AN EFFICIENT FRAMEWORK FOR CHANNEL CODING IN HIGH SPEED LINKS

Paradesi Leela Sravanthi , K. Ashok Babu
Dept of ECE, Sri Indu College of Engg, Hyderabad, India.

ABSTRACT

This paper explores the benefit of channel coding for high-speed backplane or chip to chip interconnects, referred to as the high-speed links. Although both power constrained and bandwidth-limited, the high-speed links need to support data rates in the Gbps range at low error probabilities. Modeling the high-speed link as a communication system with noise and inter symbol interference (ISI), this work identifies three operating regimes based on the underlying dominant error mechanisms. The resulting framework is used to identify the conditions under which standard error control codes perform optimally, incur an impractically large overhead, or provide the optimal performance in the form of a single parity check code. For the regime where the standard error control codes are impractical, this thesis introduces low complexity block codes, termed pattern-eliminating codes (PEC), which achieve a potentially large performance improvement over channels with residual ISI. The codes are systematic, require no decoding and allow for simple encoding. They can also be additionally endowed with a (0, n − 1) run-length-limiting property. The simulation results show that the simplest PEC can provide error-rate reductions of several orders of magnitude, even with rate penalty taken into account. It is also shown that channel conditioning, such as equalization, can have a large effect on the code performance and potentially large gains can be derived from optimizing the equalizer jointly with a pattern-eliminating code. Although the performance of a pattern-eliminating code is given by a closed-form expression, the channel memory and the low error rates of interest render accurate simulation of standard error-correcting codes impractical.

I. INTRODUCTION

The field of channel coding started with Claude Shannon's 1948 landmark paper [1]. For the next half century, its central objective was to find practical coding schemes that could approach channel capacity (hereafter called Bthe Shannon limit) on well-understood channels such as the additive white Gaussian noise (AWGN) channel. This goal proved to be challenging, but not impossible. In the past decade, with the advent of turbo codes and the rebirth of low-density parity-check (LDPC) codes, it has finally been achieved, at least in many cases of practical interest. Currently, communication bus links in various applications approach Gb/s data rates. Such links are often an important part of multi-processor interconnection [10], processor-to-memory interfaces [11], and SONET/Fibre channels [12], high-speed network switching, and local area networks [13]. It is also likely that many high-speed digital signals will be transmitted between analog and digital chips. Traditionally, system designers have addressed the need for high-speed chip-to-chip links by increasing the number of high-speed signals, which leads to an increase in the cost and complexity of the system. Therefore, the per-pin interconnection bandwidth should be increased. Improving the performance of both parallel and serial interconnects has been an important research area over the last decade [14-16]. Although each type of interconnect has some advantages and disadvantages, the general trend has been toward serial links. However at the same time, significant amount of research has been performed to improve the performance of popular, general purpose parallel buses [17]

II.CODING FOR THE AWGN CHANNEL

A coding scheme for the AWGN channel may be characterized by two simple parameters: its signal-to-noise ratio (SNR) and its spectral efficiency η in bits per second per Hertz (b/s/Hz). The SNR is the ratio of average signal power to average noise power, a dimensionless quantity. The spectral efficiency of a coding scheme that transmits R bits per second (b/s) over an AWGN channel of bandwidth W Hz is simply η=R/W b/s/Hz. Coding schemes for the AWGN channel typically map a sequence of bits at a rate R b/s to a sequence of real symbols at a rate of 2B symbols per second; the discrete time code rate is then r = R/2B bits per symbol. The sequence of real symbols is then modulated via pulse amplitude modulation (PAM) or quadrature amplitude modulation (QAM) for transmission over an AWGN channel of bandwidth W. By Nyquist theory, B (sometimes called the B Shannon bandwidth [3]) cannot exceed the actual bandwidth W. If B≈W, then the spectral efficiency is

$$\eta = R/W \approx R = B \text{ ¼ } 2r. \quad \text{(II.1)}$$

We therefore say that the nominal spectral efficiency of a discrete-time coding scheme is 2r, the discrete-time code rate in bits per two symbols. The actual spectral efficiency η=R/W of the corresponding continuous-time scheme is upper bounded by the nominal spectral efficiency 2r and approaches 2r as B→W. Thus, for discrete-time codes, we will often denote 2r by η, implicitly assuming B ≈W. Shannon showed that on an AWGN channel with a given SNR and bandwidth W Hz, Shannon showed that on an AWGN channel with a given SNR and bandwidth W Hz, the rate of reliable transmission is upper bounded by

$$R < W \log_2(1+SNR). \quad \text{(II.2)}$$

Moreover, if a long code with rate R<W log2(1+SNR) . is chosen at random, then there exists a decoding scheme such that with high probability the code and decoder will achieve highly reliable transmission (i.e., low probability of decoding error). Equivalently, Shannon's result shows that the spectral efficiency is upper bounded by

$$\eta < log\ 2(1+SNR) \quad \text{(II.3)}$$

or, given a spectral efficiency η, that the SNR needed for reliable transmission is lower bounded by

$$SNR > 2\boldsymbol{\eta}\text{-}1 \quad \text{(II.4)}$$

So, we may say that the Shannon limit on rate (i.e., the channel capacity) is

$$W \log_2(1+SNR) \text{b/s} \quad \text{(II.5)}$$

or equivalently that the Shannon limit on spectral efficiency is log2(1+SNR) b/s/Hz, or equivalently that the Shannon limit on SNR for a given spectral efficiency η is $^{2}\boldsymbol{\eta}$-1 Note that the Shannon limit on SNR is a lower bound rather than an upper bound. These bounds suggest that we define a normalized SNR parameter SNR norm as follows:

$$\text{SNR norm} = (\text{SNR}/^{2}\boldsymbol{\eta}\text{-}1) \quad \text{(II.6)}$$

Then, for any reliable coding scheme, SNR norm >1 i.e., the Shannon limit (lower bound) on SNR norm is 1 (0 dB), independent of η. Moreover, SNR norm measures the Bgap to capacity,[i.e., 10log 10SNRnorm is the difference in decibels (dB)1 between the SNR actually used and the Shannon limit on SNR given η , namely $^{2}\boldsymbol{\eta}$-1 If the desired spectral efficiency is less than 1 b/s/Hz (the so called power-limited regime), then it can be shown that binary codes can be used on the AWGN channel with a cost in Shannon limit on SNR of less than 0.2 dB. On the other hand, since for a binary coding scheme the discrete-time code rate is bounded by r<=1 bit per symbol, the spectral efficiency of a binary coding scheme is limited to $\boldsymbol{\eta} \leq 2r \leq 2$ b/s/Hz, so multilevel coding schemes must be used

if the desired spectral efficiency is greater than 2 b/s/Hz (the so-called bandwidth-limited regime). In practice, coding schemes for the power-limited and bandwidth-limited regimes differ considerably. A closely related normalized SNR parameter that has been traditionally used in the power-limited regime is Eb/N0, which may be defined as

$$Eb/No = SNR/\eta = ((2^{\eta}-1)/\eta)\ SNRnorm \qquad (II.7)$$

For a given spectral efficiency η , Eb/N0 is thus lower bounded by

$$Eb/No > ((2^{\eta}-1)/\eta) \qquad (II.8)$$

so we may say that the Shannon limit (lower bound) on Eb/N0 as a function of η is $(2^{\eta}-1)/\eta$ This function decreases monotonically with η and approaches ln2 as $\eta \to 0$ so we may say that the ultimate Shannon limit (lower bound) on Eb/N0 for any η is ln2(-1.59db). We see that as $\eta \to 0$ Eb/N0$\to$ SNRnorm ln 2, so Eb/N0 and SNRnorm become equivalent parameters in the severely power-limited regime. In the power-limited regime, we will therefore use the traditional parameter Eb/N0.

III. ALGEBRAIC CODING

The algebraic coding paradigm dominated the first several decades of the field of channel coding. Indeed, most of the textbooks on coding of this period (including Peterson [4], Berlekamp [5], Lin [6], Peterson and Weldon [7], Mac Williams and Sloane [8], and Blahut [9]) covered only algebraic coding theory. Algebraic coding theory is primarily concerned with linear (n,k,d) block codes over the binary field F2. A binary linear (n,k,d)block code consists of 2k binary n-tuples, called code words, which have the group property: i.e., the component wise mod-2 sum of any two code words is another codeword. The parameter d denotes the minimum Hamming distance between any two distinct code words, i.e., the minimum number of coordinates in which any two code words differ. The theory generalizes to linear (n,k,d)block codes over non binary fields Fq. The principal objective of algebraic coding theory is to maximize the minimum distance d for a given (n,k). The motivation for this objective is to maximize error correction power. Over a binary symmetric channel (BSC: a binary-input, binary-output channel with statistically independent binary errors), the optimum decoding rule is to decode to the codeword closest in Hamming distance to the received n-tuple. With this rule, a code with minimum distance d can correct all patterns of (d-1)/2 or fewer channel errors (assuming that d is odd), but cannot correct some patterns containing a greater number of errors.

IV. 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 may also include co-channel interference, caused, for instance, by electro-magnetic coupling (crosstalk). 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, denoted Yi and expressed as

$$Yi = Zi + Ni \qquad (IV.1)$$

where Zi denotes the received signal in the absence of noise and is the noise term. Specifically, denoting the channel's pulse response by h -k,…….,h -1,h0 ,……hm where l =k + m + 1 represents the length of the pulse response and h0 is associated with the principal signal component, and letting {Xi} denote a sequence of transmitted symbols, then

Zi= j (IV.2)

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 peak values of +/-1 V.

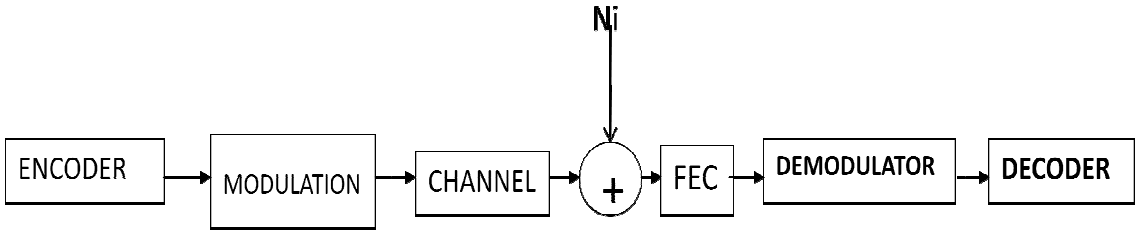

Fig. 1. Simplified model of a high-speed link. Transmit/receive equalization is reflected on the symbol-spaced pulse response.

V. COST-EFFICIENT SIGNALING SCHEMES

Recently, the noise margin on digital chip-to-chip interconnects has been decreasing for two main reasons. One reason is that supply voltages in digital complementary metal oxide semiconductor (CMOS) processes are decreasing thereby reducing the voltage available for driving I/Os. A second reason is that small signal swings are being used to reduce dynamic power dissipation on high-speed busses. It has long been known that fully-differential signals effectively reject common-mode noise and even-order distortion terms. Since common-mode noise is prevalent on matched printed circuit board (PCB) traces, differential signaling is effective for both voltage [14], [15] and current mode [16] digital chip-to-chip interfaces. Fully differential signals are now used in the Scalable Coherent Interface and RamLink [17] standards. Unfortunately, a practical problem with their implementation is that two signal paths are required for each signal. For example, using fully differential signals for a 64-bit data bus would require 128 pins on each IC package and 128 PCB traces routed between ICs. These additional costs are often prohibitive. Therefore, one important approach is to reduce the required number of pins of interconnects. A signaling scheme that has most of the advantages of fully differential signaling scheme with reduced number of signal paths is proposed in Chapter 3 to help alleviate this problem.

VI. POWER-EFFICIENT SIGNALING SCHEMES

Multi-level signaling, such as 4-level pulse amplitude modulation (4-PAM), can be used to reduce the required number of signal paths in a link or to increase the data rate of a link. Channel coding can be used to reduce the power consumption of a high-speed inter-chip link by introducing some redundancy at the transmitter. There is still a significant gap between the Shannon limit, the theoretical limit for channel capacity, and the data rates of the current state-of-the-art designs. To find a low-power scheme, channel coding can be employed as an attempt to approach the Shannon limit [1]. Finding codes that can approach Shannon limit is not a complicated task. Indeed, randomly generated codes with a large block size can be used to approach this limit [1]. The problem lies in the fact that while encoding is always a rather simple task, the decoding complexity increases exponentially with the block size, and thus quickly becomes unmanageable. On the other hand, to maintain high system performance, not only high-speed circuits but also low-loss matched transmission lines for interconnect are necessary to ensure good propagation properties such as minimum crosstalk, delay, reflection, and dispersion[17]. Achieving a highly dense system by bringing the chips closer together is only a partial solution since denser systems require denser interconnects, which in turn cause more crosstalk. Indeed, crosstalk is the dominant noise in most microstrip interconnects.

VII. POWER-EFFICIENT CIRCUIT ARCHITECTURES

Employing circuit techniques for designing the building blocks of a high-speed link is another efficient method to reduce the power and cost of high-speed links[16][17]. The potential benefits of 4-PAM signaling for increasing data rates in physical short-bus systems have been shown. Since there are several drivers in a parallel bus signaling system, the power dissipation of each driver is extremely important. Therefore, power-efficient drivers are desirable. The reported high-speed multi-level drivers have used power-inefficient unipolar architectures.

VIII. RESULTS

The simulation will vary in the parameter Message Block Length. The message block length is changed in the range: 500, 1000, 2000, 4000 and 8000.

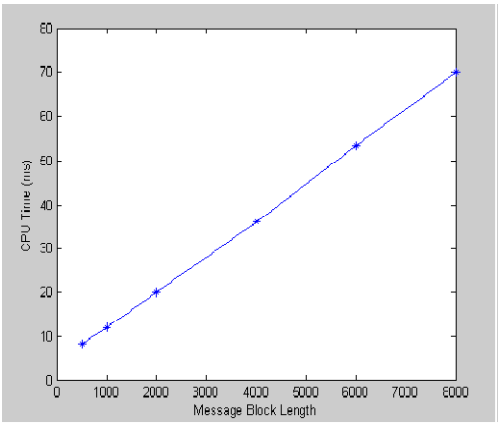

Figure 2: CPU time for generating a n ∗ 2n matrix for different message block lengths and saving it in the row wise.

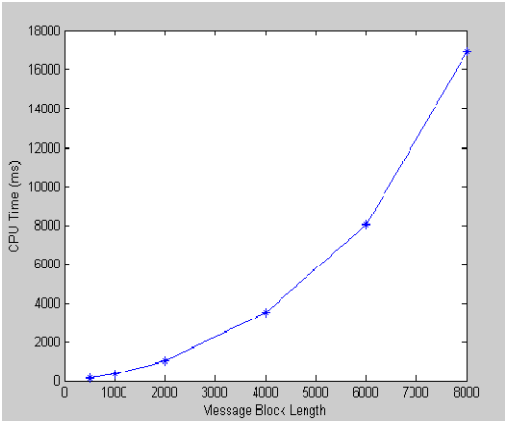

Figure 3: CPU time for generating a n ∗ 2n matrix for different message block lengths and saving it in the row and column wise.

Figure 2 and Figure 3 shows the different CPU time for generating and saving a n ∗ 2n sparse matrix in different manners in software simulation. From the graphs we could see that the CPU time has a linear increasing when saving matrix by row but has a quadratic increasing when saving the matrix by column.

Figure 4 shows the CPU time fore encoding various block length messages. We could see that when the block length increased, the encoding time has a approximate linear increasing. Actually, the encoding time is increasing by $n \cdot \ln(n)$ which is equal to the total amount of degrees. To achieve a linear encoding time, the CHANNEL code should be extent to initial code by involving appropriate pre-coding algorithms.

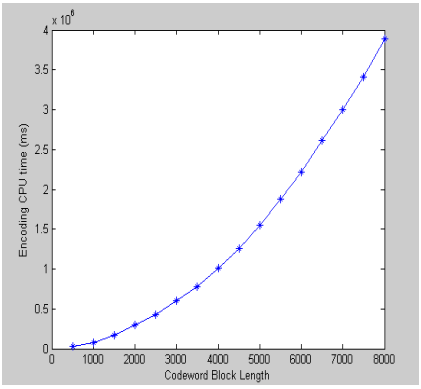

Figure 4: CPU time for encoding various block length messages.

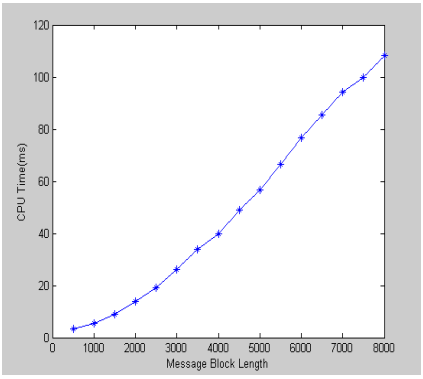

Figure 5: CPU time for decoding various block length messages symbols in the encoding process.

But the number of codeword symbols are decoded is non-fixed and depends on the binary erasure rate. Note that both the encoding time and decoding time we have shown in Figure 4 and Figure 5 does not include the time for generating the sparse matrix.

IX. CODING PERFORMANCE ANALYSIS

In this section, we are focus on the coding performance of CHANNEL codes. The simulations are based on three main parameters: Block Length, Binary Erasure Rate and Code Rate. Two sets of simulations are built. The first one is based on a binary erasure channel with arbitrary binary erasure rates. The second one is based on a AWGN channel with modified BPSK modulation method

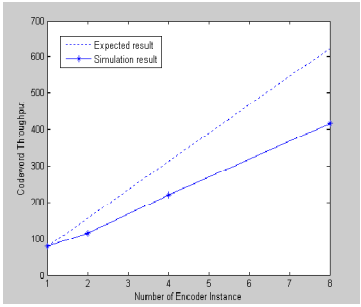

Figure 6: Throughput of parallel encoder

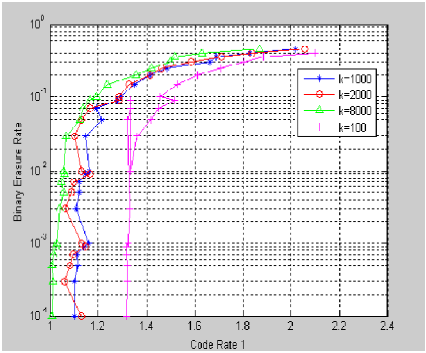

Figure 7: Average code rate for decoding messages with various block length.

In the first simulation, CHANNEL coding is performed on a Binary Erasure Channel with arbitrary binary erasure rates. The block length is changed in the range: 100, 1000, 2000 and 8000. The binary erasure rate is changed in the range: 0.0001, 0.0003, 0.0005, 0.0007, 0.0009, 0.001, 0.003, 0.005, 0.007, 0.009, 0.01, 0.03, 0.05, 0.07, 0.09, 0.1, 0.15, 0.25, 0.3, 0.35 and 0.4. The simulation shows the average code rate for decoding various block length messages. When the binary erasure rate becomes bigger, the codes with different block length have similar behaviors.

The second sets of simulation are based on AWGN channel with modified BPSK modulation method. One new parameter is used here: Threshold which implies how many received symbols will be declared as "erasures" after demodulation. The threshold is changed in the range: 0.2, 0.5, and 0.8.

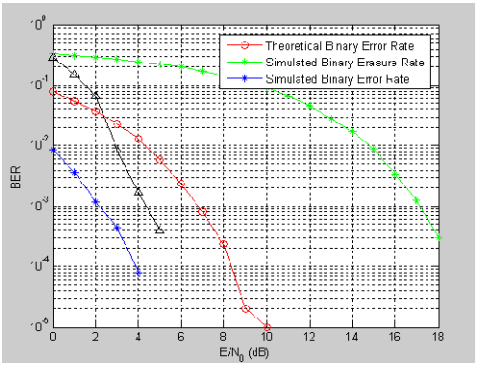

Figure 8: CHANNEL Coding over AWGN channel and modified BPSK modulation with threshold 0.3.

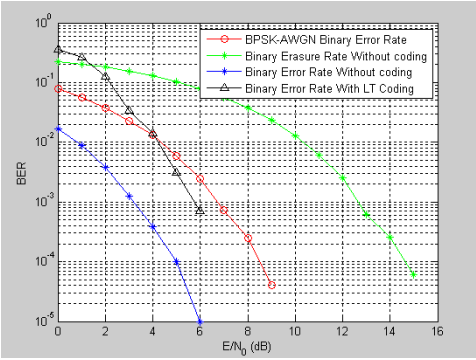

Figure 9: CHANNEL Coding over AWGN channel and modified BPSK modulation with threshold 0.5.

The bigger the threshold is, the less the received symbols will be "erasured". Figure 7, Figure 8 and Figure.9 shows the simulation results for the different sets of parameters. From these graphs we can see that the CHANNEL coding cannot correct any binary errors, whereas it actually increases the binary error rates by about ten times. However, when the binary error rate decreased, the CHANNEL code can still have some coding gains. When we compare Figure and Figure we can see that by lowing threshold in modified BPSK Modulation, we can decrease the binary error rate to get more coding gains. But at the same time we increased the binary erasure rate so the CHANNEL code will have a lower code rate. In Figure 10 we compared three 1000-bits CHANNEL codes with a (255, 223, 32) block code and the non-coding code. We can see that the block code is over-performed than the CHANNEL codes. However, considering the block code may have a longer coding time, e.g. RS code

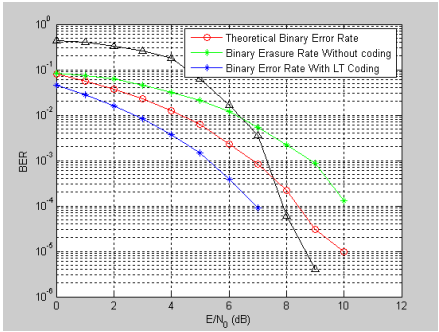

Figure 10: CHANNEL Coding over AWGN channel and modified BPSK modulation with threshold 0.8.

X. CONCLUSION:

Modeling a high-speed link as an ISI-limited system with additive white Gaussian noise allows for an abstracted framework suitable for a more theoretical approach to studying the benefit of coding for high-speed links. Possible error mechanisms are categorized according to three regimes- the large noise, the large-set-dominant and the worst-case-dominant—which are entirely specified by the system's noise level and the channel's pulse response. In the large-noise and large-set-dominant regimes, classical coding theory provides an exhaustive characterization of different error-control codes, whose hardware complexity has already been partially addressed. While the worst-case-dominant regime occurs rarely in a high-speed link, the quasi-worst-case dominant regime is shown to occur. However, further work is required on extending the pattern-eliminating properties to deal with a wider range of operating conditions. In particular, one of the remaining problems consists of identifying or developing suitable equalization or channel conditioning techniques that optimize the performance of a pattern-eliminating code. Such equalization is, in principle, significantly more power-efficient compared to that employed in current high-speed links, as the equalizer no longer needs to ensure a low error probability. The corresponding scheme could potentially yield significant

benefits for high-speed links by enabling the communication at higher data rates than those achieved previously, or by providing the same signaling speeds at greater energy efficiency.

REFERENCES

[1].C. E. Shannon, BA mathematical theory of communication,[Bell Syst. Tech. J., vol. 27, pp. 379–423 and 623–656, 1948.
[2].R. W. McEliece, BAre there turbo codes on Mars?[in Proc. 2004 Int. Symp. Inform. Theory, Chicago, IL, Jun. 30, 2004 (2004 Shannon Lecture).
[3].J. L. Massey, BDeep-space communications and coding: A marriage made in heaven,[in Advanced Methods for Satellite and Deep Space Communications, J. Hagenauer, Ed. New York: Springer, 1992.
[4].W. W. Peterson, Error-Correcting Codes Cambridge, MA: MIT Press, 1961.
[5].E. R. Berlekamp, Algebraic Coding Theory New York: McGraw-Hill, 1968.
[6].S. Lin, An Introduction to Error-Correcting Codes. Englewood Cliffs, NJ: Prentice-Hall, 1970.
[7].W. W. Peterson and E. J. Weldon, Jr., Error-Correcting Codes. Cambridge, MA: MIT Press, 1972.
[8].F. J. MacWilliams and N. J. A. Sloane, The Theory of Error-Correcting Codes. New York: Elsevier, 1977.
[9].R. E. Blahut, Theory and Practice of Error Correcting Codes. Reading, MA: Addison-Wesley, 1983.
[10]. R. Mooney, C. Dike, and S. Borkar, \A 900 Mb/s bidirectional signaling scheme," IEEE J. Solid-State Circuits, vol. 30, pp. 1538{1543, Dec. 1995.
[11]. N. Kushiyama et al., \A 500-megabyte/s data-rate 4.5 M DRAM," IEEE J. Solid-State Circuits, vol. 28, pp. 490 { 498, Apr. 1993.
[12]. Y. Ota and R. Swartz, \Multichannel parallel data link for optical communication," IEEE LTS, vol. 2, pp. 24 { 32, May 1991.
[13]. [13] M. Horowitz, C.-K. K. Yang, and S. Sidiropoulos, \High-speed electrical signaling: Overview and limitations," IEEE Micro, pp. 12{24, Jan. 1998.
[14]. R. Farjad-Rad, \A CMOS 4-PAM Multi-Gbps serial link transceiver," Ph.D. disserta tion, Stanford University, Stanford, 2000.
[15]. C.-K. K. Yang, \Design of high-speed serial links in CMOS," Ph.D. dissertation, Stan- ford University, Stanford, 1998.
[16]. S. Abdalla, \A 7.2Gb/s/pin 8-bit parallel bus transmitter using incremental signaling in 0:18¹m CMOS," Master's thesis, Univ. of Toronto, Toronto, 2002.
[17]. S. Sidiropoulos, \High performance inter-chip signalling," Ph.D. dissertation, Stanford University, Stanford, 1998.

Authors Biographies

Paradesi Leela Sravanthi was born in Hyderabad, India, in 1986. I have received the Bachelor degree from the University of JNTU, Hyderabad, in 2007. Currently I am pursuing the M.Tech with the Department of ECE branch (DECS) Engineering, Hyderabad. My research interests include Communication, signal processing, and information theory.

K. Ashok Babu, HOD – ECE of Sri Indu College of Engineering. He was born in Hyderabad and he completed his B.Tech and MTech in ECE and he has completed PhD in communication systems and has 15 years of experience at B.Tech and M.Tech. level and his constant cooperation support and providing necessary facilities throughout the M.Tech. His research interests include VLSI, ES communications and signal processing. Presently working in Sri Indu College of Engineering, Hyderabad.