A 0.4-μm CMOS 10-Gb/s 4-PAM Pre-Emphasis Serial Link Transmitter

Ramin Farjad-Rad, *Student Member, IEEE*, Chih-Kong Ken Yang, *Member, IEEE*, Mark A. Horowitz, and Thomas H. Lee, *Member, IEEE*

***Abstract*—A serial link transmitter fabricated in a large-scale integrated 0.4-μm CMOS process uses multilevel signaling (4-PAM) and a three-tap pre-emphasis filter to reduce intersymbol interference (ISI) caused by channel low-pass effects. Due to the process-limited on-chip frequency, the transmitter output driver is designed as a 5 : 1 multiplexer to reduce the required clock frequency to one-fifth the symbol rate, or 1 GHz. At 5 Gsym/s (10 Gb/s), a data eye opening with a height >350 mV and a width >100 ps is achieved at the source. After 10 m of a copper coaxial cable (PE142LL), the eye opening is reduced to 200 mV and 90 ps with pre-emphasis, and to zero without filtering. The chip dissipates 1 W with a 3.3-V supply and occupies 1.5 × 2.0 mm^2 of die area.**

***Index Terms*— Intersymbol interference (ISI), networks, pre-emphasis transmitter, serial links.**

I. Introduction

APPLICATIONS such as computer-to-computer or computer-to-peripheral interconnection are requiring gigabit-per-second rates over different distance ranges. For distances ranging 1–10 m, traditional methods of parallel buses, which require many wires, are costly and power inefficient. Optical fibers are also costly and area inefficient for these distances. Thus, low-cost, high-speed serial links using copper cables are an attractive solution for such applications [1], [2]. While other technologies, such as GaAs and bipolar, are limited in the number of transistors per die due to yield and power considerations, CMOS technology allows implementation of complex digital logic enabling more integration of the back-end processing, lowering the cost.

The maximum data rates reported over 7-m copper cable are 10 Gbps in an Si-bipolar technology [3] and 4 Gbps in 0.5-μm CMOS technology [1]. This paper describes a 10-Gbps transmitter implemented in 0.4-μm CMOS, which is intended for use with coaxial cable (PE-142LL) over a distance of 10 m.

Due to the skin-effect loss in conductors, copper cables show a low-pass frequency response that limits signaling bandwidth. Fig. 1 shows that the 10-m coaxial cable used in this work has a -3 dB bandwidth of 1.2 GHz. Furthermore, the intrinsic process speed limits the on-chip frequency. In this 0.4-μm CMOS process, the maximum on-chip operating frequency of digital logic is roughly 1 GHz.

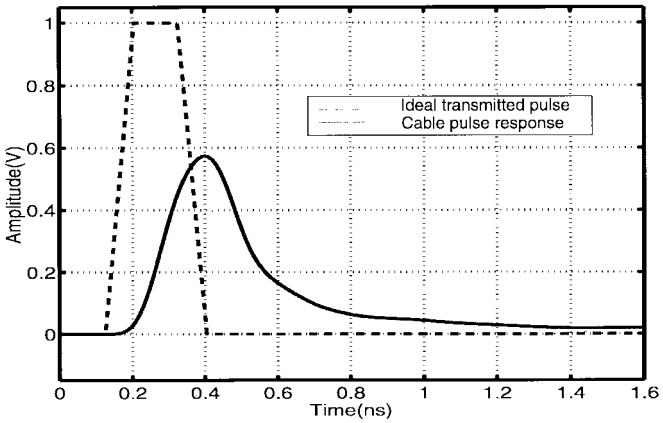


Fig. 1. Pulse response of a 10-m PE142LL coaxial cable.

Employing a 5 : 1 multiplexing scheme, a pre-emphasis technique using a three-tap finite impulse response (FIR) filter and a four-level pulse amplitude modulation (4-PAM) enables us to achieve the 10-Gbps data rate. Speed limitations due to process technology were thoroughly examined when choosing the 4-PAM modulation scheme.

Section II provides some background on the communication theoretical principles exploited in this work. Section III describes the system architecture, and Section IV presents the circuit implementation of system blocks. Measurement results are presented in Section V, followed by concluding remarks in Section VI.

II. Background

Digital communication involves two key ideas. The first is the use of orthogonal analog waveforms (basis functions) as transmission building blocks, e.g., square waves. The second uses the basis functions to associate modulation with a vector space. The number of orthogonal basis functions is the dimensionality of the vector space. A geometric arrangement of points in the vector space is called a constellation and represents all possible data symbols. A typical constellation is N-level pulse amplitude modulation (N-PAM), where the data sequence is carried by pulses of fixed shape, with N different amplitudes, each of which represents $\log_2(N)$ bits of data. For a fixed data rate, one may trade off symbol rate against constellation size.

Manuscript received September 17, 1998; revised December 24, 1998.

R. Farjad-Rad, M. A. Horowitz, and T. H. Lee are with the Center for Integrated Systems, Stanford University, Stanford, CA 94305 USA.

C.-K. K. Yang was with the Center for Integrated Systems, Stanford University, Stanford, CA 94305 USA. He is now with the University of California, Los Angeles, CA 90095 USA.

Publisher Item Identifier S 0018-9200(99)03676-8.

Channel spectral efficiency, which is measured in terms of bits per second per hertz (bps/Hz), is determined by the bandwidth of the basis waveforms. For instance, in a nonreturn-to-zero (NRZ) N-PAM communication system, the spectral efficiency is $2 \times \log_2(N)$, which increases logarithmically with the number of PAM levels (N). Note that the factor of "2" is due to the NRZ nature of the transmission. If the basis waveforms are not channel eigenfunctions, intersymbol interference (ISI) occurs. ISI can be reduced using FIR filters that invert the channel low-pass characteristic. Optimal detection can also be performed by techniques such as maximum-likelihood sequence detection or sampled matched filtering at the receiver. Coding can be used to improve system symbol error rate. For a more detailed treatment of these subjects, see [4] and [5].

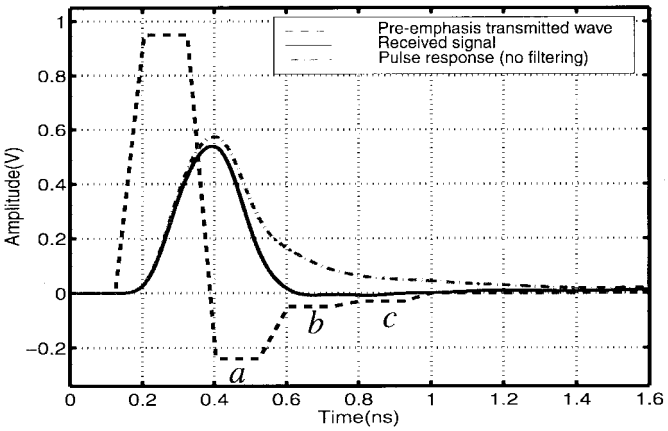


Fig. 2. Preshaped pulse at near and far ends of the channel.

III. System Architecture

In multi-gigabit/second applications, optimal detection methods demand high complexity and hence large circuit area [6]. Therefore, in this work, square pulses are used as the basis waveforms due to their simplicity in generation and the ability to perform level detection with moderate complexity for multi-gigabit/second rates [7], even though these waveforms are not optimal from an information-theoretic point of view.

For a given data rate, the 4-PAM scheme reduces the symbol rate by a factor of two compared to a conventional 2-PAM system. This symbol-rate reduction lowers not only the signal ISI in the channel but also the maximum required on-chip clock frequency. An N-PAM scheme with larger N was avoided due to limited receiver signal resolution at high speeds and the maximum transmitter output swing, both of which constrain the PAM level spacings

$$\text{Level Spacing} = \frac{X\text{mitter Swing}}{N-1}.$$

Moreover, reflection ISI of large signals due to imperfect line terminations can overwhelm subsequently transmitted low-level signals. Since this design does not attempt to equalize for reflection ISI, 4-PAM is chosen to avoid vulnerable low-amplitude symbols.

Transmitting a sequence of square symbols results in a data eye diagram. Larger eye openings correspond to better system noise immunity. Since square pulses are not channel eigenwave forms, ISI occurs, which results in severe reduction of the eye opening at symbol rates well above the channel bandwidth. The measured 1.2-GHz channel bandwidth is roughly a factor of three less than the bandwidth required to obtain a reasonable eye opening for a 5-Gsym/s 4-PAM system with square pulses, so filtering techniques must be used.

Filtering is performed by either equalizing the signal in the receiver or preshaping the outgoing pulse in the transmitter. Receiver equalization is normally more difficult to implement at gigabit/second rates than transmitter pre-emphasis. Digital receiver equalization, using FIR filters, requires high-resolution sampling analog-to-digital converters that run at gigahertz speeds, which is a quite challenging task in present CMOS technologies. Analog continuous-time equalization also needs very wide-bandwidth front-end receiver circuits that run at the same speed as the input data. The low F_t of transistors in present CMOS technologies makes the receiver equalizer design quite challenging at multi-gigabit/second rates. Input equalizers reported to date in CMOS technology all operate at data rates below 1 Gbps [8], [9]. To implement a transmit preshaping FIR filter, however, the output driver only requires adding the weighted values of the previous symbols, already known to the transmitter, to the present outgoing symbol value. Thus, the preshaping technique does not dictate the transmitter to use a faster technology to operate properly.

One method to implement the N-tap transmitter filter is what is used in [1], which requires a large area and high complexity. In this approach, all the FIR filter calculations are done by digital adders and multipliers, and a high-resolution digital-to-analog converter (DAC) generates the final pulse, which is the sum of the present symbol and N previous tap-weighted symbols. In this design, we use a completely analog technique to realize the pre-emphasis filter, where the transmitter generates the "filtered" pulse directly and independent of all previous symbols. The need for complex digital logic is removed by summing and modulating the output current in the analog domain. This method also uses minimum-resolution DAC's [$\log_2(N)$-bit DAC for N-PAM].

To design the pre-emphasis FIR filter, we have measured the pulse response of the coaxial line as shown in Fig. 1. Pulses sent through the channel experience a long tail that corrupts subsequently transmitted symbols. Simulations show that a three-tap filter with symbol-period tap spacings can reduce the amplitude of the undesired tail to $<$10% of the 4-PAM amplitude spacing ($<$3% of the total amplitude). The results of this filter for a 0.2-ns pulse (5 Gsym/s), at both the near and far ends of the channel, are shown in Fig. 2.

The on-chip frequency requirement is reduced to one-fifth the symbol rate (one-tenth the bit rate) by performing a $5:1$ multiplexing directly onto the 50-Ω line, allowing five symbols to be transmitted every cycle. The five symbols correspond to 10 bits, which include four data symbols and one symbol for line coding. In this design, line coding is performed on-chip to provide appropriate transitions for clock recovery. This method is similar to the 8 b/10 b codes used in binary transmission.

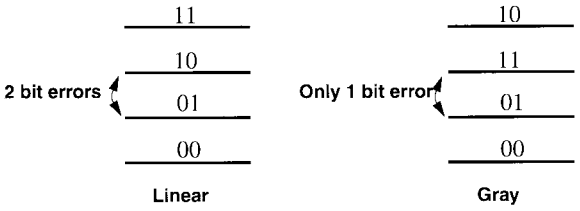


Fig. 3. Linear versus Gray-code mapping of levels.

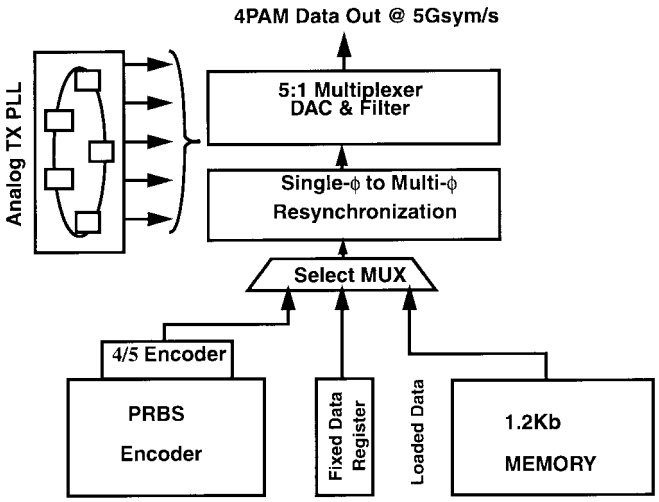


Fig. 4. Transmitter general architecture.

Since 4-PAM hard decision decoding is used in the receiver, a fixed one-to-one mapping of every two input bits to a constellation point must be chosen. Six distinct mappings exist for 4-PAM. However, only a Gray-code mapping (Fig. 3) guarantees that every nearest neighbor symbol error results in only one bit error. Thus, the expected bit error rate (BER) is reduced to that of the linear mapping.

IV. Circuit Implementation

The architecture to achieve the 10-Gb/s transmission rate is shown in Fig. 4. The multiplexing transmitter, comprising five identical drivers, uses ten different clock phases from a five-stage differential ring oscillator (Tx-PLL) to generate the output stream.

Each of the five drivers is composed of four 2-bit DAC modules (Fig. 5). The main module (DAC-M) drives the coax line with a current proportional to one of the four symbol levels, while the three other modules (DAC-Ti) implement the FIR filter.

The 2-bit DAC modules contain two differential driving legs, as shown in Fig. 6. The two driving legs are binary weighted to generate four selectable levels according to the 2-bit input data. This circuit uses D0, D1, and two clocks that are 200 ps (a symbol period) out of phase to generate a precise 200-ps current pulse. Fig. 7 shows the timing for generation of the main symbols. The resynchronizer retimes the 10-bit parallel data into five 2-bit groups. Each group has a different phase to prevent setup-and-hold-time violations for each output driver. On the rising edge of C_{ki}, the differential driving leg (Fig. 6) starts drawing a current based on the input

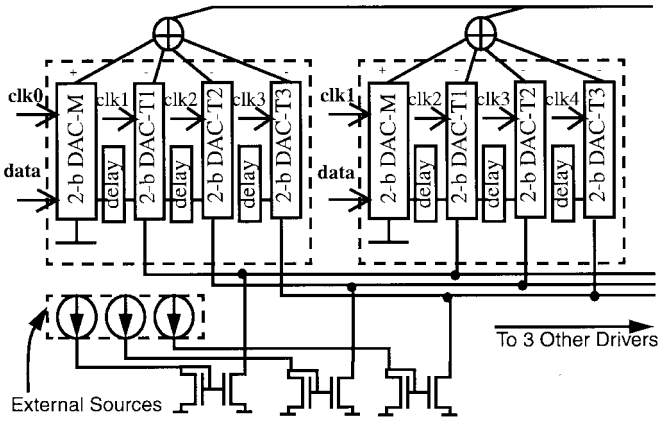


Fig. 5. Multiplexing drivers and three-tap filters.

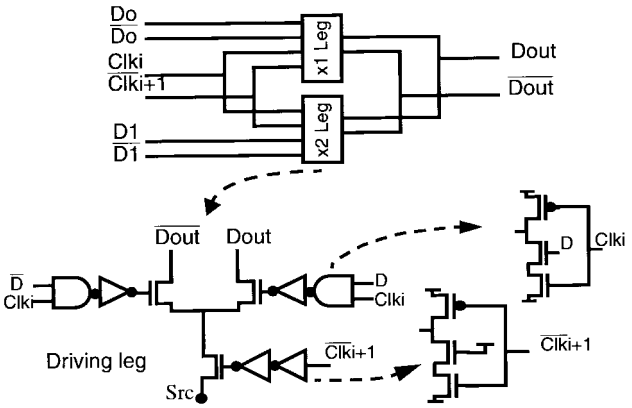


Fig. 6. A 2-bit DAC module.

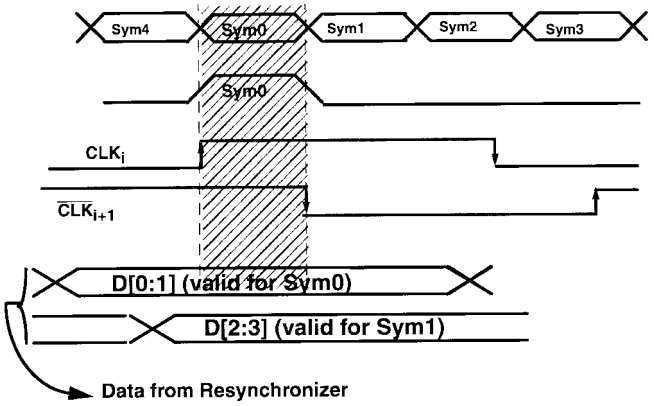


Fig. 7. Main symbol generation timing.

data value D[0; 1]. Signal $C_{\mathrm{ki+1}}$ turns off the driving leg after 200 ps at its falling edge, forming a current pulse (Sym0). The current pulses generated by the five drivers are summed at the output node, generating the 5-Gsym/s stream.

Each of the five driver blocks contains three filter modules (DAC-Ti) that use the same data but different clock inputs (Fig. 5). After the main module's current pulse, the three filter modules turn on consecutively in the next three symbol periods to cancel the tail of the main pulse. Since the preshaping of each main symbol is done at the same driver block by current-summing the modules' pulses at the output pad, no logic is

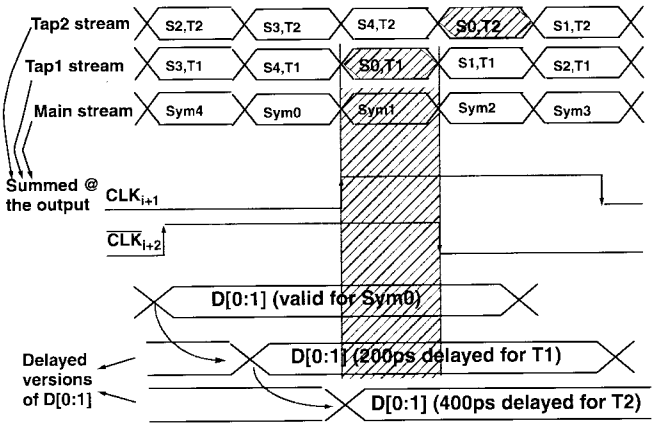


Fig. 8. Tap symbol generation timing.

needed to compute the pre-emphasized signal. Tap symbols are generated with the same mechanism as the main symbol, but instead of a complex resynchronizer, three stages with ∼200-ps delay in each driver are used to guarantee enough setup-and-hold time for the driver input data while passing from one module to the other. Fig. 8 shows this timing for the first two tap symbols.

The currents in the filter taps (tap weights) are modulated by three controllable current sources at the bottom of each module. To protect the tap currents from on-chip noise, each current source is a mirror whose input current is supplied from a clean off-chip source (Fig. 5). Because the corresponding filter modules in each of the five drivers are turned on sequentially, only one of the modules pulls current at each symbol time. Thus, each current source is shared among the five drivers.

The "Src" nodes (Fig. 6) of the main modules' legs are grounded to minimize the device size for a given output current. Smaller device sizes prevent parasitic diffusion capacitances at the output from limiting the overall bandwidth. The "Src" node of the filter modules are connected to the corresponding tap current sources.

One problem with this scheme is that phase errors due to mismatches or jitter can cause one edge to be shifted with respect to others. This shift enlarges one symbol but reduces the next, resulting in a smaller eye opening. Thus, the oscillator elements are designed for low jitter [10], and the driver buffering paths for the different clock phases are precisely matched using identical AND and NOT precharged gates (Fig. 6). The precharged topology also results in sharper transition edges for output symbols and shorter buffering delay, thus reducing the jitter due to the buffering path.

Another potential problem is that variations in the PMOS-to-NMOS strength ratio result in duty cycle error in the clocks. This effect reduces the effective width of the final output symbol, since the symbol boundaries are determined by both the falling and the rising edges of the clocks. Fig. 9 shows two extreme cases of this problem. In case (a), each pulse is longer than the optimum symbol period, thus corrupting the next symbol. In case (b), each pulse itself is shorter than a full period. To combat this problem, the control loop shown in Fig. 10 is employed. Each of the five driving blocks includes

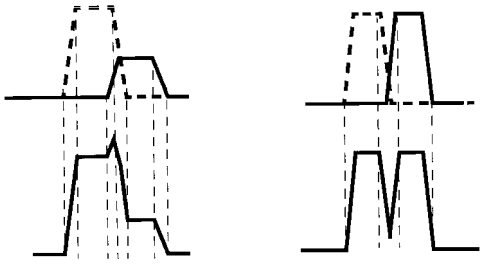


Fig. 9. Two extreme cases of reduced effective symbol width.

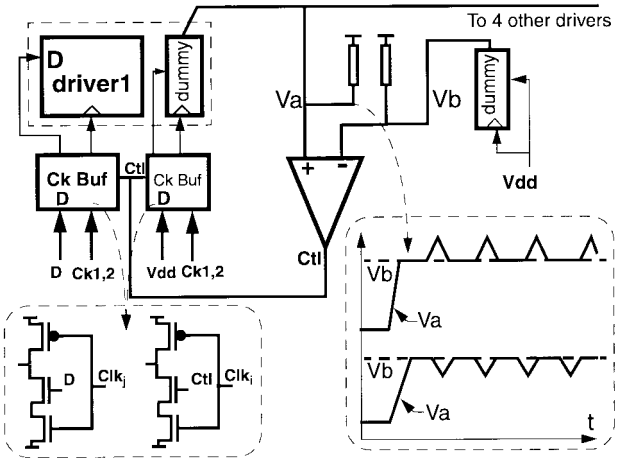


Fig. 10. Symbol-width control loop.

a dummy driver with the same topology and clock inputs as the main driver but is of a smaller size. These five dummy drivers are used to generate current pulses with fixed amplitude that are summed at node Va (in a manner similar to main symbol generation) to form one of the two waveforms shown in Fig. 10. The top waveform occurs when pulses with longer period overlap the next pulses, causing a larger average value, while the bottom waveform shows that shorter pulses result in a smaller average value. The average value at Va is compared to Vb (a reference generated by a matching dummy driver but drawing a fixed dc current) by a comparator that servos a control voltage (Ctrl) to set the symbol width error to zero for both the dummy and the main drivers.

The differential outputs are connected to 50-Ω on-chip PMOS resistors to eliminate line reflections. To achieve good linearity and almost constant 50-Ω termination for varying output voltages, the output devices must remain in saturation to ensure a high output impedance compared to the 50-Ω line. Due to short-channel effects, these devices can have a maximum output swing of 1.1 V (2.2-V differential) while maintaining a total linearity of 2% and an output impedance of 500 Ω ($\gg$50 Ω).

The transmitter also uses bondwire inductors in series with termination resistors to increase the output bandwidth by "shunt peaking." According to [11], the inductance value (L) should be $(R^2 \times C)/2$ to increase bandwidth by a factor of 1.8 with less than 3% frequency response peaking. With a total output capacitance (diffusion and interconnects) of approximately 1.4 pF and a termination resistance of 50 Ω, the

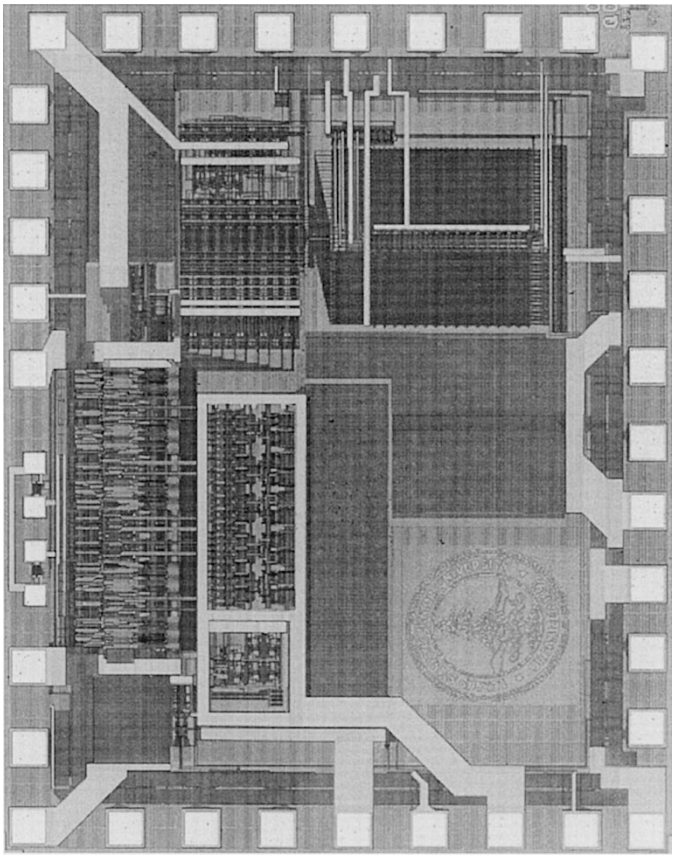


Fig. 11. Chip micrograph.

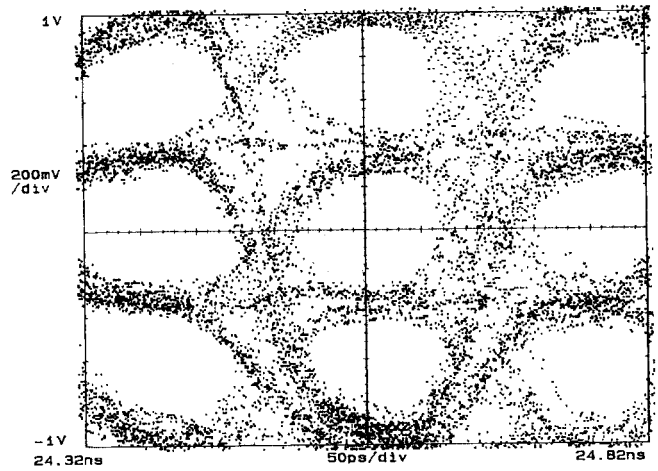
optimum peaking inductance is 1.8 nH, which corresponds to $\sim$2 mm of bondwire.

To facilitate eye-diagram generation and BER measurements, a 2^7-1 pseudorandom bit stream (PRBS) encoder is built on-chip. The 4/5-sym encoder performs line coding for the PRBS sequence. Also, a 1.2-kb memory and a 20-b data register are implemented on-chip, which enables us to load and transmit data patterns of different sizes (Fig. 4).

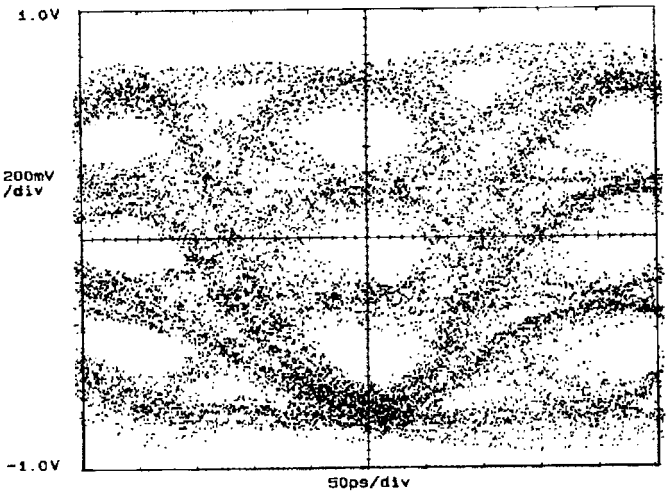
V. Experimental Results

The transmitter chip was implemented in a 0.4-μm CMOS process technology offered through LSI Logic. The die photo of the 2×1.5 mm^2 transmitter chip is shown in Fig. 11. The 4-PAM pre-emphasis driver occupies an area of 0.8×0.3 mm^2. The chip was mounted in a 52-pin ceramic quad flat package, which has internal power planes for controlled impedance and is supplied by Vitesse Semiconductor. The size of the output pads is reduced to 50×50 μm^2 to keep pad capacitance to a minimum to avoid limiting the output bandwidth. To guarantee less than 10% amplitude loss due to output *RC* filtering, the total output capacitance at the 25-Ω I/O (for a doubly terminated 50-Ω line) should not exceed 3.6 pF for a 5-Gsym/s signal. The 5 : 1 multiplexing transmitter has a total capacitance of 1.4 pF, 400 fF of which is due to the pad and metal interconnects.

The controllable filter tap weights allow channel equalization for different cable types and lengths. The transmitter achieves a symbol rate of 5 Gsym/s (10 Gb/s) with a minimum eye height of $>$350 mV and eye width of $>$110 ps over 0.3 m,



(a)



(b)

Fig. 12. Differential data eye at 10 Gb/s: (a) 0.3-m cable and (b) 10-m cable.

and an eye height of $>$200 mV and eye width of $>$90 ps over 10 m of PE142LL coaxial cable after pre-emphasis (Fig. 12). As symbols without pre-emphasis after the 10-m cable show a zero eye opening at 5 Gsym/s, it is clear that ISI mitigation is essential.

We demonstrated in [12] that a hard detector receiver in 0.4-μm CMOS can achieve a BER $< 10^{-14}$ for a 70-mV and 130-ps eye opening. The transmitter has an output jitter of 19 ps (peak to peak) and 3 ps (rms), which is negligible compared to the 200-ps symbol period. Mismatches in the voltage-controlled-oscillator stages and output driver paths increase the total phase error, effectively forming a smaller data eye. At 5 Gsym/s, a maximum phase spacing error of 13 ps ($\sim$7% of the 200-ps symbol width) was measured at transmitter output. Therefore, the 200-ps symbol period is degraded by a maximum of 41 ps due to jitter and phase spacing errors. The chip dissipates a total power of 1.5 W at 5 Gsym/s.

VI. Conclusion

Limitations in both channel bandwidth and process technology present major challenges for multi-gigabit/second com-

munication. Using signal multiplexing, transmit pulse shaping, and multilevel modulation, we can achieve bit rates of 10 Gb/s over distances of 10 m on copper coaxial cables in 0.4-μm CMOS technology despite these limitations. Multiplexing of 1:5 reduces the on-chip clock frequency to one-fifth the symbol rate, or 1 GHz. A three-tap pre-emphasis FIR filter is used to invert the channel low-frequency effects (ISI), while a 4-PAM scheme reduces the symbol rate to half that of a conventional 2-PAM system (binary transmission). This symbol-rate reduction helps lower not only the on-chip clock frequency but also the total data bandwidth, which results in smaller ISI.

Problems with the multiplexing approach are the output capacitance (which increases due to the parallelism) and requirements on the phase position accuracy of the multiple clock phase. Design choices were made to minimize the effect of these problems. A symbol-width control loop is designed to cancel the phase errors due to variations in the PMOS-to-NMOS strength ratio.

ACKNOWLEDGMENT

The authors would like to thank K. Yu, L. Sampson, S. Krishnan, B. Ellersick, B. Amrutur, A. Hajimiri, K. Falakshahi, LSI Logic, and MCC for their assistance.

REFERENCES

[1] W. J. Dally and J. Poulton, "Transmitter equalization for 4 Gb/s signaling," in *Proc. Hot Interconnects Symp.*, Aug. 1996, pp. 29–39.

[2] A. Fiedler, R. Mactagart, J. Welch, and S. Krishnan, "A 1.0625 Gbps transceiver with 2X oversampling and transmit pre-emphasis," in *ISSCC Dig. Tech. Papers*, Feb. 1997, pp. 238–239.

[3] R. Walker, K.-C. Hsieh, T. Knotts, and C.-S. Yen, "A 10Gbps Si-bipolar Tx/Rx chipset for computer data transmission," in *ISSCC Dig. Tech. Papers*, Feb. 1998, pp. 302–303.

[4] E. A. Lee and D. G. Messerschmitt, *Digital Communications*, 2nd ed. Norwell, MA: Kluwer Academic, 1994.

[5] D. A. Johns and D. Essig, "Integrated circuits for data transmission over twisted-pair channels," *IEEE J. Solid-State Circuits*, vol. 32, pp. 398–406, Mar. 1997.

[6] P. J. Black and T. H.-Y. Meng, "A 1-Gbps, four-state, sliding block Viterbi decoder," *IEEE J. Solid-State Circuits*, vol. 32, pp. 797–805, June 1997.

[7] R. Farjad-Rad *et al.*, "An equalization scheme for 4-PAM signaling over long cables," in *Proc. IEEE-CAS Mixed Signal Conf.*, July 1997, pp. 19–22.

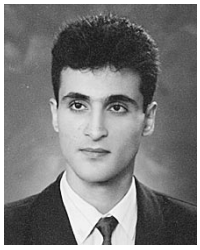
[8] L. Thon, "540 MHz 21 mW MDFE equalizer and detector in 0.25-μm CMOS," in *ISSCC Dig. Tech. Papers*, Feb. 1998, pp. 382–383.

[9] D. Xu, Y. Song, and G. T. Uehara, "A 200 MHz 9-tap analog equalizer for magnetic read channels in 0.6-mm CMOS," in *ISSCC Dig. Tech. Papers*, Feb. 1996, pp. 74–75.

[10] A. Hajimiri and T. Lee, "Jitter and phase noise in ring oscillators," in *Proc. IEEE VLSI Circuits Symp.*, June 1998, pp. 48–51.

[11] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[12] C. K. Yang, R. Farjad-Rad, and M. Horowitz, "A 0.5 μm CMOS 4 Gbps transceiver with data recovery using oversampling," *IEEE J. Solid-State Circuits*, vol. 33, pp. 713–722, May 1998.



Ramin Farjad-Rad (S'95) was born in Tehran, Iran, in 1971. He received the B.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, in 1993 and the M.Sc. degree in electrical engineering from Stanford University, Stanford, CA, in 1995, where he currently is pursuing the Ph.D. degree in electrical engineering.

He was with SUN Microsystems Laboratories, Mountain View, CA, where he worked on a 1.25-Gb/s serial transceiver for the fiber channel standard during the summer of 1995. During the summer of 1996, he was with LSI Logic, Milpitas, CA, where he examined different multi-Gb/s serial transceiver architectures. He has received three U.S. patents.

Mr. Farjad-Rad received the Bronze Medal at the 20th International Physics Olympiad, Warsaw, Poland.

Chih-Kong Ken Yang (S'93–M'98) received the B.S. and M.S degrees in electrical engineering in 1992 and the Ph.D. degree in 1998 from Stanford University, Stanford, CA.

He is an Assistant Professor at the University of California, Los Angeles. His research interests are in the area of VLSI circuit design with emphasis on high-speed interfaces.

Dr. Yang is a member of Tau Beta Pi and Phi Beta Kappa.

Mark A. Horowitz, for a photograph and biography, see p. 528 of the April 1999 issue of this JOURNAL.

Thomas H. Lee (S'87–M'87) received the S.B., S.M., and Sc.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1983, 1985, and 1990, respectively.

He was with Analog Devices Semiconductor, Wilmington, MA, until 1992, where he designed high-speed clock-recovery PLL's that exhibit zero jitter peaking. He then was with Rambus, Inc., Mountain View, CA, where he designed phase- and delay-locked loops for 500-MB/s DRAM's. In 1994, he joined the Faculty of Stanford University as an Assistant Professor, where he is primarily engaged in research into microwave applications for silicon IC technology, with a focus on CMOS IC's for wireless communications. He is the author of *The Design of CMOS Radio-Frequency Integrated Circuits* (Cambridge, U.K.: Cambridge Univ. Press, 1998).

Prof. Lee has twice received the Best Paper Award at the International Solid-State Circuits Conference. He received a Packard Fellowship in 1997.