Behzad Razavi

The Cross-Coupled Pair—Part II

Following a general overview of the cross-coupled pair (XCP) in the last issue, we begin to study specific circuit examples incorporating this topology. We deal with digital applications in this issue.

The performance of digital circuits can be improved if an XCP is tied between complementary (differential) signals. Specifically, we can add a clocked XCP or replace the PMOS devices in complementary logic with an XCP. The circuits described here exemplify the utility of these techniques.

Digital Object Identifier 10.1109/MSSC.2014.2352532
Date of publication: 12 November 2014

Sense Amplifiers

We examine sense amplifiers not necessarily because we wish to design memories but, rather, because the techniques studied here prove useful in many other applications as well. A common situation in digital (or analog) design is that a small initial imbalance, V_{XY0}, appearing between two differential nodes must be amplified, as fast as possible, to (preferably) rail-to-rail complementary signals. The circuit can be designed such that the two nodes are driven by a high impedance and travel toward the rails with a "natural" time constant [Figure 1(a)]. Assuming that X and Y are released with an initial imbalance of V_{XY0} (e.g., by means of a switch) and that $V_{in1} - V_{in2}$ is large enough to ensure $g_{m1,2} R_L | V_{in1} - V_{in2} | \approx V_{DD}$, we can ask, how much time does the circuit take to provide a certain gain, G [1]? Defining the time-dependent gain as $G = V_{XY}(t_1)/V_{XY0}$, we have

$$\frac{t_1}{\tau_1} = -\ln\left(1 - \frac{G}{A_1}\right), \qquad (1)$$

where $\tau_1 = R_L C_L$ and $A_1 = V_{XY}(\infty)/V_{XY0} \approx g_{m1,2} R_L$ (the "dc" gain). On the

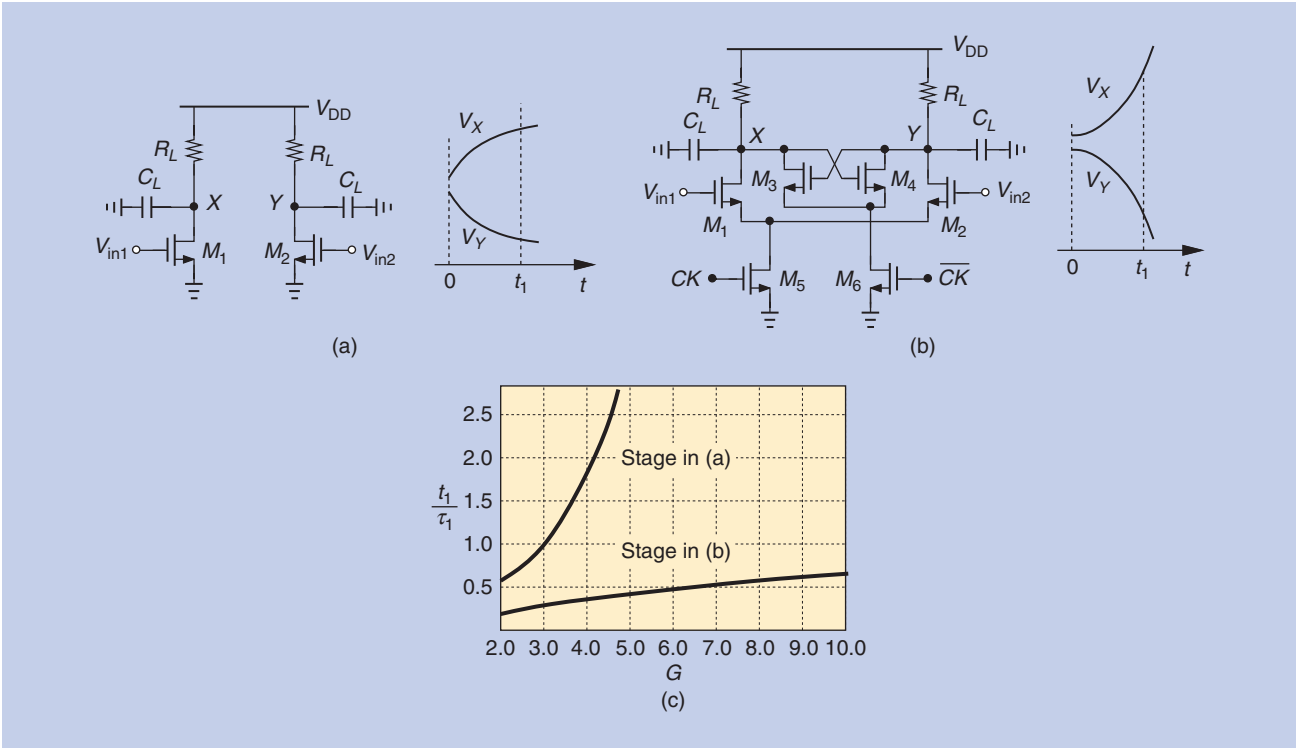

FIGURE 1: (a) A simple circuit starting with an initial difference between X and Y, (b) regenerative amplification provided by the XCP, and (c) required normalized time for obtaining a gain of G.

other hand, we can attach a clocked XCP to the differential nodes [Figure 1(b)] and disable M_1 and M_2 when the XCP turns on. Now, V_{XY} is amplified regeneratively, evolving as $V_{XY} = V_{XY0}\exp(t/\tau_{reg})$, where $\tau_{reg} = R_L C_L/(g_{m3,4}R_L - 1)$. The time necessary to obtain a certain gain, G, is given by [1]

$$\frac{t_1}{\tau_1} = \frac{1}{A_2 - 1}\ln G, \qquad (2)$$

where $A_2 = g_{m3,4}R_L$. Figure 1(c) plots t_1/τ_1 for both cases, assuming $A_1 \approx A_2 \approx 5$. We thus observe the significant advantage of regenerative amplification in sense amplifier design.

Figure 2 shows a DRAM sense amplifier dating back to 1976 [2]. In the precharge (or "equalization") mode, the NMOS devices M_3 and M_4, respectively, pull V_X and V_Y to approximately $V_{DD} - V_{TH}$ and S_1 shorts X and Y to remove residual offsets from the previous cycle (and the threshold mismatch between M_3 and M_4). In the evaluation mode, these three devices turn off, X and Y sense the single-ended memory cell level and a reference voltage, and the XCP is clocked to amplify the difference rapidly. (NMOS pull-up devices are chosen here perhaps to match the cell's common-mode level.) It is important to note that the clocking of sense amplifiers must ensure three distinct time intervals: precharge (equalization), bit line sense, and amplification.

As observed in the waveforms of Figure 2, the output voltages of the above sense amplifier cannot return to V_{DD} in the amplification mode, suffering from a degradation in their high levels. Figure 3 depicts another topology [3] employing both NMOS and PMOS cross-coupled pairs and allowing rail-to-rail swings. In this case, X and Y are precharged to V_{DD} and then connected to the memory cell. Other variants of these topologies are described in numerous papers, e.g., [5]–[7].

The problem of MOS device mismatches began to manifest itself in sense amplifiers as higher speeds and lower supplies were sought. Figure 4 shows an early example of offset cancellation within a DRAM sense amplifier [8].

In the precharge mode, CK_2 is low and the gates of M_1 and M_2 are pulled to V_{DD}, allowing V_A and V_B to assume values equal to $V_{DD} - V_{TH1}$ and $V_{DD} - V_{TH2}$, respectively. The XCP threshold mismatch is thus stored on C_1 and C_2. In the evaluation mode, first CK_2 falls to zero, initiating positive feedback around M_1 and M_2, and then CK_3 goes high, accelerating the amplification [8].

Latches

The XCP's fast amplification property also proves useful in latch design. If speed is the primary concern, current-mode logic (CML) latches with moderate voltage swings (200–300 mV single ended) can be used. Originally realized using bipolar transistors, CML latches have been inherited by CMOS technology as well. Figure 5 shows an example where M_1 and M_2 form a preamplifier and M_3 and M_4 an XCP. (The term "latch" sometimes refers to the entire circuit or just the XCP plus the load resistors.) When CK is high, the input is amplified and impressed at X and Y. When CK goes low, the XCP turns on, regenerating $V_X - V_Y$ to a final value of $I_{SS}R_D$. This condition is met if $g_{m3,4}R_D > 1$. Since the output swing, $I_{SS}R_D$, need not be as large as V_{DD}, the circuit operates faster than topologies producing rail-to-rail swings—albeit at the cost of static power.

Several variants of the CML latch have been reported. To reduce the voltage headroom consumption, the tail current source can be removed [as in Figure 1(b)] while M_5 and M_6 are biased in a current-mirror arrangement [9]. In this case, the total current flowing through the clocked devices is not constant, leading to "class-AB" operation and improving the speed. The CML latch can also incorporate inductive peaking [10, 11] so as to achieve a higher speed [Figure 5(b)]. To ensure minimal overshoot and intersymbol

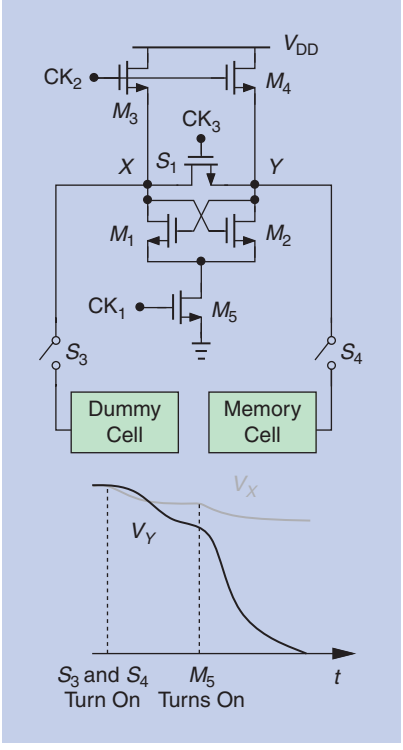

FIGURE 2: The sense amplifier reported in [2].

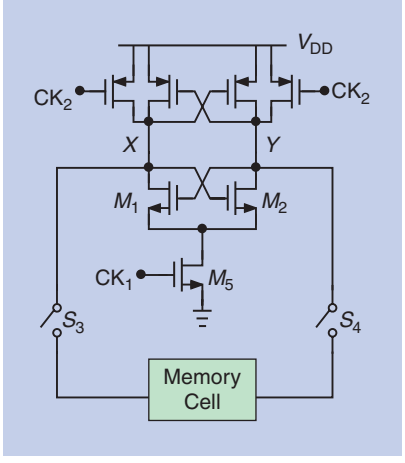

FIGURE 3: A sense amplifier with rail-to-rail outputs [3].

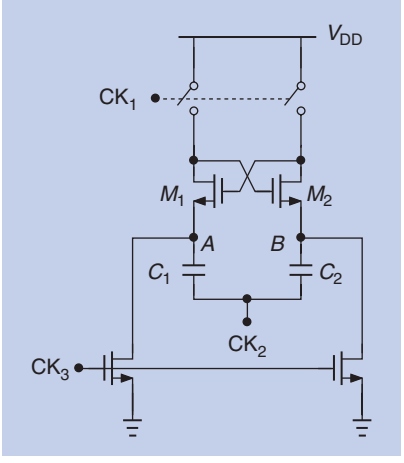

FIGURE 4: A sense amplifier with threshold mismatch cancellation [8].

interference, the damping factor of the RLC load, $\zeta = (R_D/2)\sqrt{C_L/L_1}$ (C_L is the capacitance at X or Y), should be greater than approximately 0.7. The static power and the large area consumed by this topology are justified only if ultimate speed (tens of gigahertz in 28-nm technology) is necessary.

For operation at lower speeds, the latch shown in Figure 6(a) is an attractive choice. Originating from the "cascode voltage switch logic" (CVSL) [12] and its successor [13], the latch senses and generates rail-to-rail swings, consuming no static power. If, for example, X is high and CK rises while D is also high, then the circuit reduces to that shown in Figure 6(b), where the series combination of M_1 and M_5 must "overcome" M_3. As V_X falls, M_4 turns on, enabling regeneration around the PMOS loop.

The foregoing latch merits two remarks. First, it operates with ratioed logic, demanding proper sizing of the transistors. In practice, if $W_{1,2} = W_5 = W_{3,4}$ (and the lengths are equal), the NMOS devices can robustly change the state. Second, the complementary input and output swings produce less substrate noise than do single-ended logic families, a useful property in mixed-signal design.

The latch described above can also include logic if the signals are available in complementary form [12]. Shown in Figure 7 is an example of a differential CVSL NOR gate embedded in the latch, exhibiting a smaller input capacitance than that of complementary logic. In this case, the series combination of M_3-M_5 must overcome M_7.

Questions for the Reader

1) Must we turn off M_1 and M_2 in Figure 1(b) as we activate the XCP even if the circuit is to operate as an amplifier rather than as a latch?
2) A divide-by-two circuit incorporates two instances of the latch shown in Figure 5(b). Can R_D be reduced to zero in this case?

You can share your thoughts by e-mailing me.

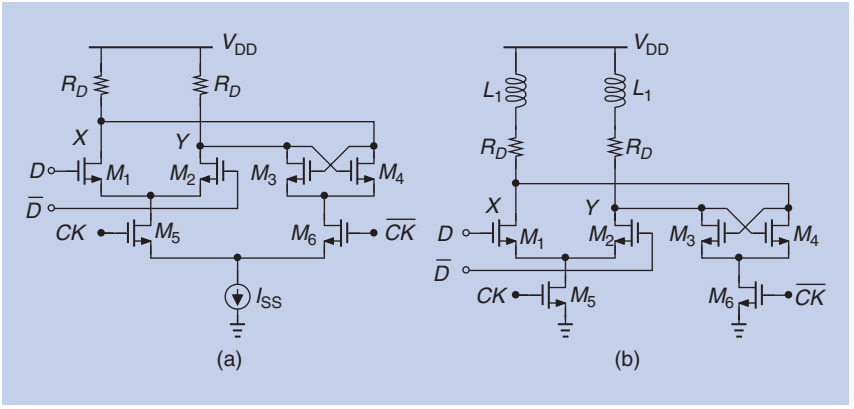

FIGURE 5: (a) A CML latch and (b) its modified version with inductive peaking and no tail current source.

Answers to Last Issue's Questions

1) Is negative capacitance the same as positive inductance? No, it is not. The former's impedance is given by $j/(C\omega)$ and its magnitude falls with frequency whereas the latter's impedance, $jL\omega$, has a magnitude that rises with frequency.
2) Can the cancellation of positive capacitance by negative capacitance be a *resonance* effect? No, it cannot. In a resonator, the phase difference between the voltage and the current changes sign at the resonance frequency. The series or parallel combination of a positive capacitance and a negative capacitance does not display such a property.
3) Why is the circuit in Figure 8 a *dynamic* latch? If used as an RS latch, the circuit allows its inputs to be low simultaneously. Since M_1 and M_2 remain off, the leakage currents at the drain nodes can corrupt the state stored by M_3 and M_4.
4) In Figure 9, M_1 and M_2 are biased and balanced by I_1 and I_2 ($I_1 = I_2$). At $t = 0$, I_{in} jumps from zero to a small positive value, I_0. We intuitively expect that V_X rises and V_Y falls. However, viewing the XCP as a resistance equal to $-2/g_m$, we obtain $V_{XY} = (-2/g_m)I_0 u(t)$, concluding that V_X should descend and V_Y should ascend! How do we explain the discrepancy between these two results?

Our intuition in fact assumes that each node bears some capacitance to ground. The differential equation governing the circuit is as follows:

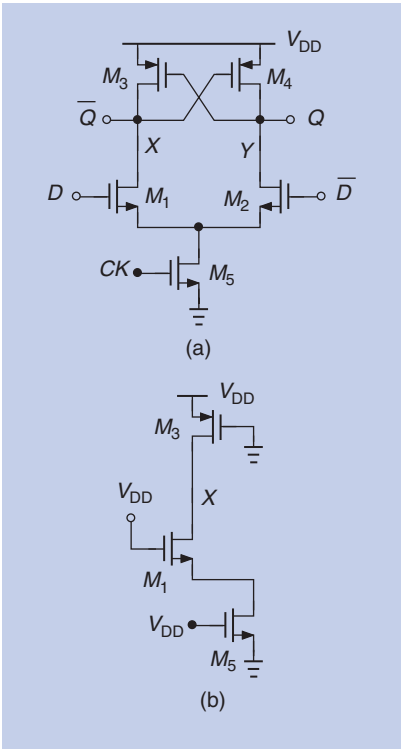

FIGURE 6: (a) A rail-to-rail latch, and (b) a simplified circuit during state change.

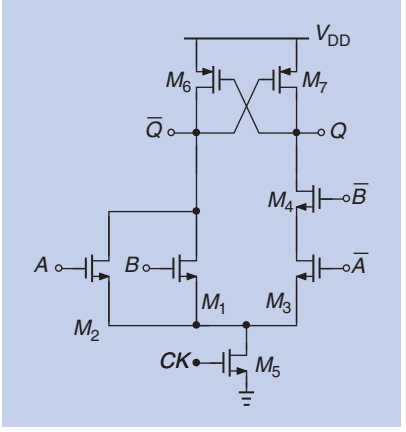

FIGURE 7: A NOR gate embedded in latch.

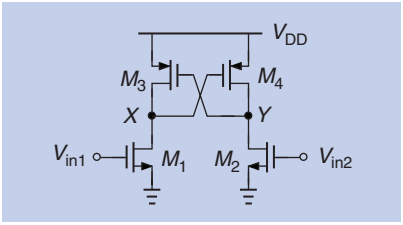

FIGURE 8: A differential buffer using an XCP.

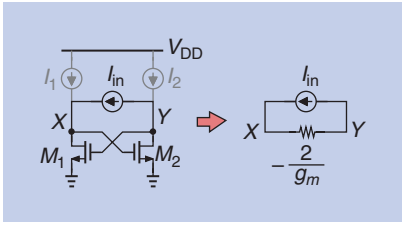

FIGURE 9: XCP operation from two perspectives.

$$C_L \frac{dV_{XY}}{dt} - g_m V_{XY} = 2I_0, \qquad (3)$$

where C_L is the total capacitance at each node. We are curious to see whether this equation's solution approaches $V_{XY} = (-2/g_m)I_0$ as $C_L \to 0$. To solve (3), we would ordinarily write $C_L dV_{XY}/(2I_0 - g_m V_{XY}) = dt$ and integrate both sides, assuming $V_{XY}(t=0) = 0$ and hence obtaining

$$V_{XY}(t) = \frac{2I_0}{g_m}\left[\exp\left(\frac{g_m}{C_L}\right)t - 1\right]. \qquad (4)$$

Unfortunately, this result does not lead to $V_{XY}(t) = -2I_0/g_m$ if $C_L \to 0$. This is because our solution has tacitly assumed that a) $2I_0 - g_m V_{XY} \neq 0$, b) $C_L \neq 0$, and c) $V_{XY}(t=0) = 0$, all of which are violated when $C_L = 0$. To solve the differential equation without these presumptions, we assume $V_{XY}(t)$ can be expressed as $\alpha \exp(\beta t) + \gamma$ and substitute for it in (3). It follows that

$$C_L\alpha\beta\exp(\beta t) - g_m\alpha\exp(\beta t) - g_m\gamma = 2I_0. \qquad (5)$$

Taking the derivative of both sides yields $(C_L\alpha\beta^2 - g_m\alpha\beta)\exp(\beta t) = 0$, i.e., $\alpha\beta(C_L\beta - g_m) = 0$ if $\beta < \infty$. This result points to different possibilities: 1) if $C_L\beta - g_m = 0$, then we can also assume $V_{XY}(t=0) = 0$ and arrive at (4); 2) if $\alpha = 0$, then (5) implies that $\gamma = -2I_0/g_m$ and $V_{XY}(t) = -2I_0/g_m$; 3) if $\beta = 0$, then (5) suggests that $\alpha + \gamma = -2I_0/g_m$ and, since $V_{XY}(0) = \alpha + \gamma$, we have $V_{XY}(t) = \alpha + \gamma = -2I_0/g_m$.

References

[1] J. T. Wu, "High-speed analog-to-digital conversion in CMOS VLSI," Ph.D. dissertation, Stanford Univ., 1988.

[2] C. N. Ahlquist et al., "A 16K dynamic RAM," in *ISSCC Dig. Tech. Papers*, Feb. 1976, pp. 128–129.

[3] S. Konishi et al., "A 64Kb CMOS RAM," in *ISSCC Dig. Tech. Papers*, Feb. 1982, pp. 258–259.

[4] J. M. Schlageter et al., "A 4K static 5-V RAM," in *ISSCC Dig. Tech. Papers*, Feb. 1976, pp. 136–137.

[5] T. Watanabe et al., "A battery backup 64K CMOS RAM with double level aluminum technology," in *ISSCC Dig. Tech. Papers*, Feb. 1983, pp. 60–61.

[6] A. Mohsen et al., "An 80ns 64K DRAM," in *ISSCC Dig. Tech. Papers*, Feb. 1983, pp. 102–103.

[7] K. Sasaki et al., "A 9ns 1Mb CMOS SRAM," *ISSCC Dig. Tech. Papers*, Feb. 1989, pp. 34–35.

[8] T. Mano et al., "Submicron VLSI memory circuits," in *ISSCC Dig. Tech. Papers*, Feb. 1983, pp. 234–235.

[9] J. Lee and B. Razavi, "A 40-Gb/s clock and data recovery circuit in 0.18-um CMOS technology," *IEEE J. Solid-State Circuits*, vol. 38, pp. 2181–2190, Dec. 2003.

[10] M. Wurzer et al., "42 GHz static frequency divider in a Si/SiGe bipolar technology," in *ISSCC Dig. Tech. Papers*, Feb. 1997, pp. 122–123.

[11] P. Heydari and R. Mohavavelu, "Design of ultra high-speed CMOS CML buffers and latches," in *Proc. ISCAS*, May 2003, pp. 208–211.

[12] L. Heller et al., "Cascode voltage switch logic: A differential CMOS logic family," in *ISSCC Dig. Tech. Papers*, Feb. 1984, pp. 16–17.

[13] L. C. Pfennings et al., "Differential split-level CMOS logic for sub-nanosecond speeds," *IEEE J. Solid-State Circuits*, vol. 20, pp. 1050–1055, Oct. 1985.