A Decoupling Technique for CMOS Strong-Coupled Structures

Li Yang and J.S.Yuan
School of Electrical Engineering and Computer Science
University of Central Florida, Orlando, FL 32816,U.S.A
{lyang, yuanj}@pegasus.cc.ucf.edu

ABSTRACT
A novel decoupling technique for CMOS gate with strong-coupled components is presented. The feedback structure is reduced to a unidirectional one, which facilitates the fast-timing simulation. Based on the waveform relaxation technique, the decoupling principles and procedures are discussed in detail. Most of the factors, which influence the operation of the strong-coupled components, are taken into account. The proposed method is validated by SPICE simulation.

Categories and Subject Descriptors
B.7.2 [**Integrated Circuits**]: Design Aids – *simulation.*

General Terms
Algorithms, Design.

Keywords
Strong coupled components, feedback loop, hysteresis, decoupling, fast-timing simulation, waveform relaxation, gate modeling, look-up table.

1. INTRODUCTION
The current trend on VLSI design gives growing demand for the fast and accurate simulation. Many papers on the modeling of CMOS gates have been published [1][2]. Nabavi-Lishi and Rumin [1] used a semi-empirical method converting a simple CMOS gate to an inerter. Chatzigeorgiou et al., [2] mapped a complex gate to the NAND or NOR gate structure and then convert it to an equivalent inverter. Both methods, however, are not applicable for CMOS gates with strong-coupled components (SCC) such as flip-flops.

Generally, CMOS digital circuits are unidirectional. This simplifies the circuit analysis because the circuit can be broken into many blocks and analyzed independently. When the feedback structure is involved, however, simple partitioning technique is no longer valid. In this paper, we propose a decoupling technique to solve the transistor-level feedback effect. The proposed analytical model clearly shows the influence of the design characteristics of the SCC gate. After decoupling, the conventional gate modeling and fast-timing simulation methods can be applied, and some circuit analyses and optimizations can be carried out directly without the complications associated with simulation.

GLSVLSI'03, April 28-29, 2003, Washington, DC, USA.

2. CHARACTERIZATION OF CMOS GATE WITH SCC
Figure 1 shows the typical cross-coupled inverters driven by another inverter. The parasitic capacitances are also shown. C_{mid} is the drain junction capacitances of inverters INV1 and INV3. It also includes the local interconnect capacitance C_{int} (not shown in Fig.1). C_L consists of the drain junction capacitance of INV2, interconnect capacitance, and fan-out loading.

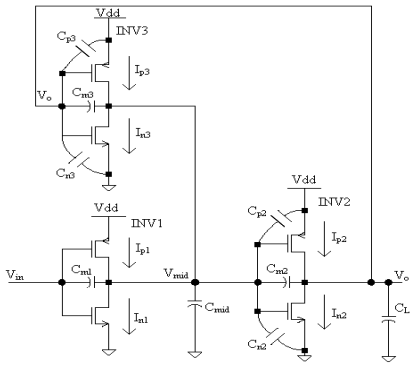

Figure 1. Schematic of cross-coupled inverters with parasitic capacitance (close-loop circuit)

Figure 2 depicts the voltage transfer curve. The hysteresis phenomenon is clearly shown. As input voltage changes, the intermediate node V_{mid} delays its voltage change because the feedback gate INV3 tries to hold the current state and resist the switch. Since INV1 is usually much stronger than INV3, INV1 finally dominates. As V_{mid} exceeds the threshold of INV2, the output V_o begins to switch. This in turn triggers the transition of INV3. INV3 then acts as the positive feedback and enhances the transition of V_{mid}. As a result, the output transition region is very narrow, as seen in Fig 2. These features also contribute to the transient effects (see Fig. 5). The overshoot/undershoot at V_{mid} node is very small, and the output transition time is relatively short. Moreover, the output transition time is insensitive to the input transition time. This insensitivity validates the use of the look-up table in the following sections.

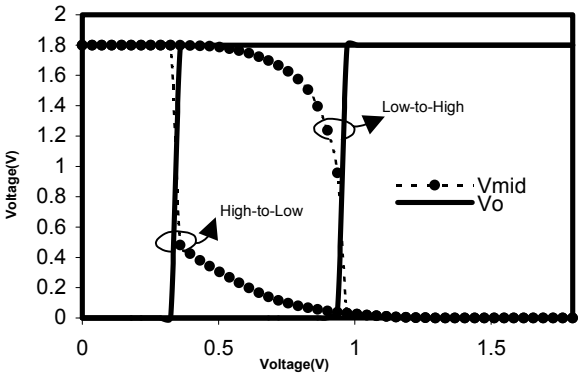

Figure 2. Voltage transfer curve of circuit in Fig. 1

3. DECOUPLING TECHNIQUE FOR CMOS GATES WITH SCC

For the circuit in Fig. 1, the differential equations at V_{mid} node and V_o node are derived using Kirchoff's current law

$$\frac{dV_{mid}}{dt} = \frac{C_{m2}+C_{m3}}{C_t}\frac{dV_o}{dt} + \frac{C_{m1}}{Ct}\frac{dV_{in}}{dt} + \frac{I_p - I_n}{C_t} \quad (1)$$

$$\frac{dV_o}{dt} = \frac{C_{m2}+C_{m3}}{C_o}\frac{dV_{mid}}{dt} + \frac{I_{p2} - I_{n2}}{C_o} \quad (2)$$

where

$$C_t = C_{m1}+C_{m2}+C_{m3}+C_{n2}+C_{p2}+C_{mid}$$
$$I_p = I_{p1}+I_{p3},\quad I_n = I_{n1}+I_{n3}$$
$$C_o = C_{m2}+C_{m3}+C_{n3}+C_{p3}+C_L$$

Note that C_t and C_o are the capacitance at V_{mid} node and V_o node, respectively.

To account for the output response after the feedback loop is broken, the characteristics of the differential equations in (1) and (2) must be maintained. In this case

- The gate-to-drain capacitance of INV2 is increased to C_{m2}+C_{m3};
- The load capacitance is increased to C_L+C_{n3}+C_{p3}; Since INV3 is usually very weak, C_{n3} and C_{p3} are negligible; and
- The intermediate node capacitance is decreased to C_{int} - C_{m3}.

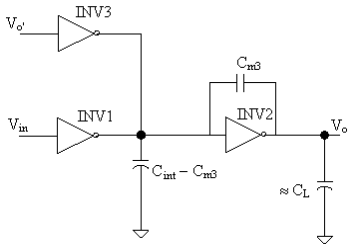

Figure 3. Gate level circuit after decoupling (open-loop circuit)

Figure 3 displays the gate level circuitry after the feedback loop is decoupled. An extra input V_o' to INV3 is needed. Assuming INV3 is very weak and $V_{o'} \approx V_o$, the differential equations (1) and (2) are essentially kept after decoupling. Practically, since V_o' is unknown, a linear output estimation is assumed. Using this assumption, only the slope k_o and a time point t_o on the output waveform are needed. In the proposed method, the slope is calculated from the look-up table, and the point t_o at $V_o = 0.9V_{dd}$ is calculated analytically.

3.1 Slope of the Output Waveform

Considering the input falling edge ($V_{in} = V_{dd} - kt$, where $k = V_{dd} / T_f$, T_f is the input fall time), the slope of the output waveform is given by $k_o = 0.8V_{dd}/\tau$, where τ is defined from $90\%V_{dd}$ to $10\%V_{dd}$. As mentioned before, τ is a weak function of the input transition time. Once the circuit configuration is given, only several output transition times are needed to set up a look-up table, and the linear interpolation is used to solve for other values. To avoid extensive simulations, a program is written to set up this table using dynamic simulation approach. The program scans different circuit configurations and adjusts the step dynamically. Note that when the feedback structure, such as flip-flop or domino circuit [3], is employed, sizes of INV2 and INV3 are usually fixed for a given technology. Only the equivalent driving gate INV1 and the loading C_L need to be scanned. Usually, 3-4 transition times are needed to provide the accuracy within 2% of the simulation results, given that the input and output transition times are in the same order for internal switching. The table setup is a one-time effort. Once done, it is used as a static resource.

3.2 t_o at the Output Waveform

Before the transition of the output voltage, two effects contribute to the output waveform. One is the gate-drain feed-forward effect. Another is the holding effect of INV3, as mentioned in Sec. 2. These two effects delayed the output transition until the input changes significantly. Therefore, the time point t_o at $V_o = 0.9V_{dd}$ can be obtained with the circuitry in Fig. 3, where V_o' is set to be a DC source around $0.95V_{dd}$ to account for the HIGH output voltage during $t \le t_o$. This method is demonstrated in Fig. 4, where the close-loop circuit represents the circuit with feedback (Fig. 1.) and the open-loop circuit represents the decoupled circuit in Fig. 3. It's clear that their output waveforms before the point (t_o, 0.9 V_{dd}) are almost identical, and so are V_{mid} waveforms. This method provides good accuracy for most circuit configurations, except when a large loading is driven by a weak gate.

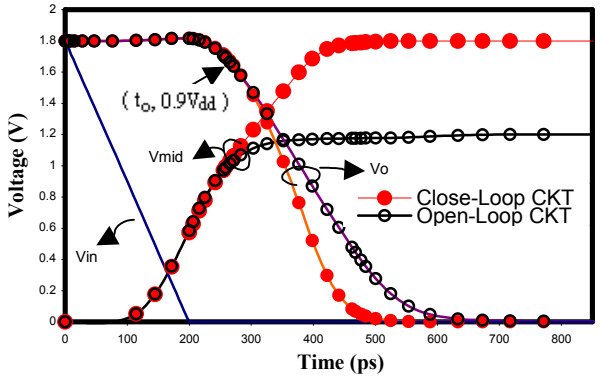

Figure 4. Voltage versus time between close-loop circuit and open-loop circuit with DC V_o'

Usually, since $C_{m2} + C_{m3} << C_t$, the coupling between (1) and (2) is very weak. Also note that the time interval of interest is

$t \in (0, t_o)$ and V_o stays close to V_{dd} during this period. Using the waveform relaxation technique (WR) [4], (1) reduces to

$$\frac{dV_{mid}}{dt} = \frac{C_{m1}}{C_t}\frac{dV_{in}}{dt} + \frac{I_p - I_n}{C_t} \tag{3}$$

In the following section, (2) and (3) are solved analytically to find t_o. By knowing the slope k_o and the time point t_o, V_o' is represented by

$$V_o = \begin{cases} V_{dd} & t \le t_d \\ V_{dd} - |k_o|(t - t_d) & t_d < t < t_d + t_f \\ 0 & t \ge t_d + t_f \end{cases} \tag{4}$$

where $t_f = V_{dd} / |k_o|$ and $t_d = t_o - 0.1V_{dd} / |k_o|$.

4. ANALYTICAL SOLUTIONS OF t_O

In order to solve (2) and (3) analytically, the n[th]-power MOS model is employed [5]. Since V_o' is a DC HIGH voltage, M_{p3} is always cut-off and its effect is neglected.

4.1 Solving V_{mid} Waveform from (3)

Region I: M_{n1} and M_{n3} are in the linear region and M_{p1} is cut-off. Since undershoot of V_{mid} is weak, V_{mid} is set to 0.

Region II: M_{n1} and M_{n3} are still in the linear region. M_{p1} enters the saturation region. As in [6], the drain current of M_{n1} is approximated as a linear function of time, depicted as $I_{n1est} = s_n(t - T_{onp1})$. The drain current of M_{n3} is approximately proportional to V_{mid}. The solution of (3) is

$$V_{mid} = R_{n3}\left[A + I_{\text{int}\,p1} - s_n\left(t - \frac{1}{E} - T_{onp1}\right) - e^{-E(t - T_{onp1})}\left(A + \frac{s_n}{E}\right)\right] \tag{5}$$

where $A = -kC_{m1}$, $T_{offn1} = \frac{V_{dd} - V_{tn1}}{k}$, $T_{onp1} = \frac{V_{tp1}}{k} + \frac{n_{p1}}{E}$,

$$E = \frac{1}{C_t R_{n3}},\ R_{n3} = \frac{V_{dsatn3}}{I_{satn3}},$$

and

$$I_{\text{int}\,p1} \approx B_{p_1}(kt - V_{tp1})^{n_{p1}} - \frac{n_{p1}k}{E}(kt - V_{tp1})^{n_{p1}-1} + \frac{n_{p1}(n_{p1} - 1)k^2}{E^2}(kt - V_{tp1})^{n_{p1}-2}$$

Due to the feedback, the V_{mid} rises slowly in this region and hence reduces the saturation region of M_{n1}. A Taylor expansion around the point $t = T_{offn1}$ up to the second order is used, for both (5) and the saturation condition of M_{n1}. The point (T_{satn1}, $V_{midsatn1}$) is then solved and used as boundary condition (BC) for the next region.

Region III: M_{p1} is still saturated and M_{n3} is still in the linear region. M_{n1} enters the saturation region. The V_{mid} is solved as

$$V_{mid} = R_{n3}(A + I_{\text{int}\,p1} - I_{\text{int}\,n1}) + u_{23} \cdot e^{-E(t - T_{satn1})} \tag{6}$$

where

$$I_{\text{int}\,n1} \approx B_{n1}[\alpha_{na} + \alpha_{nb}(t - T_{nopt} - \frac{1}{E})] + B_{n1}\alpha_{nc}\left[(t - T_{nopt})^2 - 2\frac{(t - T_{nopt})}{E} + \frac{2}{E^2}\right]$$

The constants α_{na}, α_{nb} and α_{nc} are the coefficients of the Taylor expansion of the drain saturation current of M_{n1}. The time point T_{nopt} is selected to be the middle point between T_{satn1} and T_{offn1}. Since the saturation region of M_{n1} is narrow. This Taylor expansion provides good accuracy. The integration constant u_{23} is solved straightforward from (6) using the BC from the previous region. Similarly, the point (T_{offn1}, $V_{midoffn1}$) is calculated from (6) and used as the BC for the next region.

Region IV: M_{n1} turns off. M_{p1} is still saturated and M_{n3} is still in the linear region. Two sub-regions are defined.

Region IVa. The input signal V_{in} doesn't reach its final value in this region. The V_{mid} is given by

$$V_{mid} = R_{n3}(A + I_{\text{int}\,p1}) + u_{34} \cdot e^{-E(t - T_{offn1})} \tag{7}$$

If V_{mid} at $t = T_f$ is less than V_{dsatn3}, M_{n3} is still in the linear region and the calculation continues on the Region IVb with the BC (T_f, V_{midTf}). Otherwise, M_{n3} becomes saturated before V_{in} reaches its final value. The saturation point (T_{satn3}, $V_{midsatn3}$) is used as the BC for Region Va.

Region IVb. V_{in} turns to be flat in the region. The V_{mid} is represented by:

$$V_{mid} = I_{satp1ss}R_{n3} + u_{44} \cdot e^{-E(t - T_f)} \tag{8}$$

where $I_{satp1ss}$ is the saturation current of M_{p1} in this region. The point (T_{satn3}, $V_{midsatn3}$) is calculated for Region Vb.

Region V: M_{n1} is still cut-off and M_{p1} is still saturated. M_{n3} leaves linear region and becomes saturated. Like Region IV, two sub-regions are defined.

Region Va. If V_{in} still drops in this region, V_{mid} is described by

$$V_{mid} = \frac{A - I_{satn3}}{C_t}t + \frac{B_{p1}}{C_t k \cdot (n_{p1} + 1)}(kt - V_{tp1})^{n_{p1}+1} + u_{45a} \tag{9}$$

If M_{p1} becomes linear within $t \le T_f$, the BC (T_{satp1}, $V_{midsatp1}$) is solved for Region VIa. Otherwise, the point (T_f, V_{midTf}) is used for Region Vb.

Region Vb. If V_{in} reaches static state, the voltage waveform of V_{mid} is depicted as

$$V_{mid} = \frac{I_{satp1ss} - I_{satn3}}{C_t}t + u_{x5b} \tag{10}$$

The point (T_{satp1},$V_{dsatp1ss}$) serves as the BC for the next region.

Region VI: M_{n1} is still cut-off and M_{n3} remains in saturation. M_{p1} becomes linear. The operation modes of all transistors have no further changes.

Region VIa. V_{in} still decreases in this region. The solution of (3) is

$$V_{mid} = V_{dd} - Z_6 I_6{}^{\alpha}(1 - \alpha \cdot I_6{}^{-1}) - u_{56a} \cdot e^{-I_6} \tag{11}$$

where $\alpha = \frac{1}{n_{p1} - m_{p1} + 1} - 1$, $I_{6a} = \frac{3B_{p1}(\alpha + 1)}{2C_t K_{p1} k}$,

$$I_6 = I_{6a}(kt - V_{tp1})^{\frac{1}{\alpha+1}},\ Z_6 = \frac{2(I_{satn3} - A)K_{p1}}{3B_{p1}} I_{6a}^{-\alpha}$$

The BC for the final region is (T_f, V_{midTf}).

Region VIb. The V_{in} is stable in this region. V_{mid} is represented by

$$V_{mid} = V_{dd} - V_{dsatp1ss} - \frac{\varepsilon}{\mu}\left[\frac{2}{e^{\varepsilon(t-u_{x6b})}+1} - 1\right] \qquad (12)$$

where $\varepsilon = \frac{2\sqrt{I_{satp1ss}(I_{satp1ss} - I_{satn3})}}{C_t V_{dsatp1ss}}$, $\mu = \frac{2I_{satp1ss}}{C_t V_{dsatp1ss}^2}$.

From (12), the static-state voltage of the intermediate node is depicted by

$$V_{midss} = V_{dd} - V_{dsatp1ss}(1 - \sqrt{1 - \frac{I_{satn3}}{I_{satp1ss}}}) \qquad (13)$$

It doesn't depend on the input signal.

4.2 Solving V_o for t_o from (2)

Using the results derived in Sec. 4.1, the differential equation in (2) is solved analytically for the output waveform. To keep the accuracy of the V_{mid} waveform while simplify the solving process, the piecewise-linear of V_{mid} is employed. The solving process of (2) is similar to that in Sec. 4.1 but much more simple because only one inverter is involved, and the solving process stops as long as V_o drops below $0.9V_{dd}$.

5. RESULTS AND DISCUSSIONS

The proposed decoupling method is evaluated with TSMC 0.18 μm technology. Figure 5 gives some examples for the different input transition times. To verify the decoupling method and the analytical solution provided in Sec. 4, the simulated V_{mid} waveforms are from the circuit in Fig. 3 by setting V_o' as a DC source, while the simulated V_o waveforms are from the circuit in Fig. 1. It's clear that the output voltage slopes are nearly the same in Fig. 5. The undershoot of the V_{mid} waveform is very small, even when INV1 is much stronger than INV3. Since the channel length modulation effect has been neglected in Sec. 4, the saturation current of M_{n3} has been underestimated. As a result, the proposed method has a little higher static-state voltage V_{midss} than the simulation result, as seen in Fig. 5.

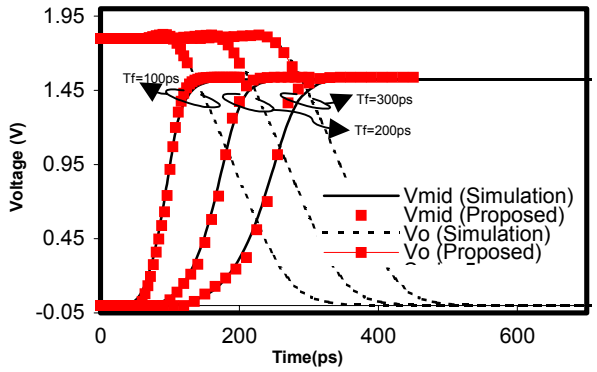

Figure 5. Comparison between simulated and calculated waveforms for different transition times

The proposed approach can be used in many CMOS gates with SCC, such as CMOS D latch and domino logic gates[3]. After decoupling, these gates are changed to be unidirectional, and the conventional gate modeling and fast-timing simulation methods can be employed. In some cases, when the fast analytical estimation is needed, the linear-estimated output V_o' is used directly. Figure 6 shows the examples of using the linear V_o' for delay analysis. It's clear that for the certain circuit configuration, the delay increases as the input transition time becomes longer because t_o is postponed, which is also seen in Fig. 5. The proposed results in Fig. 6 are very close to the SPICE simulation results.

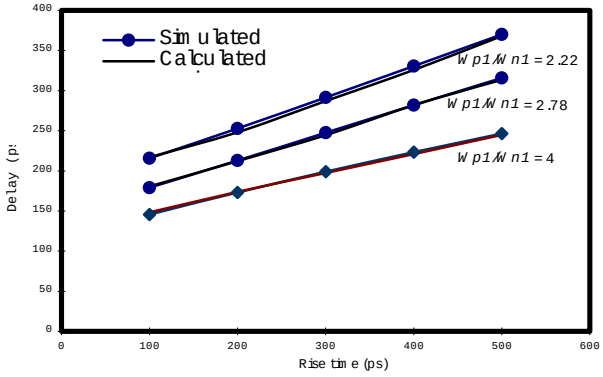

Figure 6. Propagation delay of the close-loop circuit evaluated with the linear-estimated output

6. SUMMARY

A decoupling technique for CMOS gate with SCC is presented. The characteristics of the feedback structure are analyzed. Based on the node differential equations, the feedback loop is decoupled with the waveform relaxation technique. The linear-estimated output, which is calculated from the look-up table and the analytical analysis, is used as the extra input for the decoupled gate. Good agreement with the SPICE simulation results is obtained. The proposed method can be used for analytical analysis and fast-timing simulation of CMOS SCC gates.

7. REFERENCES

[1] Nabavi-lishi, A., and Rumin, N.C., "Inverter models of CMOS gates for supply current and delay evaluation," IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems, 1994, Vol. 13, pp. 1271-1279

[2] Chatzigeorgiou, A., Nikolaidis, S., and Tasoukalas, I., "A modeling technique for CMOS gates," IEEE Trans. Computer-Aided Design, 1999, Vol. 18, pp. 557-575

[3] Weste, N.H.E. and Eshraghian, K., Principles of CMOS VLSI design: A system perspective, Addison-Wesley, 1992, 2nd edition

[4] White, J.K., and Vincentelli, A.S., Relaxation techniques for the simulation of VLSI circuits, Kluwer, Norwell, 1987

[5] Sakurai, T., and Newton, A. R., "A simple MOSFET model for circuit analysis," IEEE Trans. Electron Devices, 1991, Vol. 38, pp. 887-894

[6] Bisdounis, T., Nikolaidis, S., and Koufopavlou, O., "Analytical transient response and propagation delay evaluation of the CMOS inverter for short-channel devices," IEEE J. Solid-State Circuits, 1998, Vol. 33, pp. 302-306