A Simple Hybrid Model for Accurate Delay Modeling of a Multi-Input Gate

Arman Ferdowsi, Jürgen Maier, Daniel Öhlinger and Ulrich Schmid
Embedded Computing Systems Group (E191-02)
TU Wien, Vienna, Austria
{aferdowsi, jmaier, doehlinger, s}@ecs.tuwien.ac.at

***Abstract*—Faithfully representing small delay variations caused by transitions on different inputs in close temporal proximity is a challenging task for digital circuit delay models. In this paper, we show that a simple hybrid model, derived from considering transistors as ideal switches in a simple RC model, leads to a surprisingly accurate model. By analytically solving the resulting ODEs for a NOR gate, explicit expressions for the delay are derived. In addition, we experimentally compare our model's predictions to SPICE simulations and to existing delay models.**

***Index Terms*—Digital circuit delay models, multi input switching**

I. Introduction

Digital circuit design relies heavily on fast digital timing analysis techniques, since they are orders of magnitude faster than analog simulations, e.g., in *SPICE*. In contrast to static approaches, dynamic digital timing analysis predicts the propagation of arbitrary signal traces throughout a circuit. Going beyond the popular pure (= constant input-to-output delay) and inertial delay (= constant delay + too short pulses being removed) models [1], *single-history delay models* [2], [3] achieve an improved behavioral coverage. Indeed, as proved in [4], it is inevitable for any faithful delay model that a gate's input-to-output delay $\delta(T)$, for a given transition, depends on a parameter like the previous-output-to-input delay T.

The *involution delay model* (IDM) proposed in [3] consists of zero-time boolean gates, which are interconnected by single-input single-output involution delay channels. The latter are characterized by a delay function $\delta(T)$, which is a negative involution, in the sense that $-\delta(-\delta(T)) = T$. Unlike all other existing delay models, the IDM faithfully models glitch propagation in the simple short-pulse filtration problem, and is hence the only candidate for a faithful delay model so far.

Whereas the accuracy of IDM predictions for inverter chains or clock trees reported in [5] is impressive, this is less so for circuits involving multi-input gates. We conjecture that this is mainly due to the inherent lack of properly covering output delay variations caused by *multiple input switching* (MIS) in close temporal proximity [6], also known as the *Charlie effect* (named after Charles Molnar, who identified its causes in the 70th of the last century). Compared to the *single input switching* (SIS) case, the output transition is sped up/slowed down with decreasing transition separation time on different inputs here. Single-input, single-output IDM delay channels obviously cannot exhibit such a behavior.

Multiple approaches have been proposed to cover MIS effects in literature, ranging from linear [7] or quadratic fitting [8] over higher-dimensional model representation [9] to recent machine learning methods [10]. However, none of these naturally generalizes to multi-input involution channels.

In order to define a 2-input IDM channel, we generalize the analog first-order model matching a classic IDM channel (which can be viewed as a hybrid model with two modes) to a four-mode hybrid model: For each state of the inputs $(A, B) \in \{(0,0),(0,1),(1,0),(1,1)\}$, a system of first-order *ordinary differential equations* (ODEs) is derived, which governs the analog trajectory of the gate's output in the respective mode. At an input change, the mode is instantaneously switched, in a way that guarantees *continuity* of the output signal. Whereas similar approaches have been advocated in [11], [12], these rely on analog fitting or extraction of unique switching waveforms.

Main contributions: (1) We introduce[1] a simple hybrid ODE model of a 2-input CMOS `NOR` gate. Replacing the transistors in a simple RC circuit model by ideal switches (as in switch-level simulation) enables us (i) to analytically solve the ODE systems for every state and (ii) to derive expressions for the resulting MIS gate delay.

(2) We use an appropriately extended version of our (publicly available[2]) *Involution Tool* [5] to compare the average modeling accuracy of our hybrid model to other analog/digital simulations. The results show that—using empirically fitted parameters—our hybrid model outperforms the original IDM as well as standard models like inertial delays. Nevertheless, it turns out that one of the MIS effects cannot be properly modeled, which somewhat impairs the overall accuracy.

Paper organization: In Section II, we explain the causes for MIS delay variations and determine characteristic values for a `NOR` gate. After presenting our simple hybrid ODE model in Section III, we investigate its ability to capture MIS effects qualitatively (Section IV) and quantitatively (Section V). Some conclusions and directions of future research close the paper in Section VI.

II. Multiple Input Switching (MIS)

In this section, we will provide some basic explanations for MIS effects and quantify those by conducting analog

This research was supported by the Austrian Science Fund (FWF) project DMAC (grant no. P32431).

[1] A full version of our paper is available at arXiv [13].
[2] Accessible via https://github.com/oehlinscher/InvolutionTool.

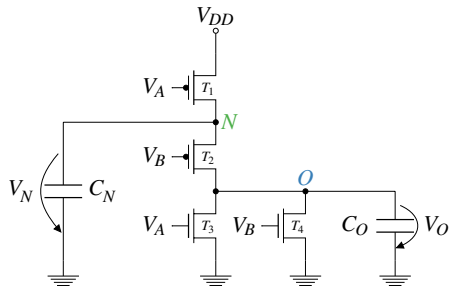


Fig. 1. Transistor level implementation of the NOR gate.

simulations using Spectre (version 19.1) and the Nangate Open Cell Library featuring FreePDK15™ FinFET models [14]. The investigated CMOS NOR gate is among the simplest multi-input gates and hence a natural target for our analysis. Its transistor-level implementation, with the parasitic capacitance C_N and the output load capacitance C_O, is shown in Fig. 1.

In the sequel, we apply the rising/falling input waveforms $f_{\uparrow/\downarrow}(t-t_A)$ on input A resp. $f_{\uparrow/\downarrow}(t-t_B)$ on input B, whereat t_A resp. t_B denote the point in time the *discretization threshold voltage* $V_{th} = V_{DD}/2$ is crossed. In the same spirit, t_O denotes the time when the output voltage V_O crosses $V_{DD}/2$. Varying t_A and t_B allows us to represent the gate delay over the relative input separation time $\Delta = t_B - t_A$.

We first consider the case of a falling output transition. In a nutshell, either transistor T_3 or T_4 starts to conduct (is closed), while one of the two pMOS transistors in series stops conducting (is opened). Consequently, the output is drained and V_O starts to decrease. It is easy to see that it makes a difference whether only one or both nMOS transistors are closed, i.e., only a single or both inputs switch, as it takes, at least theoretically, only half the time to drain the output in parallel. So the MIS causes a *speed-up* here, whose effect is clearly visible in the analog waveforms shown in Fig. 2a: When both transistors start to conduct, the output slope changes notably, leading to a reduction of the gate delay.

As the first rising input transition already induces an output transition, the relevant gate delay is $\delta_S^{\downarrow}(\Delta) = t_O - \min(t_A, t_B)$, i.e., the time difference between the threshold crossing of the output and the earlier input (see Fig. 2b). As predicted, the delay is the smallest for simultaneous transitions ($\Delta = 0$), whereat the change in delay is around 30%. Note that $\delta_S^{\downarrow}(-\infty) \neq \delta_S^{\downarrow}(\infty)$ is mainly caused by transistor T_2, which is closed in one case, connecting nodes N and O, while in the other it is open (see Section III). Although the absolute values differ, our results fit very well[3] to previous investigations in other technologies [7], [8], [9].

For rising output transitions, the behavior of the NOR is quite different. First and foremost, the gate only switches after both inputs have changed (see Fig. 2c), resulting in the gate delay $\delta_S^{\uparrow}(\Delta) = t_O - \max(t_A, t_B)$. At the transistor level, each falling input transition causes one of the nMOS to stop conducting while simultaneously one of the pMOS gets closed.

[3]We ran our simulations with an older technology library (65 nm) as well, which confirmed the delay values reported in the literature.

We emphasize that the shape of the output signal in Fig. 2c is essentially independent of Δ, only the position in time varies. This is in accordance with the fact that there is only a single path connecting the output to V_{DD}.

The SIS delays $\delta_S^{\uparrow}(\infty)$ and $\delta_S^{\uparrow}(-\infty)$ again differ (see Fig. 2d), i.e., the gate delay depends on the order of the input transitions. Taking a closer look at the schematics in Fig. 1 reveals that an early transition on A closes the topmost transistor and thus causes the internal node N between the pMOS to be charged to V_{DD}. By contrast, an early transition on B causes N to be fully discharged, which obviously prolongs the transition time.

In any case, the gate delay for $|\Delta| \to 0$ increases, i.e., the MIS effect is a *slow-down* here. The main causes are coupling capacitances between N and the input: If both inputs switch at the same time, the parasitic current (dis)charges C_N, possibly below *GND*, since both adjacent transistors are still open. After they start to conduct, the additional charge has to be compensated, which explains the increased delay. Naturally, the delay variations depend on the initial value of V_N and thus on the switching history of the gate. Note that we used the worst case ($V_N = GND$) in all our simulations.

III. Simple hybrid ODE model

In our attempt to analytically express the gate delays of a NOR gate, we replace the transistors by zero-time switches: Depending on whether the appropriate input is above (logical 1) resp. below (logical 0) $V_{th} = V_{DD}/2$, an nMOS transistor is replaced by a fixed resistor $R < \infty$ or removed ($R = \infty$), while a pMOS is handled in the opposite way. Note that this is similar to the approach used in [15], with the main difference that we added a capacitance at the internal node in the p-stack (C_N) and one at the output (C_O) (cf. Fig. 1). Thus, we end up with a system of coupled first-order differential equations.

Since the inputs (A, B) are interpreted as binary signals (logical 0/1), we need to consider 4 different RC circuits and their corresponding ODE systems $V'(t) = A \cdot V(t) + g(t)$. $V(t) \in \mathbb{R}^2$ is a two-element vector, representing the voltage V_N at the internal node N in Fig. 1, and the gate output voltage V_O, i.e.,

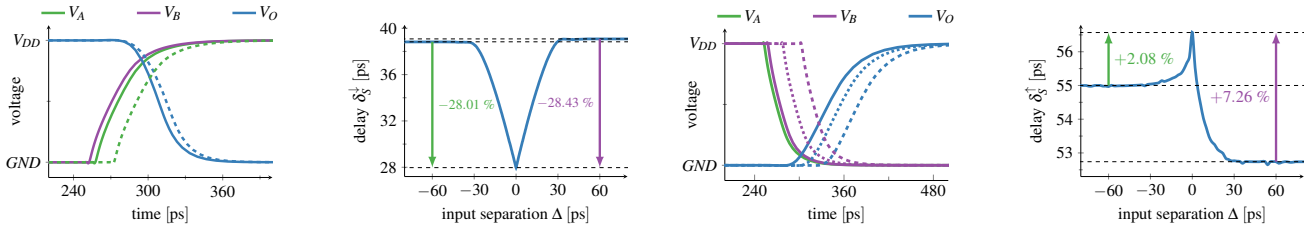
$$V(t) = \begin{pmatrix} V_N \\ V_O \end{pmatrix},$$

while the non-homogeneous term $g(t)$ is either identically zero or a constant. In the sequel, we will evaluate $V(t)$ for each input combination individually.

A. System $(1,1)$: $V_A = V_{DD}$, $V_B = V_{DD}$

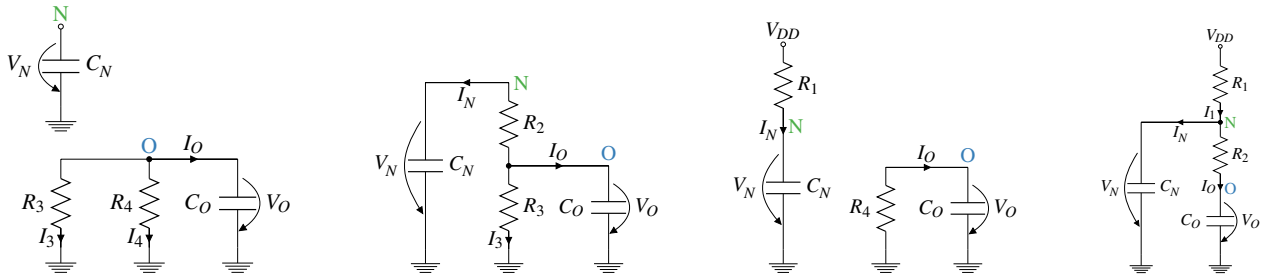
If inputs A and B are above the threshold, both nMOS transistors are conducting and thus replaced by resistors (see Fig. 3a), causing the output O to be discharged in parallel. By contrast, N is completely isolated and keeps its value. Formally, we obtain

$$C_N \cdot \frac{\mathrm{d}}{\mathrm{d}t} V_N = 0,$$
$$C_O \cdot \frac{\mathrm{d}}{\mathrm{d}t} V_O = I_O = -I_3 - I_4 = -V_O \cdot \left(\frac{1}{R_3} + \frac{1}{R_4} \right).$$



(a) Falling output transition (b) Falling output delay (c) Rising output transition (d) Rising output delay

Fig. 2. Analog simulation results for the CMOS NOR gate.



(a) system $(1,1)$ (b) system $(1,0)$ (c) system $(0,1)$ (d) system $(0,0)$

Fig. 3. First order RC approximations.

This homogeneous system can be rewritten in matrix form

$$\begin{pmatrix} \frac{\mathrm{d}}{\mathrm{d}t}V_N \\ \frac{\mathrm{d}}{\mathrm{d}t}V_O \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -(\frac{1}{C_O R_3} + \frac{1}{C_O R_4}) \end{pmatrix} \cdot \begin{pmatrix} V_N \\ V_O \end{pmatrix},$$

which leads to the general solution

$$\begin{pmatrix} V_N \\ V_O \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot e^{-\left(\frac{1}{C_O R_3} + \frac{1}{C_O R_4}\right)t}.$$

B. System $(1,0)$: $V_A = V_{DD}$, $V_B = GND$

Since T_1 and T_4 are open (see Fig. 3b), node N is connected to O, and O to *GND*. Note that both capacitances have to be discharged over resistor R_3, resulting in less current that is available for discharging C_O. More specifically, we observe $I_O = -I_3 - I_N$ and hence obtain

$$C_N \cdot \frac{\mathrm{d}}{\mathrm{d}t}V_N = I_N = -\frac{V_N - V_O}{R_2},$$
$$C_O \cdot \frac{\mathrm{d}}{\mathrm{d}t}V_O = I_O = -I_3 - I_N = -\frac{V_O}{R_3} + \frac{V_N - V_O}{R_2}.$$

The matrix form of this homogeneous system is

$$\begin{pmatrix} \frac{\mathrm{d}}{\mathrm{d}t}V_N \\ \frac{\mathrm{d}}{\mathrm{d}t}V_O \end{pmatrix} = \begin{pmatrix} -\frac{1}{C_N R_2} & \frac{1}{C_N R_2} \\ \frac{1}{C_O R_2} & -(\frac{1}{C_O R_2} + \frac{1}{C_O R_3}) \end{pmatrix} \cdot \begin{pmatrix} V_N \\ V_O \end{pmatrix}$$

which has the general solution

$$\begin{pmatrix} V_N \\ V_O \end{pmatrix} = c_1 \cdot \begin{pmatrix} \frac{1}{C_N R_2} \\ \alpha + \beta \end{pmatrix} e^{\lambda_1 t} + c_2 \cdot \begin{pmatrix} \frac{1}{C_N R_2} \\ \alpha - \beta \end{pmatrix} e^{\lambda_2 t},$$

where

$$\alpha = \frac{C_O R_3 - C_N(R_2 + R_3)}{2 C_O C_N R_2 R_3}, \tag{1}$$

$$\beta = \frac{\sqrt{(C_O R_3 + C_N(R_2 + R_3))^2 - 4 C_O C_N R_2 R_3}}{2 C_O C_N R_2 R_3}, \tag{2}$$

$$\lambda_{1,2} = -\frac{C_O R_3 + C_N(R_2 + R_3)}{2 C_O C_N R_2 R_3} \pm \beta. \tag{3}$$

C. System $(0,1)$: $V_A = GND$, $V_B = V_{DD}$

Opening transistors T_2 and T_3, as shown in Fig. 3c, decouples the nodes N and O once again. We thus get the non-homogeneous system of ODEs

$$C_N \cdot \frac{\mathrm{d}}{\mathrm{d}t}V_N = I_N = \frac{V_{DD} - V_N}{R_1},$$
$$C_O \cdot \frac{\mathrm{d}}{\mathrm{d}t}V_O = I_O = -\frac{V_O}{R_4},$$

which is in matrix representation

$$\begin{pmatrix} \frac{\mathrm{d}}{\mathrm{d}t}V_N \\ \frac{\mathrm{d}}{\mathrm{d}t}V_O \end{pmatrix} = \begin{pmatrix} -\frac{1}{C_N R_1} & 0 \\ 0 & -\frac{1}{C_O R_4} \end{pmatrix} \cdot \begin{pmatrix} V_N \\ V_O \end{pmatrix} + \begin{pmatrix} \frac{V_{DD}}{C_N R_1} \\ 0 \end{pmatrix}.$$

It is easy to check that the fundamental matrix solution to the corresponding homogeneous system is

$$\phi(t) = \begin{pmatrix} e^{-\frac{t}{C_N R_1}} & 0 \\ 0 & e^{-\frac{t}{C_O R_4}} \end{pmatrix}$$

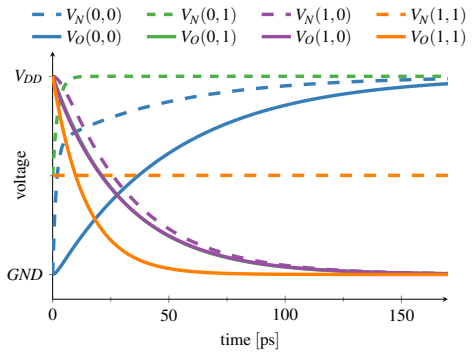


Fig. 4. Temporal evolution of the trajectories for all systems.

leading to the general non-homogeneous solution

$$\begin{pmatrix} V_N \\ V_O \end{pmatrix} = \begin{pmatrix} c_1 \cdot e^{-\frac{t}{C_N R_1}} + V_{DD} \\ c_2 \cdot e^{-\frac{t}{C_O R_4}} \end{pmatrix}.$$

D. System $(0,0)$: $V_A = GND$, $V_B = GND$

Closing both pMOS transistors, as shown in Fig. 3d, causes both capacitances to be charged over the same resistor R_1, similarly to system (1,0). Since $I_O = I_1 - I_N$, the ODE system describing the behavior is

$$C_O \cdot \frac{\mathrm{d}}{\mathrm{d}t} V_O = I_O = \frac{V_N - V_O}{R_2},$$
$$C_N \cdot \frac{\mathrm{d}}{\mathrm{d}t} V_N = I_N = I_1 - I_O = \frac{V_{DD} - V_N}{R_1} - \frac{V_N - V_O}{R_2},$$

which is a non-homogeneous system, with the matrix form

$$\begin{pmatrix} \frac{\mathrm{d}}{\mathrm{d}t} V_N \\ \frac{\mathrm{d}}{\mathrm{d}t} V_O \end{pmatrix} = \begin{pmatrix} -(\frac{1}{C_N R_1} + \frac{1}{C_N R_2}) & \frac{1}{C_N R_2} \\ \frac{1}{C_O R_2} & -\frac{1}{C_O R_2} \end{pmatrix} \cdot \begin{pmatrix} V_N \\ V_O \end{pmatrix} + \begin{pmatrix} \frac{V_{DD}}{C_N R_1} \\ 0 \end{pmatrix}.$$

By straightforward but tedious calculations, it follows that the fundamental matrix solution of the homogeneous system is

$$\phi(t) = \begin{pmatrix} \frac{1}{C_N R_2} \cdot e^{\lambda_1 t} & \frac{1}{C_N R_2} \cdot e^{\lambda_2 t} \\ (\alpha + \beta) \cdot e^{\lambda_1 t} & (\alpha - \beta) \cdot e^{\lambda_2 t} \end{pmatrix},$$

where

$$\alpha = \frac{C_O(R_1 + R_2) - C_N R_1}{2 C_O C_N R_1 R_2}, \tag{4}$$

$$\beta = \frac{\sqrt{(C_N R_1 + C_O(R_1 + R_2))^2 - 4 C_O C_N R_1 R_2}}{2 C_O C_N R_1 R_2}, \tag{5}$$

$$\gamma = -\frac{C_N R_1 + C_O(R_1 + R_2)}{2 C_O C_N R_1 R_2}, \tag{6}$$

$$\lambda_{1,2} = \gamma \pm \beta. \tag{7}$$

The general solution of the non-homogeneous system is

$$\begin{pmatrix} V_N \\ V_O \end{pmatrix} = \begin{pmatrix} \frac{c_1}{C_N R_2} e^{\lambda_1 t} + \frac{c_2}{C_N R_2} e^{\lambda_2 t} + V_{DD} \\ c_1 \cdot (\alpha + \beta) e^{\lambda_1 t} + c_2 \cdot (\alpha - \beta) e^{\lambda_2 t} + V_{DD} \end{pmatrix}.$$

E. Trajectory Comparison

Fig. 4 depicts the signals $V_{N/O}(n,m)(t)$ over time t in system (n,m). The initial values were set to $V_N(0) = V_O(0) = V_{DD}$, with the exception of $V_N(0,0)(0) = V_O(0,0)(0) = GND$ and $V_N(1,1)(0) = V_{DD}/2$. Compared to the cases where only one nMOS is closed, the output trajectory of system $(1,1)$ is much steeper. Note that this is in line with the considerations for the speed-up MIS effect in Section II.

IV. Modeling MIS Effects

In this section, we will investigate how well our simple hybrid ODE model is capable of faithfully representing the MIS effects described in Section II. It turns out that the speed-up is modeled appropriately, whereas the slow-down is only partially covered. Note that also the approaches presented in [11], [12] struggled with this effect, such that the authors finally resorted to fitting the delays for these cases.

For our analysis, we computed the delay as a function of the input separation time $\Delta = t_B - t_A$ for falling and rising output transitions, and compared it with our analog simulation results (cf. Fig. 2b and Fig. 2d). Similar to Section II, we start with a falling output transition. To compute the delay for a given Δ, we need to combine two solutions:

1): Starting in the system $(0,0)$ initially, which models a gate whose inputs have been 0 for a very long time, we switch to $(1,0)$ resp. $(0,1)$ at time $t = 0$ and compute the corresponding trajectory.

2): When in the mode entered in 1), we switch to system $(1,1)$ at time t_s, and determine t_O where $V_O(t_O) = V_{DD}/2$. The delay is extracted as $t_O - \min(t_A, t_B) = t_O$, since the earlier of the two inputs triggers the output transition. Starting in system $(1,0)$ results in $\Delta = t_s$, whereat for system $(0,1)$ we get $\Delta = -t_s$, which accounts for the reversed order of the input transitions.

The calculation of the rising output delay is carried out analogously, with the exception that we start in the system $(1,1)$ initially and switch to $(1,0)$ $(\Delta < 0)$ resp. $(0,1)$ $(\Delta > 0)$ at $t = 0$, before turning to $(0,0)$ at t_s. The sought delay is now equal to $t_O - \max(t_A, t_B) = t_O - t_s$, since it is the later of the two inputs that initiates the output change. Note carefully, however, that it is unfortunately not clear which initial value to use for V_N in the system $(1,1)$ here: As the latter does not change the value of $V_N(t)$ at all, the proper initial value would be the *actual* value of V_N in the state (m,n) the system was in *before/at* the switch to $(1,1)$ occurred.

For a quantitative comparison, we parameterized the resistances and capacitances in our model using a least square fitting approach, with the goal to match the output threshold crossing times $\delta_S^{\uparrow}(\pm\infty)$ and $\delta_S^{\uparrow}(0)$ resp. $\delta_S^{\downarrow}(\pm\infty)$ and $\delta_S^{\downarrow}(0)$ shown in Fig. 2. Interestingly, simultaneous fitting of all three data points turned out to be impossible, even for the "well-behaved" case of falling output transitions. To understand why, we derived analytic expressions for the values

$$\delta^{\downarrow}(-\infty) \approx \ln(2) \cdot C_O R_4 \quad \text{and} \quad \delta^{\downarrow}(0) = \frac{\ln(2) \cdot C_O R_3 R_4}{R_3 + R_4}$$

TABLE I
EMPIRICALLY OBTAINED PARAMETER VALUES

Parameter	Value
R_1	$37.088 \times 10^3\,\Omega$
R_2	$44.926 \times 10^3\,\Omega$
R_3	$45.150 \times 10^3\,\Omega$
R_4	$48.761 \times 10^3\,\Omega$
C_N	$59.486 \times 10^{-18}\,\mathrm{F}$
C_O	$617.259 \times 10^{-18}\,\mathrm{F}$

(as well as for $\delta^{\downarrow}(\infty)$, $\delta^{\uparrow}(-\infty)$, $\delta^{\uparrow}(0)$ and $\delta^{\uparrow}(\infty)$), by inverting the explicit formulas of our trajectories. Unfortunately, necessary simplifications that enabled these calculations induced some approximation errors, which made a direct computation of the desired parameters impossible.

Nevertheless, important insights could be gained. More specifically, since R_3 and R_4 are the on-resistors of the two nMOS transistors and should hence be roughly the same, we obtain $\frac{\delta^{\downarrow}(-\infty)}{\delta^{\downarrow}(0)} \approx \frac{R_3+R_4}{R_3} \approx 2$. Since the desired ratio according to our simulations (see Fig. 2) is $\frac{\delta_S^{\downarrow}(-\infty)}{\delta_S^{\downarrow}(0)} \approx \frac{38\,\mathrm{ps}}{28\,\mathrm{ps}}$, however, we could not even simultaneously fit these two values with reasonable choices for R_3 and R_4.

We solved the problem by adding (that is, subtracting) a pure delay $\delta_{\min} = 18\,\mathrm{ps}$, also present in the original IDM, which defers the switching to the new state upon an input transition. This results in an effective ratio of $\frac{20\,\mathrm{ps}}{10\,\mathrm{ps}} = 2$, which could finally be matched by least squares fitting, leading to the parameter values presented in Table I. To also accommodate $\delta_{\min}$ in the MIS delay computations, we just need to transform $t_O - \min(t_A, t_B) = t_O$ to $\delta_M^{\downarrow}(\Delta) = t_O + \delta_{\min}$ and, correspondingly, $\delta_M^{\uparrow}(\Delta) = t_O - t_s + \delta_{\min}$. Note that the same $\delta_{\min} = 18\,\mathrm{ps}$ was also used for rising output transitions.

Utilizing the found parameters, we can finally visualize the delay predictions of our model. Fig. 5 shows the very good fit of $\delta_M^{\downarrow}(\Delta)$ for a falling output transition compared to the analog simulation results presented in Section II. Unfortunately, a comparable coverage of the MIS effects for rising output transitions cannot be achieved (see Fig. 6): For none of the initial values $V_N \in \{GND, V_{DD}/2, V_{DD}\}$, the computed delay $\delta_M^{\uparrow}(\Delta) = t_O - t_s + \delta_{\min}$ reasonably matches $\delta_S^{\uparrow}(\Delta)$ obtained in our analog simulations. More specifically, for $V_N = V_{DD}$ and $V_N = V_{DD}/2$, our simple ODE model fails to correctly predict the case of $\Delta < 0$, i.e., switching to system $(1, 0)$ at $t = 0$ and then to $(0, 0)$ at time $t = \Delta$. In the case of $V_N = GND$, which reasonably matches $\delta_S^{\uparrow}(\pm\infty)$ and $\delta_S^{\downarrow}(\pm\infty)$ and thus has been used in Section V, it fails to model the MIS peak around $\Delta = 0$ in Fig. 2d for small negative and positive Δ.

We conclude that our simple ODE model perfectly captures the MIS effects caused by the parallel transistors, but not for the ones caused by transitors arranged in series.

V. MODELING ACCURACY

In this section, we will compare our hybrid model to inertial delays and the IDM. To be able to do so, we added our model to our Involution Tool [5]. Using the QuestaSim Foreign Language Interface (FLI) [16] allowed us to escape the Involution Tool's standard VHDL environment and to execute C code. In a second step, Python code was called from C, which finally implemented our hybrid channels.

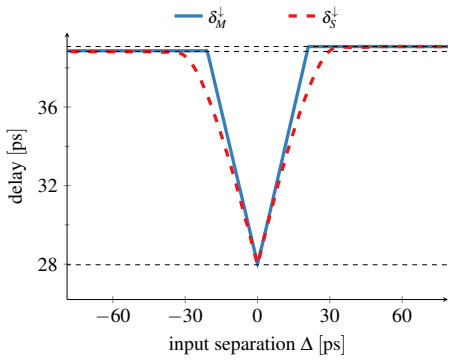


Fig. 5. Computed MIS delays for falling output transitions.

For the evaluation, we again used the 15 nm Nangate Open Cell Library featuring FreePDK15™ FinFET models [14] ($V_{DD} = 0.8\,\mathrm{V}$). Based on a Verilog description of a NOR gate, we utilized the Cadence tools Genus and Innovus (version 19.11) to perform optimization, placement and routing. Finally, we extracted the parasitic networks from the final layout to obtain *SPICE* models, which we used as golden reference in analog Spectre (version 19.1) simulations.

Using the parameter set introduced in Table I, we performed simulations for various waveform configurations, ranging from very short to broad pulses. Each simulation consisted of 500 transitions, except for the last simulation, where we generated 250 transitions. The simulations have been repeated 20 times, and the averaged results are presented in Fig. 7. The waveform configuration *100/50 - LOCAL* describes the case where transitions are created individually for each input, according to a normal distribution, with $\mu = 100\,\mathrm{ps}$ and $\sigma = 50\,\mathrm{ps}$. *GLOBAL* indicates that the transitions are not calculated separately for each input but rather for all inputs together. This option allows to test how accurately the delay models perform for large absolute values of Δ, since concurrent transitions are unlikely with this configuration.

The results are compared in terms of the deviation area, which is calculated as follows: The digitized *SPICE* traces are subtracted from the corresponding traces of the digital delay model and the absolute area is summed up. Since absolute values are meaningless, the results are normalized with the inertial delay as baseline.

For short pulses ($\mu = 100\,\mathrm{ps}, \mu = 200\,\mathrm{ps}$), the superiority of the hybrid model with $\delta_{\min}$ can be clearly seen. The deviation area is less than half that of the inertial delays. Moreover, it also outperforms the Exp-Channel, which we chose as representation for the IDM in the Involution Tool, with $\delta_{\min} = 20\,\mathrm{ps}$. Note that we had to determine the latter empirically, since there is no proper parametrization of IDM channels representing multi-input gates available. The hybrid model without pure delay performs worse, which is primarily due to the imperfect

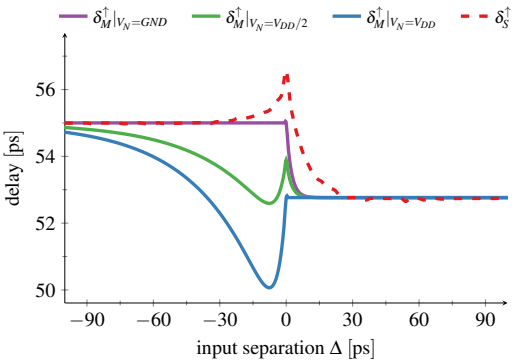


Fig. 6. Computed MIS delays for rising output transitions.

delay matching. We should also emphasize that, for the first two waveform configurations, a lot of transitions are happening within a range of $\Delta = [-40\,\text{ps}, 40\,\text{ps}]$, where the hybrid model without pure delay has deficiencies.

For broader pulses ($\mu = 2\,\text{ns}, \mu = 5\,\text{ns}$), which are covered by the last two waveform configurations, it can be seen that the hybrid model and the inertial delay model perform similar. This is due to the fact that $|\Delta| \gg 100\,\text{ps}$, where the matching is nearly perfect. The Exp-Channel shows deficiencies for broad pulses, which is caused by placing the delay channel at the output and the consequential inability to determine which input caused the transition. Since $\delta_S^{\uparrow}(\infty)$ and $\delta_S^{\uparrow}(-\infty)$ differ, the Exp-Channel simply delays the transition by their average, which explains the observed inaccuracies.

In terms of simulation runtime, our simple experiments reveal a minor overhead of the hybrid model compared to the simple inertial delay model or the Exp-Channel of 6 %, which seems acceptable in view of the increased modeling accuracy. For more robust numbers, more extensive simulation runs are necessary, which we are planning to execute in the near future.

VI. Conclusions

We introduced a simple hybrid ODE model for a two-input CMOS NOR gate, which naturally generalizes the hybrid analog model corresponding to standard single-input, single-output involution channels. The ODEs governing the switching waveforms of the output, based on the state of the inputs, have been obtained by replacing transistors with ideal switches in a simple RC model of the circuit. By analytically solving the resulting ODE systems, we obtained a digital gate delay model that faithfully reproduces all MIS effects, except in one particular situation. We also incorporated our hybrid model in the Involution Tool for digital timing analysis and compared the average accuracy for random traces in a custom circuit for different channel models. Our results show that our new hybrid model outperforms both classic involution channels and standard inertial delay channels with respect to modeling accuracy. Future work will be devoted to alternative models that fully capture all MIS effects.

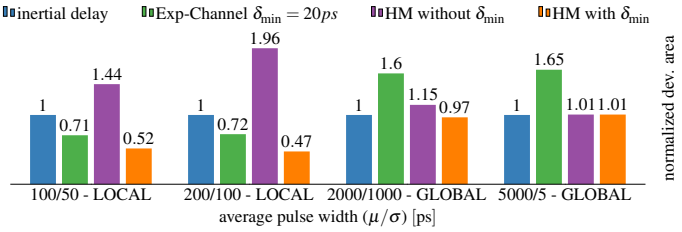


Fig. 7. Accuracy of inertial delay, Exp-Channel and hybrid model compared to analog simulations of a NOR gate. Lower bars indicate better results.

References

[1] S. H. Unger, "Asynchronous sequential switching circuits with unrestricted input changes," *IEEE Transaction on Computers*, vol. 20, no. 12, pp. 1437–1444, 1971.

[2] M. J. Bellido-Díaz, J. Juan-Chico, and M. Valencia, *Logic-Timing Simulation and the Degradation Delay Model*. London: Imperial College Press, 2006.

[3] M. Függer, R. Najvirt, T. Nowak, and U. Schmid, "A faithful binary circuit model," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 39, no. 10, pp. 2784–2797, 2020.

[4] M. Függer, T. Nowak, and U. Schmid, "Unfaithful glitch propagation in existing binary circuit models," *IEEE Transactions on Computers*, vol. 65, no. 3, pp. 964–978, March 2016.

[5] D. Öhlinger, J. Maier, M. Függer, and U. Schmid, "The involution tool for accurate digital timing and power analysis," *Integration*, vol. 76, pp. 87 – 98, 2021.

[6] L.-C. Chen, S. K. Gupta, and M. A. Breuer, "A new gate delay model for simultaneous switching and its applications," in *Proceedings of the 38th Design Automation Conference (IEEE Cat. No.01CH37232)*, 2001, pp. 289–294.

[7] A. R. Subramaniam, J. Roveda, and Y. Cao, "A finite-point method for efficient gate characterization under multiple input switching," *ACM Trans. Des. Autom. Electron. Syst.*, vol. 21, no. 1, pp. 10:1–10:25, Dec. 2015. [Online]. Available: http://doi.acm.org/10.1145/2778970

[8] J. Shin, J. Kim, N. Jang, E. Park, and Y. Choi, "A gate delay model considering temporal proximity of multiple input switching," in *2009 International SoC Design Conference (ISOCC)*, Nov 2009, pp. 577–580.

[9] J. Sridharan and T. Chen, "Modeling multiple input switching of cmos gates in dsm technology using hdmr," in *Proceedings of the Design Automation Test in Europe Conference*, vol. 1, March 2006, pp. 6 pp.–.

[10] O. V. S. Shashank Ram and S. Saurabh, "Modeling multiple-input switching in timing analysis using machine learning," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 40, no. 4, pp. 723–734, 2021.

[11] E. Melcher, W. Röthig, and M. Dana, "Multiple input transitions in cmos gates," *Microprocessing and Microprogramming*, vol. 35, no. 1, pp. 683 – 690, 1992, software and Hardware: Specification and Design. [Online]. Available: http://www.sciencedirect.com/science/article/pii/016560749290387M

[12] N. Abdallah and P. Bazargan-Sabet, "Modeling the effects of input slew rate and temporal proximity of input transitions in event-driven simulation," in *2006 Proceeding of the Thirty-Eighth Southeastern Symposium on System Theory*, March 2006, pp. 185–189.

[13] A. Ferdowsi, J. Maier, D. Öhlinger, and U. Schmid, "A Simple Hybrid Model for Accurate Delay Modeling of a Multi-Input Gate," 2021, arXiv:2111.11182 [cs.OH].

[14] M. Martins, J. M. Matos, R. P. Ribas, A. Reis, G. Schlinker, L. Rech, and J. Michelsen, "Open cell library in 15nm freepdk technology," in *Proceedings of the 2015 Symposium on International Symposium on Physical Design*, ser. ISPD '15. New York, NY, USA: ACM, 2015, pp. 171–178. [Online]. Available: http://doi.acm.org/10.1145/2717764.2717783

[15] E. Schneider, S. Holst, X. Wen, and H.-J. Wunderlich, "Data-parallel simulation for fast and accurate timing validation of cmos circuits," in *2014 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2014, pp. 17–23.

[16] Mentor Graphics Corporation, *Foreign Language Interface Manual*, 2016, Software Version 10.5c.