

Design of Ultrahigh-Speed Low-Voltage CMOS CML Buffers and Latches

Payam Heydari, *Member, IEEE* and Ravindran Mohanavelu, *Member, IEEE*

Abstract—A comprehensive study of ultrahigh-speed current-mode logic (CML) buffers along with the design of novel regenerative CML latches will be illustrated. First, a new design procedure to systematically design a chain of tapered CML buffers is proposed. Next, two new high-speed regenerative latch circuits capable of operating at ultrahigh-speed data rates will be introduced. Experimental results show a higher performance for the new latch architectures compared to the conventional CML latch circuit at ultrahigh-frequencies. It is also shown, both through the experiments and by using efficient analytical models, why CML buffers are better than CMOS inverters in high-speed low-voltage applications.

Index Terms—Broad-band circuits, current mode logic, device mismatch, environmental noise, tapered buffers, ultrahigh-speed CMOS circuits.

I. INTRODUCTION

THE rapidly growing volume of data transfer in telecommunication networks has recently drawn considerable attention to the design of high-speed circuits for gigabit communications networks. Wavelength-division multiplexing (WDM) and time-division multiplexing (TDM) were developed for use in the next-generation transmission systems. Ultramassive capacity transmission experiments have been reported using a WDM system with a per-channel data rates of 10 Gb/s for SONET OC-192 and 40 Gb/s for SONET OC-768. High-speed integrated circuit (IC) technologies with very high data rates are thus required for both WDM and TDM systems. Advances in nanometer CMOS technology has enabled CMOS integrated circuits to take over the territories thus far claimed by GaAs and InP devices.

Designing a high-speed CMOS circuit operating near f_T of the MOS device is very challenging. System blocks in a gigabit communications system need to be realized by very simple circuits utilizing minimum number of active devices. Parts of the circuit blocks that process high-speed signals in a communication transceiver should possibly abandon the use of pMOS devices due to their inferior unity-gain frequency. This, in turn, imposes additional design constraint on the ultrahigh-speed circuits.

Buffers and latches are the circuit cores of many high-speed blocks within a communication transceiver and a serial link.

As an example of a gigabit communications system, Fig. 1 depicts the block diagram of a typical optical transceiver. Front-end current mode logic (CML) tapered buffer chain, serial-to-parallel converters, clock and data recovery (CDR), multiplexers, and demultiplexers use high-speed buffers and latches extensively. A conventional CMOS inverter exhibits some drawbacks that prevent it from being vastly used in high-speed low-voltage circuits. First, a CMOS inverter is essentially a single-ended circuit. Recall that in a multigigahertz frequency range, the short on-chip wires act as coupled transmission lines. The electromagnetic coupling thus causes serious operational malfunctioning in the circuits, particularly single-ended circuits. Besides, the pMOS transistor in a static CMOS inverter will severely limit the maximum operating frequency of the circuit [1], [2].

CMOS current-mode logic style was first introduced in [3] to implement a gigahertz MOS adaptive pipeline technique. Since then, it has been extensively used to implement ultrahigh-speed buffers [4], [5], latches [5], multiplexers and demultiplexers [6], and frequency dividers [7]. CML circuits can operate with lower signal voltage and higher operating frequency at lower supply voltage than static CMOS circuits. However, CML logic style suffers from more static power dissipation than CMOS inverters. Recently, there have been efforts to alleviate this shortcoming [8], [9]. Particularly, one technique to efficiently reduce the power consumption of CML buffers is to implement the circuit in a multithreshold CMOS technology (MTCMOS) [8], and [8] actually designed a 1 : 8 2.5 Gb/s demultiplexer as a test vehicle to report a power saving of 37%. Due to their superior performance, CML buffers are the best choice for high-speed applications. As a consequence, it is an essential need to have a systematic approach to optimally design CML buffers and CML buffer chains.

This paper presents a systematic procedure of CML buffer design and introduces two new CMOS CML latch circuits. The paper is organized as follows. First, in Section II, a brief summary of static CMOS inverter is given. Then, in Section III, the large-signal behavior of a differential circuit is extensively illustrated. This will prepare us to study the design of CMOS buffer chain (Section IV). Section V discusses the performance and operation of the tapered CML buffer with consideration of device mismatch. In Section VI, we illustrate two novel CML latches in 0.18- μm CMOS process that are capable of operating at a ≥ 10 -GHz clock signal. Section VII provides various experimental results that verify the accuracy of our design approach. Finally, Section VIII provides the concluding remarks.

Manuscript received February 23, 2003; revised September 10, 2003.

P. Heydari is with the Department of Electrical Engineering and Computer Science, University of California, Irvine, CA 92697-2625 USA (e-mail: payam@ece.uci.edu).

R. Mohanavelu is with International Rectifier, El Segundo, CA 90245 USA. Digital Object Identifier 10.1109/TVLSI.2004.833663

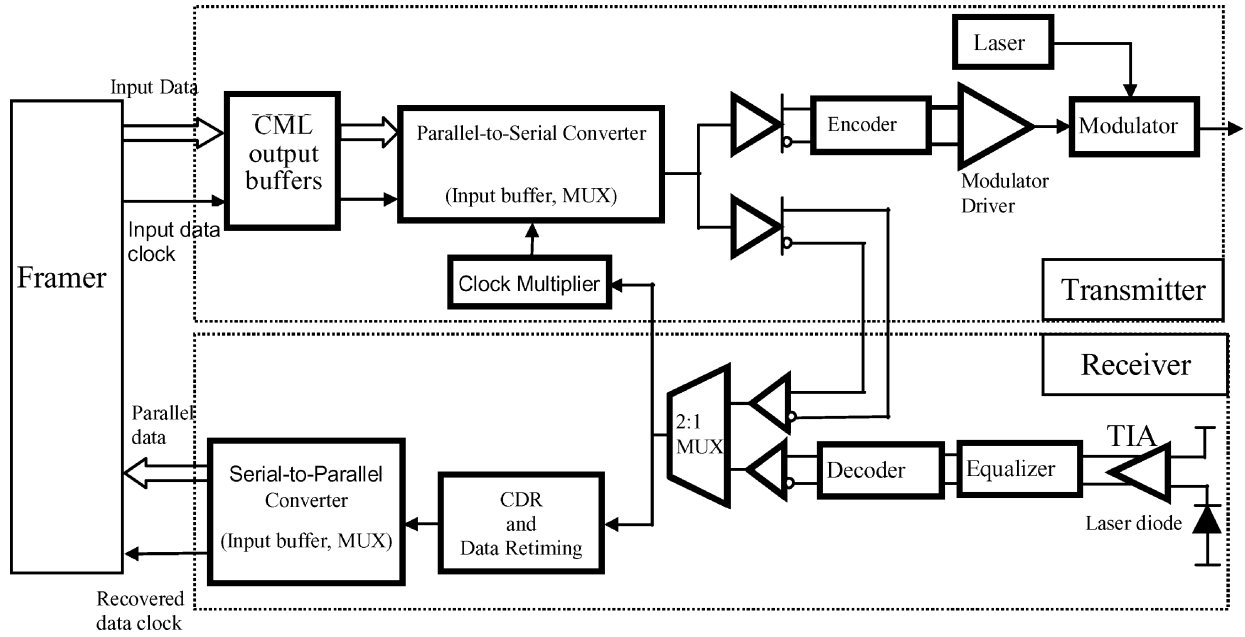


Fig. 1. System block diagram of an optical transceiver.

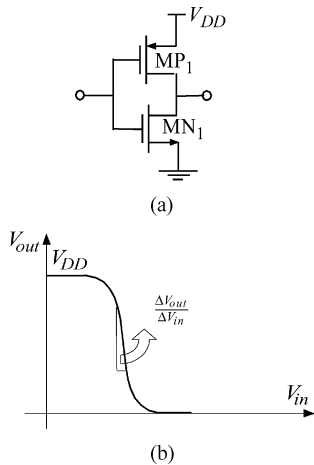


Fig. 2. (a) CMOS inverter. (b) Transfer characteristics.

II. CMOS BUFFERS

A conventional static CMOS buffer is shown in Fig. 2(a) where the input-output transfer curve is shown in Fig. 2(b).

A CMOS inverter has a number of advantages. The static power dissipation of a CMOS inverter is negligible, assuming the leakage current to be small. It exhibits the largest small-signal gain compared to any other area-efficient single-stage buffer with the same transistor sizes, and thus, is an ideal candidate for bus drivers and signal buffers in digital circuits. It shows an optimum performance with the technology scaling and has a large noise margin.

A CMOS inverter, however, suffers from a number of drawbacks that make it vulnerable in ultrahigh-speed integrated circuits. First, the use of pMOS transistor degrades the circuit maximum operation frequency (bandwidth). Secondly, like any single-ended circuit, a CMOS inverter is highly susceptible to the environmental noise sources such as power-ground

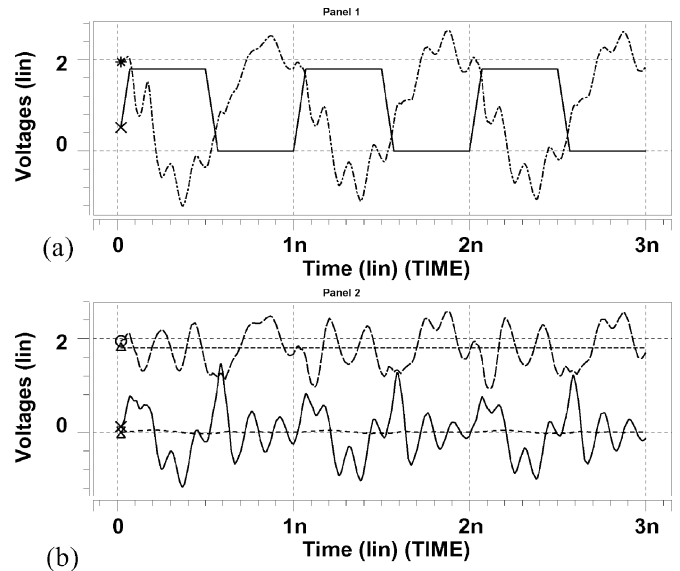


Fig. 3. (a) Input-output voltages of eight CMOS buffers switching simultaneously. (b) Power-ground bounce.

noise, substrate noise, and crosstalk. Large current surges during the voltage switching of output CMOS buffers driving large off-chip loads exacerbates the fluctuations on supply and ground rails. Noisy supply and ground wire results in noise-margin reduction, as well as a larger propagation delay for all predrivers connected to the same power and ground rail. Shown in Fig. 3(a) and (b), are the input-output voltages and the power-ground bounce noise due to simultaneous switching of eight CMOS inverters driving a large 2-pF off-chip capacitor. The gate aspect-ratios of the nMOS and pMOS devices in each CMOS inverter are $20 \mu\text{m}/0.2 \mu\text{m}$ and $40 \mu\text{m}/0.2 \mu\text{m}$, respectively. The inductor associated with the bondwires and pad to leadframe parasitics is assumed to be 2 nH. The bondwire resistance is 1 Ω . Obviously, other CMOS circuits connected

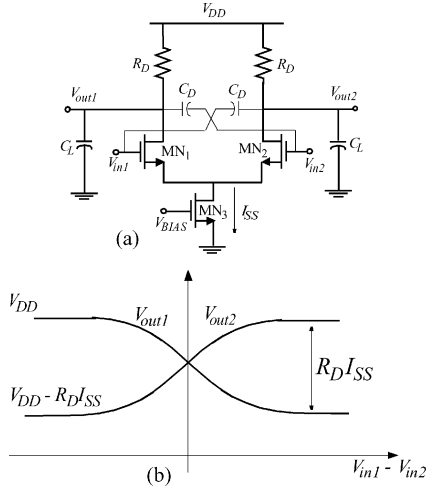


Fig. 4. (a) Neutralized CMOS differential pair. (b) Transfer characteristics.

to these noisy power and ground rails are affected by large unwanted oscillations that may cause false logic switchings. The experiment is carried out in the absence of on-chip decoupling capacitance to highlight the effect of the power–ground bounce on the performance of the off-chip CMOS drivers.

III. CML BUFFERS

A CML buffer is based on the differential architecture. Fig. 4(a) shows a basic differential architecture. The tail current, I_{SS} , provides an input-independent biasing for the circuit. The differential circuit is easily neutralized using a pair of capacitors, C_D , as indicated in Fig. 4(a), that will diminish the deleterious effects of input–output coupling through the device overlap capacitance, C_{GD} .

Various experimental simulations of CML circuits reveal that the long-channel transistor model still gives rise to a good estimation of the dynamic behavior of these circuits. The reason is because a CML circuit is a low-voltage circuit where the differential voltage swing is around the device threshold voltage.

As the differential input varies from $-\infty$ to $+\infty$, each output node of the differential pair varies from $V_{DD} - R_D I_{SS}$ to V_{DD} . Fig. 4(b) shows the voltage variations of the output nodes in terms of the differential input [10].

From Fig. 4(a), one can see that the maximum output differential voltage swing, V_{odm} , is only a function of the drain resistor and the tail current, provided that the full current switching takes place. Clearly, the maximum output swing of a CML buffer is less than that of a CMOS inverter, which makes this class of buffers an ideal choice for low-voltage integrated circuits design.

The minimum value of the input common-mode level, $V_{in,CMmin}$ is achieved when the tail current begins to operate in saturation. The input common-mode level reaches its maximum value, $V_{in,CMmax}$ when the transistors MN_1 and MN_2 are either at pinch-off or at cutoff [10]

$$V_{GS,12} + (V_{GS3} + V_{THN}) \leq V_{in,CM} \leq \min \left[V_{DD} - R_D \frac{I_{SS}}{2} + V_{THN}, V_{DD} \right] \quad (1)$$

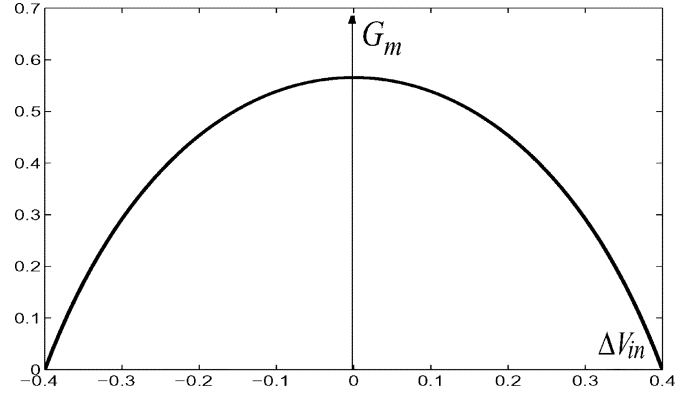


Fig. 5. Large-signal G_m as a function of the differential input.

where $V_{GS,12}$ is the common-mode overdrive voltage of transistors MN_1 and MN_2 . Similarly, the output common-mode level varies from V_{DD} (when both MN_1 and MN_2 are off, and MN_3 is in the triode region) to $V_{DD} - R_D I_{SS}/2$ (when all transistors are in saturation). The voltage transition of the output common-mode level from V_{DD} to $V_{DD} - R_D I_{SS}/2$ is determined by the subthreshold current of MN_1 or MN_2 .

The advantage of the differential CML buffer is understood by reviewing its large-signal behavior in response to a differential input signal. Assuming that the input common-mode level is bounded within the operating range specified in (1), a small voltage difference between V_{in1} and V_{in2} results in a corresponding differential current $I_{D1} - I_{D2}$, as follows [10]:

$$\Delta I_D = I_{D1} - I_{D2} = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} \Delta V_{in} \sqrt{\frac{4I_{SS}}{\mu_n C_{ox} \frac{W}{L}} - \Delta V_{in}^2}. \quad (2)$$

The differential current is an odd function of the input differential voltage, ΔV_{in} , and thus, becomes zero when the circuit is in equilibrium. Furthermore, a differential stage is more linear than a single-ended stage due to the absence of the even harmonics from the input–output characteristics. The large-signal transconductance G_m is the slope of $\Delta I_D - \Delta V_{in}$ transfer characteristics, that is

$$G_m = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} \frac{2\Delta V_{in,max}^2 - 2\Delta V_{in}^2}{\sqrt{2\Delta V_{in,max}^2 - \Delta V_{in}^2}} \quad (3)$$

where $\Delta V_{in,max} = \sqrt{2I_{SS}/(\mu_n C_{ox}(W/L))}$. The large-signal transconductance varies with the input differential voltage, as also shown in Fig. 5, where in this figure $\Delta V_{in,max} = 0.4$ V. As the input differential voltage exceeds a limit, one transistor carries the entire current I_{SS} , thereby, turning off the other transistor. $\Delta V_{in,max}$ represents the maximum input differential voltage.

An input-dependent transconductance results in a nonlinear large-signal gain. To simplify the analysis, the average value of the transconductance is utilized

$$G_{m,avg} = \frac{\int_0^{\Delta V_{in,max}} G_m(\Delta V_{in}) d(\Delta V_{in})}{\int_0^{\Delta V_{in,max}} d(\Delta V_{in})} = \sqrt{\frac{1}{2} \mu_n C_{ox} \frac{W}{L} I_{SS}}. \quad (4)$$

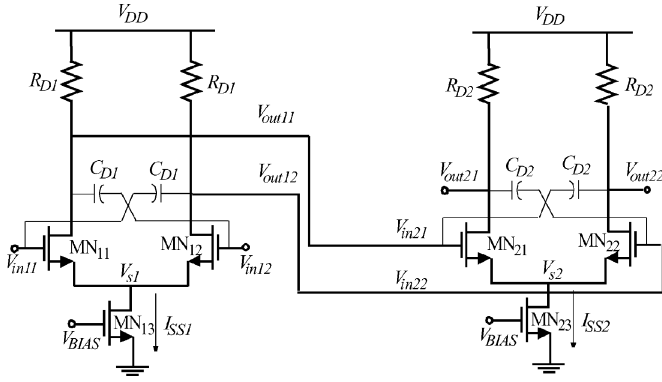


Fig. 6. Two CML buffers in cascade.

Note that $G_{m,avg}$ is $(1/\sqrt{2})g_{m,ss}$ where $g_{m,ss}$ is the small-signal transconductance of the differential pair. A differential pair architecture using a differential signaling is insensitive to common-mode fluctuations, which makes it a better choice as a buffer than a CMOS inverter, particularly in low-noise circuit design where noise mostly appears as a common-mode component. Moreover, a noninverting buffer is easily realized using a single differential stage, as opposed to the CMOS inverter where a noninverting buffer is realized by two inverters in cascade. Therefore, a noninverting differential buffer exhibits a lower propagation delay than a CMOS buffer. A differential stage will be operating as a CML buffer, if and only if a complete current switching takes place. To make sure that the current switches entirely from one side of the differential stage to the other side, the differential input voltage must be at least $\Delta V_{in,max}$.

Moreover, a differential CML buffer exhibits a higher bandwidth than a conventional CMOS inverter. This is readily proved either using the time-domain delay analysis or small-signal approximation.

IV. TAPPED CML BUFFER DESIGN

To achieve the best performance in a CML buffer, a complete current switching must take place and the current produced by the tail current flows through the ON branch only. To quantify the underlying conditions for complete current switching, one should consider that in practice, a CML buffer often drives another CML buffer (e.g., a tapered buffer chain), which means that output terminals of the driving buffer stage are connected to the input terminals of the driven stage, as shown in Fig. 6. To satisfy the current switching requirement, the differential voltage swing of the first CML buffer must exceed $\Delta V_{in2,max}$ of the following stage, i.e.,

$$R_{D1} I_{SS1} \geq \sqrt{\frac{2I_{SS2}}{(\mu_n C_{ox} (\frac{W}{L})_2)}}. \quad (5)$$

In a special case of having identical CML stages in Fig. 6, (5) results in a lower bound of $\sqrt{2}$ for the maximum small-signal voltage gain at equilibrium, $A_{v,eq}$.

Furthermore, the load resistors should be small in order to reduce the RC delay and increase the bandwidth. To guarantee a high-speed operation, nMOS transistors of the differential pair

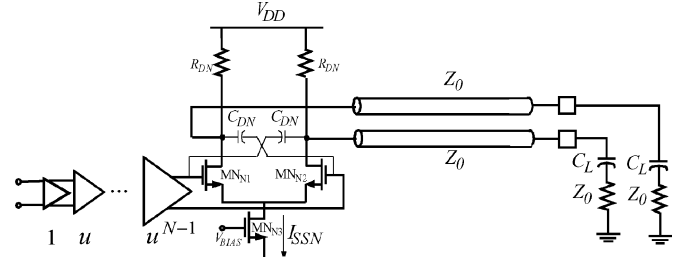


Fig. 7. Output CML buffer driving off-chip loads. The chip-package interface is electrically modeled using a lossless transmission line.

must operate only in the saturation. To satisfy this requirement for the circuit shown in Fig. 4(a), first, the input common-mode voltage must be within the interval specified in (1); and secondly

$$V_{in_k,max} - V_{THN} \leq V_{out,kj} \leq V_{DD} \quad \text{for } k = 1, 2 \quad \text{and } j = 1, 2 \quad (6)$$

which sets a maximum allowable level for the differential output swing as follows:

$$R_{Dk} I_{SSk} \leq V_{THN} \quad \text{for } k = 1, 2. \quad (7)$$

In the particular case of output drivers, a high-speed CML driver must drive a large off-chip load through the bondwire and package trace. The output driver must thus have a large current drive capability. This means that nMOS transistors of the second CML buffer in Fig. 6 must be large. A large transistor has a large gate-to-channel capacitance that seriously degrades the propagation delay and the voltage swing of the preceding predriver stage. To reduce the propagation delay of the predriver, a chain of tapered buffers is introduced between the first predriver stage and the output buffer. It is readily proved that the minimum delay is obtained by dividing the delay equally over all stages [11]. This is achieved by gradually scaling up all stages with a constant taper factor u . On the other hand, the chip package interface at very high frequencies is appropriately modeled as a transmission line that is terminated by a load impedance, which is a series RC circuit (cf. Fig. 7). The series load resistance, Z_0 , provides the high-frequency parallel matched termination to the bondwire. Fig. 7 shows the schematic of the output CML driver driven by $N - 1$ tapered CML stages along with the chip-package interface being modeled as the transmission line.

The chip bondwires exhibit high- Q inductances. Therefore, it is safe to model the chip-package interface using a lossless transmission line. To avoid potentially disastrous transmission line effects such as slow ringing and propagation delays, the bondwires are terminated both at the source using a series termination ($R_{DN} = Z_0$), and at the destination using a parallel termination (Z_0). Given a well-defined output voltage swing ($R_D I_{SS}$) and with R_D being determined by the matched termination, the tail current I_{SSN} is easily calculated. For instance, an output differential voltage swing of 0.4 V for a 50 Ω line driver requires a bias current of 8 mA. Now, using a set of constraints, we present design guidelines to design a tapered CML buffer chain and determine appropriate values for the circuit components of the CML buffer.

The propagation delay is computed using the open-circuit time constant method [12]. For instance, the delay of the simple low-voltage differential stage of Fig. 4(a) is $0.69R_D C_L$. Various HSPICE simulations on high-speed CML buffers show that the delay obtained by the open-circuit time-constant method is within 10% of the actual simulation.

Minimizing the overall propagation delay of CML buffer increases the overall operation frequency of the buffer significantly. For a slowly varying input signal, increasing the small-signal voltage gain, $g_{m,ss}R_{out}$, will further decrease the output transient variations and the output transition time. In a chain of tapered CML buffers, to attain a constant voltage swing, transistor sizes are scaled up while the drain resistances are scaled down with a constant scaling factor. This will lead us to the fact that small-signal voltage gains of all constituent stages of the buffer chain are identical

$$\begin{aligned} R_{D1} \sqrt{\mu_n C_{ox} \left(\frac{W}{L}\right)_1 I_{SS1}} &= R_{D2} \sqrt{\mu_n C_{ox} \left(\frac{W}{L}\right)_2 I_{SS2}} \\ &= \dots = R_D \sqrt{\mu_n C_{ox} \left(\frac{W}{L}\right) I_{SS}}. \end{aligned}$$

As a consequence, (5) and (7) provide us with a lower bound for the maximum small-signal voltage gain at equilibrium, that is

$$\left(A_{v,eq} = R_D \sqrt{\mu_n C_{ox} \frac{W}{L} I_{SS}} \right) \geq \sqrt{2}. \quad (8)$$

The drain resistor, R_{DN} , of the last output CML buffer is determined by the series impedance matching to bondwire's characteristic impedance. Subsequently, I_{SSN} of the last driver stage is calculated using the output differential voltage swing and R_D . The only remaining parameter in the last CML driver is the (W/L) of the source-coupled transistor pair, which is obtained from the common-mode analysis of the last CML buffer. If the common-mode input voltage lies in the allowable range given by (1), then the tail current will equally be divided between the two branches of the differential stage, i.e.,

$$\begin{aligned} \left(V_{in,k,CM} - V_{sk} - V_{THN} = \sqrt{\frac{I_{SSk}}{(\mu_n C_{ox} \left(\frac{W}{L}\right)_k)}} \right) \\ \leq V_{in,k,CM} - V_{BIAS} - 2V_{THN}, \quad \text{for } k = 1, 2, \dots, N \quad (9) \end{aligned}$$

where $V_{in,k,CM}$ is the common-mode input voltage of the k^{th} driver in the buffer chain. $V_{in,k,CM}$ is specified by the output common-mode voltage of the previous stage. The inequality in (9) guarantees that the tail current is in the saturation region. Given a **tapered buffer chain** with a constant differential voltage swing, the maximum (W/L) of the transistor pair of the k^{th} CML buffer is then calculated by solving (10)

$$V_{DD} - R_D \frac{I_{SS}}{2} - V_{BIAS} - 2V_{THN} \geq \sqrt{\frac{I_{SSk}}{(\mu_n C_{ox} \left(\frac{W}{L}\right)_k)}}. \quad (10)$$

In (10), $R_D I_{SS}$ is the constant differential output swing of a tapered CML buffer chain.

As mentioned above, in a chain of tapered CML buffers, the minimum delay is obtained by dividing the delay equally over

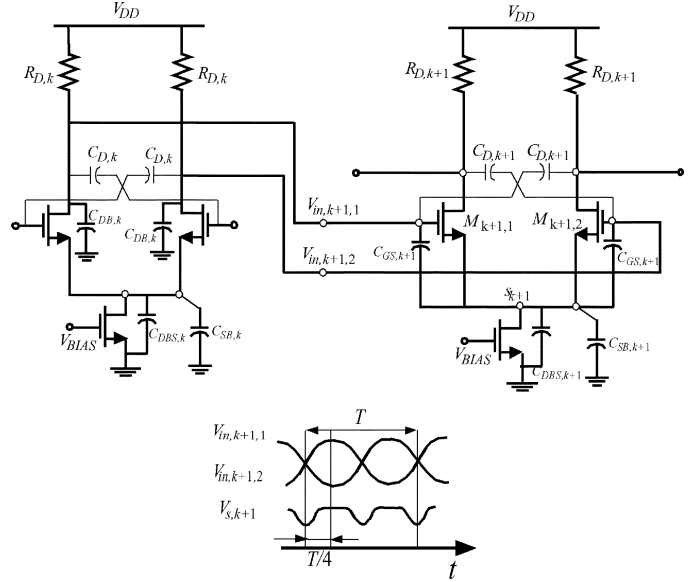


Fig. 8. The k^{th} and $(k+1)^{\text{st}}$ stages of a tapered CML buffer along with the parasitic capacitances.

all stages. However, the question is how many buffer stages are required to achieve the optimum delay. To answer this question, the propagation delay of an arbitrarily chosen CML stage in a buffer chain is first derived. Fig. 8 shows the k^{th} stage in a chain of N tapered stages driving another CML stage along with the capacitors that contribute to the delay calculation.

The common node s_{k+1} shown in Fig. 8 experiences a double-frequency variation compared to the voltage variations [10]. The input capacitance seen at the gate terminal of the $(k+1)^{\text{st}}$ stage is, therefore, slightly smaller than the gate-source capacitance $C_{GS,k+1}$. Ignoring the channel length modulation in MOS devices, and assuming the gate terminals of the $(k+1)^{\text{st}}$ stage to have fully differential voltages, the current-voltage relationship of at each gate terminal of the $(k+1)^{\text{st}}$ stage is expressed as follows:

$$I_{G,k+1,i} = \left[C_{GS,k+1} + \frac{C_{S,k+1}}{2} \frac{V_{in,k+1,i}}{\sqrt{2V_{in,max}^2 - V_{in,k+1,i}^2}} \right] \times \frac{dV_{in,k+1,i}}{dt} \quad i = 1, 2 \quad (11)$$

where $V_{in,max} = 0.5 \Delta V_{in,max}$, with $\Delta V_{in,max}$ defined earlier in Section III. Equation (11) states that the large-signal input impedance of the differential pair can be defined using a non-linear voltage-dependent effective capacitance. The value of this effective input capacitance is a function of the input voltage, thereby, varying with time. Assuming a sinusoidal input with the amplitude of $V_{in,max}$, the time average of this effective capacitance is calculated as follows:

$$\begin{aligned} C_{eff,k+1} &= C_{GS,k+1} + \frac{4}{T} \left(\int_0^{T/4} \frac{C_{S,k+1}}{2} \frac{V_{in,k+1,i}}{\sqrt{2V_{in,max}^2 - V_{in,k+1,i}^2}} dt \right) \\ &\cong C_{GS,k+1} - \frac{C_{S,k+1}}{2\pi} \ln(3 + 2\sqrt{2}) \quad (12) \end{aligned}$$

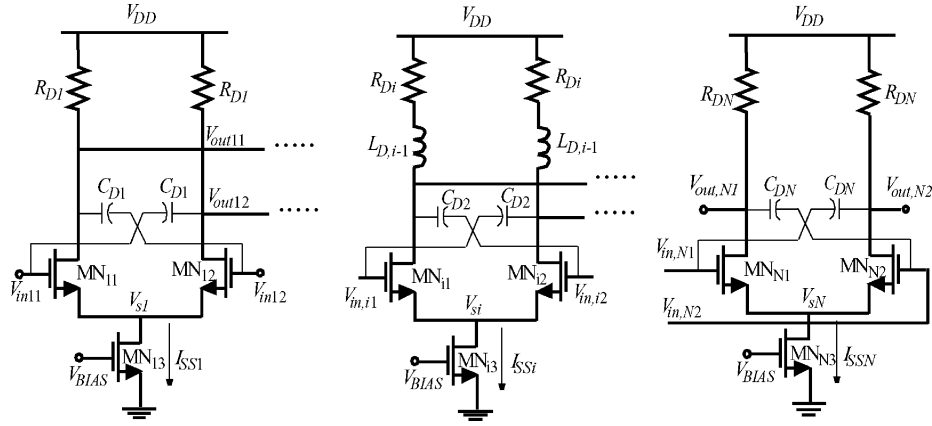


Fig. 9. Multiple stage CML buffers along with the inductive peaking.

where $\ln(x)$ represents the natural logarithm of x . In fact, it is easily shown that input capacitance of seen at the input gate terminal of the $(k+1)^{\text{st}}$ stage is less than $C_{\text{GS},k+1}$. This highlights the advantage of the differential pair in high frequency compared to the static CMOS inverters.

The 50% delay of the k^{th} stage is as follows:

$$t_{d,k} = 0.69 R_{D,k} (C_{\text{DB},k} + C_{\text{eff},k+1}). \quad (13)$$

As a generalization to the single-stage delay calculation, consider a chain of tapered CML buffers driving a lossless transmission line with a characteristic impedance of Z_0 . Suppose that the gate aspect ratio of the transistor pair of the last CML line driver is X times larger than that of the first predriver stage. The total propagation delay of the buffer chain is readily calculated

$$t_d = \sum_{k=1}^N t_{d,k} = 0.69 N R_{D1} (C_{\text{DB1}} + X^{1/N} C_{\text{eff},1}). \quad (14)$$

Interestingly, the functional dependence between delay and the number of stages (or taper factor) is similar to the one in a CMOS buffer chain first proposed in [13]. It is proved that the optimum number of stages will be the numerical solution to the following:

$$X^{1/N_{\text{opt}}} = \exp \left[1 + \frac{C_{\text{DB1}}}{X^{1/N_{\text{opt}}} C_{\text{eff},1}} \right] \quad (15)$$

or in the special case, if $C_{\text{DB1}} \ll C_{\text{eff},1}$ then, $N_{\text{opt}} = \ln(X)$ which is a well-known result.

To further increase the bandwidth (reduce the delay), the intermediate stages (including the last stage) use inductive peaking as demonstrated in Fig. 9.

In addition, the inductor in series with the drain resistor delays the current flow through the branch containing the resistor, making more current available for charging the device capacitors, and reducing the rise and fall times. From another perspective, the addition of an inductance in series with the load capacitance introduces a zero in the transfer function of the CML stage which helps offset the roll-off due to parasitic capacitances. For any intermediate CML stage, the optimized value of the inductor is easily obtained. Since each CML stage is neutralized by cross-connected capacitors, $C_{D,i}$, the equivalent half-circuit model corresponding to the i^{th} intermediate stage is roughly modeled by the circuit shown in Fig. 10(a).

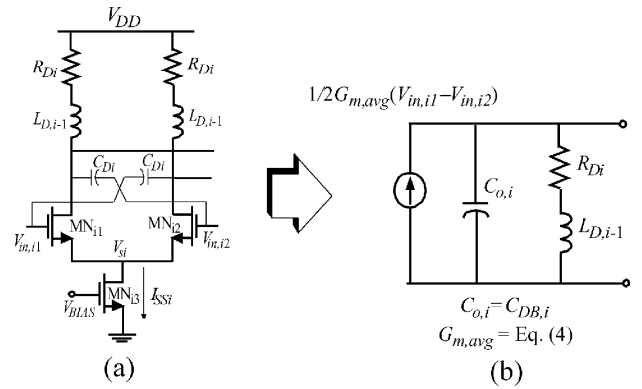


Fig. 10. (a) CML stage. (b) Equivalent circuit for the half-circuit model.

The equivalent circuit shown in Fig. 10(b) is a second-order circuit that exhibits an overshoot in its magnitude response. A straightforward calculation reveals that to achieve a maximally flat frequency response, we must have [12]

$$L_{D,i-1} = \frac{R_{D,i}^2 C_{o,i}}{1 + \sqrt{2}} \quad (16)$$

which results in an increased bandwidth that is about 1.7 times larger than the unpeaked case [12]. Inductance values are scaled with the same taper factor as the drain resistors, to retain a constant delay per each stage.

V. DEVICE MISMATCH

The analysis undertaken in Sections III and IV assume that all devices are identically matched. In practice, the inaccuracies in manufacturing process introduce device mismatches. Mismatches cause three major effects on the performance of the circuits, and in particular the CML buffers [10]: 1) dc offset; 2) finite even-order distortion; and (3) lower common-mode rejection. Details about each of these effects can be found in [10].

Focusing on the multistage tapered CML buffer shown in Figs. 7 and 9, the most significant effect of the dc offset is to drive the conducting transistors of latter stages of the tapered CML buffer into triode region. This observation suggests that the last stages of the tapered buffer are exposed to the performance degradation more seriously than the first stages within the CML tapered buffer. For example, the input offset voltage

of the M^{th} CML stage of an N -stage tapered CML buffer is added to the amplified replicas of the offset voltages from previous stages, i.e.,

$$V_{OS,M} = \sum_{k=1}^{M-1} \left[\left(\prod_{i=k}^{M-1} A_{v,i} \right) V_{OS,\text{in}k} \right] + V_{OS,\text{in}M} \quad \text{for } M = 2, 3, \dots, N \quad (17)$$

where $V_{OS,\text{in}k}$ represents the input offset voltage of the k^{th} stage, $A_{v,i}$ is the small-signal voltage gain of the i^{th} stage. At this point, we establish an analogy between the offset and device noise. In the noise analysis of integrated circuits, the effect of all noise sources in the circuit are referred back to the input, and is represented by input referred noise sources [10]. The input-referred noise sources indicate how much the input signal is corrupted by the circuit's noise. On the other hand, the output-referred noise does not allow a fair comparison of the performance of different circuits because it depends on the gain (see [10]).

Similar to the device noise analysis, the overall offset voltage for a chain of N tapered buffers is referred back to the input and is represented by a voltage source, $V_{OS,\text{in}}$, i.e.,

$$\begin{aligned} V_{OS,\text{in}} &= \frac{\sum_{k=1}^{N-1} \left[\left(\prod_{i=k}^{N-1} A_{v,i} \right) V_{OS,\text{in}k} \right] + V_{OS,\text{in}N}}{\prod_{i=1}^{N-1} A_{v,i}} \\ &= V_{OS,\text{in}1} + \frac{V_{OS,\text{in}2}}{A_{v,1}} + \frac{V_{OS,\text{in}3}}{A_{v,1}A_{v,2}} + \\ &\quad \dots + \frac{V_{OS,\text{in}N}}{A_{v,1} \dots A_{v,N-1}}. \end{aligned} \quad (18)$$

Interestingly, (18) resembles the Friis equation proposed in [14] for the overall noise-figure of N electronic systems in cascade.

Discussions in Section IV showed that the voltage gains of all N CML stages are identical, which simplifies (18)

$$V_{OS,\text{in}} = V_{OS,\text{in}1} + \frac{V_{OS,\text{in}2}}{A_v} + \frac{V_{OS,\text{in}3}}{A_v^2} + \dots + \frac{V_{OS,\text{in}N}}{A_v^{N-1}}. \quad (19)$$

The input offset voltage $V_{OS,\text{in}k}$ ($k = 1, \dots, N$) is directly proportional to the equilibrium overdrive voltage, transistor dimension mismatch, and load resistor mismatch [10]. The number of stages is determined by (15), and cannot be changed. Equation (19) states the input-referred noise voltage is inversely proportional to the voltage gain. An effective way of decreasing the offset voltage $V_{OS,\text{in}}$ is thus to set the voltage gain to its *maximum* allowable value, while ensuring that (9) will be satisfied.

The current tails of the tapered CML buffer are designed using current mirrors. The transistor mismatch results in the current mismatch in current tails [10]. This current mismatch is inversely proportional to (W/L) of the current tail, which sets a design constraint for the dimension of the reference transistor in the current mirror.

As mentioned earlier, the device mismatch causes the common-mode rejection of each CML stage to be reduced. This, in fact, degrades the superior performance of the CML

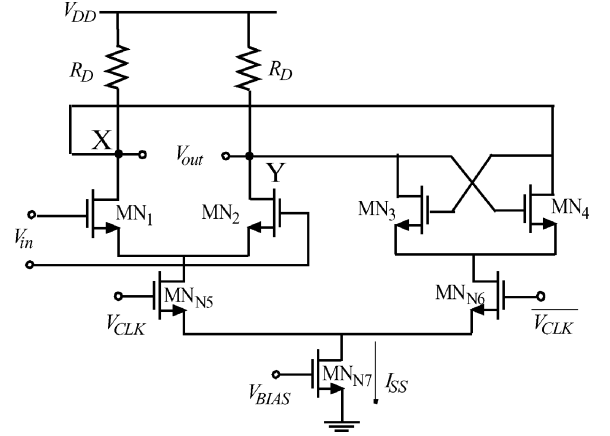


Fig. 11. Circuit schematic of a CMOS CML buffer.

buffers in the presence of the power-ground bounce, because the common-mode power-ground bounce, and crosstalk noise are converted to the differential output components, distorting the output differential signal. Furthermore, the common mode to differential mode conversion gain increases with frequency due to the parasitic capacitances of the MOS device [10]. In a tapered CML buffer, the bias currents of subsequent CML stages are scaled up, while the drain resistances are scaled down. Interestingly, both of these phenomena lead to a decrease of the common-mode to differential-mode conversion gain.

VI. ULTRAHIGH -SPEED LATCH DESIGN

A CML latch consists of an input tracking stage, M_{N1} and M_{N2} , utilized to sense and track the data variation and a cross-coupled regenerative pair, M_{N3} and M_{N4} , being employed to store the data. Fig. 11 demonstrates a CMOS CML latch circuitry.

The track and latch modes are determined by the clock signal inputs to a second differential pair, M_{N5} and M_{N6} . When the signal V_{CLK} is "HIGH," the tail current I_{SS} entirely flows to the tracking circuit, M_{N5} and M_{N6} , thereby allowing V_{out} to track V_{in} . In the latch-mode, the signal V_{CLK} goes low, the tracking stage is disabled, whereas the latch pair is enabled storing the logic state at the output.

Like CML buffers, a CML latch operates with relatively small voltage swings which is $2V_{\text{THN}}$ peak-to-peak differential-mode. Fig. 11 allows us to implement high-speed latch circuit. However, there are several shortcomings involved in the design of the regenerative latch in Fig. 11, that lead to a complete operation failure at very high speed datarates (≥ 10 Gb/s) when the circuit is realized in $0.18 \mu\text{m}$ CMOS technology. The primary limitation is that a single tail current is used for both tracking and latch circuits. Consequently, the bias operations of tracking and latch circuits are tightly related. This will severely limit the allowable transistor sizes for a reliable latch operation. At ultrahigh-speed datarates (≥ 10 Gb/s) the parasitic capacitances of transistors, M_{N1} and M_{N2} , degrade the required minimum small-signal gain for a proper tracking operation [(8)]. Therefore, the tail current must be sufficiently high to achieve a wider range of linearity and a larger transconductance. On the other

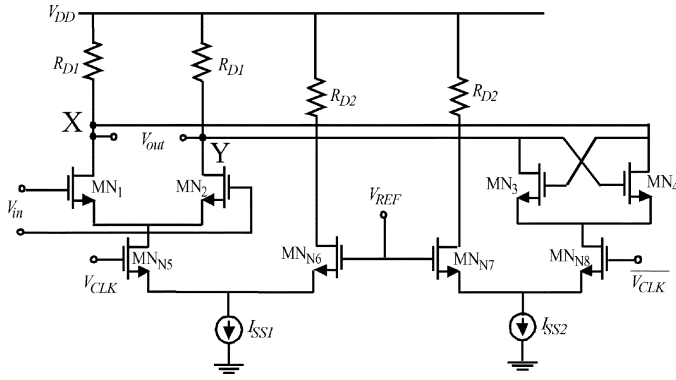


Fig. 12. Circuit schematic of the new CMOS CML latch circuit.

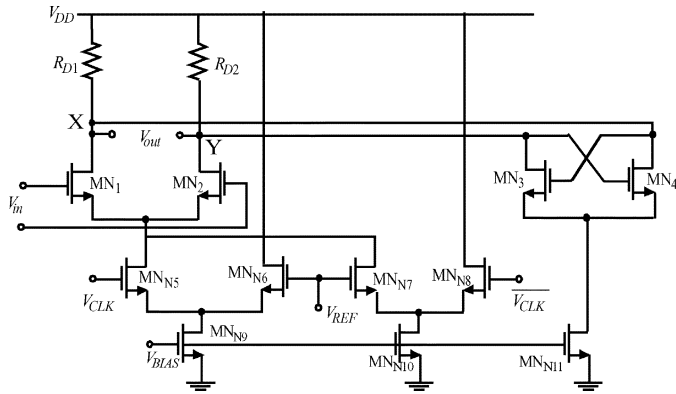
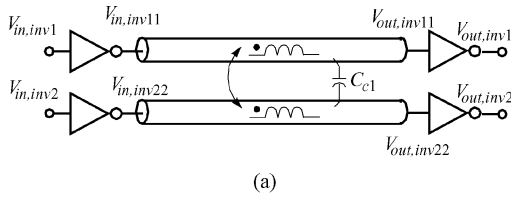
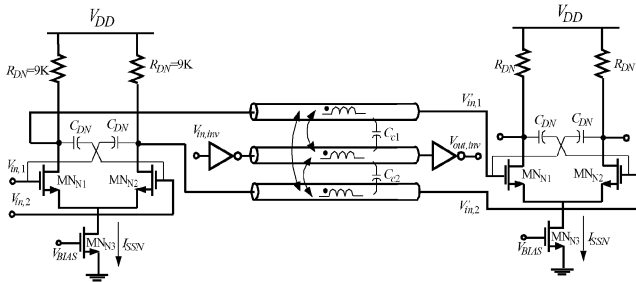


Fig. 13. Circuit schematic of another novel CMOS CML latch circuit capable of operating at ultrahigh-speed data rates.



(a)



$$(W/L)_{N1} = (W/L)_{N2} = 5\mu\text{m}/0.2\mu\text{m}$$

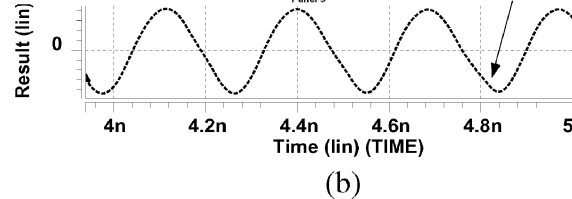
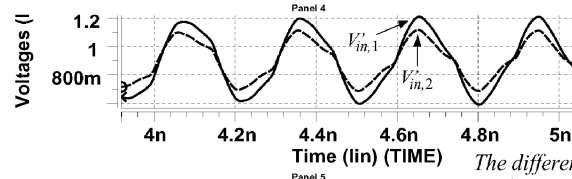
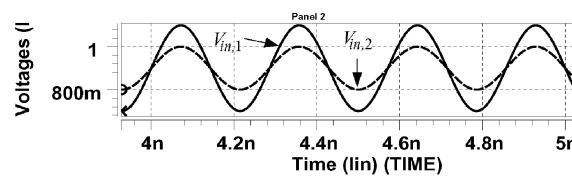
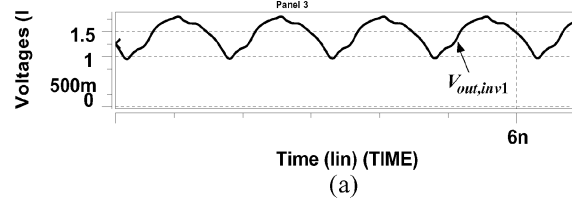
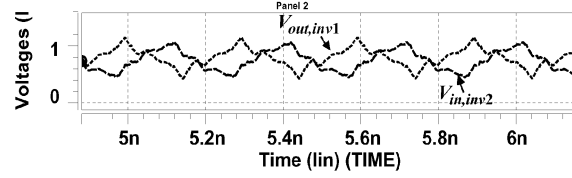
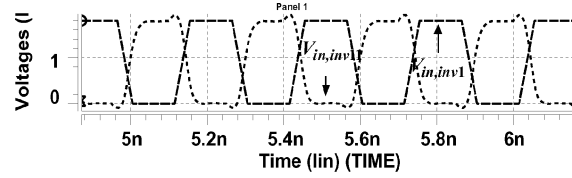
$$(W/L)_{N3} = 20\mu\text{m}/0.2\mu\text{m}$$

(b)

Fig. 14. (a) CMOS inverters driving two adjacent coupled interconnects that are terminated by CMOS inverters. (b) Two interconnects driven by a CML buffer and coupled to another interconnect which is driven by CMOS inverter.

hand, the latch circuit does not need a large bias current at ultrahigh-frequencies.

To address the aforementioned problems, the regenerative CML latch is modified so that the latch circuit and the tracking circuit use two distinct tail currents. Fig. 12 shows the new CML latch circuit.



(b)

Fig. 15. (a) Input and output waveforms of Fig. 14(a). (b) Input and output signals of Fig. 14(b).

As observed in Fig. 12, the tracking stage and the latch stage are now separately optimized for a correct latch operation at ultrahigh-speed. Note that it is important for the source coupled pair transistors to have high gain. This is obviously achieved with larger W/L for each transistor of the cross-coupled pair. However, this technique greatly limits the driving capability. Therefore the CML latch is followed by a CML buffer to recover the logic level.

There is still one underlying problem that causes a speed limitation on the proposed circuit as well as the conventional counterpart. During each transition from the amplification mode (when V_{CLK} is "HIGH") to the latching mode (when V_{CLK} is "LOW"), the current tail of the cross-coupled pair must first recharge the capacitances of the cross-coupled pair as it starts drawing current from the output nodes, X and Y, and changing the logic state. This will increase the minimum achievable clock period at which the latch circuit works properly.

An alternative to the proposed circuit shown in Fig. 12, is depicted in Fig. 13, where the latch transistor always draws current from the nodes X and Y and there is no need for the charge to be

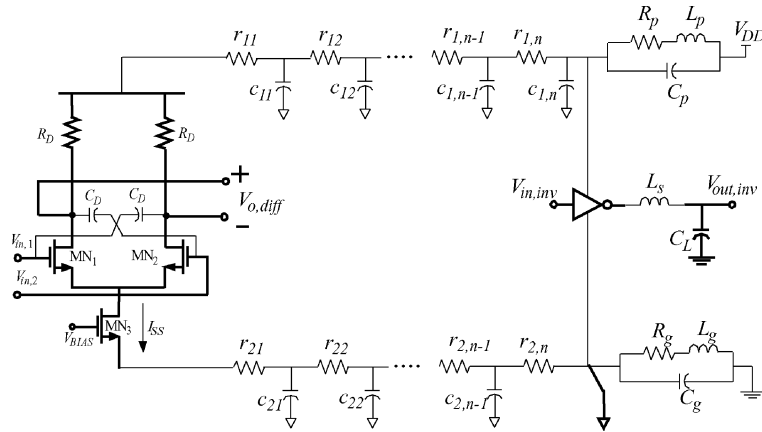


Fig. 16. CML buffer along with on-chip power/ground wires and chip-package interface circuit model.

built up during the latching phase. There are several advantages associated with the circuit of Fig. 13.

Firstly, the new CML latch circuit in Fig. 13 does not suffer from the current spiking seen at the drain of the clock transistors. This phenomenon becomes clear by studying the circuit in the tracking mode when the input signal V_{CLK} is “HIGH”. During the tracking interval, transistor MN7 is switched on drawing a portion of the tail current and reducing the current spikes. On the other hand, the cross-coupled pair MN3-MN4 is always enabled, hence no current spike takes place during the transition from tracking to latching mode. Experiments in Section 6.4 verify the above observation.

Secondly, an enabled cross-coupled pair during the tracking mode directly contributes to smaller rise and fall times for the output voltages at nodes X and Y. Recall that a cross-coupled pair exhibits a negative resistance that lowers the equivalent resistance at each node, X and Y, for a fixed output voltage swing, thereby decreasing rise and fall times of the output voltages.

The new latch circuit, however, consumes more power than the circuits shown in Figs. 11 and 12.

VII. EXPERIMENTAL RESULTS

In this section the performance of the CML buffer is evaluated by performing experiments on single stage as well as multiple stages of the buffer. Experiments are set up to show the performance of the novel CML latches depicted in Figs. 12 and 13 at 20 GHz data rate. First, the noise susceptibility of the CML buffer is experimentally compared with CMOS inverter. Next, the accuracy of (15) is verified by running HSPICE simulation on a chain of CML buffers. The effect of inductive peaking on the bandwidth and speed enhancement will be investigated. Finally, the performance of CML latch circuits demonstrated in Figs. 12 and 13 are compared against the conventional CML latch shown in Fig. 11.

A. Noise Performance

A CML buffer exhibits a superior noise performance compared to a conventional CMOS inverter, particularly because environmental noise sources (e.g., crosstalk, power-ground noise) appear as common-mode signals. This will be experimentally verified by performing the following experiment.

First, crosstalk noise is emulated using parallel interconnects located within close proximity of each other, as depicted in Fig. 14(a) and (b).

To have a performance comparison, we place, first, a CMOS inverter, and then, a CML buffer at the outputs of coupled interconnects, one at a time [Fig. 14(a) and (b)]. To highlight the superiority of noise performance of the CML buffer, the middle line in Fig. 14(b) will be driven by a CMOS inverter. The noise coupled from this line to its neighboring lines is the same and has the large amplitude. The input signal frequency for all CMOS inverters is 3.3 GHz, while it is 3.5 GHz for CML buffers. As a consequence, this experiment also shows the performance of CML buffer in the presence of harmonic distortion. All circuits are designed using 0.18 μm standard CMOS process.

Fig. 15(a) and (b) demonstrate the output signals of CMOS inverter and CML buffer, respectively. The experiment is set up to demonstrate the worst-case scenario in which the noise fluctuation and the voltage waveform are 180° out of phase. The plots in this slide show the input and output voltage waveforms of the CMOS inverters. The first plot in Fig. 15(a) shows the input waveforms. The second plot shows the outputs at the output terminals of transmission lines. These signals are also the inputs to the following CMOS inverters. The third plot indicates the output of last inverter stage. Similarly, the first plot in Fig. 15(b) shows the two inputs given to the input terminals of the first CML buffer. The second plot shows the outputs $V'_{in,1}$ and $V'_{in,2}$ at the output terminals of transmission lines. These signals are also the inputs to the following CML buffer. The third plot indicates the output of the last CML stage.

As observed in Fig. 15(a), the output voltage $V_{out,inv1}$ of the CMOS inverter in Fig. 14(a) does not have a rail-to-rail swing because of the crosstalk noise effect from the other adjacent line. In fact, this CMOS inverter is incapable of generating a logic “LOW”. On the other hand, the functionality of a CML buffer remains intact in the presence of the coupling noise from a neighboring line, as seen in Fig. 15(b).

A CML buffer also shows a better performance in the presence of power/ground noise than a CMOS inverter. Noise on power and ground wires have very small degrading effects on the differential output voltage. Fig. 16 demonstrates a circuit that emulates the actual scenario where on-chip power/ground

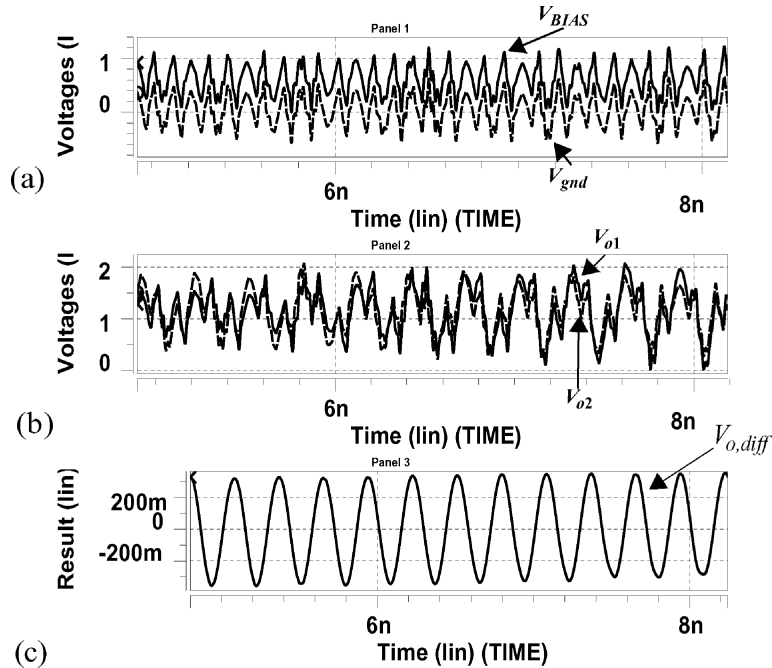


Fig. 17. (a) On-chip ground, input bias voltage of the tail current. (b) Single-ended output voltages of the CML buffer. (c) The differential output voltage of the CML buffer.

wires are modeled using distributed RC circuits; and the chip-package interface parasitics including parasitics associated with bondwires and package traces are modeled using (R_p, L_p, C_p) and (R_g, L_g, C_g) . A static CMOS inverter driving an off-chip load generates Power/ground fluctuations. Shown in Fig. 17(a)–(c) are the on-chip power/ground waveforms, the single-ended outputs and the differential output of the CML buffer. The differential architecture is capable of filtering the common-mode noise and generates a clean differential output with a maximum of approximately 0.4 V.

B. Tapered CML Buffer Experiment

Similar to a CMOS tapered buffer, a single CML buffer might not be sufficient to drive an off-chip load. There are, however, more design tradeoffs involved in the design of a CML tapered buffer than in a CMOS tapered buffer. A superior high-frequency performance in a CML buffer is guaranteed only if the design guidelines explained thoroughly in Section III to be taken into consideration.

Fig. 18(a) plots propagation delay as a function of number of CML stages for different values of X , where X is the ratio between the off-chip load impedance and the load impedance of the first predriver stage. In practice, X is between 30–100. The optimum number of buffer stages will thus be between 3 and 4 versus number of stages for tapered CMOS buffer designed in $0.18 \mu\text{m}$ CMOS process. The delay variation in terms of the number of stages for CML tapered buffer and CMOS tapered buffer are almost identical. However, the total propagation delay of a CML buffer chain for a given value of X is less than that of CMOS buffer chain, which is in accordance with what is expected. Remember that 50% propagation delay of a CMOS inverter is inversely proportional to nMOS and pMOS transconductance parameters and directly proportional to the load capac-

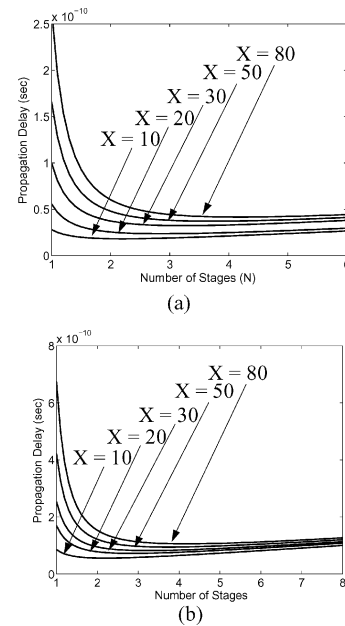


Fig. 18. (a) Delay versus number of stages for CML tapered buffer chain. (b) Delay versus number of stages for CMOS tapered buffer chain.

itance [1]. According to (13), the propagation delay of a CML buffer is directly proportional to the load capacitance (similar to a CMOS inverter) and the drain resistance. A larger threshold voltage and a lower drift velocity associated with a pMOS transistor cause the propagation delay of a CMOS inverter to be larger than that of a CML buffer that uses the same transistor size [Fig. 18(a) and (b)].

C. Inductive Peaking

The inductive peaking was proposed as an efficient and simple circuit technique to speed up the buffer's response.

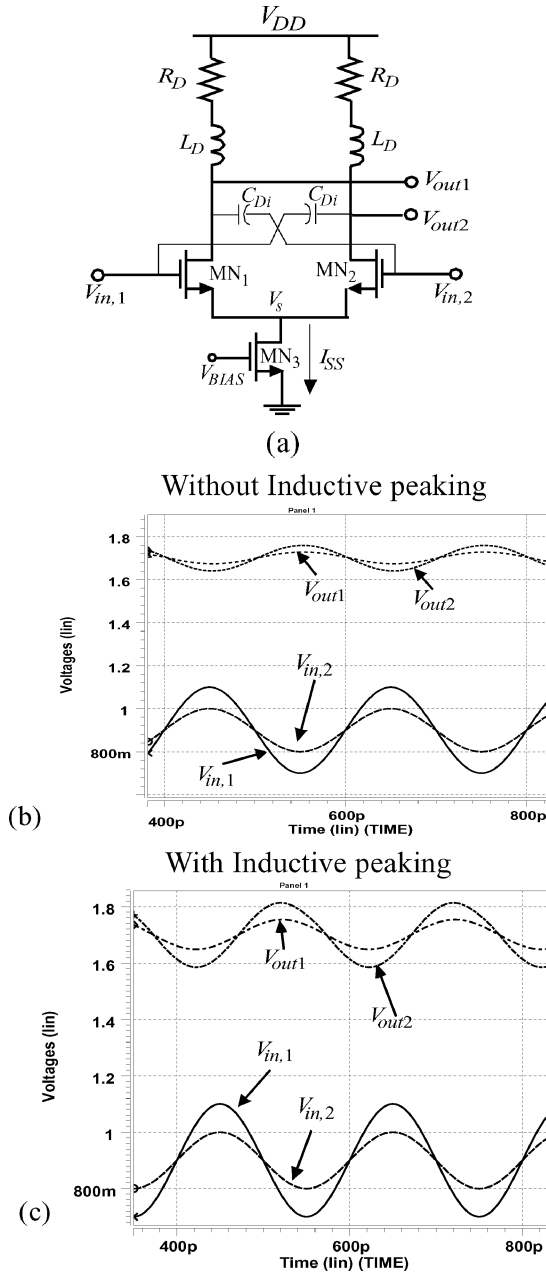


Fig. 19. (a) Single stage CML buffer with inductive peaking. (b) Input and output waveforms of a CML buffer without inductive peaking. (c) Input and output waveforms of a CML buffer with inductive peaking.

Fig. 19(b) and (c) demonstrate the differential output voltage of a CML buffer without and with the inductive peaking [as depicted in Fig. 19(a)], respectively. The inductance value is 2 nH and signals are running at 5 GHz which is the frequency set forth in SONER/SDH OC-48. The output voltages of CML buffer in the presence of inductance will have larger amplitude and as a result faster rise and fall times.

D. CML Latch

The performance comparison of the latch circuits are made by separately incorporating these latches in an ultrahigh-speed flip-flop that retimes the input data with a rate of 20 Gb/s using a half-rate clock signal at 10 GHz that is locked to the input data.

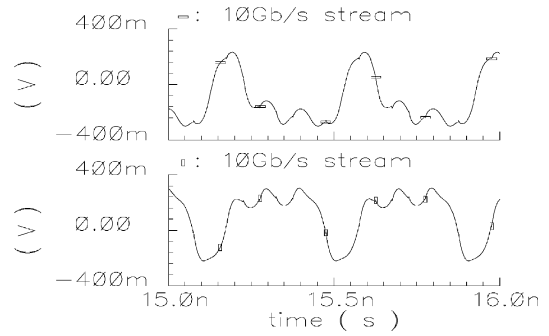


Fig. 20. Voltage waveforms at the output of a flip-flop consisting of two conventional latch circuits of Fig. 11.

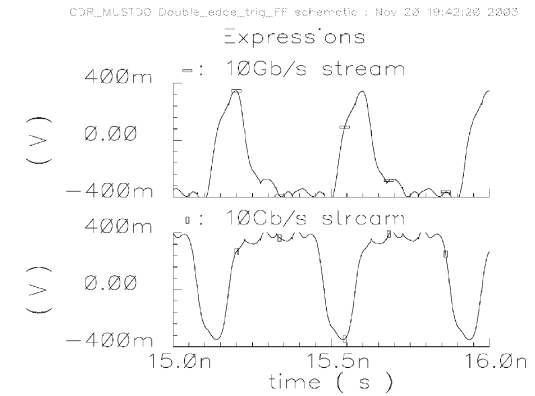


Fig. 21. Voltage waveforms at the output of a flip-flop consisting of two new latch circuits of Fig. 12.

The actual outputs are differential 10 Gb/s data streams demultiplexed from a 20 Gb/s data stream. Four latches are used to create a double-edge triggered flip-flop. The first latch of the flip-flop drives a single latch while second latch of the flip-flop drives a CML buffer. To perform a meaningful and sound comparison, all latches are designed to be identical in terms of the current levels, transistor sizes and drain resistors. The proposed CML latch circuit in Fig. 12 has a superior performance compared to the one shown in Fig. 11 at ultrahigh-frequencies (≥ 10 GHz) for the input datarate. Figs. 20 and 21 demonstrate outputs of the master-slave flip-flop circuits consisting of latch circuits of Figs. 11 and 12 at 20 GHz datarate, respectively. The output nodes of the flip-flop that is made up of conventional CML latches generate large ringings that can yield operation failure (*cf.* Fig. 20). The ringings are largely reduced at the output voltages of the flip-flop consisting of latch shown in Fig. 12. Besides, the output signal transients are smaller compared to those in the conventional flip-flop circuit. Fig. 22 shows the output voltages of the flip-flop based on the latch circuit of Fig. 13. The output voltages exhibit even smaller rise and fall times and sharper transition edges compared to both (20) and (21).

As illustrated in Section V, the latch circuit in Fig. 13 also diminishes the current spikes at the tail current. This observation is verified by comparing the current waveforms of Figs. 23–25 for the latch circuits shown in Figs. 11–13, respectively. While the tail currents MN_{N5} – MN_{N6} of the latch Fig. 11, and the tail currents MN_{N5} and MN_{N8} of the latch in Fig. 12 exhibit spikes,

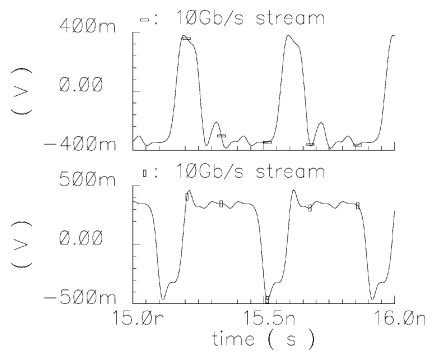


Fig. 22. Voltage waveforms at the output of a flip-flop consisting of two novel latch circuits of Fig. 13.

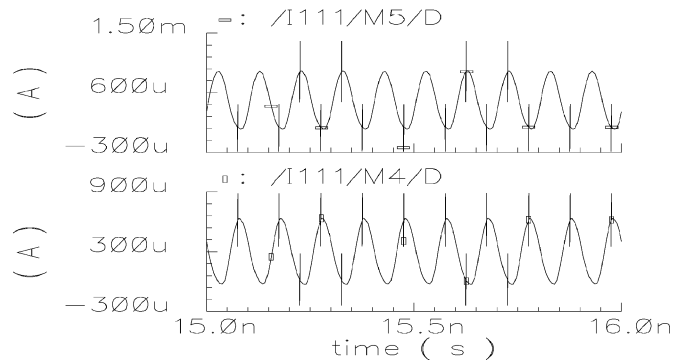


Fig. 23. Current waveforms of the tail currents MN_{N5} and MN_{N6} of the conventional latch in Fig. 11.

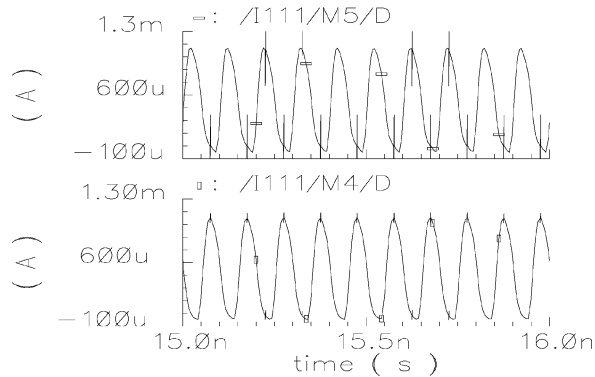


Fig. 24. Current waveforms of the tail currents MN_{N5} and MN_{N8} of the new latch in Fig. 12.

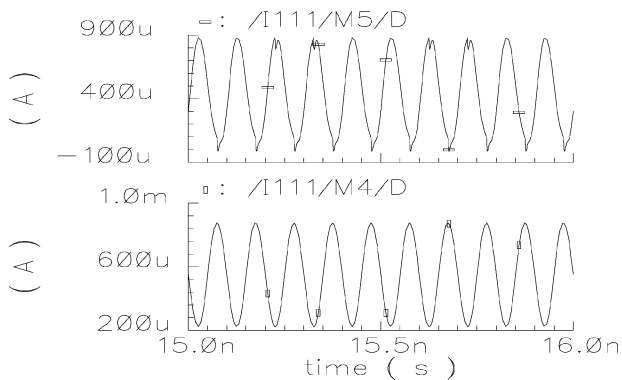


Fig. 25. Current waveforms of the tail currents MN_{N5} and MN_{N8} of the new latch in Fig. 13.

the tail currents MN_{N5} and MN_{N8} of the latch circuit in Fig. 13 does not show any spike.

VIII. CONCLUSIONS

In this paper we investigated important problems involved in the design of a CML buffers and latches. A new design procedure to systematically design a chain of tapered CML buffers was proposed. We proved that the differential architecture of a CML buffer makes it functionally robust in the presence of environmental noise sources (e.g., crosstalk, power/ground noise). Two new 20 GHz regenerative latch circuits were introduced. Experimental results show higher performance for the new latch architectures compared to the conventional CML latch circuit. It was also shown, both through the experiments and by using efficient analytical models, why CML buffers are better than CMOS inverters in high-speed low-voltage applications.

ACKNOWLEDGMENT

The authors thank Jazz Semiconductor, Inc., Newport Beach, CA, for providing the device and simulation data, and in particular, M. Racanelli and P. Colestock for their help and support.

REFERENCES

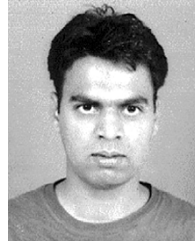
- [1] J. Rabaey, *Digital Integrated Circuits: A Design Perspective*. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [2] B. Razavi, "Prospects of CMOS technology for high-speed optical communication circuits," *IEEE J. Solid-State Circuits*, vol. 37, pp. 1135–1145, Sept. 2002.
- [3] M. Mizuno, M. Yamashina, K. Furuta, H. Igura, H. Abiko, K. Okabe, A. Ono, and H. Yamada, "A GHz MOS adaptive pipeline technique using MOS current-mode logic," *IEEE J. Solid-State Circuits*, vol. 31, pp. 784–791, June 1996.
- [4] K. Irvani, F. Saleh, D. Lee, P. Fung, P. Ta, and G. Miller, "Clock and data recovery for 1.25 Gb/s ethernet transceiver in 0.35 μm CMOS," in *Proc. IEEE Custom Integrated Circuits Conf.*, May 2001, pp. 261–264.
- [5] H.-T. Ng and D. J. Allstot, "CMOS current steering logic for low-voltage mixed-signal integrated circuits," *IEEE Trans. VLSI Syst.*, vol. 5, pp. 301–308, Sept. 1997.
- [6] A. Tanabe, M. Umetani, I. Fujiwara, K. Kataoka, M. Okihara, H. Sakuraba, T. Endoh, and F. Masuoka, "0.18- μm CMOS 10-Gb/s multiplexer/demultiplexer ICs using current mode logic with tolerance to threshold voltage fluctuation," *IEEE J. Solid-State Circuits*, vol. 36, pp. 988–996, June 2001.
- [7] H.-D. Wohlmut, D. Kehrer, and W. Simburger, "A high sensitivity static 2:1 frequency divider up to 19 GHz in 120 nm CMOS," in *Proc. IEEE Radio Frequency Integrated Circuits (RFIC) Symp.*, June 2002, pp. 231–234.
- [8] M. H. Anis and M. I. Elmasry, "Self-timed MOS current mode logic for digital applications," in *Proc. IEEE Int. Conf. ASIC/SOC*, 2002, pp. 193–197.
- [9] —, "Self-timed MOS current mode logic for digital applications," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 5, May 2002, pp. 113–116.
- [10] B. Razavi, *Design of Analog CMOS Integrated Circuits*. New York: McGraw-Hill, 2001, pp. 101–134.
- [11] S.-M. Kang and Y. Leblebichi, *CMOS Digital Integrated Circuits: Analysis and Design*. New York: McGraw-Hill, 1999.
- [12] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
- [13] N. Hedenstierna and K. O. Jeppson, "CMOS circuit speed and buffer optimization," *IEEE Trans. Computer-Aided Design*, vol. CAD-6, pp. 270–281, Mar. 1987.
- [14] H. T. Friis, "Noise figure of radio receivers," *Proc. IRE*, vol. 32, pp. 419–422, July 1944.



Payam Heydari (S'98–M'00) received the B.S. degree in electronics engineering and the M.S. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 1992 and 1995, respectively, and the Ph.D. degree in electrical engineering at the University of Southern California, Los Angeles, in 2001.

During the summer of 1997, he was with Bell-labs, Lucent Technologies, where he worked on noise analysis in deep submicron VLSI circuits. He worked at IBM T. J. Watson Research Center on gradient-based optimization and sensitivity analysis of custom integrated circuits during the summer of 1998. Since August 2001, he has been an Assistant Professor of electrical engineering at the University of California, Irvine, where his research interests are design of high-speed analog, RF, and mixed-signal integrated circuits, and analysis of signal integrity and high-frequency effects of on-chip interconnects in high-speed VLSI circuits.

Dr. Heydari received the Best Paper Award at the 2000 IEEE International Conference on Computer Design (ICCD). He also received the Technical Excellence Award from the Association of Professors and Scholars of Iranian Heritage in California in 2001. He serves as a Member of the Technical Program Committees of the IEEE Design and Test in Europe (DATE), the International Symposium on Physical Design (ISPD), the International Symposium on Quality Electronic Design (ISQED), and the International Symposium on Low-Power Electronics and Design (ISLPED).



Ravindran Mohanavelu (M'01) received the B.S. degree from the Indian Institute of Technology (IIT) Kanpur, in 1998, and the M.S. degree in material science and engineering and from The Iowa State University, Ames, in 2000. He is currently working toward the M.S. degree in electrical engineering from the University of California, Irvine.

He is currently working as a design engineer at International Rectifier, El Segundo, CA. His research interests include broadband integrated circuit design.