Experimental Results on the Current-Mode WTA-MAX Circuit with Multi-Chip Capability

Teresa Serrano-Gotarredona and Bernabé Linares-Barranco

National Microelectronics Center (CNM), Ed. CICA, Av. Reina Mercedes s/n, 41012 SPAIN.
Phone: 34-5-4239923, Fax: 34-5-4624506, E-mail: bernabe@cnm.us.es

Abstract

This paper presents experimental results on the current-mode WTA-MAX circuit with multi-chip capability presented recently [1], [2]. The circuit technique used is based on current replication, transport and comparison. Traditional techniques rely on the matching of an array of \( N \) transistors, where \( N \) is the number of system inputs. This implies that when \( N \) increases, the size of the circuit and the distance between transistors also increases, resulting in transistor matching degradation and loss of precision in the overall system performance. Furthermore, when multi-chip systems are required, the transistor matching in even worse and performance is drastically degraded. The circuit presented in this paper does not rely on the proper matching of \( N \) transistors, but on the precise replication of currents, performed by current mirrors with a limited number of outputs. Thus \( N \) can increase without degrading the precision, even if the system is distributed among several chips. Also, the different chips constituting the system can be of different foundries without degrading the overall system precision. Experimental results are presented that attest these facts.

I. Introduction

Recently [1]-[2], a current-mode WTA topology was proposed which in principle seemed to present no precision degradation when distributing the WTA among different chips. At that time only simulation results were reported and a dubious stability proof was provided. In this paper we provide experimental results that support the early simulations, and a more rigorous stability analysis is also included. The circuit schematic is shown in Fig. 1. The circuit consists of \( N \) equal cells plus an \( N \)-output current mirror. Each cell consists of a two-output current mirror, a current comparator (implemented with a digital inverter), and a MOS transistor acting as a switch. The reason why this circuit works as a Winner-Take-All and how it can be split into several chips has been shown in [1], [2]. In the next Sections we will provide rigorous stability proofs and conditions for the circuit and show experimental test results from prototypes fabricated in 2.9\( \mu m \) and 1.0\( \mu m \) CMOS processes.

II. System Stability Analysis

A. Coarse Analysis

Let us assume that the dynamics of each cell can be modelled by the following first-order nonlinear differential equation

\[
C_c \frac{d v_{xj}}{dt} + G_c (v_{xj}(t) - v_{xw}) + I_j = I_{xj}(t)
\]

where \( C_c \) is the total capacitance available at node \( v_{xj} \), \( G_c \) is the total conductance at this node, and \( v_{xw} \) is the inverter trip voltage. Let us also assume that the output current of a cell is \( I_{xj}(t) = \int U(v_{xj} - v_{xw}(t)) \), where \( U(.) \) is a continuous and differentiable approximation to the step function, \( U(x) = I(1 + e^{-x}) \), with \( e \) being positive and non-zero but close to zero. Now consider eq. (1) for two nodes, \( j \) and \( w \). Let \( w \) be the node that eventually should become the winner.

Eq. (2) becomes the following solution

\[
v_{xj}(t) - v_{xw}(t) = \frac{I_{xj} - I_{xw}}{G_c} + \left[ v_{xj}(0) - v_{xw}(0) \right] - \frac{I_{xj} - I_{xw}}{G_c} e^{\frac{t}{\tau_c}} \tag{3}
\]

where \( \tau_c = G_c/C_c \). After a few time constants \( \tau_c \), the difference between the two node voltages will remain constant and equal to their difference at the equilibrium point. Therefore, we can obtain the expression for \( v_{xw}(t) \), applying eq. (3) would obtain \( v_{xj}(t) \) for the rest of the nodes.

Consider now eq. (1) for node \( w \). Since \( I_x = \Sigma I_{xj} \), it becomes

\[
v_{xw}(t) = v_{xj} - I_{xj} - \frac{I_{xj}}{G_c} = \Sigma I_j U(v_{xj} - v_{xj}) \tag{4}
\]

Since \( v_{xj} \) is given by eq. (3), after a few time constants \( \tau_c \) becomes

\[
\frac{C_c}{G_c} \frac{d v_{xj}}{dt} = v_{xj} - v_{xw} - \frac{I_{xj}}{G_c} + \sum J_j U(v_{xj} - v_{xj} - \frac{I_{xj}}{G_c}) \tag{5}
\]

This first order differential equation has stable equilibrium points if \( d^2 v_{xj}/dt^2 < 0 \). By deriving eq. (5) with respect to \( v_{xj} \) results \( C_c d^2 v_{xj}/dt^2 = -G_c - \Sigma J_j U(.) \). Since \( G_c, I_x \), and \( U(.) \) are always positive, this derivative is always negative for all possible values of \( v_{xw} \) (including the equilibrium point). Consequently, eq. (5) represents the dynamics of a stable system. This discussion assumes that the \( N \)-output current mirror presents no delay. This is not very realistic. Appendix A shows that the circuit is still stable when assuming that the \( N \)-output current mirror presents a delay modeled by first-order dynamics.

B. Fine Analysis

Performing electrical simulations of the circuit in Fig. 1, verifies that the previous stability analysis is a good approximation as long as the equilibrium point does not lie in the transition region of any of the \( N \) sigmoidal functions \( U(.) \). This can only be guaranteed if the two largest inputs \( I_j \) and \( I_w \) are sufficiently different. If two or more inputs \( I_j \) are maximum and very similar, the equilibrium point of the system will be in the transition region of some sigmoidal functions \( U(.) \).

In this case, transistor parasitic elements that have been neglected in the previous stability analysis may render unstable behavior. Consequently, some kind of compensation is necessary.
A.

Under unstable conditions the system exhibits the following characteristics (observed through electrical simulations with Hspice):

- Only the cells $j$ whose sigmoid functions $U(\cdot)$ must be in their transition region at the equilibrium point are unstable. The rest of the cells behave as if the system had reached its equilibrium point.
- The unstable cells present oscillations (presence of complex conjugate poles).
- In the case of $\alpha = 1$, and with two or more equal maximum inputs, the steady-state oscillating waveforms at these cells become the same function of time, regardless of their initial conditions.

This last observation suggests that a stability analysis could be performed by simply considering one cell in the system, which represents the parallel connection of all unstable cells, as shown in Fig. 2(a). On the other hand, since the unstable cells have the equilibrium point in the transition region of their sigmoids $U(\cdot)$, we can linearize these sigmoids for the stability analysis. Therefore, let us consider the small signal equivalent circuit shown in Fig. 2(b), where the circuitry comprised by dashed lines represents the parallel of all cells with equal and maximum input. The rest of the circuitry models the $N$-output current mirror (or set of current mirrors) responsible for distributing the global current $I_T$ among the $N$ cells. The minimum set of dynamic elements needed for the system to present unstable oscillating behavior are parasitic capacitors $C_A$, $C_p$, and $C_e$ (observed through electrical simulation). Appendix B presents a small signal analysis of the circuit in Fig. 2(b), and shows that the stability condition for this circuit is approximately

$$AM < \frac{C_A}{C_p} \left( \frac{g_{ms}}{C_e} + \frac{g_{sd}}{C_p} \right)$$

where $M$ is the number of cells with equal and maximum input. This condition is not easy to satisfy since $A$ must be large for proper operation, $M$ may become large, and it is not trivial to make the right hand side of eq. (6) very large.

Stability compensation can be achieved by introducing capacitor $C_A$, as shown in Fig. 2(c). Appendix C presents the small signal analysis of this circuit and shows that the stability condition for this circuit is

$$C_A > \frac{g_{ms}}{g_{ms} + g_{sd} + C_p}$$

Note that now the stability condition does not depend on gain $A$, and is easier to fulfill. However, now capacitor $C_A$ degrades the settling speed of the system. Capacitor $C_A$ acts as a Miller capacitance. Since the DC-gain from node $v_{ij}$ to node $v_{ij}$ is approximately $-A$ (i.e. the negative of the slope of $U(\cdot)$), there will be an effective Miller capacitance of value $(A + 1) C_A$ parallel with the original $C_e$ capacitor. If the sigmoid is not in its transition region $A = 0$, but if the sigmoid in its transition region $A$ can be very large. Therefore, for compensated cells eq. (5) must be changed to

$$\frac{A + 1}{A} C_A + \frac{1}{C_e} = \frac{1}{C_p}$$

III. Experimental Results

A WTA-MAX system with $N=10$ competing cells has been designed and fabricated in two different technologies. The first prototype has been integrated in a double-metal single-poly 1.0µm CMOS technology (ES2), and the other in a double-metal double-poly 2.5µm CMOS process (METIET). Both technologies were available through the European silicon foundry service, EUROCHIP. We used a simple 3-transistor current mirror for the 2-output NMOS current mirror of each cell. However, we used active input current mirrors [5] for the $N$-output PMOS current mirror and for the extra NMOS assembling current mirrors [1, 2].

The DC transfer curves of the system have been measured for different input current levels and for different system configurations. Fig. 3 shows thirty transfer curves when the competing cells are inside the same chip. Each curve is obtained by randomly selecting a pair of input cells, $i$ and $j$, applying a constant input current $I_0$ to the first, and sweeping the input current of the second $I_j$ from $0.9 \times I_0$ to $1.1 \times I_0$. The figure represents the two inverter output voltages, $v_{ij}$ and $v_{ij}$, versus the current $I_j$. For each pair of cells, $i$ and $j$, we measure the value of $I_j$ at the point where $v_{ij} = v_{ij}$. Let us call this value $I_{ij}$. Thirty curves were measured for each value of $I_0$, resulting in thirty values of $I_{ij}$. The difference between the mean of these thirty $I_{ij}$ values and $I_0$ is a measure of the systematic error of $I_0$. The difference between the mean of these thirty $I_{ij}$ values and $I_0$ is a measure of the systematic error of $I_0$. The variance of the thirty $I_{ij}$ values represents the random error of $I_0$. Let us call it $\sigma (I_0)$. In the case of Fig. 3, corresponding to a WTA inside one single chip fabricated in the ES2 1.0µm CMOS technology with $I_0 = 100$µA, we measured a random deviation of $\sigma (I_0) = 1.04\%$ and a systematic error of $\epsilon (I_0) = 0.03\%$. Table 2 contains the measured total error (defined as $\sigma (I_0) + \epsilon (I_0)$) for three decades of change in $I_0$. The table shows results for the cases of WTAs are inside one chip, assembled using two chips of different technologies. Note that the precision degradation is very small when the system is distributed among two chips, regardless of whether the chips are of the same technology or not. This is the main advantage of this WTA-MAX circuit with respect to others reported in literature. This is shown in Table 2 which depicts the simulation results of another WTA [4]. The input
Table 1. Current Mode WTA Precision Measurements

<table>
<thead>
<tr>
<th>Technology</th>
<th>chips</th>
<th>10pA</th>
<th>100pA</th>
<th>500pA</th>
<th>1mA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES2_1.0µm</td>
<td>1</td>
<td>2.00%</td>
<td>1.07%</td>
<td>0.58%</td>
<td>0.56%</td>
</tr>
<tr>
<td>ES2_1.0µm</td>
<td>2</td>
<td>2.35%</td>
<td>1.03%</td>
<td>0.59%</td>
<td>0.57%</td>
</tr>
<tr>
<td>MIElEC_2.4µm</td>
<td>1</td>
<td>1.94%</td>
<td>0.98%</td>
<td>0.70%</td>
<td>0.69%</td>
</tr>
<tr>
<td>MIElEC_2.4µm</td>
<td>2</td>
<td>2.15%</td>
<td>1.17%</td>
<td>0.96%</td>
<td>0.87%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Technology</th>
<th>chips</th>
<th>1.5V</th>
<th>2.5V</th>
<th>3.5V</th>
<th>4.5V</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES2_1.0µm</td>
<td>1</td>
<td>0.39%</td>
<td>0.40%</td>
<td>0.41%</td>
<td>1.73%</td>
</tr>
<tr>
<td>MIElEC_2.4µm</td>
<td>2</td>
<td>2.62%</td>
<td>9.65%</td>
<td>1.31%</td>
<td>out of range</td>
</tr>
</tbody>
</table>

Table 2. WTA Precision Computations (obtained through Spice simulations) for the Circuit reported in [4]

Table 3. Measured delay times for one chip WTAs

<table>
<thead>
<tr>
<th>I_{IN}</th>
<th>\Delta I_{IN}</th>
<th>ES2_1.0µm</th>
<th>MIElEC_2.4µm</th>
</tr>
</thead>
<tbody>
<tr>
<td>10µA</td>
<td>2µA</td>
<td>1.68µs</td>
<td>6.50µs</td>
</tr>
<tr>
<td>100pA</td>
<td>3.2µA</td>
<td>2.35µs</td>
<td>1.86µs</td>
</tr>
<tr>
<td>100pA</td>
<td>20µA</td>
<td>470ns</td>
<td>480ns</td>
</tr>
<tr>
<td>10pA</td>
<td>100pA</td>
<td>235ns</td>
<td>230ns</td>
</tr>
<tr>
<td>100pA</td>
<td>154ns</td>
<td>134ns</td>
<td>375ns</td>
</tr>
<tr>
<td>200pA</td>
<td>150ns</td>
<td>104ns</td>
<td>136ns</td>
</tr>
</tbody>
</table>

Table 4. Measured delay times for a two-chips WTA transport, and comparison of currents. This maintains good precision for circuits with a large number of input and when the circuit is distributed among several chips. Stability analysis of the proposed circuit has been addressed and stability conditions derived. Delay measurements were performed for I_{IN} values of 10µA, 50µA, 100µA, and 500µA, and for 4.5µs, equal to 0.2µs and I_{IN}. Table 3 shows the measured delay times for those cases where the system is inside one single chip. Table 3 shows the delay times measured when a WTA is assembled using 2 chips of the ES2_1.0µm process.

IV. Conclusions

A WTA-MAX circuit design technique based on current-mode signal processing has been proposed. The precision of the circuit relies on the proper replication,
Let us assume the N-output current mirror of Fig. 1(c) has the first-order dynamics defined by the small signal equivalent circuit depicted in Fig. 5. Current $I_{o}$ represents each of the N outputs of this current mirror, and $Z_{ij}$ its input. The dynamics of this current mirror in time-domain are given by

$$I_o(t) + T_p I_o(t) = \sum_j I_{oj}(t) , \quad T_p = \frac{C_z}{g_m} . \quad (9)$$

By substituting eq. (1) and its derivatives into eq. (9) and subtracting them for two nodes j and k yields,

$$\tau_p C_z [v_x(t) - v_x(t)] + \left( C_z + \tau_p C_z \right) [v_{xj}(t) - v_{xk}(t)] + G_z [v_{xj}(t) - v_{xk}(t)] = I_{o} - I_j$$

The solution to this differential equation is

$$v_{xj}(t) - v_{xk}(t) = \frac{I_o - I_j}{C_z} e^{-\tau_p C_z t} , \quad \tau_p = \frac{C_z}{g_m} . \quad (10)$$

where $K_1$ and $K_2$ are determined by initial conditions. Consequently, after a few time constants $\tau_p$ and $\tau_g$, $T_p C_z v_{xj}(t) - T_p C_z v_{xk}(t) = v_{xj}(t) - v_{xk}(t) \approx 0$. Therefore, the expression $v_{xj}(t) - v_{xk}(t)$ remains constant in time and $\tau_p C_z$ is effectively the common time constant of this circuit. Consequently, eq. (9) reduces to a first-order differential equation:

$$I_{o} = \sum_j I_{oj}(t) , \quad \tau_p C_z = \frac{C_z}{g_m} .$$

This term contributes to the constant term and to the $v_{xj}(t)$ term of eq. (12). Summing over all cells obtains $\sum_f I_U(t) = -S_{in}(t) + K$, where $K$ and $S$ are constants and $S > 0$ (because $I_{o}, U_{o}(t) > 0$). Therefore, the poles of eq. (12) are the roots of $C_z s^2 + (C_z + \tau_p C_z) s + (C_z + S) = 0$ which always have a negative real part. Consequently, eq. (12) converges always to its unique equilibrium point.

The worst case occurs for very large values of $M$, for which eq. (20) reduces to

$$C_A > \frac{g_m}{E_{mp} g_c + S_{mn} C_g} .$$

Since all coefficients $a, b, c, d$ are positive, the roots of this polynomial have negative real parts if $bc - ad > 0$. Considering that parasitic capacitances $C_p, C_c$, and $C_g$ are approximately of the same order of magnitude and that $E_{mn} > E_{mn} G_c$, the stability condition simplifies to

$$AM < \frac{C_p + S_{mn} C_g}{E_{mp} g_c} . \quad (20)$$

VIII. Appendix C

The frequency domain KCL equations of the linear circuit of Fig. 2(c) are

$$I_o = V_{s1} \left( \frac{C_g + S_{mn}}{g_m} \right) + C_t C_g = \frac{1}{1 + \frac{C_t}{S_{mn}}} I_o \quad (14)$$

Routine analysis yields the following third-order polynomial

$$as^3 + bs^2 + cs + d = 0$$

$$a = \frac{C_p C_g}{E_{mp}} + C_t C_g = \frac{1}{1 + \frac{C_t}{S_{mn}}} \quad (15)$$

Since all coefficients $a, b, c, d$ are positive, the roots of this polynomial have negative real parts if $bc - ad > 0$. Considering that parasitic capacitances $C_p$, $C_c$, and $C_g$ are approximately of the same order of magnitude and that $E_{mn} > E_{mn} G_c$, the stability condition simplifies to

$$AM < \frac{C_p + S_{mn} C_g}{E_{mp} g_c} . \quad (20)$$

The worst case occurs for very large values of $M$, for which eq. (20) reduces to

$$C_A > \frac{g_m}{E_{mp} g_c + S_{mn} C_g} . \quad (21)$$