Transactions Briefs

Joint Equalization and Coding for On-Chip Bus Communication

Srinivasa R. Sridhara, Ganesh Balamurugan, and Naresh R. Shanbhag

***Abstract*—In this paper, we propose using joint equalization and coding to improve on-chip communication speeds by signaling at rates beyond the rate governed by resistance–capacitance (*RC*) delay of the interconnect. Operating beyond the *RC* limit introduces inter-symbol interference (ISI). We mitigate the effects of ISI by employing equalization. The proposed equalizer employs a variable threshold inverter whose switching threshold is modified as a function of past output of the bus. We demonstrate even higher speedups by combining equalization with crosstalk avoidance coding. Specifically, simulation results for a 10-mm 32-bit bus in 0.13-μm CMOS technology show that 1.28$\times$ speedup is achievable by equalization alone and 2.30$\times$ speedup is achievable by joint equalization and coding.**

***Index Terms*—Coding, crosstalk avoidance, delay, equalization, interconnection networks, on-chip buses, system-on-chip (SOC).**

I. Introduction

PROPAGATION delay of global buses acts as performance bottleneck in many system-on-chip (SOC) and network-on-chip (NOC) designs [1]–[5]. Fig. 1 from the 2003 International Technology Roadmap of Semiconductors (ITRS) [1] shows the delay scaling trend with technology. While gate delay reduces with scaling, global wire delay increases. Repeater insertion mitigates the delay to some extent at the cost of additional power dissipation and chip area [1], but the gap between logic and interconnect delay is still expected to increase with scaling.

Coding, and in particular, cross-talk avoidance codes (CAC) [6]–[10], has emerged as an attractive technique for reducing delay *and* power. CAC reduces delay and power by eliminating worst-case coupling transitions in a bus. For example, the worst-case delay of a wire in a bus can be modeled as $(1 + 4\lambda)\tau_0$ [6], where λ is the ratio of coupling capacitance to bulk capacitance and τ_0 is the delay of a crosstalk-free wire. In CAC, bus delay is reduced from $(1 + 4\lambda)\tau_0$ to $(1 + p\lambda)\tau_0$ by reducing the maximum coupling between adjacent wires, where $p = 1$, 2, or 3. The best CACs [9] have shown to provide a speedup of 2.09$\times$, and an energy savings of 12% simultaneously. However, CACs have a large area overhead due to the need for additional bus wires.

Manuscript received September 19, 2006. This work was supported by the Gigascale System Research Center (GSRC), one of five research centers funded under the Focus Center Research Program (FCRP), a Semiconductor Research Corporation program, and Texas Instruments, Inc.

S. R. Sridhara was with the University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA. He is now with DSP Solutions R&D Center, Texas Instruments, Dallas, TX 75243 USA (e-mail: sridhara@ti.com).

G. Balamurugan is with the Circuits Research Laboratory, Intel Corporation, Hillsboro, OR 97124 USA (e-mail: ganesh.balamurugan@intel.com).

N. Shanbhag is with the Coordinated Science Laboratory and the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: shanbhag@uiuc.edu).

Digital Object Identifier 10.1109/TVLSI.2007.915484

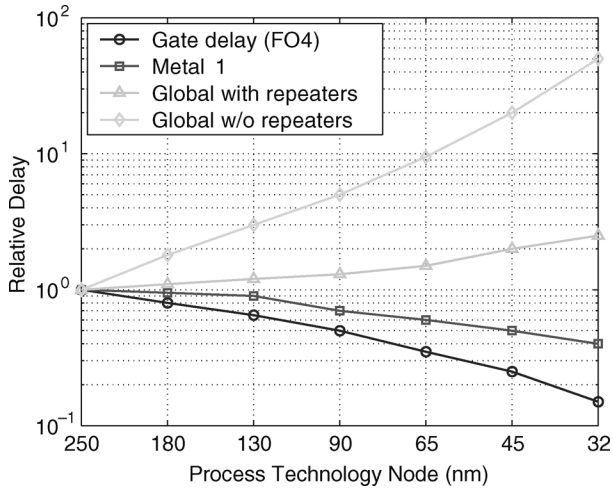


Fig. 1. Gate and wiring delay versus feature size [1].

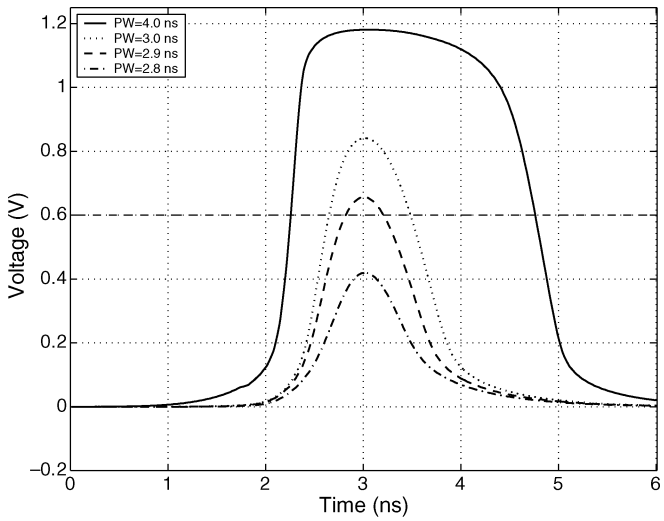


Fig. 2. Reduction of peak amplitude due to ISI for a lone 1 in a 10-mm interconnect with pulse width PW.

On-chip signaling to date has focused on satisfying the resistance–capacitance (*RC*) delay limit, which can be approximated by the Elmore delay of *RC* network composed of driver, interconnect, and receiver. Clocking a bus faster than the *RC* limit leads to inter-symbol interference (ISI) [11]. ISI spreads out data pulses and reduces their peak magnitude resulting in bit errors at the receivers. In order to gain more insight into the behavior of ISI, consider the transmission of a lone pulse, i.e., a single 1 among a series of 0's or a single 0 among a series of 1's over an *RC*-dominated interconnect. Fig. 2 plots the received pulses for a lone 1 with various pulse widths PW for a 10-mm interconnect with CMOS inverters as driver and receiver in a 0.13-μm CMOS process. We observe that the peak amplitude reduces with reduction in PW. Further, the peak amplitude goes below the decision threshold of 0.6 V for $PW < 2.9$ ns indicating that errors will occur. Thus, current day systems employ larger pulse widths in order to ensure reliable signaling in the presence of DSM noise and clock jitter, thereby limiting the achievable data rate.

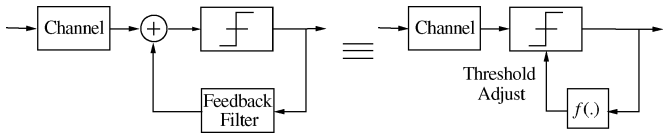
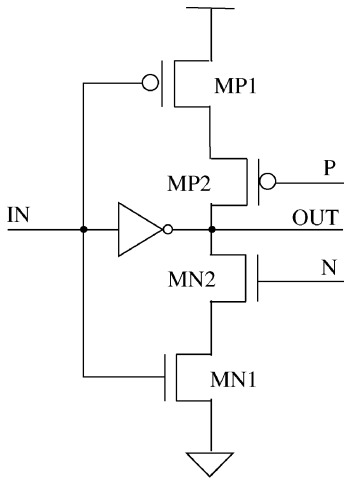


Fig. 3. DFE.



P	N	V_{th}
GND	VDD	V_{th}^{0}
GND	GND	V_{th}^{+}
VDD	VDD	V_{th}^{-}

Fig. 4. Variable threshold inverter.

In this paper, we propose the use of equalization for signaling beyond the *RC* limit and demonstrate its utility in the context of global signaling in 0.13-μm CMOS process. For a 32-bit, 10-mm bus, we show that equalization alone provides a speedup of 1.28× with 1% area overhead over an unequalized and uncoded bus. We further propose the use of joint equalization and coding where a speedup of 2.3× is achieved with 69% area overhead, where most of the area overhead is due to coding. To the best of our knowledge, this is the first work that proposes equalization and joint equalization and coding for on-chip signaling. Note that several chip input/output (I/O) signaling systems have begun to employ equalization for data rates greater than 3.125 Gb/s [12] and that equalization is expected to remain the workhorse for future I/O systems. Equalization and, more recently, joint equalization and coding have already been employed to enhance data rates in long-haul and macro networks. This work extends the applicability of equalization into the on-chip domain and is part of the general trend of retargetting communication-theoretic techniques to on-chip systems in order to tradeoff delay/power with reliability.

II. Equalization

In communication and I/O signaling systems, equalizers are implemented using filters. These filters mitigate ISI by canceling the effect of past and future bits on the current bit. Fig. 3 shows a decision-feedback equalizer (DFE) [11], where the effect of past bits (or decisions) on the current received pulse is canceled by employing a feedback filter. The output of the subtracter is fed to a slicer in order to determine the transmitted bit. This operation is equivalent to modifying the threshold of the slicer based on the past bits as shown in the figure. Such a DFE requires multiply-and-accumulate circuits and is impractical for on-chip buses. Instead, we modify the threshold based on past bits using a variable threshold inverter shown in Fig. 4.

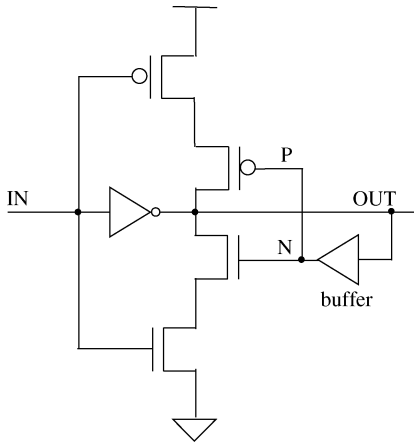


Fig. 5. Equalizer for a single wire.

In Fig. 4, the threshold voltage of the inverter is controlled by using signals P and N. When $P = \mathrm{GND}$ and $N = V_{\mathrm{DD}}$, the threshold voltage is V_{th}^{0}, which is the nominal threshold voltage of the inverter. When $P = \mathrm{GND}$ and $N = \mathrm{GND}$, the pull down path is off and the threshold voltage increases to V_{th}^{+}. Similarly, when $P = V_{\mathrm{DD}}$ and $N = V_{\mathrm{DD}}$, the pull up path is off and threshold voltage decreases to V_{th}^{-}. As shown in Fig. 4, a weak inverter is required to ensure that the output of the inverter is never floating. The relative sizing of transistors MP1, MP2, MN1, and MN2 determines the values of V_{th}^{+} and V_{th}^{-}. Further, there exists a tradeoff between the range of variability in the threshold voltage and susceptibility of the inverter to DSM noise due to change in the threshold voltage. We will address this issue in greater detail in Section II-B.

A. Equalizer for a Single Wire

For a single wire, worst-case delay occurs whenever the wire switches. Therefore, the threshold voltage of the inverter should be adjusted such that it anticipates a transition on the wire. If $\mathrm{OUT} = V_{\mathrm{DD}}$ ($\mathrm{IN} = \mathrm{GND}$), then the threshold voltage should be lowered ($P = V_{\mathrm{DD}}, N = V_{\mathrm{DD}}$) such that the delay of a $\mathrm{GND} \rightarrow V_{\mathrm{DD}}$ transition on the wire is reduced. Similarly, if $\mathrm{OUT} = \mathrm{GND}$, then the threshold voltage should be increased ($P = \mathrm{GND}$, $N = \mathrm{GND}$) such that the delay of a $V_{\mathrm{DD}} \rightarrow \mathrm{GND}$ transition on the wire is reduced. Therefore, the control signals P and N are

$$P = N = \mathrm{OUT}. \tag{1}$$

Fig. 5 shows the equalizer circuit for a single wire. A buffer is used to drive P and N nodes from OUT. The buffer is designed such that its delay is greater than delay from IN to OUT. Thus, the OUT node is able to charge or discharge before P and N signals change.

B. Equalizers for Buses

In a bus, the worst-case delay occurs when both neighbors of a switching wire switch in the opposite direction. In other words, the middle wire in a set of three wires will have the worst-case delay when the wires transition from 101 to 010 or vice versa. Therefore, the equalizer for buses needs to utilize the past bits of all three wires when adjusting the threshold voltage for the middle wire. This is achieved by modifying the threshold voltage for the middle wire only when the past output of the set of three wires is either 101 or 010. When the past output is 101, the threshold voltage is raised to V_{th}^{+} and when the past output is 010, the threshold voltage is lowered to V_{th}^{-}. If the past output is neither 101 nor 010, then the threshold voltage is V_{th}^{0}.

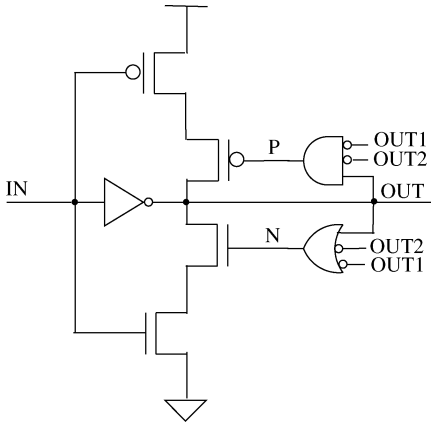


Fig. 6. Equalizer for buses.

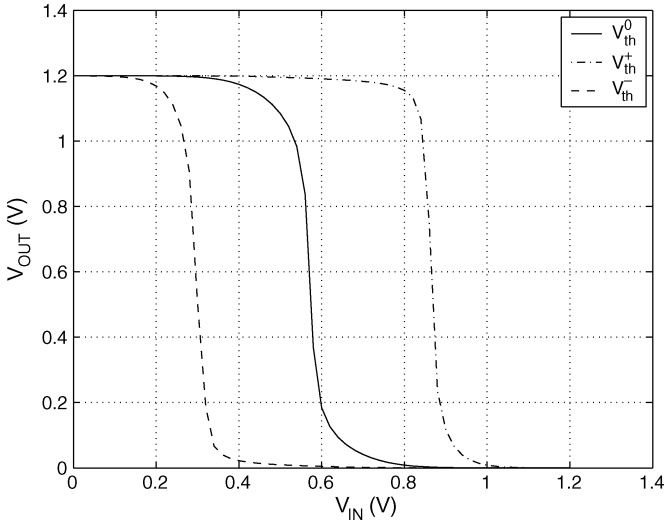


Fig. 7. DC transfer characteristics of the equalizer.

The previous settings of threshold voltage can be obtained by setting $P = V_{\rm DD}$ when the past output is 010 and $P = {\rm GND}$ otherwise, and by setting $N = {\rm GND}$ when the past output is 101 and $N = V_{\rm DD}$ otherwise. Therefore

$$P = \overline{\rm OUT1} \cdot {\rm OUT} \cdot \overline{\rm OUT2} \tag{2}$$

$$N = \overline{\rm OUT1} + {\rm OUT} + \overline{\rm OUT2} \tag{3}$$

where ${\rm OUT1}$ and ${\rm OUT2}$ are the outputs of the two adjacent wires.

Fig. 6 shows the equalizer circuit for buses. As shown in Fig. 6, the control signals P and N are obtained by using AND and OR gates, respectively. Once again, the delays of the control logic gates are designed to be greater than IN-to-OUT delay so that the OUT node is able to charge or discharge before P and N change. In an equalized bus, all nonboundary wires will have the equalizer as receivers, while the boundary wires can use CMOS inverters. This is because a boundary wire has only one neighbor and, hence, experiences much lower crosstalk.

Modifying the threshold voltage of the receiver increases its susceptibility to noise, in general. We argue next that the robustness of the proposed equalizer is not reduced in any way if crosstalk noise is the dominant noise source. For example, in a crosstalk noise dominated scenario, the middle wire in a set of three wires will experience a glitch whenever a 000 $\rightarrow$ 101 transition occurs. If the receiver threshold for such a wire is lower than the nominal, then the potential for an error increases. The proposed equalizer lowers (raises) the threshold when the past output is 010 (101). A glitch will be introduced if and only if the middle wire remains quiet and one or both of the adjacent wires make a transition from 0 $\rightarrow$ 1 (1 $\rightarrow$ 0). However, such a glitch will make the middle wire voltage to increase (decrease) beyond $V_{\rm DD}$ (GND). Hence, such a glitch will not affect the robustness of the circuit. However, the modification of threshold voltage can make the bus prone to errors due to other sources of noise. Thus, the proposed technique is an example of the fundamental tradeoff between delay and reliability.

The proposed equalizer can reduce the bus delay to $(1 + 3\lambda)\tau_0$ at best. This is because the threshold remains at the nominal level for all past outputs except 010 and 101. Therefore, when the past output is 110, the threshold voltage is V_{th}^0. If a 110 $\rightarrow$ 101 transition occurs, then the middle wire has a delay of $(1 + 3\lambda)\tau_0$. Since the proposed equalizer mitigates the effect of ISI only after the worst-case transition has occurred, its performance is bounded by $(1 + 3\lambda)\tau_0$.

The worst-case bus delay can be also reduced to $(1 + 3\lambda)\tau_0$ by employing crosstalk avoidance codes referred to as forbidden overlap codes (FOC) [10]. In FOC, a codeword having bit pattern 010 does not transition to another codeword having the bit pattern 101 and vice versa. This is achieved by employing additional bus wires and codec circuits. For example, 40 wires are required to implement a coded 32-bit bus [10]. Compared to FOC, equalization requires very small area overhead. Equalization does not require additional wires and the receiver circuits require small area overhead compared to FOC as we show in Section IV.

III. Joint Equalization and Coding

Crosstalk avoidance codes can also reduce the bus delay to $(1 + 2\lambda)\tau_0$ by either avoiding opposing transitions on adjacent wires [8] or avoiding specific data patterns [7]. Crosstalk avoidance codes proposed in [8] reduce the worst-case delay by ensuring that a transition from one codeword to another does not cause adjacent wires to transition in opposite directions. This constraint on data transitions is known as forbidden transition condition [9]. A forbidden transition code (FTC) satisfies the forbidden transition condition.

The forbidden transition condition ensures that a transitioning wire has its neighbors either quiet or switching in the same direction. When at least one of the neighbors switches in the same direction, the delay of the wire is less than $(1+2\lambda)\tau_0$. The worst-case delay of $(1+2\lambda)\tau_0$ occurs only when a transitioning wire has both its neighbors quiet. This implies that the worst-case delay with FTC can occur in all possible states of the past outputs of the neighbors. Therefore, the equalizer for the coded bus needs to consider only the past bit of the wire. Hence, the equalizer for a coded bus is same as the equalizer for a single wire shown in Fig. 5. Thus, crosstalk avoidance coding provides speedup and simplifies the equalizer by eliminating opposing transitions in adjacent wires.

IV. Simulation Results

In this section, we present simulation results to demonstrate the improvement in delay achieved by employing equalization and coding. The achieved improvements vary with bus length L, ratio of coupling capacitance to bulk capacitance λ, and the process technology. We quantify the improvements in a standard 0.13-μm CMOS technology for various values of L and λ.

We consider a metal four bus with L in the range of 6–14 mm with minimum width of 0.2 μm and minimum spacing of 0.2 μm. For a given bus geometry, the value of λ depends on the metal coverage in upper and lower metal layers. We vary λ between the following two extreme scenarios. First, 100% metal coverage is assumed in metal layers 3 and

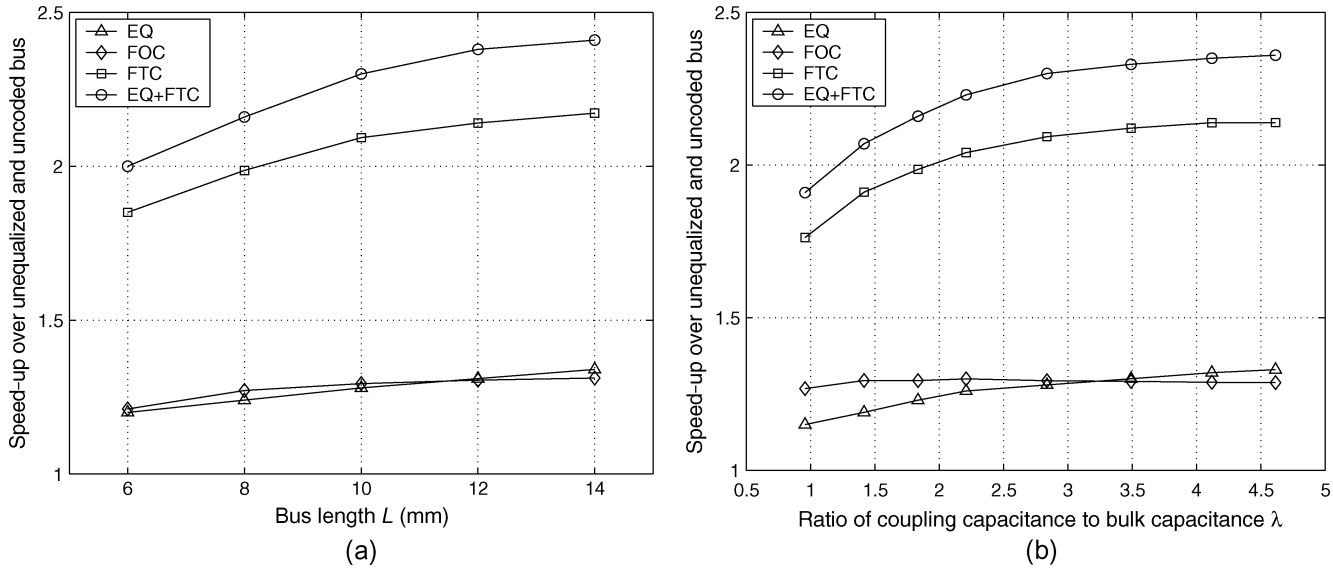


Fig. 8. Speedup over unequalized and uncoded 32-bit bus: (a) as a function of bus length L at $\lambda = 2.8$ and (b) as a function of λ at $L = 10$ mm.

5, resulting in $\lambda = 0.95$. Second, all the bulk capacitance is assumed to be from metal 4 to the substrate, resulting in $\lambda = 4.6$.

Encoder and decoder for CAC are synthesized using Synopsys Design Compiler with a 0.13-μm CMOS standard cell library and optimized for speed. The overhead required in terms of area and delay is obtained from synthesized gate level netlists.

The baseline for comparison is an unequalized and uncoded bus with optimally sized CMOS inverters as drivers and receivers. The equalizer sizing is obtained by a simulation-based methodology using HSPICE that optimizes the bus delay. Buses are modeled as *RLC* transmission lines in HSPICE. Fig. 7 shows the dc transfer characteristics of the equalizer designed for a 10-mm bus. A wide range variability in threshold voltage is achievable by appropriately sizing the transistors of the variable threshold inverter. If the output of wire has a slow ramp due to large capacitance, then significant speedup is achievable by appropriately modifying the threshold.

Fig. 8 plots the speedup achieved by equalization (EQ), forbidden overlap coding (FOC), forbidden transition coding (FTC), and joint equalization and FTC (EQ+FTC). Speedup is defined as the ratio of the delay of unequalized and uncoded bus to the delay of the proposed bus. Note that the delay of the proposed bus includes the delay of equalizer, encoder, and decoder. The speedups achieved increase in all cases with L as shown in Fig. 8(a). The effectiveness of equalization improves with L as longer buses have slower ramps. The effectiveness of coding improves with L as the coding delay forms smaller portion of the total bus delay for longer buses.

The speedups achieved increases with λ, as shown in Fig. 8(b). This is because the benefits of employing equalization and coding increase with crosstalk, which in turn increases with λ. For 10-mm bus at $\lambda = 2.8$, equalization, FOC, FTC, and joint equalization and FTC achieve speedups of 1.28×, 1.29×, 2.09×, and 2.30×, respectively. As expected, equalization and FOC provide approximately the same speedup as both reduce the bus delay to about $(1 + 3\lambda)\tau_0$. However, equalization provides speedup at the cost of a small area overhead, while coding requires large area overhead as shown next.

Fig. 9 shows the tradeoff between speedup and area overhead for a 10-mm 32-bit bus. The area overhead for equalization is due to more complex receiver circuits. However, coding requires additional wires along with encoder and decoder circuits. This wiring overhead along with codec circuits translates into 28% area overhead over the uncoded

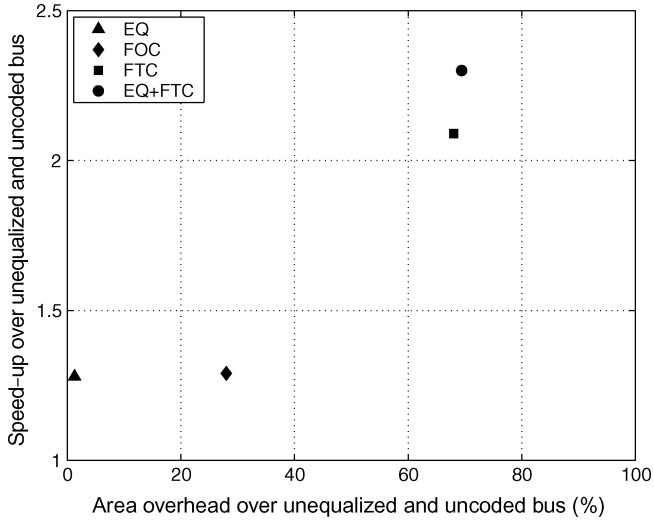


Fig. 9. Speedup versus area overhead for a 10-mm 32-bit bus with $\lambda = 2.8$.

bus for FOC as shown in Fig. 9. The area overhead for equalization is about 1%. Thus, equalization provides the same speedup as FOC but at a much lower area overhead. Higher speedups are achieved by trading off more area via FTC. The use of FTC alone provides a speedup of 2.09× with 68% area overhead. The speedup improves to 2.3× when FTC is combined with equalization and the area overhead increases to 69%. Thus, equalization is an effective technique to improve throughput.

V. Conclusion

We have proposed joint equalization and coding as a method to counter the trend of increasing gap between gate and interconnect delay. We have shown that equalization provides 1.28× speedup at the cost of small power overhead, while joint equalization and coding provides 2.30× speedup at the cost of large area overhead and no power penalty. Technology scaling will improve the performance of the proposed techniques. The increasing value of coupling capacitance will increase crosstalk and, hence, benefits of employing equalization and crosstalk avoidance coding will improve. Further, codec delay,

area, and power dissipation will reduce with scaling and, hence, the performance of the coded bus will improve.

The equalizer proposed in this paper employs just the past bits of three adjacent wires to adjust the threshold of the receiver. More complex equalizers can be envisaged that utilize several past outputs and/or outputs of nonimmediate neighbors. Technology scaling will enable and necessitate the use of such receivers in the future to achieve high-speed on-chip bus communication.

REFERENCES

[1] Semiconductor Industry Association, San Jose, CA, "International technology roadmap for semiconductors," 2003 [Online]. Available: http://public.itrs.net

[2] R. Ho, K. W. Mai, and M. A. Horowitz, "The future of wires," *Proc. IEEE*, vol. 89, no. 4, pp. 490–504, Apr. 2001.

[3] L. Benini and G. D. Micheli, "Networks on chips: A new SoC paradigm," *IEEE Computer*, vol. 35, no. 1, pp. 70–78, Jan. 2002.

[4] D. Pamunuwa, L.-R. Zheng, and H. Tenhunen, "Maximizing throughput over parallel wire structures in the deep submicrometer regime," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 11, no. 4, pp. 224–243, Apr. 2003.

[5] A. Nalamalpu, S. Srinivasan, and W. P. Burleson, "Boosters for driving long onchip interconnects—Design issues, interconnect synthesis, and comparison with repeaters," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 21, no. 1, pp. 50–62, Jan. 2002.

[6] P. P. Sotiriadis and A. Chandrakasan, "Reducing bus delay in submicron technology using coding," in *Proc. ASP-DAC*, 2001, pp. 109–114.

[7] C. Duan, A. Tirumala, and S. P. Khatri, "Analysis and avoidance of cross-talk in on-chip buses," in *Proc. Hot Interconnects*, 2001, pp. 133–138.

[8] B. Victor and K. Keutzer, "Bus encoding to prevent crosstalk delay," in *Proc. ICCAD*, 2001, pp. 57–63.

[9] S. R. Sridhara, A. Ahmed, and N. R. Shanbhag, "Area and energy-efficient crosstalk avoidance codes for on-chip buses," in *Proc. ICCD*, 2004, pp. 12–17.

[10] S. R. Sridhara and N. R. Shanbhag, "Coding for reliable on-chip buses: Fundamental limits and practical codes," in *Proc. VLSI Des.*, 2005, pp. 417–422.

[11] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 2000.

[12] J. L. Zerbe *et al.*, "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.

Improved Accuracy Pseudo-Exponential Function Generator With Applications in Analog Signal Processing

Cosmin Popa

***Abstract*—A new CMOS current-mode pseudo-exponential circuit based on the n-order Taylor series expansion will be presented. The most important advantage of the circuit with respect to the previously reported similar ones is the smaller value of the total computing error (under 0.3 dB), for a maximal output range of the proposed function generator greater than 40 dB. The total error could be very easily reduced by increasing the number of terms considered in the Taylor expansion. The circuit also presents the advantage of the independence of the output current on technological parameters. The frequency response is improved due to the strong inversion operation of all MOS transistors and to the current-mode operation of the circuit. The circuit area is relatively small due to the exclusively utilization of MOS transistors. The SPICE simulations confirm the theoretical estimated results. The proposed exponential function generator is designed in 0.12-μm CMOS technology and it consumes a reasonable power (less than 0.3 mW) for obtaining the previous mentioned computing error and has a low-voltage operation (a minimal accepted supply voltage under 1.2 V). The total silicon occupied area of the exponential function generator with third-order approximation is about 5.9 μm $\times$ 7.9 μm.**

***Index Terms*—Analog signal processing, computational circuits, function generator, second-order effects, total computing error.**

I. INTRODUCTION

Exponential circuits represent important building blocks for telecommunication applications, medical equipment, hearing aid, disk drives, and, especially, for analog signal processing.

The exponential function could be easily obtained in the bipolar technology from the exponential characteristic [1], [2] of the bipolar transistor. The nonzero value of the base current, especially for p-n-p transistors and the temperature dependence of the bipolar transistor parameters (the thermal voltage is linear increasing with temperature and the saturation current has an exponential dependence on temperature) introduces relatively large errors in obtained exponential characteristic.

In CMOS technology, the exponential law is available only for the weak inversion operation of the MOS transistor. The great disadvantage of the computational circuits using MOS transistors in weak inversion [3] is the poor circuit frequency response caused by the much smaller drain currents available for charge and discharge of the parasitic capacitances of the MOS transistors. Thus, circuits realized in CMOS technology that require a good frequency response can be designed using only MOS transistors working in strong inversion (usually in saturation). In order to obtain the exponential function using the square characteristic of the MOS transistor in saturation, the original idea is to approximate the exponential function with its nth-order expansion (the polynomial series). The approximation error is proportional with the number of terms neglected in the expansion.

In this paper, an original CMOS VLSI current-mode pseudo-exponential circuit will be presented based on an excellent approximation of the exponential function by its n-order limited Taylor series expansion. The obtained exponential function has the important advantage of the independence of the output current on technological parameters, thus the total computing error being strongly reduced. The circuit accuracy could be additionally increased by considering superior-order terms from the exponential function expansion, a tradeoff between accuracy and circuit complexity being necessary to be made. The performances of the original proposed implementation of the exponential function generator are better than the performances of the previous reported similar circuits: the maximal output range is greater than 40 dB, comparing with 15 dB for [4], 35 dB for [5], and approximately 30 dB for [6] and [7], while the total computing error could be decreased under 0.3 dB for the new circuit, less then 0.5 dB for [4] and [5] and 1 dB for [6]. The circuit from [8] has a voltage-mode operation and the [9] circuit has a digital implementation, having, both of them, the disadvantage of an important reduction of the computing rate; the current-mode operation of the original proposed exponential function generator, correlated with the strong inversion biasing of all MOS active devices assures a very good circuit frequency response for a reasonable power consumption (less than 0.3 mW) and a low-voltage operation (a minimal accepted supply voltage under 1.2 V).

Manuscript received December 16, 2005; revised March 21, 2007 and May 6, 2007.

The author is with the Faculty of Electronics, Telecommunications and Information Technology, University Politehnica of Bucharest, Bucharest 060042, Romania (e-mail: cosmin_popa@yahoo.com).

Digital Object Identifier 10.1109/TVLSI.2007.915495