

A Synchronous High-Speed, High-Accuracy, Loser-Take-All Circuit

Bogdan M. Wilamowski and Don L. Jordan,

University of Wyoming, Department of Electrical Engineering, Laramie, Wyoming 82071, U.S.A.
wilam@ieee.org, dlj@trib.com.

I. Abstract

This paper presents an accurate current mode synchronous Loser-Take-All circuit based on a simple regenerative pair. The regenerative pair and the related resetting transistors lead to a circuit requiring only four transistors, seven for deep binary tree structures. It achieves its high speed through regenerative feedback. The basic circuit requires no current mirrors and has only parasitic losses, resulting in high accuracy. This paper also includes simulation results for a single LTA circuit with two inputs and for a sixty-four input, six-layer, binary tree circuit.

II. Introduction

The MAX or Winner-Take-All (WTA) and MIN or Loser-Take-All (LTA) circuits are important parts of many neuro-fuzzy systems. MIN and MAX circuits are the prime components of fuzzy logic. WTA circuits are used in many neural layer architectures such as Learning Vector Quantization (LVQ), Adaptive Resonance Theory (ART), Kohonen feature maps, and many others.

Günay and Sánchez-Sinencio [1] provide a detailed overview and comparison of several CMOS WTA circuits. However, this paper limits its consideration to asynchronous circuits nor does it mention several newer concepts [2], [3], [4]. Asynchronous WTA circuits can find the winner as the signals change due to their ability to process and compare signals in continuous mode. On the other hand, these circuits tend to have rather limited accuracy for multiple inputs due to signal interaction and the difficulty in matching transistors which are not near to each other. Usually, one must limit the number of inputs to ten or fewer to obtain acceptable resolution.

Some applications, such as LVQ for graphics compression, require selecting the winner out of hundreds of signals. Demosthenous, et al., [5] developed a binary tree approach that solves this problem with a circuit using three current mirrors and one latch per synchronous WTA stage. This leads to a complex circuit with sixteen

transistors per cell, twelve when eliminating one current mirror by using complimentary design techniques, and slower speeds.

This paper proposes an alternative approach. Take the DeMorgan transform [2] of the inputs, then compare them using synchronous LTA circuits. The basic circuit presented here uses a regenerative pair as the LTA network, thus it requires only four transistors, two for the regenerative pair and two to reset the circuit. Its design eliminates the need for current mirrors which results in improved speed and accuracy when compared against the WTA approach. It also has the advantage that it does not require extensive circuitry to identify the input with the minimum current.

III. The Basic LTA Circuit

The LTA circuit shown in Fig. 1 has behavior similar to that of flip-flops and of sense amplifiers found in dynamic memories. Fig. 2 shows the post-layout SPICE simulation results of the basic LTA circuit for a $2\mu\text{m}$, N-Well process (MOSIS SCNA20) with $6\mu\text{m}$ wide by $2\mu\text{m}$ long transistors and a 5V supply. Transistors M1 and M2 have merged sources and the M1-M2 and M2-M4 pairs have merged drains to minimize parasitic capacitance. Each current input includes an additional capacitance of 5fF to ground in every cell to account for parasitics from the wiring and the current sources. The *RESET* pulse has a delay of 5ns and a fall time of 5ns.

The simulation results show that the circuit has five distinct operating intervals: RESET, DECISION, SWITCHING, CHARGING, and LATCH. Consider the case where I_b is the larger current. The RESET interval occurs while the *RESET* line has a high level on it, turning transistors M3 and M4 on and discharging the capacitances at their drains. The DECISION interval commences during the falling edge of *RESET* when transistors M3 and M4 start to come out of their ohmic regime and transistors M1 and M2 go into subthreshold conduction. M1 and M2 commit to one of the inputs while in the subthreshold regime. Consequently, the circuit will

make a commitment even if the inputs are equal due to inherent noise.

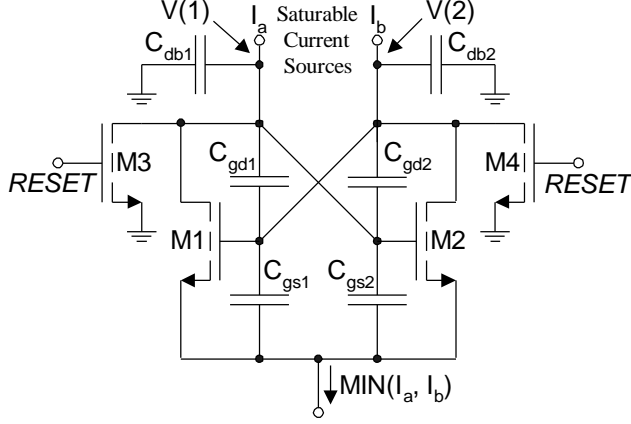


Fig. 1. Proposed LTA Circuit Showing the Booster Section and Lumped Parasitic Capacitances.

The input currents charge the parasitic capacitances until the voltage at the gate of M1 equals its threshold voltage and the SWITCHING interval begins. The capacitance charged by I_a dumps through M1, producing a spike in the output current. The CHARGING interval begins when M1 enters its ohmic regime. I_b continues to charge the gate of M1 and the gate-to-drain capacitance of M2 (C_{gs1} , C_{gd1} , and C_{gd2}), producing a plateau in the output due to the displacement current, and driving M1 even deeper into its ohmic operating mode until the current source saturates. The convergence of the output indicates the start of the LATCH interval, which lasts until the next RESET interval begins with a rising edge on the RESET line. M2 now blocks I_b during the LATCH interval while M1 provides an ohmic path to any subsequent layer for I_a . The blocking action of the cut-off transistors requires that the current sources be able to saturate. Typical current sources have this capability.

Simulation results indicate that the time the LTA circuit takes to converge depends more upon the larger current input than upon the smaller and that it is fairly immune to the difference between the inputs. The smaller current only has time to charge partially the capacitances associated with its input before the LTA cell begins to carry it to the following cell. The DECISION plus SWITCHING intervals take about 9.9ns to complete while the CHARGING interval takes another 22.7ns for the output current to converge for the low level input currents found in Fig. 2. This results in a total elapsed time of 32.6ns for the single cell. A high current example of $\sim 300\mu\text{A}$ versus $0.1\mu\text{A}$ converged in only 5ns and

demonstrated a large dynamic range. Note that the circuit performs correctly even in the presence of a large ohmic drop across the reset transistors. Also, note that the voltages at the current sources are strongly binary, even for the high current conditions, with the maximum input near the positive supply and the losing output seeing only an ohmic drop.

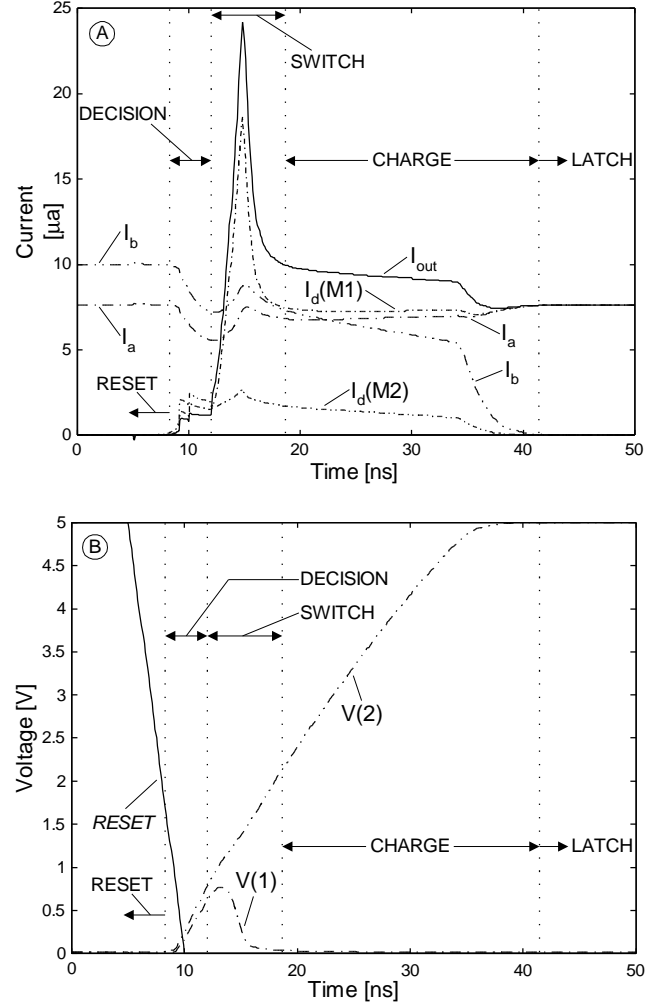


Fig. 2. Simulation Results for the LTA Cell for $I_a = 7.5\mu\text{A}$, $I_b = 10.0\mu\text{A}$. (A) Current Waveforms, (B) Voltage Waveforms.

IV. Extending the Basic LTA for Multiple Inputs

Fig. 3 shows how to connect the LTA circuit in a binary tree for processing a large number of inputs while Fig. 4 shows the SPICE simulation results for a six-layer, sixty-four input, binary tree. The minimum input to this circuit, I_1 , was about $49\mu\text{m}$. The next larger input, I_{63} , was only $49.3\mu\text{m}$ while the largest input, I_{59} , was $130\mu\text{m}$. This example used a single RESET signal for all layers with a

delay time of 5ns and a fall-time of 5ns. The binary tree converged in 51ns, the time indicated in Fig. 4 from the beginning of the DECISION interval to when I_{OUT} tracks with I_1 . I_1 has noticeable droop during the convergence period due to the relatively poor quality of the single-transistor current sources used in the example, which contribute significantly to the overall time to converge.

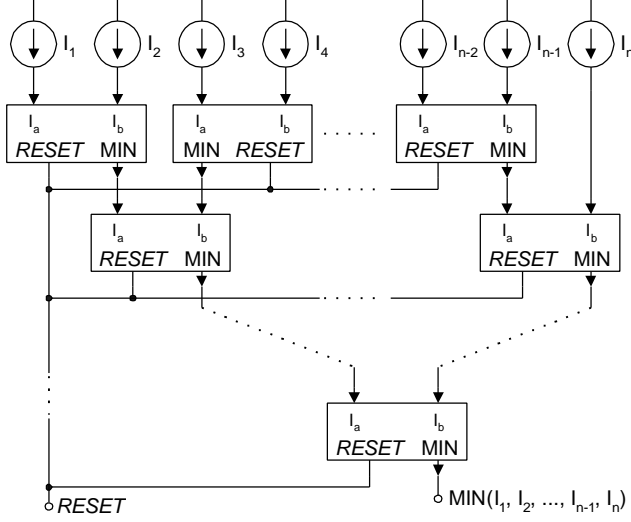


Fig. 3. Binary Tree LTA Circuit for Multiple Inputs.

The high current present in the output during the SWITCHING and CHARGING intervals serve to speed the next LTA layer by injecting supplemental charge into its input capacitances. The spike does not affect the accuracy of the comparison in the subsequent layer as each LTA cell in a given layer contributes the same amount of charge into the cells that follow. Furthermore, charge-sharing during the discharge of the parasitic capacitors prevents any premature commitment of the LTA transistors since the charge dump comes from those charged by a previously “losing” input. Equation (1) shows the amount of charge contributed by I_b to the output.

$$\begin{aligned} Q_{I_b} &= V_{dd} \cdot (C_{gs1} + C_{gd1} + C_{gd2}) \\ &\equiv \text{constant for a given layer} \end{aligned} \quad (1)$$

Distinguishing the input with the minimum current simply requires identifying which input has the lowest voltage. All inputs with larger currents will be in saturation and near the positive supply. Only the input with the minimum current will be low — offset from the voltage at the output of the binary tree only by the ohmic drops of the tree layers.

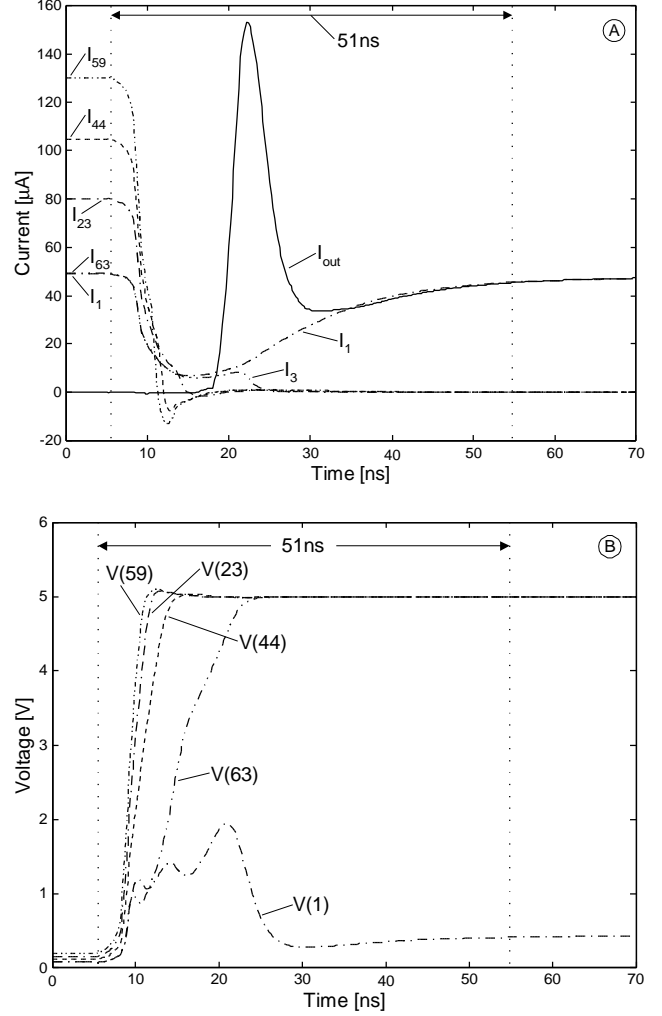


Fig. 4. Simulation Results for the Six-Layer, 64-Input Binary Tree LTA. (A) Current Waveforms, (B) Voltage Waveforms.

The LTA cells in this example use the optional “booster” shown in Fig. 5. Deep binary tree structures need the booster for reliable operation. Consider an input that initially passes through the topmost layers of the tree before it becomes the larger input to a cell. The voltage at the input to that cell increases as the current charges the parasitic capacitances, but this also increases the voltage at the sources of transistors in the preceding cells. Eventually, the first cell cuts off as the gate-to-source voltage of its ohmic transistor drops below the threshold voltage, thus keeping the rejecting cell from charging sufficiently. The booster section detects when a following cell rejects the current passed by the regenerative pair by monitoring the voltage at their sources. M5 and M6 identify when the voltage exceeds a preset level, V_{REF} . M6 then connects M7 to the output. M7 quickly saturates since it only charges the parasitic capacitances of the cell

in the next layer. Boosters do not affect the accuracy of the binary tree since they only activate in branches rejected by lower layers. The booster also insures that any “winning” input sees at most a saturated current source and an ohmic transistor between it and the positive supply.

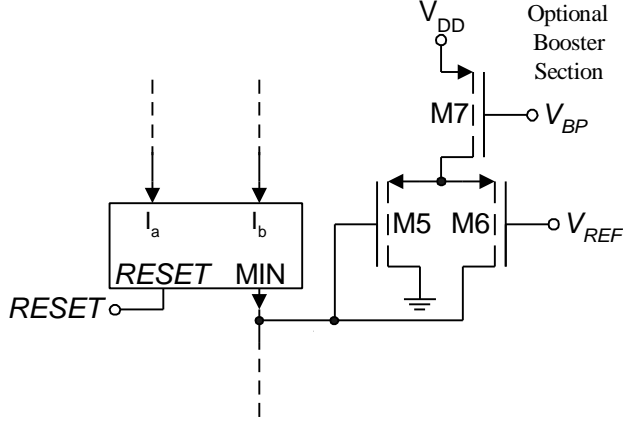


Fig. 5. Application of the Optional Booster Section.

V. LTA Design Considerations

Parasitic capacitances and the maximum input current levels usually dominate the speed of the LTA, indicating the use of the smallest gates possible in the regenerative pair. On the other hand, ohmic drop, threshold voltage (unless one uses boosters), and the supply voltage limit the maximum number of layers. The drain-to-source voltage of the ohmic transistor must not exceed the threshold voltage, thus turning on the other transistor in the regenerative pair (neglecting subthreshold conduction), for the output current to equal the minimum input. Equation (3) shows the upper current limit on the minimum input imposed by this restriction.

$$I_{IN_MAX} < K' \cdot \frac{W}{L} \cdot V_{DS} \cdot \left(V_{GS} - V_{TH} - \frac{V_{DS}}{2} \right) \quad (2)$$

Let $V_{DS} = V_{TH}$, therefore...

$$I_{IN_MAX} < K' \cdot \frac{W}{L} \cdot V_{TH} \cdot \left(V_{GS} - \frac{3}{2} \cdot V_{TH} \right) \quad (3)$$

$V_{GS} \approx V_{DD}$, V_{TH} adjusted for body effect.

The greatest restriction on the number of layers without needing boosters stems from the voltage drops across the individual layers. For example, an eight-input, three-layer binary tree performs correctly when using a 5V supply, but a sixteen-input, four-layer binary tree requires either a higher supply voltage or the addition of boosters in the upper layers.

VI. Conclusion

The LTA circuit presented here appears to achieve the objectives of high speed and high accuracy. It accomplishes its high accuracy by eliminating current mirrors in the circuit and with its simple design of only four transistors, which allows for compact designs with the transistors residing in close proximity to each other. The circuit can process a large number of inputs using a binary tree structure and an additional booster section of three transistors per cell. Its high speed comes from regenerative feedback, small parasitic capacitances resulting from the simple design, and the fact that the larger current, not the smaller, charges the capacitances.

In addition, one can easily extend the concept to rank incoming currents by successive elimination of the losers. Simply detect the input with the lowest voltage, record its location, then supply that input with a large current during the next cycle.

VII. References

- [1] Z. S. Günay and E. Sánchez-Sinencio, "CMOS Winner-Take-All Circuits: A Detail Comparison," 1997 IEEE International Symposium on Circuits and Systems, June 9-12, 1997, Hong Kong, pp. 41-44.
- [2] S. Siskos, S. Vlassis, I. Pitas, "Analog Implementation of Fast Min/Max Filtering," *IEEE Trans. Circuits Syst. II*, vol 45, no. 7, July 1998, pp. 913-918.
- [3] T. Serrano, B. Linares-Barranco, "A Modular Current-Mode High-Precision Winner-Take-All Circuit," *IEEE Trans. Circuits Syst. II*, vol 42, no. 2, February 1995, pp. 132-134.
- [4] B. Sekerkiran, U. Çilingiroglu, "Precision Improvement in Current-Mode Winner-Take-All Circuits Using Gain-Boosted Regulated-Cascade CMOS Stages," *IEEE IJCNN*, 1998, pp. 553-556.
- [5] A. Demosthenous, S. Smedley, J. Taylor, "A CMOS Analog Winner-Take-All Network for Large-Scale Applications," *IEEE Trans. Circuits Syst. I*, vol. 45, no. 3, March 1998, pp. 300-304.