An efficient scheme for the evaluation of coding performance of high speed links incorporating joint error behaviour based on Monte Carlo Method

V. Rama Krishna Sharma[1], KVBL Deepthi[2], G. Santhosh Kumar[3]

[1]Associate Professor, ECE Department, GNITC, Hyderabad, India
[2]Assistant Professor, ECE Department, MVSR Engineering College, Hyderabad, India
[3]Assistant Professor, ECE Department, GNITC, Hyderabad, India

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 analyses error correlation through concepts of error region, channel signature, and correlation distance. This framework provides a deeper insight into joint error behaviours 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.

Keywords: channel signature, integrated circuit interconnections, inter symbol interference (ISI),High speed links, correlation distance, error region.

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 wire line, 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 underlying 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 inter-symbol 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 (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. Modelling 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 system'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 non overlapping blocks of symbols and assumes independence in the error behavior across distinct blocks, rather than individual symbols. The computational mechanics are
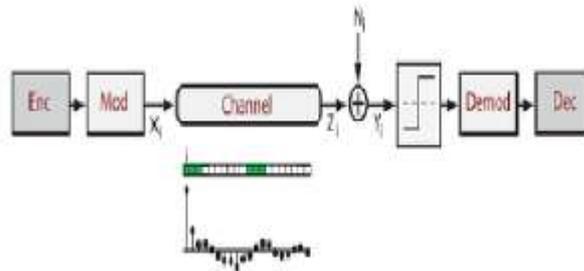

Figure. 1. Simplified model of a high-speed link. Transmit/receive equalization is reflected on the symbol-spaced pulse response

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.

System Model

A simplified model of a high-speed link is shown in Fig. 1.The bit stream, 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 ay 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 $h_{-k} \dots h_{-1}, h_0 \dots \dots h_m$, 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

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

The noise term, representing the combined thermal noise and timing jitter, is assumed to be Gaussian with the standard deviation of σ=3mv relative to the X_l peak values of ± 1 V [5].

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

$$p_i = p(\{y_i < 0 | X_i = 1\} \cup \{Y_i > 0 | X_i = -1\}). \qquad (3)$$

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 $h_{-k}......h_{-1},h_0.......h_m$. and the probability distribution of the noise. Efficient methods of computing the marginal error i 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 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 behaviour. For instance, the performance of a –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 behaviour 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 behaviour 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 of ISI values that are responsible for the majority of errors at the receiver. More formally, letting 0<f≤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 ε_f=(-∞,v] such that

$$P(Z_i \in \varepsilon_f | E_i , X_i=1) \geq f$$

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 are responsible for the majority of the errors, and at the same time small enough so that every event in is likely to cause an error.

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 $h_{-k}......h_{-1},h_0.......h_m$, 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* $s= (s_{-k} .... s_0 .... s_m)$, which takes values from the set $\{ -1 ,0 ,1\}^l$ and is defined as

$$S=(\text{sign}(h_{-k}).........\text{sign}(h_{-1}),\text{sign}(h_0)............\text{sign}(h_m)).$$

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 $(X_{i+k}.........X_i...........X_{i-m})$=p, where

$$P=(-s_{-k}..........-s_{-1},1,-s_1.........-s_m).$$

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 P nests if P and some shifted versions of ±P, each shifted by at most l-1 symbols relative to the first pattern align exactly on the overlapping symbols. For instance, the pattern p=1 1 -1 -1 1 1 nests simultaneously in three locations, as shown below

$$1\ 1\ -1\ -1\ 1\ 1 = p$$
$$-1\ -1\ 1\ 1\ -1\ -1 = -p$$
$$1\ 1\ -1\ -1\ 1\ 1 = p$$

Suppose that the channel is such that the worst-case sequences ±p nests 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 ±p, a relatively large proportion will also contain additional nestings of ±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 compared to the assumption that errors occur independently across distinct symbols. The maximal degree of error correlation is
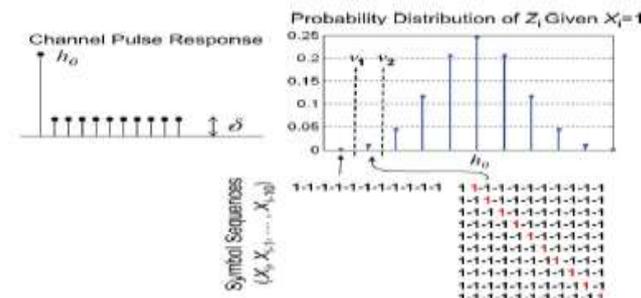

Figure. 2: Simplified pulse response [left] and the corresponding ISI distribution [right]

incurred by channels whose worst-case patterns nest fully, that is, simultaneously on every symbol. One example p = 1 1 1 1 1 1 ....1 , which has the property that any symbol sequence with a run length of 2l causes worst-case ISI to occur on at least l+1 symbols. Another maximally-nested worst-case pattern is the alternating pattern p=1 -1 1 -1 1 -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 p=1-1 -1 -1 -1 -1 -1 -1 -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, $\in_f$ = (-∞ , v1] , where v1 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 $\in_f$=(-∞,v2] , where v2 > v1 as shown in the figure. By allowing the error-prone sequences to deviate p from by one symbol, it becomes possible to nest two error-prone sequences, as illustrated below

$$1\ 1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1$$
$$1\ -1-1\ -1-1-1\ -1-1\ -1\ -1\ -1$$

 In particular, focusing on symbol , let be the event where the surrounding symbols are given by the error-prone pattern $X_i....X_{i-l+1}$=1 1 -1 -1 -1 -1 -1 -1 -1 -1. As shown above, if occurs, then not only does belong to , but so does as well, regardless of the remaining symbol .As a result, if is sufficiently large compared to the marginal error probability ,this observational one is sufficient to guarantee that the proportion of code words with two errors is increased compared to the independent-errors approximation.

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 behaviour of the system by Monte Carlo simulation, the behaviour at high error probabilities is not necessarily indicative of the system's behaviour 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 behaviour 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 ±p cause the same ISI. The problem then decouples into considering the error region$\in_f$, which specifies the maximal deviation k that may be responsible for an error, and the channel signature s , which determines the patterns ±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 behaviours. However, the previous discussion provides insight into the error behaviour of high-speed links.

B. Error Correlation in High-Speed Links

In practical channels, the direct link between the channel sig-nature, error region and joint error behaviour holds only asymptotically. For instance, in the worst-case-dominant conditions, the problem reduces to the nesting properties of 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 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 behaviour, 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.
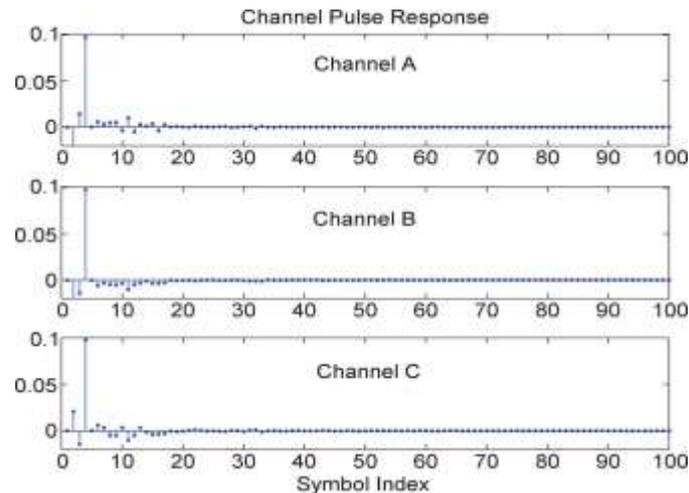
Figure. 3. Equalized pulse response of the standard 802.3ap B32 [10] channel operating at 10 Gb/s. Channel A: Actual channel. Channel B: All-positive signature. Channel C: Randomly generated signature.
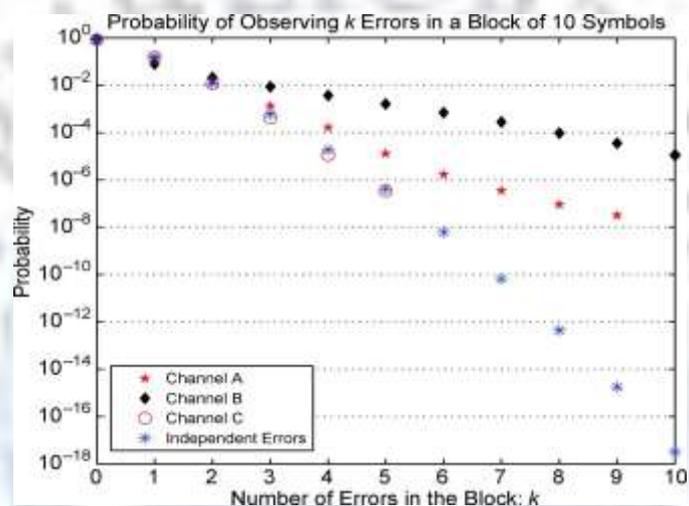

Figure. 4. Joint error statistics for the channels of Fig. 3. Margining by 36 mV increases the error rate to 1.78×10^{-2} , rendering the higher-order error events observable through Monte Carlo simulation (2.5×10^{-7} symbol blocks). The figure illustrates the extent to which channel signature affects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

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. This behaviour 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).

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 behaviour 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

Page | 119

accurately accounting for the effect of the worst-case ISI on joint error behaviour 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 σ=3mv, considering as dominant all interference coefficients greater than σ, 2σ, 3σ and 4σ yields the corresponding error-region parameters of f=7.7×10^{-3},f=0.82,f=0.97 and f=1-2.3×10^{-3}, respectively. The most accurate simulation results were obtained for the cut-offs of 3σ and 4σ. The former case yields the error statistics displayed in the figure, which shows the proportion of code words with a given number of errors in a block of ten symbols. The results are compared to 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.
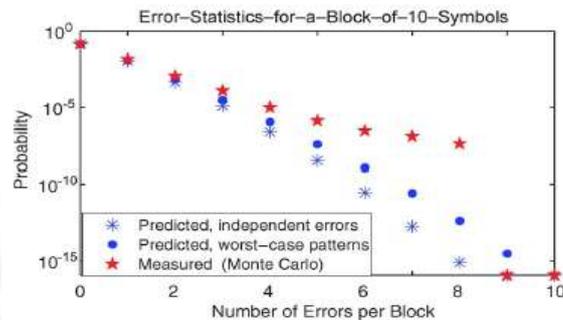

Figure 5: error statistics for the channel of fig 3(a) margined by 36 mv.

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 behaviours 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 any two symbols effectively become independent can. 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, to |i-j|, to vanish when |i-j|≥l for received signals (and |i-j|≥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 behaviour, is contained within a length of approximately eight to ten symbols. The shape of the error- 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.
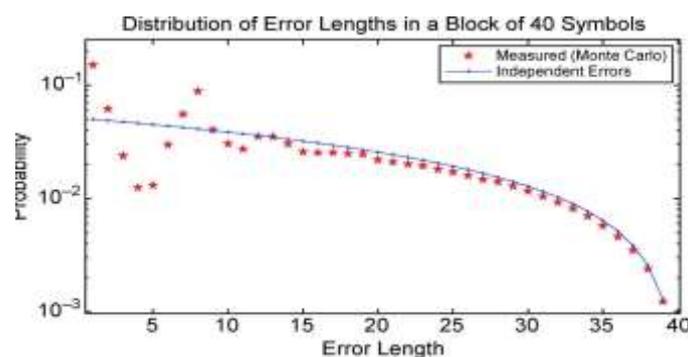

Figure. 6. Distribution of error lengths in a block of 40 symbols with two symbol errors, for the channel of Fig. 3(a) with a threshold bias of 36 mV. The distribution of error lengths is obtained through Monte Carlo simulation (2.5×10^7) symbol blocks) and normalized relative to the total proportion 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.

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 behaviour of a specific high-speed link channel at high error rates need not be indicative of its behaviour 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. 7Specifically, 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.

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 from 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 behaviour 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 n1, n2 , ...,nk 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 $l+n_j-1$, where 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^{l+n-1}. Then, the probability of observing i errors in the jth block, denoted by $p^{(j)}(i)$, is computed considering the possible error patterns and taking into account the noise.

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

$$p_m = \sum_{\substack{\text{all compositions} \\ (m_1,\ldots,m_k)}} \tilde{p}^{(1)}(m_1)\tilde{p}^{(2)}(m_2)\ldots\tilde{p}^{(k)}(m_k). \quad (7)$$
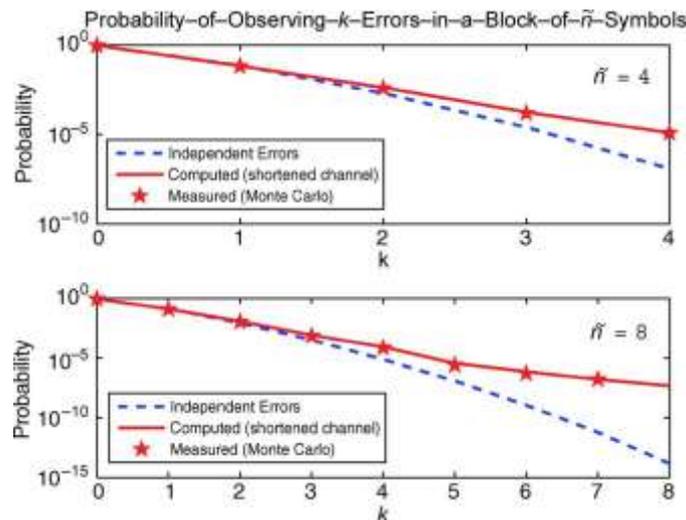
Figure. 7. Error statistics for blocks of 4 and 8 symbols for the channel of Fig. 3(a) with a threshold bias of 36 mV. The results computed analytically based on a shortened channel with l=16 by considering all 2^{n+l-1} symbol patterns match closely the Monte Carlo measurements (2.5×10^7 symbol blocks) for the actual (unshortened) channel. The corresponding result based on the independent-errors assumption is also included.

In addition, if the error statistics are equal across the k blocks, which occurs if $n_1 = n_2 = \cdots ... = n_k$ 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 $p_0, p_1 ....p_n$, the performance ofa -error-correcting code, for example, is given by $WER_t = 1 - (p_0 + \cdots ...... + p_t)$.

For the high-speed link channel of Fig. 3(a), letting l=16 yields the block error statistics p^i of Fig. 7 for blocks of n=4 [top] and n=8 [bottom] symbols. As expected, the results match closely the statistics captured through Monte Carlo simulation. aplying 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 n=4 and 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 , as the number of possible symbol patterns of length l+n-1is typically large. Based on the block size 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 $p^{(i)}$ 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 highspeed 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/un biasing procedures, as some form of joint statistics becomes required to partially describe the system's behaviour.

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 behaviour at low error rates. Rather than un biasing 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.
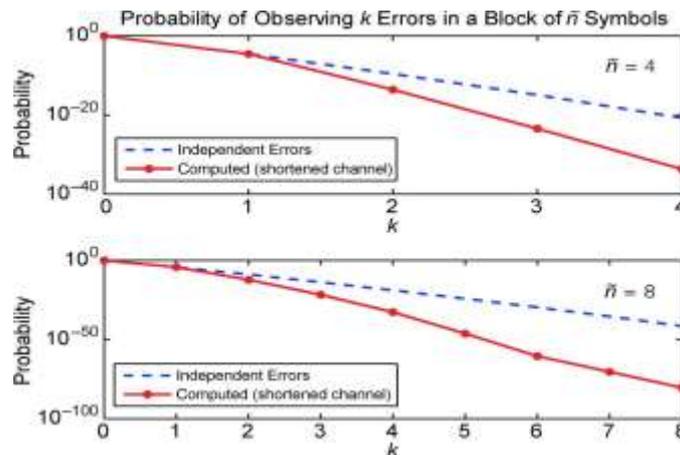
Figure.8. Error statistics for blocks of 4 and 8 symbols for the channel of Fig. 3(a) without the threshold bias of 36 mV. The results computed analytically based on a shortened channel with l=16 by considering all 2^{n+l-1} symbol patterns are compared to the results based on the independent-errors assumption.

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

$$P = 1 \ -1 \ 1 \ 0 \ -1 \ -1 \ -1 \ -1 \ 1 \ 1 \ -1 \ -1 \ -1 \ 0 \ -1 \ 1 \ -1$$

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 behaviour 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 behaviours.

Conclusion

Decoupling the effect on the joint error behaviour 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 system's joint error behaviour. 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 trade-off 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 he 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]. 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.

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

[3]. J. Zerbe, C. Werner, V. Stojanovic´, 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.

[4]. V. Stojanovic´, 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.

[5]. N. Blitvic and V. Stojanovic´, "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.

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

[7]. 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.

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

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