

A Transceiver for High-Speed Global On-Chip Data Communication

Daniël Schinkel, Eisse Mensink, Eric Klumperink, Ed van Tuijl and Bram Nauta
IC-Design Group, University of Twente
P.O.Box 217, 7500 AE Enschede, The Netherlands
E-mail: d.schinkel@utwente.nl

Abstract— Global on-chip data communication is becoming a concern as the gap between transistor speed and interconnect bandwidth increases with CMOS process scaling. In this paper we show how a special form of equalization, pulse-width pre-emphasis, can significantly increase the data rate for a given length of uninterrupted interconnect, especially if used in combination with low-ohmic termination and twisted differential interconnects. To validate these techniques, a bus-transceiver test chip in a 0.13 μm , 1.2V, 6M copper CMOS process has been designed. The chip uses 10mm long differential interconnects with wire widths and spacing of only 0.4 μm . With transceivers operating in conventional mode, the chip achieves only 0.55Gb/s/ch. The achievable data rate increases to 3Gb/s/ch (consuming 2pJ/bit) if pulse-width pre-emphasis and low-ohmic termination are turned on.

Keywords— on-chip communication; interconnect; data bus; intersymbol interference (ISI); pulse-width; pre-emphasis; transceivers

I. INTRODUCTION

On-Chip communication is getting more attention, as (global) interconnects are rapidly becoming a speed, power and reliability bottleneck for digital CMOS systems [1]. Technological advances such as copper interconnects and low-k dielectrics are by itself not sufficient to let the interconnect bandwidth keep up with the advances in transistor speeds.

From a circuit design perspective, a general solution to the limited interconnect bandwidth is the use of repeaters, which make the repeated wire delay linear with length instead of the quadratic dependency of an unrepeatable wire [2]. However, the number of repeaters should be kept to a minimum as they cost area and power and make floorplanning more difficult as portions of active area all over the chip have to be reserved for large repeater circuits. Furthermore, the classical approach to repeater insertion [2], using plain buffers/inverters as

repeaters has serious limitations for global communication. With plain non-clocked buffers as repeaters, delay optimization requires closely spaced repeaters and delay variations due to crosstalk and due to process variations will accumulate and limit the achievable data rate.

These arguments motivate the search for more advanced solutions that can increase the data rate for a given length or can increase the unrepeatable wire length for a given data rate, preferably in combination with a decrease in crosstalk sensitivity and power consumption.

The authors of [3] propose to use low-swing signaling over differential 10mm aluminum interconnects, but with the requirement of clocked switches along the wire, increasing the already troublesome clock load. In [4] it is proposed to use 16 μm wide differential wires (20mm long) and exploit the LC regime (transmission-line behavior) of these wires, but at the expense of a significant increase in power consumption and interconnect area. Both papers achieve 1Gb/s/ch in a 0.18 μm CMOS technology.

Low-ohmic termination, optimized interconnect dimensioning and equalization are other techniques to increase the achievable data rate, as discussed in [5]. In a recent paper [6], we have shown that a novel form of pre-emphasis (pulse-width pre-emphasis) is a suitable form of equalization. With the pre-emphasis used in combination with resistive termination and twisted differential interconnects, 3Gb/s/ch is achieved over 10mm of uninterrupted interconnects.

In this paper we discuss the motivations behind the presented transceiver, with a focus on pulse-width pre-emphasis. Section II discusses the design choices for the interconnect itself. Section III describes the pulse-width pre-emphasis technique. Section IV explains the implementation and section V shows the results. Finally, the conclusions are given in section VI.

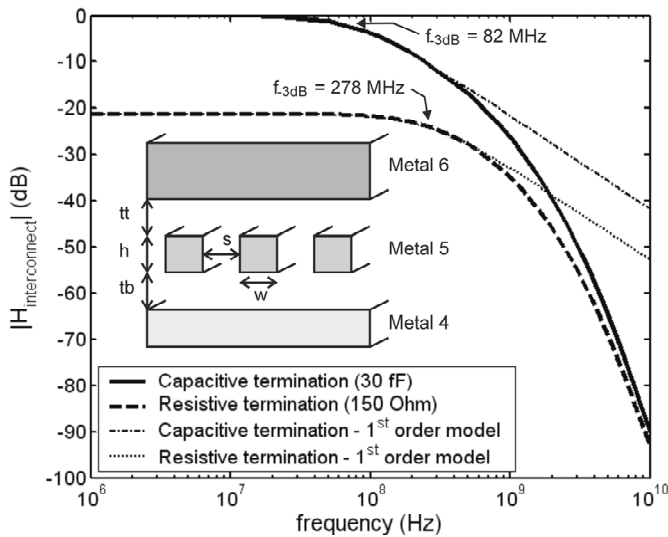


Figure 1. Interconnect structure and transfer functions

II. INTERCONNECT TOPOLOGY

The interconnects, as shown in Figure 1, are modeled with a 3D EM-field solver and a distributed RLC model is extracted (0.15k Ω /mm, 0.35nH/mm and 0.27pF/mm). The bus is placed in metal 5 as it is assumed that the thick top-metal is reserved for clock and power routing. In the EM-field solver, metal 4 and metal 6 plates approximate the effect of other high-density interconnects. The dimensions of the interconnects are optimized for highest bandwidth per cross-sectional area. Analysis and simulations show that the bandwidth per cross-area peaks when all dimensions (w , s , h , tt and tb) are equal. This results in both a width and a spacing of 0.4 μ m.

Figure 1 also shows the simulated interconnect transfer function [5]. For these long and narrow interconnects the effect of inductance is negligible, as the inductance only begins to play a role for frequencies where $\omega L \geq R$, which in this case is true for frequencies >68 GHz (with a heavily attenuated transfer). A significant part of the transfer function can be approximated by a first-order RC model, as visible in the figure. Note that the bandwidth increases 3 times with low-ohmic resistive termination instead of (conventional) capacitive termination [7].

Differential interconnects are used to reduce the overall crosstalk [8]. Furthermore, to cancel neighbor to neighbor crosstalk between channels in a bus, one twist is placed at 50% of the length in the even channels and two twists are placed at 25% and 75% in the uneven channels, as shown in Figure 2. These positions are nearly ideal (not completely ideal due to layout constraints) for crosstalk cancellation provided that symmetric line termination is used [9].

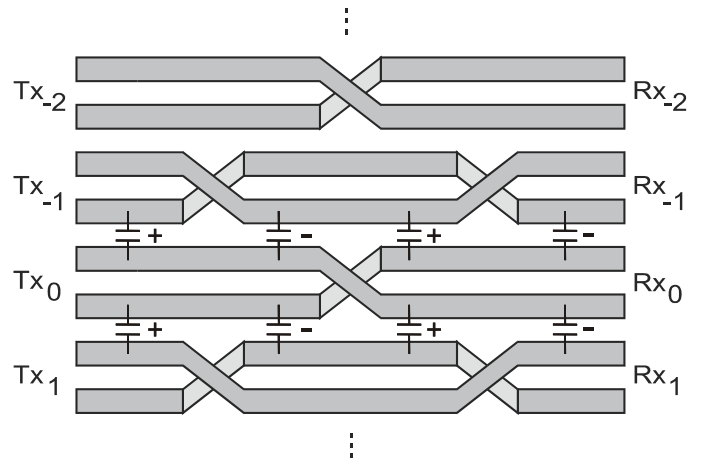


Figure 2. Twisted differential bus

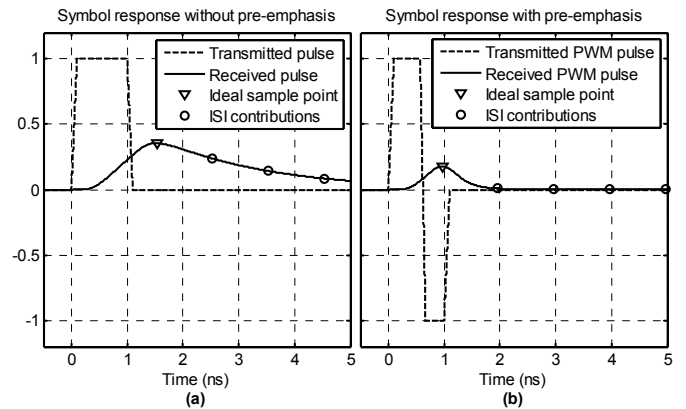


Figure 3. Symbol response of a (capacitively terminated) 10mm interconnect with 1ns symbol period

III. PULSE-WIDTH PRE-EMPHASIS

The dominance of the first-order roll-off makes an on-chip interconnect very suitable for simple pre-emphasis transmission schemes (e.g. 2 taps FIR). Conventional pre-emphasis schemes (overdrive signaling) used for inter-chip communication [10] are less suitable for on-chip implementation. The large drive impedance of simple current-summing transmitters degrades interconnect bandwidth, while low-ohmic voltage transmitters require additional low-impedance voltage levels and their performance is degraded by slew-rate.

As a robust alternative, the use of pulse-width (PW) pre-emphasis is proposed in this paper. As shown in Figure 3, PW pre-emphasis can greatly reduce the amount of inter-symbol interference (ISI), by using the second part of the symbol-time to compensate for the remaining line charge.

A way to analyze ISI is to look at the response of a channel to a single isolated symbol, as explained in [5]. For first-order low-pass channels, PW pre-emphasis can completely cancel the ISI as shown in Figure 4.

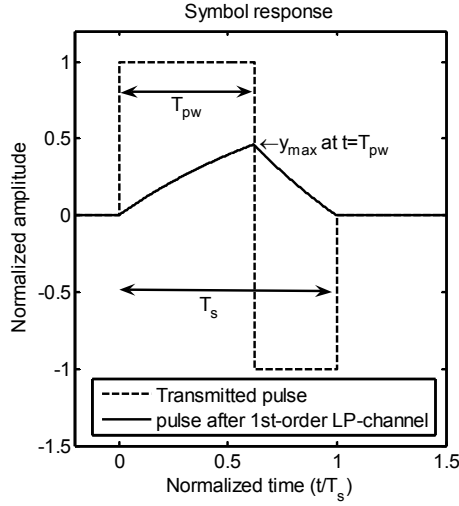


Figure 4. PW pre-emphasis for a first-order channel

The required pulse-width for zero ISI is a function of the symbol-time (T_s) and of the time constant of the channel (τ_{ch}) and can be found by writing the symbol response as a summation of three step responses:

$$x(t) = step(t) - 2step(t - T_{pw}) + step(t - T_s) \quad (1)$$

The step response terms of a first-order channel are simple exponential functions (valid from the start of the step):

$$y(t) = \left(1 - e^{-\frac{t}{\tau_{ch}}}\right)_{t \geq 0} - 2 \left(1 - e^{-\frac{t - T_{pw}}{\tau_{ch}}}\right)_{t \geq T_{pw}} + \left(1 - e^{-\frac{t - T_s}{\tau_{ch}}}\right)_{t \geq T_s} \quad (2)$$

The response will be zero for $t \geq T_s$, meaning no ISI, on the following condition:

$$e^{-\frac{t}{\tau_{ch}}} \left(-1 + 2e^{\frac{T_{pw}}{\tau_{ch}}} - e^{\frac{T_s}{\tau_{ch}}}\right) = 0 \rightarrow \frac{T_{pw}}{\tau_{ch}} = \ln\left(\frac{1}{2} + \frac{1}{2}e^{\frac{T_s}{\tau_{ch}}}\right) \rightarrow$$

$$PW = \frac{T_{pw}}{T_s} = \ln\left(\frac{1}{2} + \frac{1}{2}e^{\frac{T_s}{\tau_{ch}}}\right) \frac{\tau_{ch}}{T_s} \quad (3)$$

So the ideal pulse-width is a function of the ratio between T_s and τ_{ch} . The maximum value of the response is found at $t=T_{pw}$ (which is hence the ideal detection instant) and can be found by substituting (3) into the first term of (2):

$$y_{\max} = y(T_{pw}) = 1 - \left(\frac{1}{2} + \frac{1}{2}e^{\frac{T_s}{\tau_{ch}}}\right)^{-1} \quad (4)$$

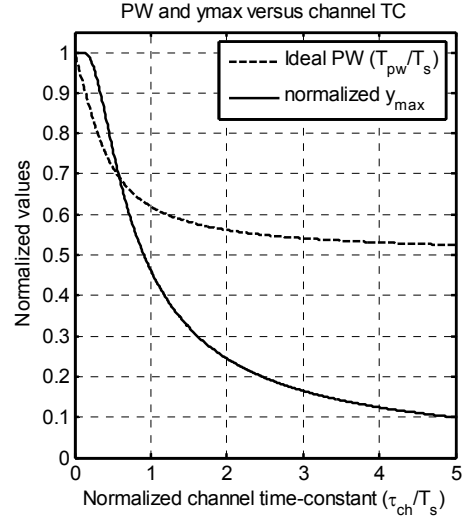


Figure 5. PW pre-emphasis properties for a first-order channel

For symbol times much smaller than the channel time constant, a high amount of de-emphasis is needed and the ideal pulse-width approaches 50% while y_{\max} approaches zero as shown in Figure 5. As an example, the ideal pulse-width for 2Gb/s data rate, with a channel-corner frequency of 80MHz ($\tau_{ch}/T_s = 2\text{ns}/0.5\text{ns} = 4$) is 53% and the receiver swing is only 12% of the transmitter swing.

To analyze the achievable data rate with higher-order channel transfer functions, a high-level numerical eye-diagram analysis can be carried out. A symbol response can be derived from the step-response of an actual distributed wire and the ideal sample instant and the amount of ISI can be analyzed for various pulse-widths, similar to the procedure described in [5]. The higher-order effects of the actual line-transfer give a slight deviation from the results shown in Figure 5 at high ratios of τ_{ch}/T_s and a residue of ISI will be present. At moderate ratios, the results are not very different and (3) can be used to estimate the ideal pulse-width.

IV. IMPLEMENTATION

The schematic of the PW pre-emphasis transmitter is shown in Figure 6, together with some signal waveforms. The PW modulated signal is generated with a clock with adjustable duty-cycle that selects either Data or not(Data). The not(Data) is delayed by half a clock-cycle to increase the timing margin. In the prototype IC the duty-cycle is controlled by an external current source, to provide programmability. $I_{\text{bias}}=0$ results in conventional binary signaling. At a 3GHz clock an I_{bias} of 80 μA , 200 μA or 400 μA results in transmitted symbols with pulse-widths of respectively 75%, 58% or 52%.

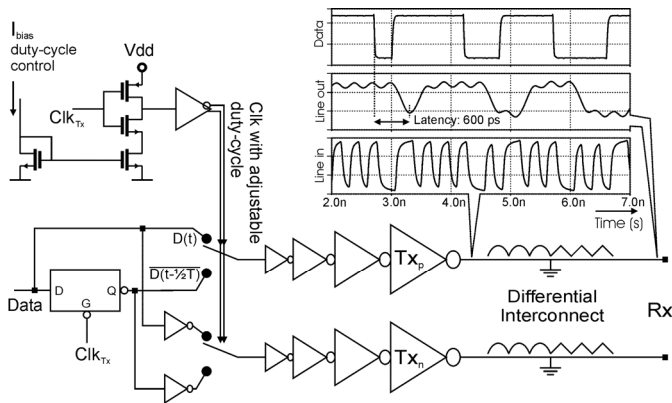


Figure 6. Transmitter schematic and signal waveforms

In a product application, the I_{bias} can be fixed at design time, as the transmission scheme is robust towards (circuit and wire) parameter deviations. The line-driver inverters are scaled to have an R_{out} of about 60Ω and the size of the differential transmitter is about $300\mu\text{m}^2$.

The schematic of the receiver is shown in Figure 7. The input inverters use transmission gates as selectable feedback resistors. In this way, either conventional termination or (active) resistive ($R_{in} \approx 150\Omega$) termination can be selected. A clocked comparator followed by a dynamic latch samples the received data. The size of the prototype differential receiver is about $1000\mu\text{m}^2$ (non-optimized).

The complete design is optimized for low mismatch and to function over all process corners. Dynamic latches, low-Vt transistors and small fan-outs (≤ 3) are used to meet the target data-rate of 3Gb/s even at the slow process corner. The simulated latency of the transceiver is about 650ps at 3Gb/s, composed of 180ps for the transmitter, 420ps for the channel and 50ps for the receiver.

Figure 8 shows the test chip micrograph. The chip has been fabricated in a standard 1.2V, 6M, $0.13\mu\text{m}$ CMOS process with copper interconnects. A 7 channel 10mm differential bus is surrounded by Gnd/Vdd-connected metal stripes and is laid out in a serpentine fashion to fit in a 1mm by 1mm area. A single-ended bus is placed below the differential bus, providing some intra-bus cross-talk.

An external single-channel 3.2Gb/s pattern generator/analyzer is used for the data generation and BER measurement. Large on-chip delay lines (chains of flip-flops) provide all bus-channels with pseudo-independent data. The phase of the RxClk can be adjusted externally to adapt to the eye position and measure its width.

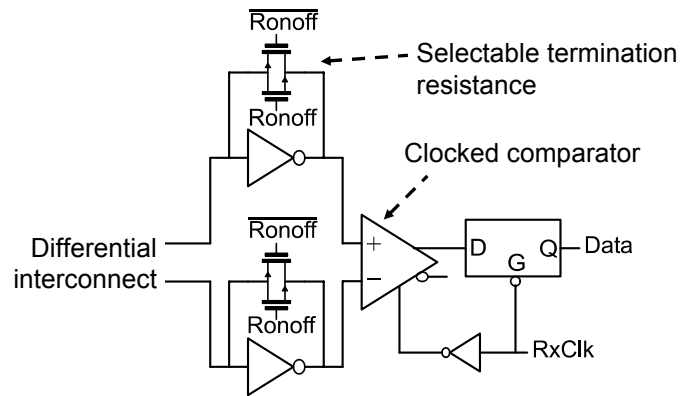


Figure 7. Receiver schematic

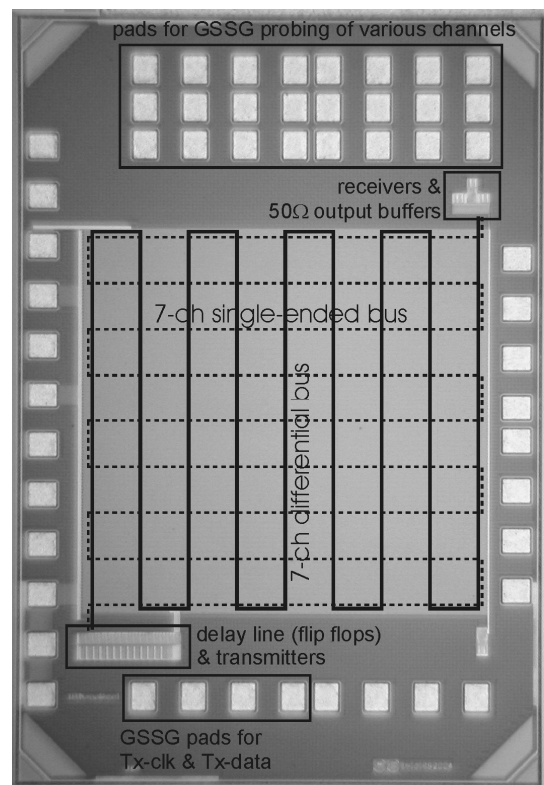


Figure 8. Chip micrograph

V. EXPERIMENTAL RESULTS

The measured line parameters are $0.19\text{k}\Omega/\text{mm}$ and $0.25\text{pF}/\text{mm}$ which agree with simulations given the tolerance bounds of the process.

Figure 9 shows the measured eye-diagrams at the input of the clocked comparator, both with and without the use of PW pre-emphasis and resistive termination, with data-rates at the edge of immeasurable BER ($< 1\text{e-}12$). Note that the achievable data rate increases 4 times by PW pre-emphasis, 3 times by resistive termination and 6 times by the combination of both.

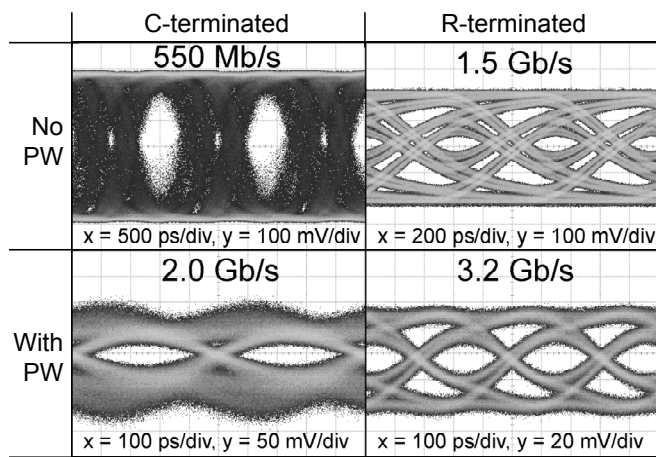


Figure 9. Eye-diagrams for various transceiver settings. The output buffers compress the vertical scale; on-chip signals are 6 to 9dB larger

At 3.2Gb/s, the eye-opening at the receiver side is so small that offset and memory effects in the clocked comparator lead to measurable BER ($5e-9$). At 3Gb/s, error free operation is possible for all 10 measured samples (with nominal biasing). At 2.5 Gb/s, the design is very robust and the BER remains immeasurable with large external parameter deviations: $1.0V < V_{dd} < 1.5V$ (nominal 1.2V); $34\% < TxClk \text{ duty cycle} < 62\%$ (nominal 50%); $130\mu A < I_{bias} < 400\mu A$ (nominal $200\mu A$); $-130ps < RxClk \text{ skew} < +130ps$.

Figure 20.7.6 illustrates crosstalk from a non-twisted neighboring interconnect on both single-ended (SE) halves of an interconnect with one twist at a data rate of 2.5Gb/s. The reduction in crosstalk on the differential voltage (due to the twist) is apparent and the differential signal eye is completely open.

At 3Gb/s, the total power consumption (Tx+Rx) for a single channel is 6mW.

VI. CONCLUSIONS

Experimental results have verified that pulse-width pre-emphasis can increase the achievable data rate, with a factor 4 if used with conventional termination and with a factor of 6 if it is combined with low-ohmic receiver termination.

The presented pulse-width pre-emphasis technique is a simple and robust equalizing approach, suitable for advanced deep-submicron processes. The technique can also be applied to inter-chip or wire-line [11] communication.

The Tx and Rx circuits are well suited for power-management, as the speed-enhancing, but power-consuming techniques can be easily turned on and off dynamically.

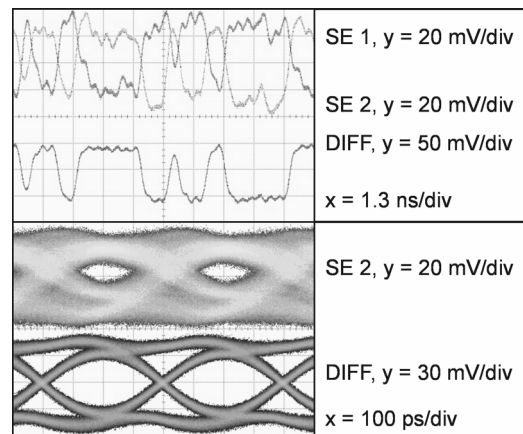


Figure 10. Effect of crosstalk on single-ended (SE) and differential (twisted) interconnect at 2.5Gb/s.

ACKNOWLEDGEMENT

The authors would like to thank Philips Research for chip fabrication and thank the Dutch Technology Foundation (STW, project TCS.5791) for funding.

REFERENCES

- [1] R. Ho, K. W. Mai, M. A. Horowitz, "The Future of Wires," Proc. IEEE, pp. 490-504, April 2001.
- [2] H. Bakoglu, "Circuits, Interconnections and Packaging for VLSI," Addison-Wesley, 1990.
- [3] R. Ho, K. Mai, M. Horowitz, "Efficient On-Chip Global Interconnects," Dig. Symp. VLSI Circuits, pp. 271-274, June 2003.
- [4] R. Chang, et. al., "Near Speed-of-Light Signaling Over On-Chip Electrical Interconnects," J. Solid State Circuits, vol. 38, pp. 834-838, May 2003.
- [5] E. Mensink, D. Schinkel, E.A.M. Klumperink, A.J.M van Tuijl, "Interconnects and On-Chip Data Communication Techniques," Proceedings of SAFE & ProRISC 2004, pp. 556-561, November 2004.
- [6] D. Schinkel, E. Mensink, E. A. M. Klumperink, A. J. M. van Tuijl and B. Nauta, "A 3Gb/s/ch Transceiver for RC-limited On-Chip Interconnects," IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, pp. 386-387, Feb. 2005.
- [7] E. Seevinck, P. van Beers, H. Ontrop, "Current-Mode Techniques for High-Speed VLSI Circuits with Application to Current Sense Amplifier for CMOS SRAM's," J. Solid State Circuits, vol. 26, pp. 525-536, April 1991.
- [8] Y. Massoud, et. al., "Modeling and Analysis of Differential Signaling for Minimizing Inductive Cross-Talk," Proc. of Design Automation Conf., pp. 804-809, June 2001.
- [9] E. Mensink, D. Schinkel, E. A. M. Klumperink, A. J. M. van Tuijl and B. Nauta, "Optimally-Placed Twists in Global On-Chip Differential Interconnects," Proc. 31th Eur. Solid-State Circ. Conf (ESSCIRC), pp. 475-478, Sept. 2005.
- [10] R. Farjad-Rad, et. al., "A 0.4- μm CMOS 10-Gb/s 4-PAM Pre-Emphasis Serial Link Transmitter," J. Solid State Circuits, vol. 34, pp. 580 -585, May 1999.
- [11] J.H.R. Schrader, E.A.M. Klumperink, J.L. Visschers and B. Nauta, "CMOS Transmitter using Pulse-Width Modulation Pre-Emphasis achieving 33dB Loss Compensation at 5-Gb/s," IEEE Symp. on VLSI Circuits 2005; Dig. Tech. Papers, pp 388-391, June 2005.