CURRENT-MODE LOGIC

Recent advances in integrated wireless and optical transceivers demand very high-speed circuits. High-speed prescalars, counters, multiplexers, demultiplexers, phase/frequency detectors A/D, D/A, and timing extraction of digital signals are the main processing steps employed in such systems. The design technique utilized in implementing most of these functions is extensively based on using current-mode logic (CML) circuits (1–4). CML circuits are commonly used in such applications because they have the advantage of excellent threshold voltage control and an inherent differential operation. Differential CML techniques are widely used in high-speed bipolar circuits and recently in CMOS and BiCMOS high-speed design (5,6). CML and CML-like schemes become inevitable as they improve noise rejection from supply and substrate in typical mixed-signal blocks such as integrated frequency synthesizer.

Basic functions such as multiplexing blocks, full adders, fast comparators, and fast flip-flops can be implemented in a very compact way using two-level series-gated topologies. In addition, almost skew-free complementary outputs are provided, which tends to be a prerequisite for circuits at frequencies beyond a few gigahertz. The switching speed of series-gated bipolar circuits is related to a large number of device and circuit parameters in such stacked logic implementations. The maximum bit rate of these circuits is limited by the output time constants of the series-gated latches and series-gated current switches. The speed limitation caused by these time constants is minimized by keeping the internal single-ended signal swing in the latches small, that is, between 250 mVp-p and 300 mVp-p.

Increasing the operating speed and/or reducing the power consumption require low voltage swings across the loads. This is commonly achieved by adopting a differential-mode operation, wherein the need for exact reference voltages in CML is eliminated, in addition to reducing crosstalk and other unwanted common-mode signals within the circuit.

The propagation delay for the series-gated CML circuits can directly be obtained by SPICE simulations. However, a circuit optimization procedure based on SPICE simulations requires a very large number of simulation runs to cover the two-level series-gated CML circuit design space. Therefore, it is difficult to optimize the design of a two-level series-gated CML circuit over a large design space within a limited time with SPICE. If a simple yet reasonably accurate model of the propagation delay time of series-gated CML circuits is available for a large design space, then the designer could use the model to select the circuit and device parameters that correspond to the minimum propagation delay time. So, the delay model assists the designer in narrowing the design space in its early phase, instead of relying entirely on the designer intuition. The circuit simulations can be started based on clear guidelines, set by this approximate model, to further optimize the design.

This article introduces the reader to some basic principles of CML design. First, the static dc behavior is derived along with the relevant key circuit parameters. Second, the performance of CML gates is quantified in terms of speed, propagation delay time, power dissipation, and circuit complexity. An analytical model calculating propagation delay times for two-level series-gated CML high-speed bipolar circuits is developed with emphasis on using the SPICE parameters file for the available process. The analytical delay model accounts for different transistor sizes of the two levels. Moreover, high-current effects are also considered in the presented model. Exploiting these two features, the model has been successfully applied in optimizing the design of a variety of two-level series-gated CML circuits. A comparison with the results obtained by SPICE is presented, to verify the applicability of the presented model.

CML CIRCUIT PERFORMANCE ESTIMATION

The CML gate is basically a differential bipolar amplifier stage. True differential inputs can be used, especially when speed is the target. The unbuffered differential stage, or the CML gate, is very sensitive to capacitive loading, therefore, emitter-follower buffer stages are often used in emitter-coupled logic (ECL). Originally, the CML gate was designed with a large logic swing of 800 mV. System requirements for higher speed operation have led to the use of logic swings as low as 190 mV under fully differential operation.

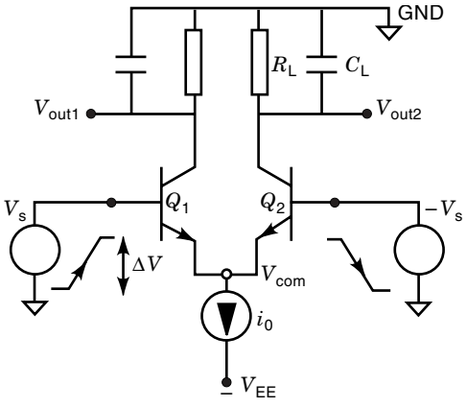


Figure 1. Schematic of a single-level CML circuit.

Evaluating the delay of CML circuits and optimizing their performance can be simply obtained by running a circuit simulator such as SPICE many times, which becomes very time consuming and impractical for large circuits. Moreover, physical insight into the crucial circuit and device parameters affecting gate performance is difficult to achieve when relying purely on circuit simulations. Therefore, an accurate analytical propagation delay model is a key to various optimization tasks at different design phases. Several attempts have been made in the literature to derive analytic/closed form delay expressions for CML bipolar circuits, in order to provide engineering insight into the relative importance of process, device, and circuit parameters (7–13).

A short switching time and a low power dissipation are obviously contradictory requirements, because decreasing the power dissipation implies reducing the currents available for charging of the parasitic capacitances. For the CML basic circuit in Fig. 1, the power dissipation is given by:

$$P = i_0 \cdot V_{EE} \tag{1}$$

The single-ended voltage swing (ΔV) is related to the tail current (i_0) by:

$$R_L = \frac{\Delta V}{i_0} \tag{2}$$

The propagation delay is approximated by

$$t_{PD} = R_L C_{eff} = \frac{\Delta V C_{eff}}{i_0} = \frac{\Delta V V_{EE} C_{eff}}{P} \tag{3}$$

where V_{EE} is the supply voltage; P is the power consumption; and C_{eff} is the effective capacitance representing the junction and interconnection capacitances. The above expressions indicate that for a given P, the circuit should have a small output swing to reduce the propagation delay. However, the lower limit on the voltage swing is the transition width ΔT, which is discussed later.

Static Transfer Curve

The static transfer characteristic for CML shown in Fig. 2 can be derived using a simple transistor model. Assuming the transistors are geometrically equal, their saturation currents are equal, then

$$V_{in} = V_T \ln\left(\frac{i_1}{I_s}\right) - V_T \ln\left(\frac{i_2}{I_s}\right) \tag{4}$$

that simplifies to

$$V_{in} = V_T \ln\left(\frac{i_1}{i_2}\right) \tag{5}$$

which i_1, i_2 are the collector currents of transistors Q_1, Q_2, respectively. I_s, V_T are the device saturation current and thermal voltage, respectively. The ratio between the two currents is

$$\frac{i_1}{i_2} = \exp(V_{in}/V_T) \tag{6}$$

Noting that the sum of the two currents must always remain equal to i_0,

$$i_0 = i_1 + i_2 \tag{7}$$

The voltage at the output nodes (see Fig. 2) will be

$$V_{out1} = -\frac{\Delta V}{(1 + \exp(-V_{in}/V_T))} \tag{8}$$

$$V_{out2} = -\frac{\Delta V}{(1 + \exp(V_{in}/V_T))} \tag{9}$$

Gain

For signals to propagate in logic networks, the individual elements (gates and flip-flops) are usually designed to have a gain greater than 1. Cascading of the logic elements produces a swing that increases through the elements until some equilibrium value is reached; logical "1" and "0" levels. Assuming

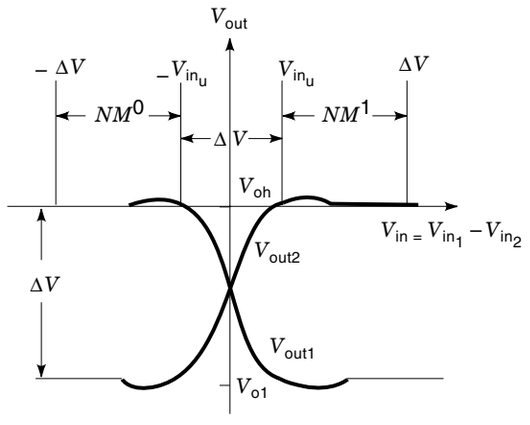


Figure 2. CML dc-transfer characteristic and pertinent noise margins.

constant V_{BE}, the gain (G) is a function of V_{in}, ΔV, and temperature.

$$G = \frac{dV_{out}}{dV_{in}} = \frac{\Delta V}{V_T} \cdot \frac{\exp(V_{in}/V_T)}{[1 + \exp(V_{in}/V_T)]^2} \tag{10}$$

As can be seen, the maximum gain occurs when the input is zero.

$$G_{max} = \frac{\Delta V}{4V_T} \tag{11}$$

Thus, for a gain greater than unity ΔV must be greater than ($4V_T$).

ac Noise Margin (Transition Width)

The difference between V_{in} at the unity gain point and V_{in} that makes the output $= -\Delta V/2$ will be defined as the ac noise margin. For the single input circuit, the input voltage for $V_{out} = -\Delta V/2$ is equal to V_{ref}. The ac noise margin is then

$$\Delta T = 2V_T \cdot \ln\left(\frac{\Delta V}{V_T} - 2\right) \tag{12}$$

which determines the lowest voltage swing needed for proper switching between the two output levels.

MODELING AND OPTIMIZATION OF CML

The advantages of CML and CML-like circuits are the high-speed nonsaturating operation of current switches. The propagation delay (t_{PD}) of CML circuits depends on the power dissipation or the switching current per gate, as shown in Fig. 3. At low power dissipation, R_LC_L is the dominant factor limiting the delay while at high power dissipation, R_bC_{jc} plays a more important role than the others (see Fig. 3). The minimum achievable delay is restricted by τ_f where τ_f is the forward transit time, R_L is the pull-up resistor and R_b is the base resistance, including the intrinsic, extrinsic, and the contact regions. C_{jc} is the base-collector capacitance, C_L is the parasitic wiring and load capacitance.

It turns out that the delay time can be expressed approximately as a linear combination of the time constants of the circuit under study, with each time constant being weighted by a factor that is determined by the circuit topology. The value of any weighting factor K_i is obtained by partial differentiation of the delay expression with respect to the involved time constant elements. This sensitivity analysis approach is done by changing the value of the time constant, running the circuit simulator (SPICE), and recording the corresponding change in delay time.

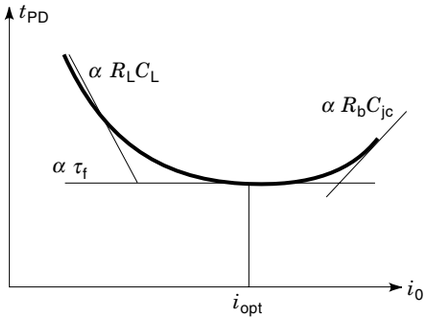


Figure 3. CML propagation delay versus tail current. The delay is proportional to different time constants at different current level.

$$\frac{\partial^2 t_{PD}}{\partial R_i \partial C_i} = 0 \cdots + K_i + \cdots + \cdots 0 \tag{13}$$

Therefore, determining each weighting factor means running SPICE four times. Additionally, the linearity range of each weighting factor should be well checked over a certain range of the respective time constant.

More accurate expressions have been reported, based on a Gummel–Poon transistor model (SPICE model) (12). Starting with 24 weighting constants ($K_1 - K_{24}$), eight is enough to describe the delay with a 5% error. The linearity of these expression has been checked to prove its validity over a wide range of transistor and circuit parameters. However, it should be noted that different transistor geometries may result in significantly different time constants. Therefore, the product of the weighting factors and the time constants should be considered for a specific device.

Large-scale integration requires low power dissipation per gate to keep the total power consumption per chip at a practical level. Therefore, a low switching current per gate is accordingly needed, and a large pull-up resistor is required if the voltage swing is fixed at a constant value to have large enough noise margins. Consequently, the R_LC_L time constant associated with the pull-up resistor should be decreased to achieve higher speed operation. From Fig. 3, it is obvious that there is an optimum value of the switching current (i_0) or the load resistor (R_L) for a given output swing (ΔV).

Two-Level Series-Gated CML

The logic flexibility of CML gates is very high because differential stages can be stacked easily. Up to three levels of series gating are possible with $V_{EE} = 5$ V. The high potential of series gating can clearly be seen in the implementation of OR, NOR, XOR and XNOR gates shown in Fig. 4. The main advantage of using series-gating principle is the flexibility of designing variety of logic gates with differential inputs and outputs. Notice that the 2-input OR/NOR gate of Fig. 4(a) and the 3-input OR/NOR gate of Fig. 4(b) have differential input and output, yet the circuits are not fully symmetric as the XOR/XNOR gate of Fig. 4(c). The delay time of XOR and XNOR gates, using series gating is about 50 to 70% higher than that of a simple inverter.

One of the advantages of the analytical propagation delay expressions based on the sensitivity analysis approach, is that it allows the designer to study the contribution of each time constant on the switching speed. The maximum toggling frequency of flip-flops based on two-level series-gated circuits can be predicted by relating the performance to that of the constituent XOR gates (14). It has been found that the behavior of the propagation delay of XOR gates is quite linear and an agreement to 10% with SPICE simulations has been obtained. However, the delay expression of that approach is valid only for circuits with transistors of the *same* size.

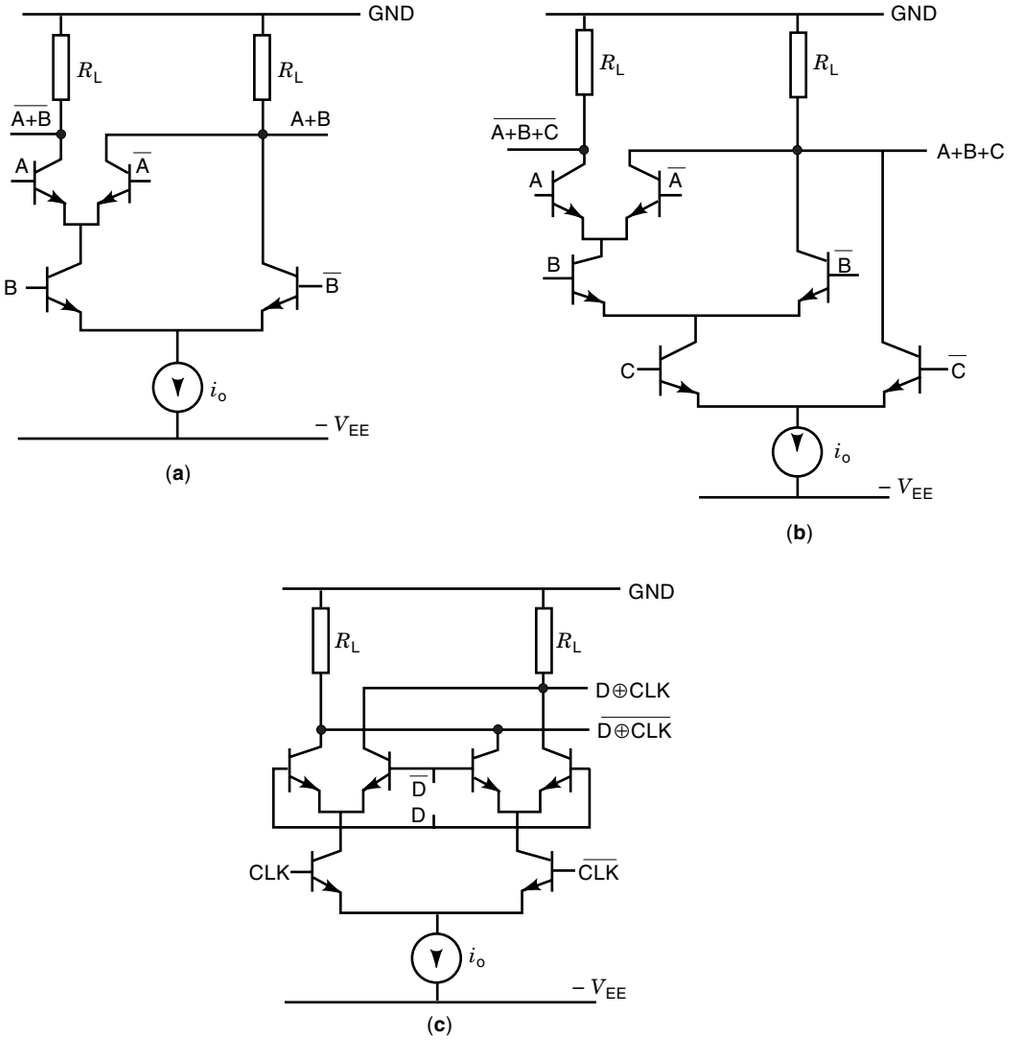


Figure 4. Schematic of two-level series-gating CML technique for (a) 2-input OR/NOR gate; (b) 3-input OR/NOR gate; (c) XOR/XNOR gate.

ANALYTICAL DELAY MODEL OF TWO-LEVEL CML

The delay will be calculated based on a definition originally proposed by Ashar (7) and later refined by Tien (8). A link is made between the concept of network function and that of charge-control in developing this kind of analytical propagation delay model. The transistor is modeled by a set of charge-control equations and the remaining passive elements are modeled by a set of circuit equations. Solving these two sets of equations would lead to the calculation of the delay. The inherent symmetrical nature characterizing differential CML circuit topologies is utilized in reducing the set of circuit equations which is essential for more complicated topologies (e.g., 2-level series-gated circuitry).

The model assumes that all the active and passive circuit elements are linearized. The emitter-base diode resistance and the diffusion capacitance are linearized by considering their average during the switching period. Other parasitic capacitances are also linearized according to the same rule. High-current effects implemented in the Gummel–Poon BJT SPICE model are included in the delay model by modifying the expression for the forward transit-time, as illustrated in (16).

Formulation of the Delay Model

The output response of digital switching circuits due to an input delta function (impulse) is both a delayed and distorted function, as shown in Fig. 5. For a linear network, the propagation delay (t_{PD}) is the averaged time by which the response of the delta function input is delayed. If $F(s)$ is the transfer function of the circuit, then the propagation delay t_{PD} becomes (7):

$$t_{PD} = \frac{\int_0^\infty t f(t)\,dt}{\int_0^\infty f(t)\,dt} = \mathrm{Lim}_{s\to 0}\frac{\int_0^\infty t f(t)e^{-st}\,dt}{\int_0^\infty f(t)e^{-st}\,dt} = \left.\frac{-\dfrac{dF(s)}{ds}}{F(s)}\right|_{s=0} \tag{14}$$

The method itself is independent of the BJT model, while the resultant delay expression is limited by the accuracy of the model used.

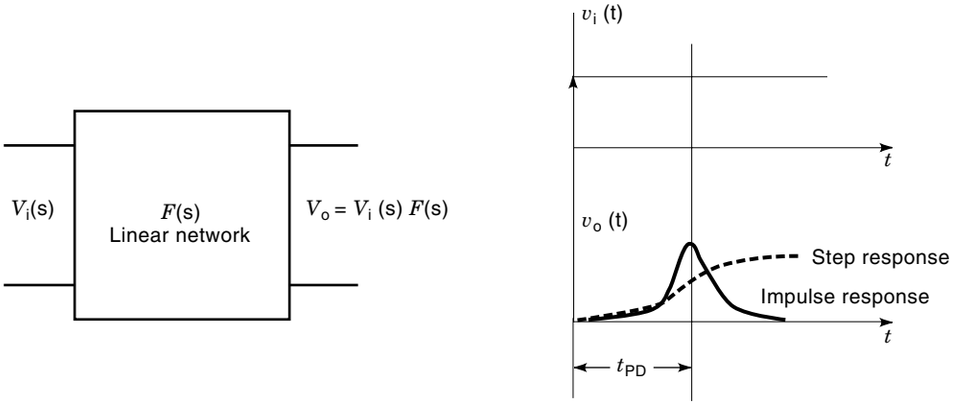


Figure 5. Output response of a digital switching circuit.

The four basic series-gated CML-based configurations (all inputs and outputs are differential) are shown in Fig. 6 (17,18): (a) is an XOR characterized by two inputs (upper-level input: D, $\overline{D}$, lower-level input CLK, $\overline{CLK}$) and one output Q, $\overline{Q}$; (b) is an MUX that has two upper-level inputs (D_1, $\overline{D_1}$, D_2, $\overline{D_2}$), which are multiplexed (selected) by the lower-level input (CLK, $\overline{CLK}$) and appear on the output Q, $\overline{Q}$; (c) is an DEMUX that has one upper-level input D, $\overline{D}$, one lower-level input CLK, $\overline{CLK}$, and two demultiplexed (deserialized) outputs Q_1, $\overline{Q_1}$, Q_2, $\overline{Q_2}$; and (d) is a D-latch that has an upper-level input D, $\overline{D}$, a lower-level input CLK, $\overline{CLK}$, and an output Q, $\overline{Q}$. Apart from minor differences among the four

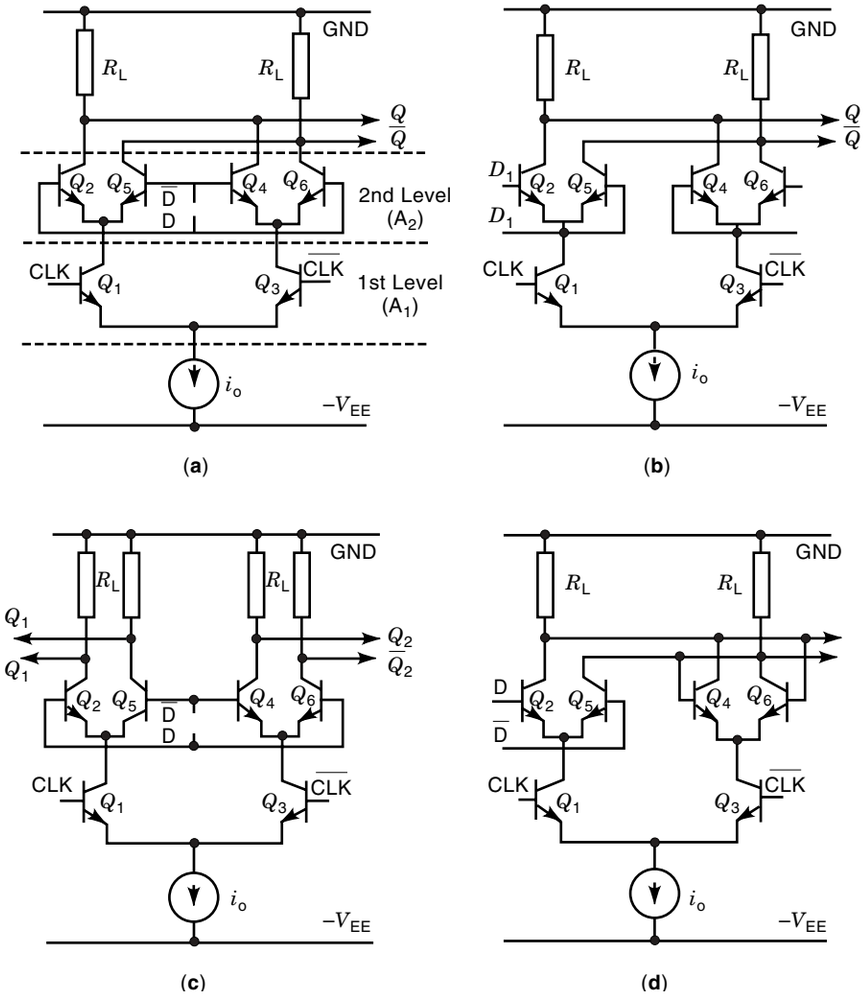


Figure 6. Schematics of 2-level series-gated CML-based circuits (a) XOR; (b) 2:1 MUX; (c) 1:2 DEMUX; (d) D-latch.

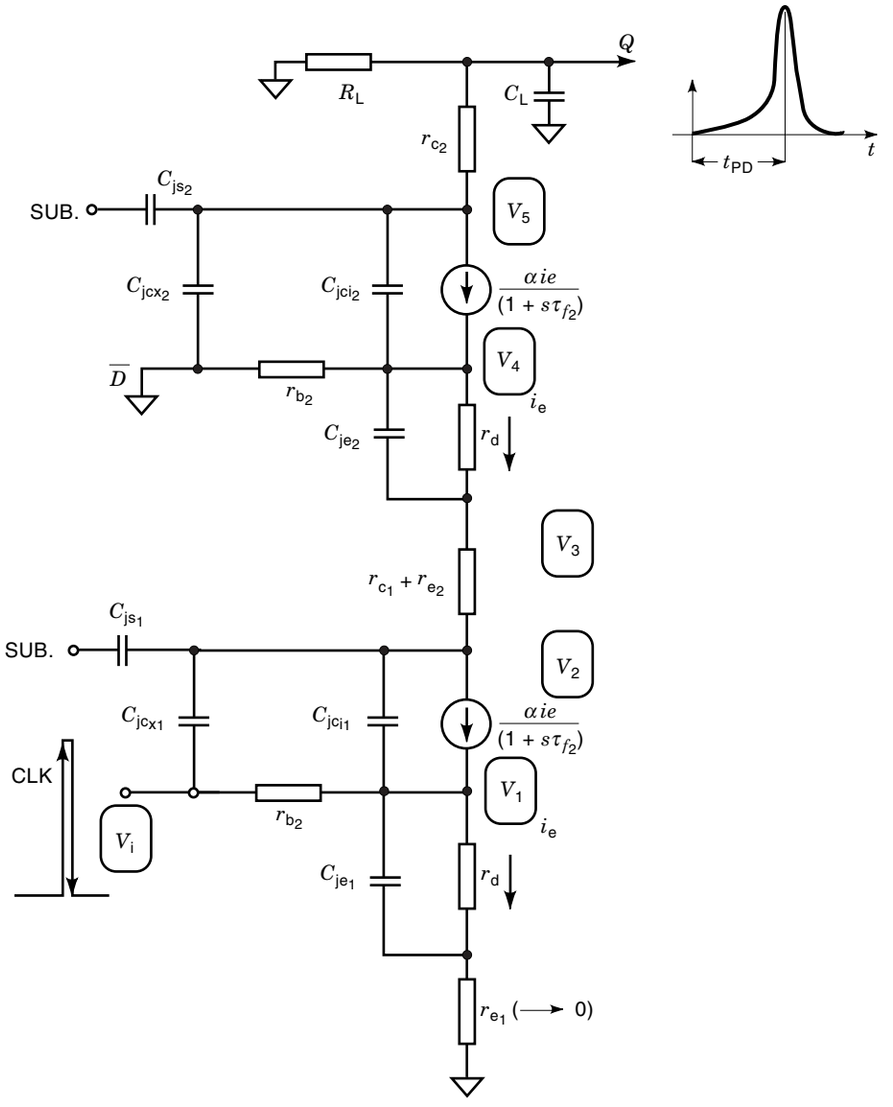


Figure 7. Equivalent circuit used in delay model for a series-gated CML-based XOR circuit.

topologies, the propagation delay from the lower level (*CLK*, $\overline{CLK}$) to the output (Q, $\overline{Q}$) would yield similar expressions. As such, the XOR topology will be taken as a prototype in the subsequent analysis. In addition, the D-latch topology will also be studied, as it is used extensitely in frequency dividers.

Basic Two-Level XOR CML Circuit

The basic two-level series-gated XOR CML circuit is shown in Fig. 6(a). Transistors of the first-level Q_1 and Q_3 are identical and have an emitter area of A_1. Transistors of the second-level Q_2, Q_4, Q_5, and Q_6, are identical and have an emitter area of A_2. The linearized equivalent circuit diagram is also shown in Fig. 7, which is used in deriving the subsequent circuit equations.

The following remarks can be made (see Fig. 7):

1. The resistance of the base–emitter junction r_d has been linearized and taken as:

$$r_d = \frac{2.4V_T}{i_0} \tag{15}$$

where $V_T = kT/q$ and the factor 2.4 is due to a first-order linearization scheme (16).

2. The collector current-controlled current source of the BJT is assumed to be a single-pole function at a frequency of f_T to account for transit-time and diffusion capacitance

$$\alpha(s) = \frac{\alpha_0}{1 + s\tau_f} \tag{16}$$

where α_0 is the low frequency common-base current gain.

3. All the BJT device parasitic capacitances are considered in addition to the use of an arbitrary device geometry for the two levels. All parasitic capacitances, collector-substrate capacitance C_{js}, extrinsic base-collector capacitance C_{jcx}, intrinsic base-collector capacitance C_{jci} and emitter-base capacitance C_{je} are voltage-dependent as:

$$C(V_r) = \frac{C(0)}{(1 + V_r/\phi)^m} \tag{17}$$

where $C(0)$ is the zero-bias capacitance, ϕ and m are process-dependent coefficients, and V_r is the magnitude of the reverse bias across the parasitic capacitance. These are available in the SPICE parameters file of any technology. For most CML circuits, the low and high logic levels of the input and output are the same so that C_{jci} and C_{jcx} are reasonably correct if we use their zero-bias values in our delay calculations. On the other hand, C_{js} is greatly reduced from its zero-bias value when the substrate is reversely biased by the most negative supply voltage (i.e., $V_r = V_{EE}$).

The five nodal equations (V_1, V_2, V_3, V_4, V_5) representing the first-level internal base, internal collector, second-level internal emitter, internal base, and internal collector node voltages describe the XOR circuit completely.

$$\begin{bmatrix} a_{11}+sb_{11} & a_{12}+sb_{12} & a_{13}+sb_{13} & a_{14}+sb_{14} & a_{15}+sb_{15} \\ a_{21}+sb_{21} & a_{22}+sb_{22} & a_{23}+sb_{23} & a_{24}+sb_{24} & a_{25}+sb_{25} \\ a_{31}+sb_{31} & a_{32}+sb_{32} & a_{33}+sb_{33} & a_{34}+sb_{34} & a_{35}+sb_{35} \\ a_{41}+sb_{41} & a_{42}+sb_{42} & a_{43}+sb_{43} & a_{44}+sb_{44} & a_{45}+sb_{45} \\ a_{51}+sb_{51} & a_{52}+sb_{52} & a_{53}+sb_{53} & a_{54}+sb_{54} & a_{55}+sb_{55} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{bmatrix} = \begin{bmatrix} V_i/r_{b1} \\ sC_{jci1}V_i \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{18}$$

where

$$\begin{aligned} a_{11} &= \frac{1-\alpha}{r_d} + 1/r_{b1} & b_{11} &= C_{je1} + C_{bc1} \\ a_{12} &= 0 & b_{12} &= -C_{jci1} \\ a_{21} &= \alpha/r_d & b_{21} &= -C_{jci1} \\ a_{22} &= 1/r_{c1} & b_{22} &= C_{T1} \\ a_{23} &= -1/r_{c1} & b_{23} &= 0 \\ a_{32} &= -1/r_{c1} & b_{32} &= 0 \\ a_{33} &= -(1/r_d + 1/r_{c1}) & b_{33} &= -C_{je2} \\ a_{34} &= 1/r_d & b_{34} &= C_{je2} \\ a_{43} &= -\frac{1-\alpha}{r_d} & b_{43} &= -C_{je2} \\ a_{44} &= 1/r_{b2} + \frac{1-\alpha}{r_d} & b_{44} &= C_{jci2} + C_{je2} \\ a_{45} &= 0 & b_{45} &= -C_{jci2} \\ a_{53} &= -\alpha/r_d & b_{53} &= 0 \\ a_{54} &= -\alpha/r_d & b_{54} &= -C_{jci2} \\ a_{55} &= \frac{1}{R_L + r_{c2}} & b_{55} &= C_{T2} \end{aligned} \tag{19}$$

The other coefficients are equal to zero. The propagation delay (from a lower-level input to the output) is then calculated from the transfer function $F(s)$, which links the output voltage V_5 to the input impulse function applied at V_i. After using standard matrix methods and relations for solving linear networks, $F(s)$ can be written as (18):

$$F(s) = \frac{N(s)}{M(s)} \tag{20}$$

where $N(s)$ and $M(s)$ are the output co-factor determinant and characteristic determinant of the XOR circuit. Therefore, the delay expression t_{PD} yields:

$$t_{PD} = \left.\frac{-\frac{dF(s)}{ds}}{F(s)}\right|_{s=0} = \frac{-\frac{dM(s)}{ds}}{M(s)} - \left.\frac{-\frac{dN(s)}{ds}}{N(s)}\right|_{s\to 0} = D_1 + D_2 \tag{21}$$

$$\begin{aligned} t_{PD} = {} & \tau_{f1} + \tau_{f2} + \tau_{b1}C_d + r_{b1}C_{je1} + r_{b1}\left(2 + \frac{r_{c1}}{r_d}\right)C_{jci1} \\ & + r_d(C_{jci1} - C_{jcx1}) \\ & + r_d\left(1 + \frac{r_{c1}}{r_d}\right)C_{T1} + 2r_{b2}C_{jci2} \\ & + r_{b2}\left(2 + \frac{r_d}{r_{b2}} + \frac{r_d}{r_{c1}}\right)C_{je2} + (R_L + r_{c2}) \\ & + C_{T2} + R_L C_{js2} + \tau_L \end{aligned} \tag{22}$$

where $C_d = \tau_f/R_L$ is the bipolar transistor diffusion capacitance and $C_T = C_{js} + C_{jci} + C_{jcx}$ represents the total parasitic capacitance at the internal collector node. The last term in Eq. (22) accounts for the extra delay due to the load capacitance C_L (node Q in Fig. 7). Since node (Q) is a single-time constant node, the propagation delay contribution (50% point) becomes

$$\tau_L = 0.69 R_L C_L \tag{23}$$

The factor (0.69) is introduced to give the delay when the output reaches 50% of its maximum. Without this factor, the calculated delay would have been that when the output reaches 67% of its maximum (7). All parasitic capacitances should be weighted according to Eq. (17), in order to account for their voltage dependence. For this case study (V_{EE} = 5 V), the weighting factors shown in Table 1 have been used for the CML circuit. The input state can be either High (V_H) or Low (V_L). The logic level of the input states for the upper level (U.L.) and lower level (L.L.) are listed in the table.

It is noteworthy that the propagation delay t_{PD} is, in fact, the average of the Rise (Low/High) and Fall (High/Low) delay components as shown, in comparison with SPICE simulation results given in the next section.

Table 1. Weighting Factors of Device Parasitics

	CML
Input levels	U.L.: $V_H = 0$ U.L.: $V_L = -\Delta V$ L.L.: $V_H = -V_{BE}$ L.L.: $V_L = -V_{BE} - \Delta V$
Parasitic element	value
Lower level (L.L.)	
C_{js1}	$0.56C_{js1}(0)$
C_{jc1}	$0.88C_{jc1}(0)$
C_{je1}	$1.24C_{je1}(0)$
Upper level (U.L.)	
C_{js2}	$0.54C_{js2}(0)$
C_{jc2}	$C_{jc2}(0)$
C_{je2}	$1.24C_{je2}(0)$

HIGH-CURRENT EFFECTS

In high-speed CML circuits, bipolar transistors operate at high collector-current densities to minimize the delay through the circuit (3,4). As the transit-time is one of the most significant parameters characterizing the high-frequency properties (f_T–i_c curve) in bipolar transistors, a precise description of this parameter, at high-collector current densities j_c, is crucial. At high current densities, high-current effects such as base push-out, lateral spreading, space-charge-limited current flow, and quasi-saturation increase the transit-time and, hence, decrease f_T. The transit-time deterioration in the high-current region depends on physical device parameters, geometry, collector current, and base-collector voltage. High-current effects are modeled in SPICE (15) by an empirical expression as

$$\tau_f(\text{high\_current}) = \tau_f \left[1 + X_{tf} \exp\left(\frac{V_{bc}}{1.44V_{tf}}\right)\left(\frac{I_{cc}}{I_{cc}+I_{tf}}\right)^2\right] \tag{24}$$

where X_{tf}, V_{tf}, and I_{tf} are SPICE fitting parameters controlling the total fall-off of f_T, the change in f_T with respect to base-collector voltage V_{bc}, and the change in f_T with respect to current. I_{cc} is the collector terminal current in the absence of the high-current effects, which corresponds to that of Ebers–Moll model (15). Despite its empirical nature, Eq. (24) can describe the high-current effects with good accuracy if optimizing schemes are used to extract high-current fitting parameters. This is quite feasible if operation is to be expected near the onset of the high-current region, which is the case in typical high-speed CML circuits. In CML circuits, I_{cc} is related to the collector current i_c by (16):

$$I_{cc} = 1.5\left(\frac{i_c^2}{I_{kf}}\right)\left\{1 + \left[1 + \left(\frac{2I_{kf}}{3i_c}\right)^2\right]^{0.5}\right\} \tag{25}$$

I_{kf} is the forward knee current modeling the onset of the current gain β roll-off due to high-current effects. It is important to distinguish between I_{cc} and i_c before using Eq. (24) in the delay model. Fig. 8 plots the transit-time obtained from Eq. (24) versus i_c using Eq. (25) for different transitor areas. On the same graph, the transit-time is calculated if i_c and I_{cc} are assumed equal. This illustrates the large error between the two cases, which makes this assumption generally invalid.

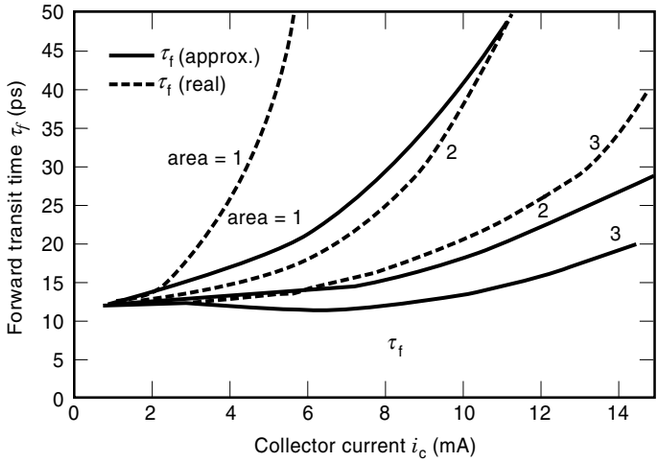


Figure 8. Transit-time versus current for two cases; approximate ($i_c = I_{cc}$), real ($i_c \neq I_{cc}$).

In addition, high-current effects start impacting the transit-time at current values earlier than I_{kf}. This means that neglecting these effects based on operating at current levels below I_{kf} is unreliable. The high-current effects are generally two-dimensional, which are strongly dependent on the technology and the device structure. These remarks suggest that the inclusion of the high-current effects may become inevitable if accurate modeling is sought.

During the switching event, the curent of "ON" transistor changes from its maximum value i_0 to a final value of zero after switching. Note that V_{bc} is also time-dependent changing from ΔV before switching to $-\Delta V$. Therefore, according to Eq. (24) the transit-time becomes time-dependent and nonlinear differential equations arise in the delay calculations. The situation is simplified by using an average value for the transit-time in the high-current region. The study showed that in typical CML circuits with a low-voltage swing, the average occurs when i_c is about 0.8 to 0.9 i_0 and V_{bc} is zero (16).

Circuit Complexity

The delay formula in Eq. (22) is only valid for the case of zero CML fanout. For the case of CML fanout > 0, the extra delay (D_{FO}) due to fanout of N identical CML circuits is given by (18):

$$D_{FO} = NR_L\left[0.5C_{D2} + C_{je2} + C_{jcx2} + C_{jci2}\left(1 + 0.5\frac{R_L + r_{c2}}{r_d + r_{e2}}\right)\right]$$
$$D_{FO} = 0.5N\tau_{f2} + NR_L\left[C_{je2} + C_{jcx2} + C_{jci2}\left(1 + 0.5\frac{R_L + r_{c2}}{r_d + r_{e2}}\right)\right] \tag{26}$$

The factor (0.5) accounts for the average value of the diffusion capacitance and Miller's effect during any transition. The net CML propagation delay is the sum of t_{PD} and D_{FO} from Eq. (22) and Eq. (26).

MODEL VERIFICATION

The delay model has been applied in studying a two-level XOR circuit under different operating conditions (18). This includes: (1) identical device areas, (2) arbitrary device areas, (3) different current or power levels, (4) different load capacitance, and (5) different fanouts. The verification of the delay expression was carried out using the results of SPICE simulations for the same circuit when it operates under the same conditions. Nevertheless, two propagation delay components (i.e., Rise and Fall) have been extracted and the average delay was used in the verification. It has also been found that this average delay depends on the slewing rate (Rise/Fall-time) of the input pulse. The delay model does not cover this input condition as it assumes an impulse input driving the circuit under study. Apart from a constant multiplying factor ($0.9 < \text{MF} < 1$), the delay expression t_{PD} can give all the information and reveals the same behavior versus various circuit and device parameters as that which may be obtained from SPICE simulation results. Once t_{PD} is adjusted (multiplied) by the MF, the accuracy of this corrected delay formula can be

Table 2. 0.8 μm BiCMOS Electrical Parameters

Parameter	Unit	Value
Emitter Area	$\mu m \times \mu m$	0.8×8
τ_f	ps	12
r_b	Ω	170
r_c	Ω	60
r_e	Ω	10
C_{jcx}	fF	28
C_{jci}	fF	10
C_{js}	fF	60
β_f		98
I_k	mA	10
I_{tf}	mA	160
V_{tf}	V	2.5
X_{tf}		750

within 10% in most cases. One needs to run SPICE only once to determine the value of the factor MF. The electrical parameters used in the propagation delay model are based on a 0.8 μm BiCMOS process and are listed in Table 2.

Figure 9 shows the propagation delay predicted by our delay model versus the tail current for various transistor emitter areas (A_1, A_2 are the area scaling factors of the lower and upper level, respectively). The emitter stripe width is always taken to be the minimum (i.e., 0.8 μm) while the emitter length (L_E) is determined by $8A$ μm. So, for a unity area scaling factor ($A = 1$), the corresponding emitter area is 8×0.8 μm^2. In Fig. 9 all transistors are assumed to have identical sizes. The model results are in good agreement with that of SPICE simulations whereby the input rise time is 150 ps. The value of MF used in the model is found to be 0.9 for best agreement with SPICE for the XOR and D-latch configurations based on 0.8-μm BiCMOS process listed in Table 2.

The XOR delay load sensitivity of CML configuration is plotted in Fig. 10 for various fanouts. The delay model results are in good agreement with SPICE simulation in this case of identical device sizes for upper and lower levels ($A_1 = A_2$). Figure 11 shows the XOR propagation delay versus the emitter area of the second level (A_2), while the first-level emitter area (A_1) can be: (1) half A_2, (2) identical to A_2, and (3) twice A_2. Optimization by SPICE leads to a 5% reduction in the minimum delay, whereas the delay model achieves 6.6% by optimizing the device areas at room temperature, as shown in the figure. In a temperature range -40°C to $+80$°C, the improvement percentage changes to 6% to 4.4% for SPICE and 7.5% to 6.5% for the delay model, respectively. The dominant temperature effect is attributed to V_T and its associated parameter r_d in Eq. (15). Worst- and best-case variations in the BiCMOS process parameters file of Table 2 have only $\pm$ 0.2% change on the 5% factor achieved by device-area optimization.

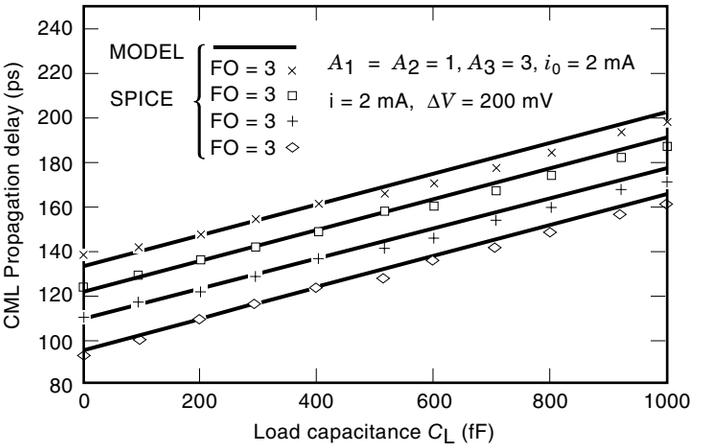


Figure 10. XOR propagation delay versus load capacitance for different fanout.

MODEL APPLICATIONS

Optimizing the XOR Circuit

The delay formula in Eq. (22) can be used in predicting the optimum load resistance R_L or tail current i_0, in addition to the optimum sizing of A_1, A_2 for the maximum switching speed (i.e., minimum propagation delay). This is feasible by partially differentiating Eq. (22) w.r.t. the respective design parameter. The choice of R_L involves a trade-off between the C_D and $1/r_d$ terms on the one hand and R_L and r_d terms on the

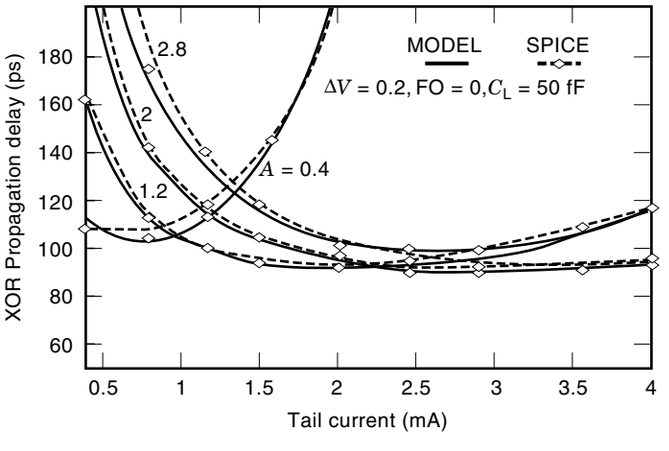


Figure 9. XOR delay versus current for identical-size devices.

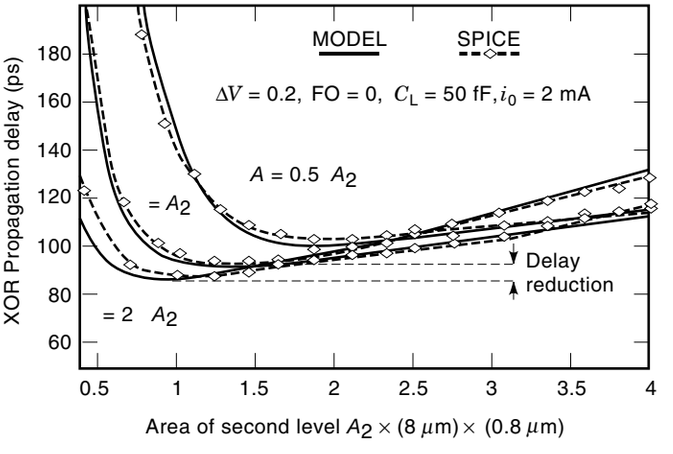


Figure 11. XOR delay versus area for different-size devices.

other hand. If high-current effects are neglected, the optimum R_L for two-level series-gated CML circuits can be expressed as

$$R_{\text{Lopt}} = \sqrt{\frac{r_{b1}\tau_{f1} + K_v r_{b1} r_{c1} C_{jci1}}{\frac{1}{K_v}\left[2C_{jci1} + C_{js1} + C_{je2}\left(1 + \frac{r_{b2}}{r_{c1}}\right)\right] + C_{T2} + C_{js2} + 0.69C_L}} \tag{27}$$

where $K_v = 2.4\ \Delta V/V_T$. Equation (27) yields the optimum current, which is given by

$$i_{0\text{opt}} = \frac{\Delta V}{R_{\text{Lopt}}}$$

The optimum sizing (A_1) for the lower level can be deduced from Eq. (22) and still be close to the actual optimum design point. If

$$\frac{\partial t_{PD}}{\partial A_1} = 0 \tag{28}$$

then

$$A_{1\text{opt}} = \sqrt{\frac{r_{bo} r_{co}\left(C_D + \frac{r_c}{r_d} C_{jci}\right)}{r_d[2r_c(C_{js} + C_{jci}) + r_b C_{je}]}} \tag{29}$$

where r_{bo} and r_{co} are the base and collector resistances of a normalized emitter area of 1 μm^2. However, there is no opti-

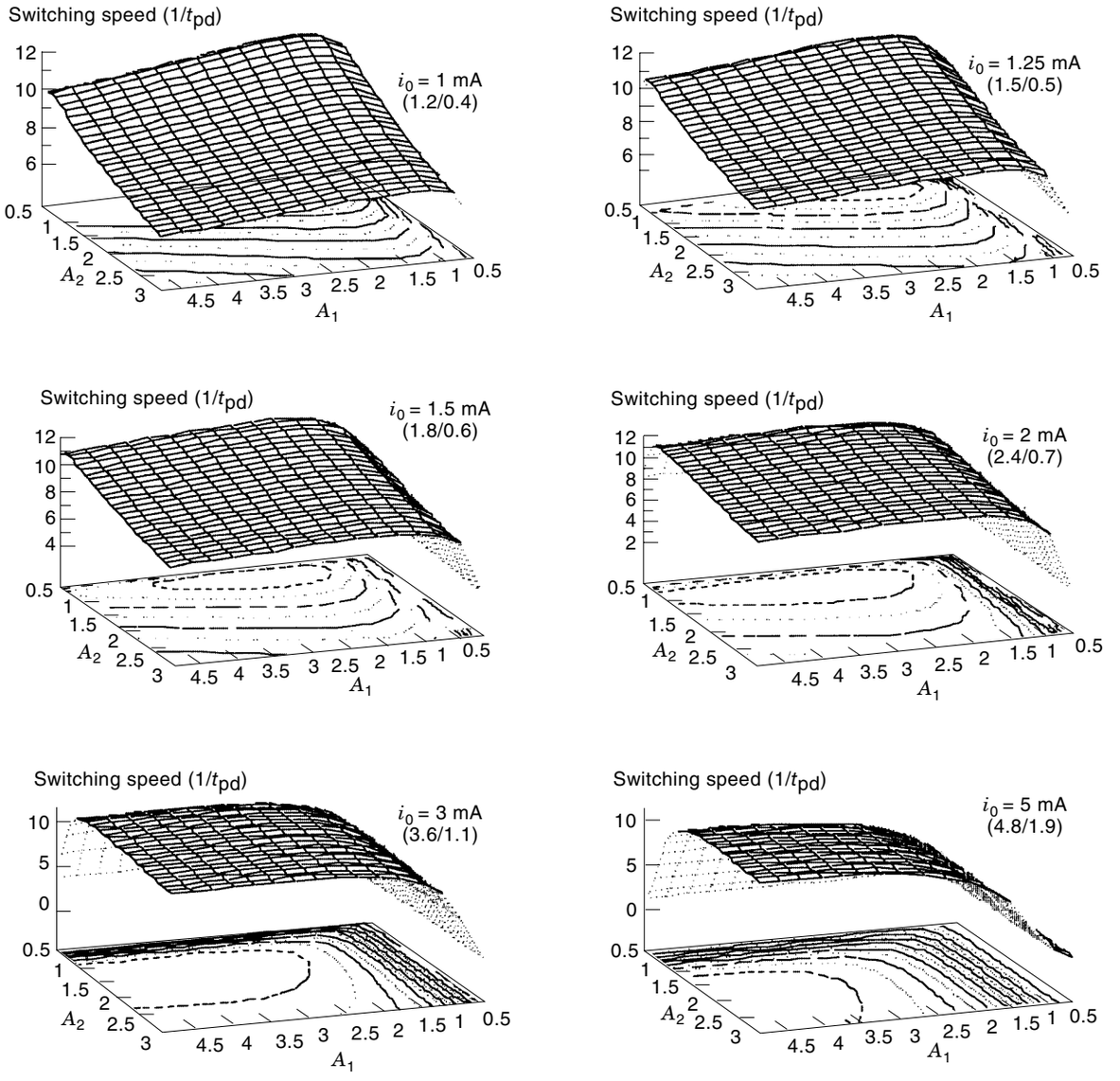
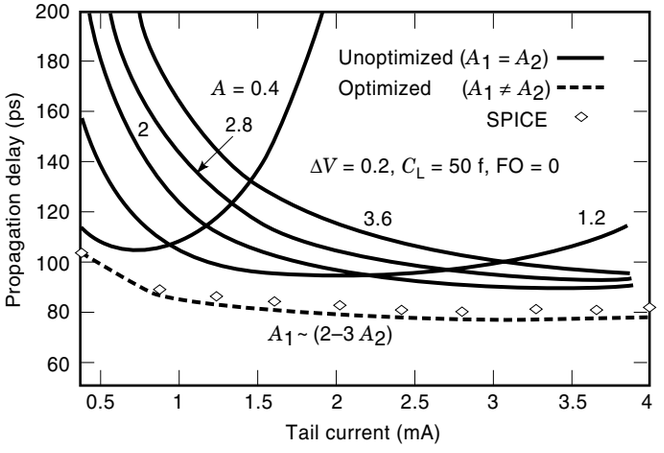


Figure 12. XOR Switching speed (1/delay) versus transistor areas of a series-gated CML circuit for different operating currents (1, 1.25, 1.5, 2, 3, and 5 mA).

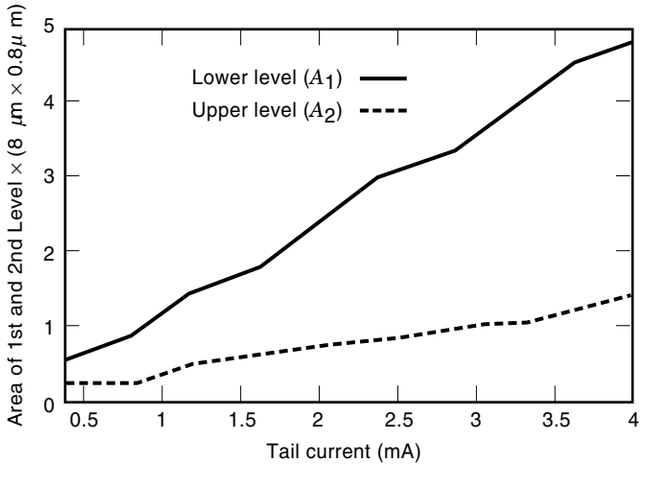
mum area for the upper level A_2 if high-current effects are neglected, since the delay t_{PD} will always increase with an increase in area A_2 as:

$$\frac{\partial t_{PD}}{\partial A_2} = r_d C_{jeo} + R_L C_{To} + R_L C_{jso} \tag{30}$$

where C_{jeo}, C_{To} and C_{jso} are the transistor parasitic capacitances of a normalized emitter area of 1 μm^2. If high-current effects are included, the delay formula becomes so complicated that an explicit form for the optimum device areas tends to be infeasible. Thus, the only direct option is to take advantage of Eq. (22) in scanning the design space (A_1, A_2) and consequently determine the optimal design point. The model application in optimizing CML series-gated high-speed circuits is demonstrated in the series of 3-D plots shown in Fig. 12, which shows the optimum design point (A_1/A_2) as the tail current increases from 1 to 5 mA. It is obvious that the optimum design (minimum propagation delay) always requires individual sizing for the devices of each level. Nevertheless, the ratio of the transistor emitter-area is found to be around 3 (i.e., the lower-level device area is three times that of the upper-level) for this specific example.



(a)



(b)

Figure 13. (a) XOR delay versus current for optimized and nonoptimized devices; (b) device area for each level versus the current under the optimized condition.

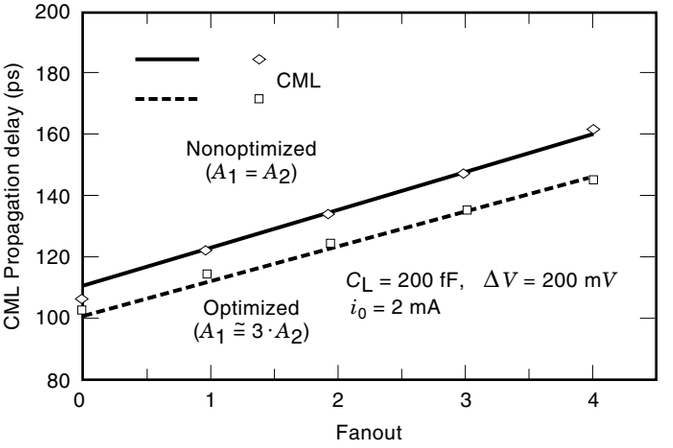


Figure 14. XOR propagation delay of a CML-based configuration versus fanout for the optimized and nonoptimized cases (MODEL: lines, SPICE: dots).

Figure 13(a) illustrates the advantage of individual device sizing of the two levels A_1 and A_2 compared with that of the identical sizing case. The delay model offers an optimized sizing design tool, which is at least one order of magnitude faster than other optimization techniques using circuit simulators. The areas of the optimized devices of the two levels are also shown in Fig. 13(b). The delay model is also verified by SPICE under optimized device areas and both agree well with less than 7 percent error.

Figure 14 is the XOR propagation delay time based on CML versus fanout for nonoptimized and optimized circuit design. A fanout of 0 to 4 upper-level XORs is equivalent to a fanout of 0 to 8 basic CMLs as the upper level input of an XOR is essentially composed of two basic CMLs.

Optimizing Static Frequency Dividers

One of the applications of the propagation delay model is to predict the maximum toggle frequency of static and dynamic frequency dividers. The D-latch schematic diagram is shown in Fig. 15(a), its representation in Fig. 15(b) and the block diagram of the toggle flip-flop is also shown in Fig. 15(c). It is interesting to notice that a CML D-latch can be viewed as an XOR, whose fanout is equal to one XOR (i.e., two basic CMLs). The first basic CML is the regenerative pair Q_4 and Q_6, whose bases are directly loading the output nodes of the same latch (master) while the second basic CML is the input CML of the second latch (slave). Since the minimum static divider configuration demands two cascaded D-latches (Master–Slave), N should be greater than or equal to unity. The minimum period of the input signal can be no less than twice the D-latch propagation delay time. The maximum toggling frequency is then given by:

$$f_{max} = \frac{1}{2(t_{PD} + D_{FO})} \tag{31}$$

Table 3 compares the maximum toggle frequency of a CML static divider predicted by the delay model to that obtained

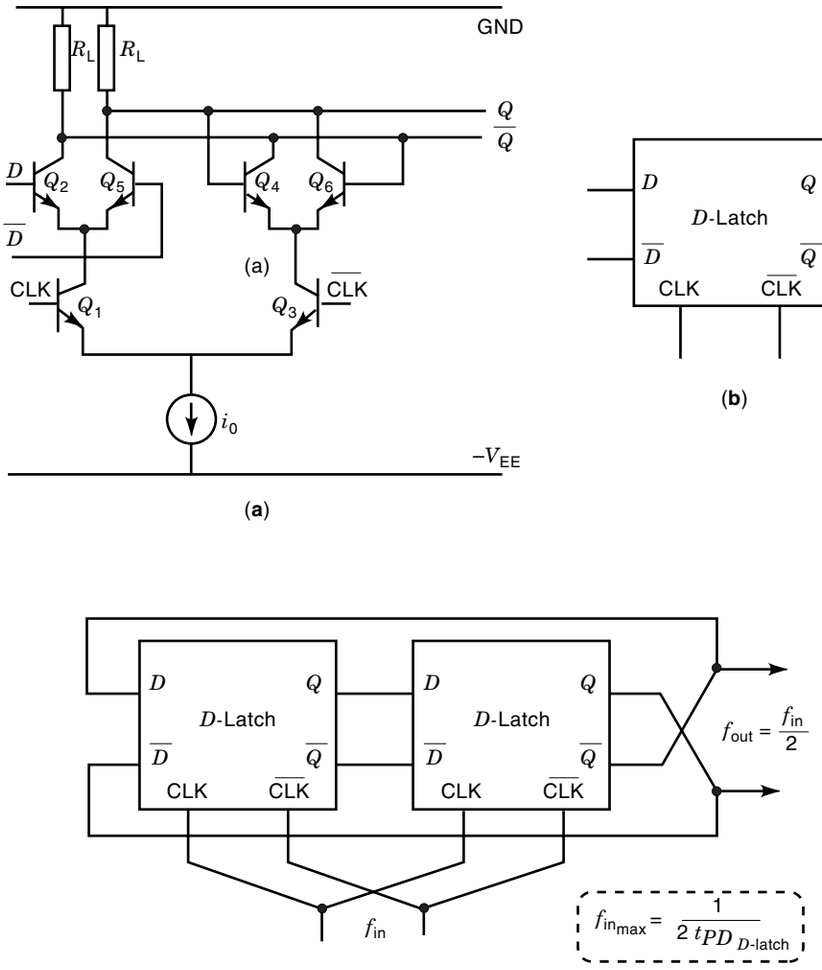


Figure 15. CML-based D-latch (a) schematic; (b) symbol; (c) block diagram of static frequency divider.

from SPICE simulations for both optimized and nonoptimized cases. The input and output voltage swings are assumed to be 200 mV each. The tail current is 2 mA and the load capacitance is 200 fF. Beyond f_{max} the divider ceases to operate correctly and the output frequency is no longer half that of the input signal. Though the improvement percentage is almost insignificant, the model results are in good agreement with SPICE simulation results.

SUMMARY

The basic principles of CML design were discussed. Both dc and transient behavior have been studied along with some relevant key circuit parameters. CML performance has been quantified in terms of speed, propagation delay time, power dissipation, and circuit complexity. An analytical model calculating propagation delay times for two-level series-gated CML high-speed bipolar circuits has been developed. The analytical delay model accounts for different transistor sizes of the two levels. High-current effects were also considered in the presented model. Exploiting these two features, the model has been successfully applied in optimizing the design of a variety of two-level series-gated CML circuits. A comparison with the results obtained by SPICE was presented to verify the applicability of the presented model. The delay model aids the designer in narrowing the design space of series-gated circuits, instead of relying entirely on the designer's intuition. Circuit simulations can be conducted based on clear guidelines set by this approximate model to further optimize the design.

Table 3. CML Static Frequency Divider Results

	SPICE	MODEL
Nonoptimized $A_1 = A_2 = 1$	$f_{max} = 3.8$ GHz $T_{min} = 263$ ps	3.7 GHz 270 ps
Optimized $A_1 = 2, A_2 = 1$	$f_{max} = 3.93$ GHz $T_{min} = 254$ ps	3.96 GHz 252 ps

BIBLIOGRAPHY

1. T. Stetzler, et al., A 2.7–4.5 V single chip GSM transceiver RF integrated circuit, *IEEE J. Solid-State Circuits,* **SC-30**: 1421–1429, 1995.

2. T. Aytur and B. Razavi, A 2-GHz, 6-mW BiCMOS frequency synthesizer, *IEEE J. Solid-State Circuits,* **SC-30**: 1457–1462, 1995.
3. H. Rein, Silicon bipolar integrated circuits for multigigabit per second lightwave communications, *IEEE J. Lightwave Technol.,* **LT-8**: 1371–1378, 1990.
4. K. Runge, et al., Silicon bipolar integrated circuits for multi-Gb/s optical communication systems, *IEEE J. Selected Areas Commun.,* **9**: 636–646, 1991.
5. M. Mizuno, et al., A GHz MOS adaptive pipeline technique using MOS current-mode logic, *IEEE J. Solid-State Circuits,* **SC-31**: 784–791, 1996.
6. R. Romanczyk and B. Leung, BiCMOS Circuits for high speed current-mode D/A converters, *IEEE J. Solid-State Circuits,* **SC-30**: 923–934, 1995.
7. K. G. Ashar, The method of estimating delay in switching circuits and the figure of merit of a switching transistor, *IEEE Trans. Electron. Devices,* **ED-11**: 497–506, 1964.
8. P. K. Tien, Propagation delay in high speed silicon bipolar and GaAs HBT digital circuits, *Int. J. High Speed Electron.,* **1**: 101–124, 1990.
9. J. M. Stork, Bipolar transistor scaling for minimum switching delay and energy dissipation, *IEDM Technol. Dig.,* 550–553, 1988.
10. M. Y. Ghannam, et al., An analytical model for the determination of the transient response of CML and ECL gates, *IEEE Trans. Electron. Devices,* **ED-37**: 191–201, 1990.
11. A. T. Yang and Y. Chang, Physical timing modeling for bipolar VLSI, *IEEE J. Solid-State Circuits,* **SC-27**: 1245–1254, 1992.
12. W. Fang, et al., Accurate analytical delay expressions for ECL and CML circuits and their applications to optimizing high-speed bipolar circuits, *IEEE J. Solid-State Circuits,* **SC-25**: 572–583, 1990.
13. H. J. Greub, et al., High-performance standard cell library and modeling technique for differential advanced bipolar current tree logic, *IEEE J. Solid-State Circuits,* **26**: 749–762, 1991.
14. W. Fang, et al., An analytical maximum toggle frequency expression and its application to optimizing high-speed ECL frequency dividers, *IEEE J. Solid-State Circuits,* **SC-25**: 920–930, 1990.
15. P. Antognetti and G. Massobrio, *Semiconductor Device Modeling with SPICE,* New York: McGraw-Hill, 1988.
16. K. M. Sharaf and M. I. Elmasry, An accurate analytical propagation delay model for high-speed CML bipolar circuits, *IEEE J. Solid-State Circuits,* **SC-29**: 31–45, 1994.
17. K. M. Sharaf and M. I. Elmasry, Analysis and optimization of series-gated CML and ECL high-speed bipolar circuits, *IEEE J. Solid-State Circuits,* **SC-31**: 202–212, 1996.
18. R. Gu, K. Sharaf, and M. Elmasry, *High-Performance Digital VLSI Circuit Design,* Norwell, MA: Kluwer, 1995.

K. M. SHARAF
M. I. ELMASRY
University of Waterloo