 1 - (p_0 + \dots + p_t)$.

For the high-speed link channel of Fig. 3(a), letting $\tilde{l} = 16$ yields the block error statistics $\tilde{p}(i)$ of Fig. 7 for blocks of $\tilde{n} = 4$ [top] and $\tilde{n} = 8$ [bottom] symbols. As expected, the results match closely the statistics captured through Monte Carlo simulation. Applying the proposed simulation technique to a codeword of $n = 16$ symbols and discounting the code constraints results in performance estimates of Fig. 8. At error rates of interest, the estimates based on block sizes of $\tilde{n} = 4$ and $\tilde{n} = 8$ symbols yield improvements of four and six orders of magnitude, respectively, over the independent-errors approximation. Though both estimates still fall short of capturing the full extent of error correlation for a system operating under these conditions, the proposed estimation method provides a simple and powerful alternative to the independent-errors approximation.

Finally, for practical codeword lengths, the computational complexity of the above method is determined by the shortened channel length $\tilde{l}$, as the number of possible symbol patterns of

length $\tilde{l} + \tilde{n}_i - 1$ is typically large. Based on the block size $\tilde{n} = 4$, the runtime for the previous example is on the order of one minute on a 1.8 GHz processor with 2 GB of memory, for codeword lengths of up to 100 symbols. Precomputing the block-wise $\tilde{p}(i)$ statistics further reduces the runtime and allows for the use of larger blocks in systematic code-space explorations.

B. Case Against Biased Monte Carlo Techniques

While Monte Carlo simulation is the standard tool for estimating the performance of complex systems, its use in high-speed links is limited due to the extremely low error probabilities in the operating regimes of interest. A potential workaround consists of biasing the system parameters to widen the error region, increase the error probability and consequently reduce the sample size required for an accurate estimate. The difficulty with this approach lies in the fact that the resulting estimate needs to be subsequently unbiased in order to represent the quantity of interest in the actual system. A common framework for biased Monte Carlo simulation is importance sampling [14]. In systems with ISI, the dimensionality of the problem is known to significantly inhibit the sample-size reduction potential of biased Monte Carlo techniques. Moreover, the ISI increases the complexity of the biasing/unbiasing procedures, as some form of joint statistics becomes required to partially describe the system's behavior.

Instead, in high-speed links, it is common practice to introduce a simple bias into system parameters such as noise or system margins and observe trends in the error behavior at low error rates. Rather than unbiasing the corresponding estimates, the resulting trends provide an idea of how a system may generally behave at low error rates. An immediate problem with this approach relates to the fact that the quality of common approximations is artificially improved when the error region is extended beyond the tails of the underlying probability distributions. For instance, the independent-errors approximation typically performs drastically worse at the error rates found in actual high-speed link systems, compared to those used in biased Monte Carlo simulations.

---

[9]MATLAB functions for counting the number of k-integer partitions and computing the corresponding partition tables are available in [11], [12]. An introductory treatment of compositions and integer partitions is available in [13], among others.
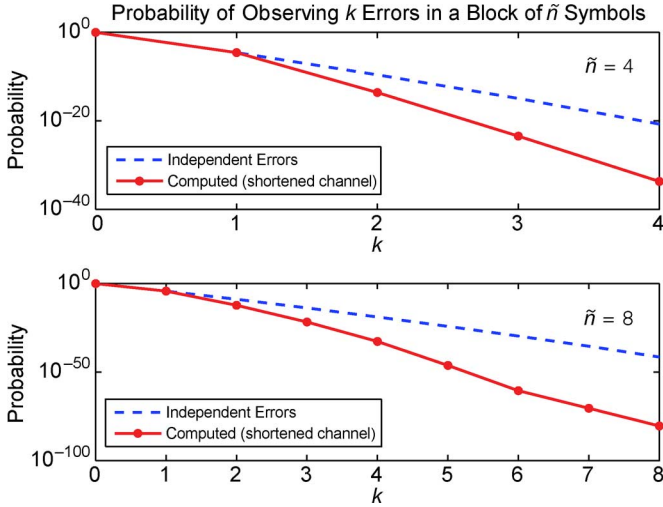
Fig. 9. Error statistics for blocks of 4 and 8 symbols for the channel of Fig. 3(a) without the threshold bias of 36 mV. The results computed analytically based on a shortened channel with $\bar{l} = 16$ by considering all $2^{\tilde{n}+\bar{l}-1}$ symbol patterns are compared to the results based on the independent-errors assumption.

TABLE I
UNCODED ERROR RATES ON AN EQUALIZED ROGERS LINK

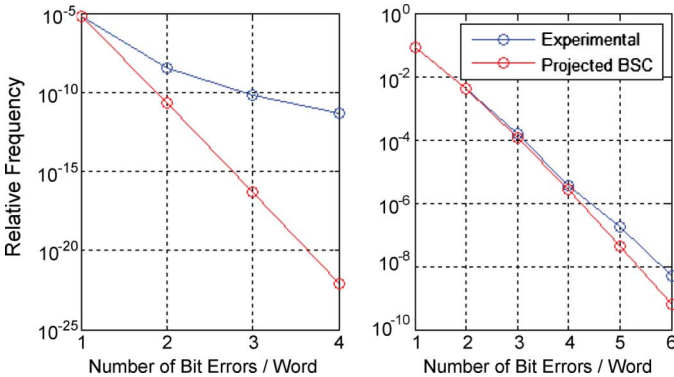| Data Rate (Gb/s)	Uncoded Bit Error Rate
5.00	3×10^{-12}
5.25	1×10^{-10}
5.50	8×10^{-10}
5.75	3×10^{-9}
6.00	2×10^{-7}
6.25	2×10^{-6}
6.50	1×10^{-4}
6.75	1×10^{-3}

Fig. 10. Testing the error independence hypothesis at signaling rates of 6 Gb/s [left] and 6.75 Gb/s [right]. For a codeword of 40 bits, the plots display the frequencies of the observed error weights, normalized with respect to the total number of received codewords of nonzero weight.

However, a much deeper issue stems from the fact that, as described in Section III, the change in the effective error region alters the error correlation and therefore the coded error statistics. Fig. 9 recomputes the statistics of Fig. 7, but without the threshold bias of 36 mV. As the error region now corresponds to the tail of the ISI distribution, the error-causing patterns are

close to the worst-case pattern. Noticing that the corresponding worst-case pattern, given by[10]

$$\mathbf{p} = 1\text{-}1\ 1\ 0\text{-}1\text{-}1\text{-}1\text{-}1\ 1\text{-}1\ 1\text{-}1\ 0\text{-}1\ 1\text{-}1$$

nests only at a few locations and only for large shifts, it follows that an occurrence of an error renders errors on surrounding symbols less likely. As expected, the higher-order error statistics fall below those predicted by the independent-errors assumption, contrary to the trend observed in the biased system of Fig. 7.

The ramifications of this observation for the high-speed link channel of Fig. 3(a) are that a simple error-control code can yield a large performance improvement, even though biased Monte Carlo simulation suggests otherwise. Based on the results of Section III, the reverse of this behavior can also occur for a different type of channel, which indicates that biased Monte Carlo simulation is an unreliable method of evaluating the performance of a given code over a *specific* high-speed link channel. Note, however, that biased Monte Carlo simulation remains a useful analysis tool for evaluating the performance of *hypothetical* link systems, as demonstrated in Sections III and IV where such methods illustrate the general effect of error region and channel signature on joint error behaviors.

V. HARDWARE IMPLEMENTATIONS OF COMMON CODES AND EXPERIMENTAL RESULTS

In addition to the proposed simulation framework, a set of hardware measurements provides a more complete characterization of the performance of common error-control codes in high-speed links. The underlying experiments, obtained as part of an earlier thesis [15], examine both the error statistics in an uncoded link and the performance of algebraic error control codes in a coded high-speed link. The latter focus on systematic binary linear block codes due to their inherent implementation simplicity.

The tests are performed on a 20-in Rogers backplane with no counterboring. The signal processing is performed on the Virtex-II Pro X FPGA. The RocketIO transceivers employ linear equalizers in the form of transmit pre-emphasis and receive high-pass filter. Equalization settings are explored exhaustively to yield the lowest uncoded error rates, shown in Table I for the signaling rates of interest. The PRBS has a period of $2^{31} - 1$ symbols. The following sections analyze the results of individual experiments.

1) Testing the Error Independence Hypothesis: Given a prior lack of suitable performance estimation methods for coded high-speed links, it is common practice to focus on marginal error statistics and ignore the effect of error correlation on code performance. For instance, in [1] and [2], the improvements achieved by the proposed codes are computed based on the assumption that the errors on distinct symbols are independent. Fig. 10 explores the validity of this assumption for a high-speed link operating at two different signaling rates, whose corresponding error rates are available in Table I. The

[10]The worst-case pattern shown corresponds to the significant part of the channel's pulse response, with the signs of the weak interference coefficients set to zero.
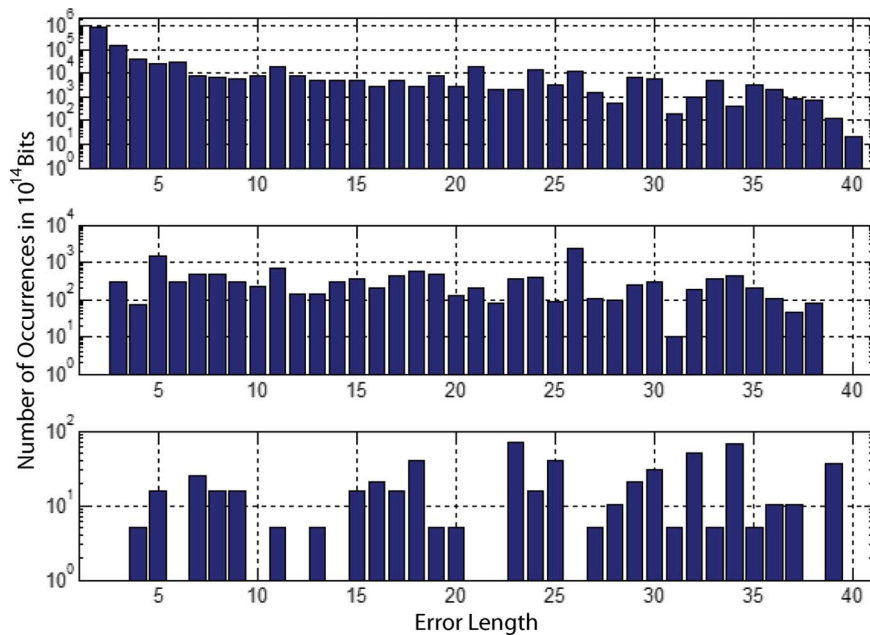
Fig. 11. Distribution of error lengths for different error weights. The error lengths of all 40-bit words with multiple errors is illustrated for a single test run, normalized to 10^{14} bits. The graphs indicate the error length distribution of error patterns of weight two [top], three [middle], and four [bottom].

figure displays the proportion of blocks of 40 symbols with k errors, where $k = 1, 2, \ldots, 40$, relative to the total number of blocks with at least one error. Note that the quantity k is frequently referred to as the *error weight* of a codeword. The results are compared to the statistics predicted for a channel where errors occur independently across symbols.

For the signaling rate of 6 Gb/s, corresponding to symbol error probability of 2×10^{-7}, assuming the error independence across symbols yields probabilities that differ from the measured results by several orders of magnitude. The measured behavior mirrors the error correlation observed in the simulation results of Fig. 4 and is indicative of considerable nesting in the system's error-prone sequences. However, compared to the earlier, simulated high-speed link, the significantly lower error rate of the realistic high-speed link exacerbates the error correlation. For instance, the proportion of codewords with exactly two errors differs from the predicted behavior by five orders of magnitude. The same degree of discrepancy is observed for the signaling rate of 5.5 Gb/s (not shown). From the point of view of coding over this particular high-speed link, the large, positive[11] error correlation indicates that codes with substantial error-correcting capabilities are required in order to ensure error rates below 10^{-12}.

While increasing the signaling rate to 6.75 Gb/s still leads to error correlation, the degree of correlation is lessened and the error-independence assumption yields relatively accurate results. The reduced error correlation is suggestive of a change in the channel's signature due to an increased data rate and different equalization settings. In addition, the dramatic increase in the corresponding error rate, now 1×10^{-3}, points to a significantly larger error region. Since the error expression no longer strongly depends on the tails of the ISI distribution, common approximations generally become more accurate.

2) Error Lengths in High-Speed Links: As discussed in Section III-C, the distribution of error lengths within a block of consecutive symbols provides spatial information on the error correlation in a given system. The measurements of observed error lengths for the high-speed link operating at 6.25 Gb/s are shown in Fig. 11. Taking into account statistical significance, the deviations from the independent-errors case in the measured distribution of error lengths predominantly occur at short lengths, further corroborating the insights of Section IV on the short-term nature of error correlation in high-speed links. The increase in the frequency of short error lengths is indicative of a high degree of nesting in the error-prone sequences.

In systems with strongly positive error correlation, burst-error correcting codes are frequently employed due to their ability to handle large error weights while requiring low overhead compared to standard error-correcting codes. However, due to the short-term nature of the error correlation in a high-speed link, such codes may not be appropriate. Specifically, the overhead of a burst-error correcting code is determined by the maximum error length that the code is required to handle. While the positive error correlation indicates that the errors are more likely to occur in clusters, the short-term nature of that correlation suggests that the behaviors of individual clusters are not significantly linked. It follows that the distribution of individual clusters across the codeword is relatively uniform, which implies that the distribution of error lengths in codewords with higher weights is approximately uniform as well, as evidenced by the plots of Fig. 11 for higher error weights. Based on these observations, burst-error correcting codes are not expected to perform well in high-speed links, which is further quantified in the following section.

3) Comparison of Algebraic Error-Control Codes in a High-Speed Link: The performance results for five different types of error-control codes, namely, Hamming codes em-
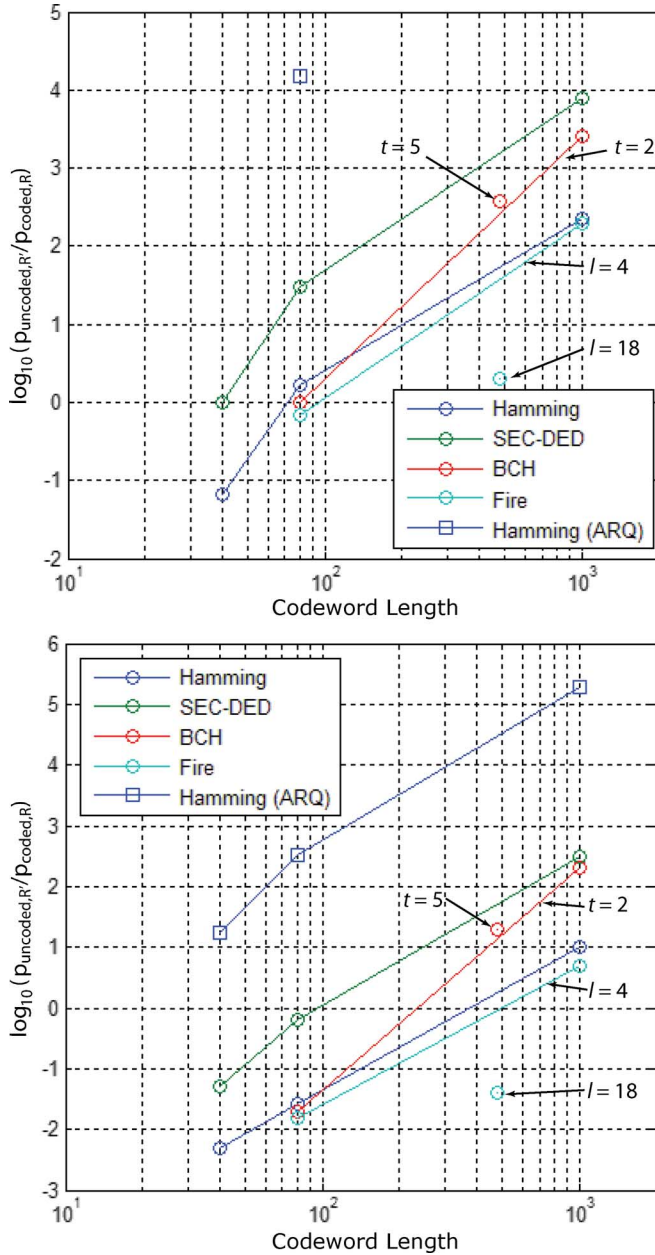
[11]Referring to an increase in the frequency of higher-order error events.

Fig. 12. Code performance results at 6 Gb/s [top] and 6.25 Gb/s [bottom].

ployed for either forward error correction or error detection with automatic repeat request (ARQ), Single-error-correcting Double-error-detecting (SEC-DED) with ARQ, Bose-Chaud-huri-Hocquenghem (BCH), and Fire codes, are shown in Fig. 12 at 6 Gb/s and 6.25 Gb/s. The values of the abscissa indicate the codeword lengths tested,[12] while the ordinate indicates the achieved improvement measured as $\log_{10}(p_{\text{uncoded,R}'}/p_{\text{coded,R}})$, where $p_{\text{coded,R}}$ is the bit error

[12]The block lengths are set according by the hardware constraints, specifically, by the range of clock frequencies available for the coded and uncoded data rates. For codes whose natural codeword length does not match the available codeword lengths, such as the Hamming codes whose codeword length is a power of two, code shortening is used. Code shortening reduces the codeword length, but preserves the error control capability of a code at the expense of a reduced code rate.

rate of the coded system operating at rate R and $p_{\text{uncoded,R}'}$ is that of the uncoded system, operating at the equivalent information rate $R' < R$. Note that additional tests for individual codes, as well as a deeper analysis of the coding performance as a function of the code rate, are available in [15].

Several general tendencies are apparent in the code performance measurements, starting with the fact that, for a fixed class of codes, the code performance invariably improves as the codeword length increases. This is a consequence of the bandwidth-limited nature of a high-speed link. For some fixed error-correcting capability, increasing the codeword length reduces the effective coding overhead. While the probability of observing multiple-bit errors increases with codeword length, and thus magnifies the probability of exceeding the code's error-correcting capability, the corresponding losses in performance do not overpower the gains in the reduced signaling rate. Note in addition that in systems where the errors are caused by specific symbol sequences, the probability of observing a certain number of errors in a codeword need not increase with codeword length at the same rate as it does in systems where the errors occur independently from one symbol to another.

Among the classical forward error-correcting codes, the more powerful BCH codes are preferable to Hamming codes, though the resulting reduction of roughly two orders of magnitude in the error rate falls short of their theoretical promise. The performance of the low-overhead $t = 2$ BCH codes matches that of the more powerful $t = 5$ BCH codes, a direct consequence of the bandwidth-limited nature of the communication channel. The Fire forward error correcting codes, designed to correct for error bursts, yield the weakest performance. While Fire codes are able to correct a much larger number of errors per codeword than both the Hamming and BCH codes, they only handle error lengths up to a certain maximum l. The large error lengths incurred by codewords with higher error weight significantly reduce the effectiveness of the low-overhead $l = 4$ Fire codes, while the bandwidth limitations inhibit any benefit provided by the more powerful $l = 18$ codes. It follows that in the realm of forward error correction, the results point to higher-order alphabets. Specifically, coding over a higher-order alphabet provides means of dealing with short-range error correlation. In a more pictorial way, the error-correcting capability of the code is no longer limited by the number of errors nor the error length, but by the number of "pockets" of errors that occur through the codeword, which more adequately corresponds to both the predicted and the observed error behaviors in a high-speed link.

The highest performance measured was that of standard codes employed for error detection, where a detection event is followed by a retransmission of the erroneous codeword. In particular, a Hamming code with an 80-symbol codeword, which generates no improvement when used in forward error correction, now yields over four orders of magnitude of improvement over the lower-rate uncoded system at 6 Gb/s and over two orders of magnitude at 6.25 Gb/s. When used for error detection, the BCH codes (not displayed) virtually eliminate all error.[13] The magnitude of improvement achieved by switching

[13]For the $t = 2$ BCH code, no errors went undetected in a total of 8.6×10^{12} 40-symbol codewords, and similarly for the larger codewords, at both data rates. Further details are available in [15].

from forward error correction to error detection and retransmission reflects the fact that, for some fixed overhead, constraining the set of possible codewords allows the decoder to accurately detect many more error events than those that can be reliably corrected.

Note that the tests measure the proportion of errors detected at the receiver, which equals the proportion of eliminated errors assuming successful retransmissions. As a retransmission mechanism and protocol, a variety of automatic repeat request (ARQ) schemes of varying complexity is available (see [16, Ch. 21]). Note that the ARQ schemes rely on the existence of a back-channel for the transmission of control information, realizable through forward common-mode signaling in a high-speed link [17].

4) ARQ Over Channels With ISI: In the context of evaluating the performance of ARQ in a high-speed link, the final experiment consists of isolating error-causing symbol patterns and evaluating their potential to cause errors upon repeated retransmissions. One example is the 80-bit pattern below, which causes multiple recurring errors when modulated and transmitted through the channel

11011000000000000000011111010000110000000 . . .

00000000000111001101100000000000000000000.

In particular, an experiment where the pattern is isolated and repeated through the link at 6 Gb/s yields an error rate of 10^{-2}! This suggests a considerable degree of nesting in the system's error-prone sequences. Comparing the above pattern to the toy example of Fig. 2, the presence of long sequences of logic zeros in the error-prone pattern points toward a channel signature that is heavily skewed towards one polarity.

Considering the performance of an ARQ scheme in these conditions, the fact that an error-prone codeword is affected by the same ISI at every retransmission seems to suggest a need for a possibly large number of retransmissions. Instead, problematic codewords can be encoded with extra overhead upon retransmission in order to break up the error-prone patterns. In the above example, a run-length limiting code (RLL) can be used to break up long single-polarity symbol sequences. A more general method of prohibiting error-causing patterns is provided by the pattern-eliminating codes [9].

VI. Conclusion

Decoupling the effect on the joint error behavior of the magnitudes of the pulse response coefficients from the corresponding sign signature yields a deeper insight into the nature of error correlation in high-speed links. It shows that error correlation is not to be ignored when evaluating the performance of coded high-speed links and, through the concept of correlation distance, yields a simple statistical method that takes into account varying degrees of short-term correlation, predominant in high-speed links. The newly-developed insights also caution against naïve use of biased Monte Carlo methods in high-speed links, as changes to the error region for the purpose of error inflation can drastically alter the systems's joint error behavior.

Both the analytical framework and the experimental results confirm the effects of noise and ISI on dominant error mechanisms in coded link performance and demonstrate the tradeoff between error-correcting capability and coding overhead. In particular, low-rate codes often outperform the more powerful high-rate codes, but the conclusion is not general. Furthermore, though several forward error correcting codes provide error rate reductions of two to three orders of magnitude, they fall short of the improvements predicted for channels with no ISI. Specifically, the positive error correlation significantly increases the probability of observing multiple errors per codeword and inhibits the performance of forward error correcting codes.

Rather than focusing on standard binary forward error correcting codes, the results point to alternative coding techniques that may be more suitable for high-speed links. The simplest binary codes employed for error detection are shown to yield improvements of over five orders of magnitude assuming eventual correct retransmission. While standard retransmission schemes may require further modification in order to become suitable for high-speed link environments, in particular addressing the issues of latency and error recurrence, the availability of a backchannel renders the technique an interesting possibility. Remaining instead in the realm of forward error correction, the need for more powerful overhead-efficient codes points to higher-order alphabets, such as in the Reed-Solomon codes. Such codes have the additional advantage of providing immunity against error bursts that span short distances, rendering them particularly suitable for combating the short-term error correlation observed in high-speed links. Straying from the established path to consider novel approaches to coding and equalization, a further exploration of the pattern-eliminating properties of codes, begun in [9], may provide simple, energy-efficient methods for dealing with residual ISI. In particular, this approach provides a systematic way of dealing with error recurrence in retransmission schemes. Finally, a fresh look at familiar equalization techniques may render simple binary codes a tool of choice in energy-efficient high-speed link design. Taking into account the effect of equalization on the channel signature and favoring equalization methods that reduce error correlation, or even introduce negative error correlation, enhances the performance of standard error control codes, thus allowing for the use of simpler codes with reduced overhead. All of these techniques, whose potential benefits to high-speed links remain virtually unexplored at this time, point to a rich area of research at the fringe of physical limits of operation and standard design paradigms.

References

[1] L. E. Thon and H.-J. Liaw, "Error-correction coding for 10 Gb/s backplane transmission," in *DesignCon*, San Jose, CA, 2004.

[2] D. Carney and E. Chandler, "Error-correction coding in a serial digital multi-gigabit communication system: Implementation and results," in *DesignCon*, San Jose, CA, 2006.

[3] V. Stojanović and M. Horowitz, "Modeling and analysis of high-speed links," in *IEEE Custom Integrated Circuits Conf.*, Sep. 2003, pp. 589–594.

[4] J. Zerbe, C. Werner, V. Stojanović, F. Chen, J. Wei, G. Tsang, D. Kim, W. Stonecypher, A. Ho, T. Thrush, R. Kollipara, M. Horowitz, and K. Donnelly, "Equalization and clock recovery for a 2.5—10 Gb/s 2-PAM/4-PAM backplane transceiver cell," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2121–2130, Dec. 2003.

[5] V. Stojanović, A. Amirkhany, and M. A. Horowitz, "Optimal linear precoding with theoretical and practical data rates in high-speed serial-link backplane communication," in *IEEE Int. Conf. Commun.*, Jun. 2004, vol. 9, pp. 2799–2806.

[6] B. K. Casper, M. Haycock, and R. Mooney, "An accurate and efficient analysis method for multi-Gb/s chip-to-chip signaling schemes," in *IEEE Symp. VLSI Circuits*, Jun. 2002, pp. 54–57.

[7] B. Ahmad, "Performance specification of interconnects," in *DesignCon*, San Jose, CA, 2003.

[8] N. Blitvic and V. Stojanović, "Statistical simulator for block coded channels with long residual interference," in *IEEE Int. Conf. Commun. (ICC)*, Glasgow, U.K., Jun. 24–28, 2007, pp. 6287–6294.

[9] N. Blitvic, L. Zheng, and V. Stojanović, "Low-complexity pattern-eliminating codes for ISI-limited channels," in *IEEE Int. Conf. Commun. (ICC)*, Beijing, China, May 19–23, 2008, pp. 1214–1219.

[10] IEEE P802.3ap Task Force Channel Model Material [Online]. Available: www.ieee802.org/3/ap/public/channel_model1

[11] D. Terr, partitiontable.m MATLAB, Central File Exchange, ID 5154.

[12] J. D'Errico, Partitions of an integer MATLAB, Central File Exchange, ID 12009.

[13] M. Bóna, *A Walk Through Combinatorics, An Introduction to Enumeration and Graph Theory*. Singapore: World Scientific Publishing, 2002.

[14] P. J. Smith, M. Shafi, and H. Gao, "Quick simulation: A review of importance sampling techniques in communications systems," *IEEE J. Sel. Areas Commun.*, vol. 15, no. 4, pp. 597–613, May 1997.

[15] M. Lee, "Channel-and-circuits-aware energy-efficient coding for high-speed links," S.M. thesis, Massachusetts Inst. Technol., Cambridge, U.K., 2006.

[16] S. Lin and D. J. Costello, *Error Control Coding*, 2nd ed. Upper Saddle River, NJ: Pearson Prentice-Hall, 2004.

[17] A. Ho and V. Stojanovic *et al.*, "Common-mode backchannel signaling system for differential high-speed links," in *Symp. VLSI Circuits Digest Tech. Papers*, Jun. 2004, pp. 352–355.

Natasa Blitvic (S'05) received the S.M. degree in electrical engineering and computer science for her work on coding techniques in high-speed links, in 2008, from the Massachusetts Institute of Technology, Cambridge, where she is currently working toward the Ph.D. degree in the field of probability theory. She received the B.Sc. A. Honors degree, Co-op, in electrical engineering from the University of Ottawa, Ottawa, ON, Canada, in 2005. As part of her undergraduate degree, she completed four research internships at the Communications Research Centre, Ottawa, ON, Canada. She holds an NSERC graduate fellowship.

Her interests include free probability and random matrices, with applications to engineering and operations research. She is also interested in communication theory in the context of nonstandard communication environments.

Maxine Lee received the B.S. degree in electrical science and engineering and the M.Eng. degree in electrical engineering and computer science from Massachusetts Institute of Technology, Cambridge, in 2005 and 2006, respectively. Her graduate work focused on hardware analysis of error-control coding in high-speed links. She is currently working toward the J.D. degree at Fordham University School of Law.

She is a patent agent at Kramer Levin Naftalis & Frankel LLP, where she prepares and prosecutes domestic and foreign patent applications in a variety of technical areas.

Vladimir Stojanovic (M'96) received the Dipl. Ing. degree from the University of Belgrade, Serbia, in 1998, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 2000 and 2005, respectively.

He is an Assistant Professor in the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge. His current research interests include design, modeling, and optimization of integrated systems, from standard VLSI blocks to CMOS-based electrical and optical interfaces. He is also interested in design and implementation of digital communication techniques in high-speed interfaces and high-speed mixed-signal IC design. He was also with Rambus, Inc., Los Altos, CA, from 2001 to 2004. He was a Visiting Scholar with the Advanced Computer Systems Engineering Laboratory, Department of Electrical and Computer Engineering, University of California, Davis, during 1997–1998.