An Accurate Hybrid Delay Model for Multi-Input Gates

Arman Ferdowsi, Ulrich Schmid, and Josef Salzmann
Embedded Computing Systems Group (E191-02)
TU Wien, Vienna, Austria
{aferdowsi, s, jsalzmann}@ecs.tuwien.ac.at

***Abstract*—Accurately modeling the delay of multi-input gates is challenging due to variations caused by switching different inputs in close temporal proximity. This paper introduces a hybrid model for a CMOS NOR gate, which is based on replacing transistors with time-variant resistors. We analytically solve the entangled differential equations and derive expressions for the gate delays, which also paved the way to an empirical parametrization procedure. By comparison with Spice simulation data, we show that our model indeed faithfully represents all relevant multi-input switching effects. Using an implementation in the Involution Tool, we also demonstrate that it surpasses the few alternative models known so far in terms of accuracy.**

I. Introduction

Digital timing analysis techniques are essential for modern circuit design. *Dynamic* digital timing analysis traces the propagation of *individual* transitions of a signal throughout a circuit, like analog simulations e.g. via *SPICE* do, albeit much faster. An interesting application domain for high-accuracy dynamic timing analysis are Spiking Neural Network (SNN) hardware implementations [Fur16, BVMRRVB20], for example. Neurons in SNNs communicate via discrete-value continuous-time signals, where analog values are effectively encoded via the delay between successive spike events. Besides rate-based approaches, where a single analog value is represented by the frequency of a whole spike train, *inter-spike interval* (ISI) time encoding uses the time interval between two spikes to encode this information in a more energy-efficient way. Obviously, for the latter, it is crucial that the actual implementation of such a communication link preserves delays between successive signal transitions as accurately as possible.

Single-history delay models like [1], [2], which generalize the popular pure (= constant input-to-output delay) and inertial delay (= constant delay + too short pulses being removed) models [3], have proved their potential with respect to improved behavioral coverage and accuracy here. Particularly relevant for us is the *involution delay model* (IDM) proposed in [2], which consists of zero-time boolean gates that are interconnected by single-input single-output involution delay channels. IDM channels 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 short-pulse filtration problem [4], and is hence the only candidate for a faithful delay model known so far. The IDM is accompanied by a publicly available timing analysis framework (the *Involution Tool* [5]), which allows to compare the accuracy of different delay models. In particular, it allows to randomly generate input traces for a given circuit, and to evaluate the accuracy of IDM predictions compared to *SPICE*-generated transition times and to other digital delay models.

In [5], it has been shown that the accuracy of IDM predictions for single-input, single-output circuits like inverter chains or clock trees is very good, but less so for circuits involving multi-input gates. It was argued that this is primarily due to the IDM's 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, output transitions are sped up/slowed down with decreasing transition separation time on different inputs. Clearly, single-input, single-output delay channels cannot exhibit such a behavior at all.

Related work: MIS effects have of course been addressed in the literature before, with approaches ranging from linear [7] or quadratic [8] fitting over higher-dimensional macromodels [9] and model representations [10] to recent machine learning methods [11]. The resulting models are either empirical or statistical and unsuitable as a basis for dynamic digital timing analysis, however.

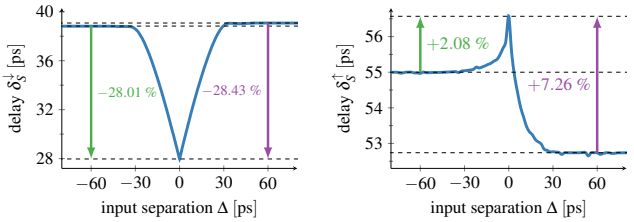
In [12], Ebergen, Fairbanks and Sutherland studied the performance of micropipelines, the control logic of which is often made up of a chain of RendezVous elements. Their analysis rests on a delay model for the RendezVous elements, which relies on the Charlie effect exhibited by the constituent Muller C-gate. Rather than providing a model that explains this MIS phenomenon, however, they just take it for granted and use it for analyzing the token propagation performance in a chain of RendezVous elements.

The only attempt to develop a delay model that captures MIS effects and is suitable for dynamic timing analysis known to us is the DATE'21 paper by Ferdowsi et. al. [13], where the authors proposed a hybrid model of a 2-input CMOS `NOR` gate based on replacing transistors by ideal zero-time switches in the simple RC model shown in Fig. 2a. Since a `NOR` gate consists only of 4 transistors, this results in a hybrid model with 4 modes, one for each possible digital state of the inputs $(A, B) \in \{(0,0),(0,1),(1,0),(1,1)\}$. Each mode is described by a simple system of constant-coefficient first order *ordinary differential equations* (ODEs), which are switched (continuously) upon an input state transition. Whereas this leads to a surprisingly accurate delay model, it fails to faithfully model one of the MIS effects, namely, for a rising output transition.

Main contributions: (1) Like in [13], we also use a CMOS `NOR` gate for our primary[1] target circuit. Instead of just replacing all transistors by zero-time switches, however, we assume that the two serial (pMOS) transistors are switched off in zero time, but are switched on by following a simple time evolution function $\sim 1/t + R^{on}$. Rather than including the state of these resistors in the ODE systems and switching those continuously upon an input transition (which would blow up the system dimension), we use properly chosen first-order ODEs with *time-varying coefficients* instead.

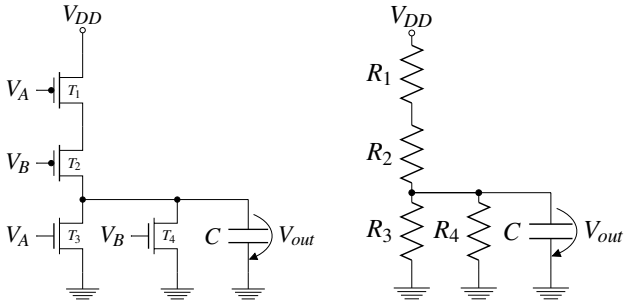
(2) We analytically solve the ODEs for all modes, and by composing the trajectory functions of different modes, like $(0,0) \to (0,1) \to (1,1)$, we also determine accurate analytic approximations for all MIS gate delays.

[1]We stress, however, that our results actually also apply literally to the "dual" NAND gate and even to some configurations of general `AOI`(and-or-inverter) gates. Moreover, part of our current work is devoted to applying our approach to other gates like Muller C-gates as well.



(a) Falling output delay (b) Rising output delay

Fig. 1: MIS effect in the CMOS NOR gate, taken from [13].



(a) Transistor level [13] (b) Hybrid Model

Fig. 2: Implementations of a CMOS NOR gate.

(3) We provide a procedure for parametrizing our new delay model for a given circuit. More specifically, given certain SIS and MIS delay values, we use insights gained from the *analytic* gate delay approximations derived in (2) for guiding the process of empirically fitting our many parameters to match these delay values. To validate our model, we apply it to a CMOS NOR gate both in a 15 nm technology and in a 65 nm technology. It turns out that the delay predictions of our model match the real delays very well, in particular, also in that MIS case where [13] fails.

(4) We implemented our new delay model in the Involution Tool [5], and experimentally compared the average modeling accuracy and the simulation times of our model to other analog/digital simulations for one of the circuits studied in [5]. As expected, our model outperforms all alternative models considered, without an undue performance penalty.

Paper organization: In Section II, we briefly summarize the results of the analog simulations used for quantifying MIS delays for a 15 nm CMOS NOR gate reported in [13], which provide our baseline. Section III introduces our hybrid ODE model. In Section IV, we analytically solve the entangled ODEs and derive gate delay expressions from these solutions. Section V provides our parametrization procedure, which is applied to both our 15 nm and some 65 nm CMOS NOR gate simulation data for model validation. Using this parametrization, we finally quantify the average modeling accuracy using the Involution Tool in Section VI. Some conclusions and directions of future research are provided in Section VII.

II. Multiple Input Switching (MIS)

In this section, we provide a summary of the MIS effects in CMOS NOR gates reported in [13]. Let t_A, t_B, and t_O denote the points in time when the analog trajectories of input signals A, B, and the output signal O cross the *discretization threshold voltage* $V_{th} = V_{DD}/2$, respectively. Varying t_A and t_B allows to study the gate delay ($t_O - t_A$ resp. $t_O - t_B$, depending on the particular output state) over the relative *input separation time* $\Delta = t_B - t_A$.

In the case of a falling output transition, either the nMOS transistor T_3 or T_4 starts to conduct (is closed), while one of the two pMOS transistors in series stops conducting (is opened). Obviously, closing both T_3 and T_4 leads to an accelerated discharge of the capacitance C and thus a (substantial) *speed-up* MIS effect. Fig. 1a shows the relevant gate delay $\delta_S^{\downarrow}(\Delta) = t_O - \min(t_A, t_B)$, i.e., the time difference between the threshold crossing of the output and the earlier input, extracted from analog simulations. The delay for simultaneous transitions ($\Delta = 0$) is indeed around 30% smaller than on the outskirts.

For rising output transitions, the behavior of the NOR is quite different. Each falling input transition causes one of the nMOS to stop conducting while simultaneously one of the pMOS gets closed. Since there is only a single path connecting the output to V_{DD}, the shape of the output signal is essentially independent of Δ; only the position in time varies. Since the gate only switches after both inputs have changed, the gate delay is $\delta_S^{\uparrow}(\Delta) = t_O - \max(t_A, t_B)$. The resulting MIS effect is a (moderate) *slow-down*, i.e., the gate delay increases for $|\Delta| \to 0$.

III. Our Hybrid Model

In our attempt to accurately model the gate delays of our NOR gate, we introduce a model (see Fig. 2b) that replaces the two serial (pMOS) transistors by time-varying resistors and the two parallel (nMOS) by ideal switches as in the model proposed in [13] (called *ideal switch model* in the sequel). The values $R_i(t)$, $i \in \{1, \ldots, 4\}$ thereby vary between some fixed on-resistance R_i and the off-resistance ∞. This results in a hybrid model with 4 different modes, corresponding to the 4 possible input states $(A, B) \in \{(0,0), (0,1), (1,0), (1,1)\}$.

Whereas one could maintain $R_1(t), \ldots, R_4(t)$ explicitly in every ODE system, and switch between those systems continuously upon an input transition, the resulting *full-state model* would increase the dimension of the ODE systems by 4 and render finding an analytic solution hopeless. We emphasize that the availability of *analytic* formulas for the trajectories turned out to be instrumental to guide the process of model parametrization (see Section V). We therefore use a different approach, namely, incorporating these resistors only in the coefficients of a simple first-order ODE, which hence become non-constant. We will relate our model to these alternative models at the end of this section.

Applying Kirchhoff's rules to Fig. 2b leads to $C\frac{\mathrm{d}V_{out}}{\mathrm{d}t} = \frac{V_{DD}-V_{out}}{R_1(t)+R_2(t)} - \frac{V_{out}}{R_3(t) \;||\; R_4(t)}$. This can be transformed to the non-homogeneous ordinary differential equation (ODE) with non-constant coefficients

$$\frac{\mathrm{d}V_{out}}{\mathrm{d}t} + \frac{V_{out}}{C\,R_g(t)} = U(t), \tag{1}$$

using $\frac{1}{R_g(t)} = \frac{1}{R_1(t)+R_2(t)} + \frac{1}{R_3(t)} + \frac{1}{R_4(t)}$ and $U(t) = \frac{V_{DD}}{C(R_1(t)+R_2(t))}$. Note that the entire voltage divider in Fig. 2b is equivalent to an ideal voltage source $U_0 = V_{DD}\frac{R_3(t)||R_4(t)}{R_1(t)+R_2(t)+R_3(t)||R_4(t)}$ and a serial resistor $R_g(t)$ sourcing C. Consequently, $CU(t) = U_0/R_g(t)$ in (1) is the short-circuit current, and $CU(t) - V_{out}/R_g(t)$ the current actually sourced into C. It is well-known that the general solution of (1) is

$$V_{out}(t) = V_0\, e^{-G(t)} + \int_0^t U(s)\, e^{G(s)-G(t)}\mathrm{d}s, \tag{2}$$

TABLE I: State transitions and modes. $\uparrow$ and $\uparrow\uparrow$ (resp. $\downarrow$ and $\downarrow\downarrow$) represent the first and the second rising (resp. falling) input transitions. $+$ and $-$ specify the sign of the switching time $\Delta = t_B - t_A$.

Mode	Transition	t_A	t_B	R_1	R_2	R_3	R_4
$T^{\uparrow}$	$(0,0) \to (1,0)$	0	$-\infty$	$on \to off$	on	$off \to on$	off
$T^{\uparrow\uparrow}_{+}$	$(1,0) \to (1,1)$	$-\lvert\Delta\rvert$	0	off	$on \to off$	on	$off \to on$
$T^{\uparrow}_{+}$	$(0,0) \to (0,1)$	$-\infty$	0	on	$on \to off$	off	$off \to on$
$T^{\uparrow\uparrow}_{-}$	$(0,1) \to (1,1)$	0	$-\lvert\Delta\rvert$	$on \to off$	off	$off \to on$	on
$T^{\downarrow}_{-}$	$(1,1) \to (0,1)$	0	$-\infty$	$off \to on$	off	$on \to off$	on
$T^{\downarrow\downarrow}_{+}$	$(0,1) \to (0,0)$	$-\lvert\Delta\rvert$	0	on	$off \to on$	off	$on \to off$
$T^{\downarrow}_{+}$	$(1,1) \to (1,0)$	$-\infty$	0	off	$off \to on$	on	$on \to off$
$T^{\downarrow\downarrow}_{-}$	$(1,0) \to (0,0)$	0	$-\lvert\Delta\rvert$	$off \to on$	on	$on \to off$	off

where $V_0 = V_{out}(0)$ denotes the initial condition and $G(t) = \int_0^t (C\, R_g(s))^{-1} \mathrm{d}s$.

Crucial for our model is choosing a suitable time evolution of $R_i(t)$ in the *on-* and *off-mode*, which must facilitate an analytic solution of (2) while being reasonably close to the physical behavior of a transistor. It turned out that it is sufficient to consider the very simple Shichman-Hodges transistor model [14], which states a quadratic dependence of the output current on the input voltage. Approximating the latter by $d\sqrt{t - t_0}$ in the operation range close to the threshold voltage V_{th}, with d and t_0 some fitting parameters, leads to the *continuous resistance model*

$$R_i^{on}(t) = \frac{\alpha_i}{t - t^{on}} + R_i;\ t \geq t^{on}, \tag{3}$$

$$R_i^{off}(t) = \beta_i (t - t^{off}) + R_i;\ t \geq t^{off}, \tag{4}$$

for some constant slope parameters α_i [Ω s], β_i [Ω/s], and on-resistance R_i [Ω]; t^{on} resp. t^{off} represent the time when the respective transistor is switched on resp. off.

In our first attempt, we used the above model for all transistors in our circuit. Besides considerably complicating the analysis, it eventually turned out that this leads to an infeasible model, however. Intuitively, the reason is that the nMOS transistors are too weak to dominate the case of falling output transitions (which they should), no matter which $\alpha > 0$ and $\beta < \infty$ are chosen. Since the ideal switch model of [13] (which corresponds to choosing $\alpha = 0$ and $\beta = \infty$) *is* able to cover some MIS effects, except for the rising output transition case, we started to conjecture that the crucial handle to fix this problem is to gradually switch on only the pMOS transistors according to (3), and set $\alpha_3 = \alpha_4 = 0$ and $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \infty$. And indeed, our results prove that we were right.

For simplicity, we will subsequently use the notation $R_1 = R_{p_A}$, $R_2 = R_{p_B}$ with the abbreviation $2R = R_{p_A} + R_{p_B}$ for the two pMOS transistors T_1 and T_2, and $R_3 = R_{n_A}$, $R_4 = R_{n_B}$ for the two nMOS transistors T_3 and T_4.

Table I shows all possible state transitions and the corresponding resistor time evolution mode switches. Double arrows in the mode switch names indicate MIS-relevant modes, whereas $+$ and $-$ indicate whether A switched before B or the other way around. For instance, assume the system is in state $(0,0)$ for quite some time in the past, i.e., A and B switched to 0 at time $t_A = t_B = -\infty$. This causes R_1 and R_2 to be in the *on-mode*, whereas R_3 and R_4 are in the *off-mode*. Now assume that, at time $t_A = 0$, A is switched to 1. This switches R_1 resp. R_3 to the *off-mode* resp. *on-mode* at time $t_1^{off} = t_3^{on} = t_A = 0$. The corresponding mode switch is $T^{\uparrow}_{-}$ and reaches state $(1,0)$. Now assume that B is also switched to 1, at some time $t_B = \Delta > 0$. This causes R_2 resp. R_4 to switch to *off-mode* resp. *on-mode* at time $t_2^{off} = t_4^{on} = t_B = \Delta$. The corresponding mode switch is $T^{\uparrow\uparrow}_{+}$ and reaches state $(1,1)$; note carefully that the delay is Δ-dependent and hence MIS-relevant.

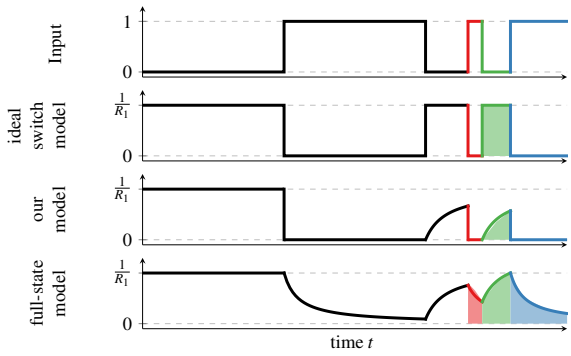


Fig. 3: Comparison of resistance mode switching of a single transistor in our model, in the ideal switch model of [13], and in the full state model.

Our goal will be to analytically compute the output voltage trajectories $V_{out}^{MS}(t)$ given by (2) for every mode switch MS listed in Table I. Due to the particular modes of the resistors in each transition, different expressions for $G(t)$ and $U(t)$ will be obtained. Table II summarizes the relevant formulas for every mode switch.

It is apparent that changing the mode of a resistor upon a state transition (e.g., switching from (4) to (3)) creates a *non-continuous* jump of the resistance value (to ∞ in our example), even for the two serial transistors T_1 and T_2 where the continuous resistance model is used. Relating *our model* (OM) to the *ideal switch model* (IM) and the *full-state model* (FM), which assures continuous resistance changes, will reveal that the differences are likely to be small, however.

In fact, recall that in the ideal switch model all mode switches change the resistance value between R and ∞ in zero time. Nevertheless, [13] reveals that most MIS effects are accurately captured, so discontinuous resistor values per se are not a problem here. And indeed, (1) reveals that $V_{out}(t)$ is not so much determined by the instantaneous value of $1/R_g(t)$ but rather by its *integral*: Intuitively, the unitless monotonically increasing function $G(t)$ in (2) reflects how often the average time constants $\tau = \mathtt{avg}(\tau(t)) = \mathtt{avg}(C\ R_g(t))$ has passed since $t = 0$. The first term of (2) hence denotes the decay of the initial output value w.r.t. the average time constant τ. The second term represents the current that could be sourced into C and its respective decay. Note that $G(s) - G(t) \in [-G(t), 0]$ is increasing for s, meaning that the smaller s the more $U(s)$ is reduced.

Consequently, a qualitative measure for the modeling accuracy is the difference of the *area* under the single-transistor plots shown in Fig. 3.Let A^0 be the area in a time interval where the input is 0, like the green segment, and A^1 the area in a time interval where the input is 1, like the red and the blue segment. We observe the following:

(i) $A^1_{FM} \geq A^1_{OM} = A^1_{IM}$ implies that both our model and the ideal switch model switch the output slower than the full-state model in this case. Similarly, $A^0_{IM} \geq A^0_{FM} \geq A^0_{OM}$ reveals that our model also switches the output slightly slower than the full-state model, which in turn switches slower than the ideal switch model.

(ii) The absolute differences of those areas is only weakly dependent on α and decrease with increasing β.

As a consequence, one may expect that an appropriately parametrized version of our model can provide a modeling accuracy comparable to the considerably more complex full-state model.

It is worth mentioning that our model has only one state variable, namely, V_{out} (due to the load capacitance C), whereas the ideal

TABLE II: Integrals $I_1(t)$, $I_2(t)$, $I_3(t)$ and $U(t)$ for every possible mode switch; $\Delta = t_B - t_A$, and $2R = R_{p_A} + R_{p_B}$.

Mode	$I_1(t) = \int_0^t \frac{\mathrm{d}s}{R_1(s)+R_2(s)}$	$I_2(t) = \int_0^t \frac{\mathrm{d}s}{R_3(s)}$	$I_3(t) = \int_0^t \frac{\mathrm{d}s}{R_4(s)}$	$U(t) = \frac{V_{DD}}{C(R_1(t)+R_2(t))}$
$T_-^{\uparrow}$	$\int_0^t (1/(\beta_1 s + \frac{\alpha_2}{s-\infty} + 2R))\mathrm{d}s$	$\int_0^t (1/(\frac{\alpha_3}{s} + R_{n_A}))\mathrm{d}s$	$\int_0^t (1/(\beta_4(s-\infty) + R_{n_B}))\mathrm{d}s$	$\frac{V_{DD}}{2RC(1+\frac{\beta_1 t}{2R})}$
$T_+^{\uparrow\uparrow}$	$\int_0^t (1/(\beta_1(s+\Delta) + \beta_2 s + 2R))\mathrm{d}s$	$\int_0^t (1/(\frac{\alpha_3}{s+\Delta} + R_{n_A}))\mathrm{d}s$	$\int_0^t (1/(\frac{\alpha_4}{s} + R_{n_B}))\mathrm{d}s$	$\frac{V_{DD}}{C((\beta_1+\beta_2)t+2R+\beta_1\Delta)}$
$T_+^{\uparrow}$	$\int_0^t (1/(\frac{\alpha_1}{s-\infty} + \beta_2 s + 2R))\mathrm{d}s$	$\int_0^t (1/(\beta_3(s-\infty) + R_{n_A}))\mathrm{d}s$	$\int_0^t (1/(\frac{\alpha_4}{s} + R_{n_B}))\mathrm{d}s$	$\frac{V_{DD}}{2RC(1+\frac{\beta_2 t}{2R})}$
$T_-^{\uparrow\uparrow}$	$\int_0^t (1/(\beta_1 s + \beta_2(s+\Delta) + 2R))\mathrm{d}s$	$\int_0^t (1/(\frac{\alpha_3}{s} + R_{n_A}))\mathrm{d}s$	$\int_0^t (1/(\frac{\alpha_4}{s+\Delta} + R_{n_B}))\mathrm{d}s$	$\frac{V_{DD}}{C((\beta_1+\beta_2)t+2R+\beta_2\Delta)}$
$T_-^{\downarrow}$	$\int_0^t (1/(\frac{\alpha_1}{s} + \beta_2(s-\infty) + 2R))\mathrm{d}s$	$\int_0^t (1/(\beta_3 s + R_{n_A}))\mathrm{d}s$	$\int_0^t (1/(\frac{\alpha_4}{s-\infty} + R_{n_B}))\mathrm{d}s$	$\frac{V_{DD}t}{C\beta_2(t^2+(\frac{2R}{\beta_2}-\infty)t+\frac{\alpha_1}{\beta_2})}$
$T_+^{\downarrow\downarrow}$	$\int_0^t (1/(\frac{\alpha_1}{s+\Delta} + \frac{\alpha_2}{s} + 2R))\mathrm{d}s$	$\int_0^t (1/(\beta_3(s+\Delta) + R_{n_A}))\mathrm{d}s$	$\int_0^t (1/(\beta_4 s + R_{n_B}))\mathrm{d}s$	$\frac{V_{DD}t(t+\Delta)}{C(2Rt^2+(\alpha_1+\alpha_2+2\Delta R)t+\alpha_2\Delta)}$
$T_+^{\downarrow}$	$\int_0^t (1/(\frac{\alpha_2}{s} + \beta_1(s-\infty) + 2R))\mathrm{d}s$	$\int_0^t (1/(\frac{\alpha_3}{s-\infty} + R_{n_A}))\mathrm{d}s$	$\int_0^t (1/(\beta_4 s + R_{n_B}))\mathrm{d}s$	$\frac{V_{DD}t}{C\beta_1(t^2+(\frac{2R}{\beta_1}-\infty)t+\frac{\alpha_2}{\beta_1})}$
$T_-^{\downarrow\downarrow}$	$\int_0^t (1/(\frac{\alpha_1}{s} + \frac{\alpha_2}{s+\Delta} + 2R))\mathrm{d}s$	$\int_0^t (1/(\beta_3 s + R_{n_A}))\mathrm{d}s$	$\int_0^t (1/(\beta_4(s+\Delta) + R_{n_B}))\mathrm{d}s$	$\frac{V_{DD}t(t+\Delta)}{C(2Rt^2+(\alpha_1+\alpha_2+2\Delta R)t+\alpha_1\Delta)}$

switch model of [13] also incorporates the voltage at the node between the two pMOS transistors (due to a parasitic capacitance). Since the charge stored in this location does affect the gate delay, a MIS effect is caused in the cases of a falling input transitions $T_+^{\downarrow\downarrow}$ and $T_-^{\downarrow\downarrow}$ (recall Section II). At first sight, one might suspect that ignoring this parasitic capacitance in our model could impair its ability to capture these MIS effects. Our results in Section V reveal, however, that this is not the case.

Finally, we want to stress that our model can be readily applied to other gates as well. This is particularly true for the CMOS `NAND` gate, which is obtained from the `NOR` gate in Fig. 2b by replacing nMOS transistors with pMOS and vice versa, and swapping V_{DD} and *GND*. Of course, now the nMOS transistors are the serial ones (where the continuous resistance models must be used). All that is needed to make the results of our analysis matching is to invert the inputs, e.g., by replacing state $(0,1)$ by $(1,0)$ and to consider $V_{DD} - V_{out}(t)$ as the output trajectory.

We are also confident that our principal modeling approach can also be applied in the case of other gates like `XOR` (exclusive-or) and `AOI` (and-or-inverter) in general, since they all involve just serial transistors. In fact, it might even be possible to analyze MIS effects in Muller C-gates, which is particularly interesting given applications such as [12]. Working this out in detail is part of our current/future work, however.

IV. Analytic Solutions

In order to verify that the ODE model introduced in Section III faithfully covers all the MIS effects described in Section II, we first derive analytic expressions for the trajectories of $V_{out}^{MS}(t)$ for every mode switch MS listed in Table I. Then, we determine the following MIS delays for an arbitrary input separation time $\Delta = t_B - t_A$, which can be directly compared to our analog simulation results in Fig. 1a and Fig. 1b:

- Compute $V_{out}^{T_-^{\uparrow}}(\Delta)$, and use it as the initial value for obtaining $V_{out}^{T_+^{\uparrow\uparrow}}(t)$; the sought gate delay is the time until the latter crosses the threshold voltage $V_{DD}/2$.
- Compute $V_{out}^{T_-^{\downarrow}}(\Delta)$, and use it as the initial value for obtaining $V_{out}^{T_+^{\downarrow\downarrow}}(t)$; the sought gate delay is the time until the latter crosses the threshold voltage $V_{DD}/2$.

Fortunately, a closer look at Table I and Table II shows a symmetry between the pairs of modes $(T_-^{\uparrow}, T_+^{\uparrow})$, $(T_+^{\uparrow\uparrow}, T_-^{\uparrow\uparrow})$, $(T_-^{\downarrow}, T_+^{\downarrow})$, and $(T_+^{\downarrow\downarrow}, T_-^{\downarrow\downarrow})$. Therefore, it is sufficient to derive analytic expressions for the case $\Delta \geq 0$ only , i.e., to consider the pairs of modes $(T_-^{\uparrow}, T_+^{\uparrow\uparrow})$ resp. $(T_-^{\downarrow}, T_+^{\downarrow\downarrow})$ listed above. The corresponding formulas for $\Delta < 0$ can be obtained from those by exchanging α_1 and α_2 as well as R_{n_A} and R_{n_B}, respectively.

A. Rising input transitions

In order to compute $V_{out}^{T_-^{\uparrow}}(t)$, consider the corresponding integrals $I_1(t)$, $I_2(t)$, and $I_3(t)$, as well as $U(t)$ in Table II. Since $\beta_1 = \beta_2 = \infty$ and $\alpha_3 = \alpha_4 = 0$, we obtain

$$I_1(t) = I_3(t) = U(t) = 0, \qquad I_2(t) = \frac{t}{R_{n_A}}.$$

Since $G(t) = (I_1(t) + I_2(t) + I_3(t))/C$, we get $e^{\pm G(t)} = e^{\frac{\pm t}{CR_{n_A}}}$ and $\int_0^t e^{G(t)} U(t) = 0$. With $V_0^{\uparrow} = V_{out}^{T_-^{\uparrow}}(0)$ as our initial value, (2) finally provides

$$V_{out}^{T_-^{\uparrow}}(t) = V_{out}^{T_-^{\uparrow}}(0) e^{\frac{-t}{CR_{n_A}}}. \tag{5}$$

Similarly, for the mode $T_+^{\uparrow\uparrow}$, we obtain

$$I_1(t) = U(t) = 0, \qquad I_2(t) = \frac{t}{R_{n_A}}, \qquad I_3(t) = \frac{t}{R_{n_B}},$$

such that $e^{\pm G(t)} = e^{\pm(\frac{1}{CR_{n_A}} + \frac{1}{CR_{n_B}})t}$ and $\int_0^t e^{G(s)} U(s) ds = 0$. Consequently, we obtain

$$V_{out}^{T_+^{\uparrow\uparrow}}(t) = V_{out}^{T_-^{\uparrow}}(\Delta) e^{-(\frac{1}{CR_{n_A}} + \frac{1}{CR_{n_B}})t}, \tag{6}$$

where $V_{out}^{T_-^{\uparrow}}(\Delta)$ can be computed via (5).

Due to the symmetry mentioned before, we can immediately conclude the following result for negative Δ:

$$V_{out}^{T_-^{\uparrow\uparrow}}(t) = V_{out}^{T_+^{\uparrow}}(\Delta) e^{-(\frac{1}{CR_{n_A}} + \frac{1}{CR_{n_B}})t}, \tag{7}$$

where $V_{out}^{T_+^{\uparrow}}(\Delta) = V_{out}^{T_+^{\uparrow}}(0) e^{\frac{-|\Delta|}{CR_{n_B}}}$.

B. Falling input transitions

In this case, we first need to compute $V_{out}^{T_-^{\downarrow}}(t)$. Again plugging $\beta_1 = \beta_2 = \infty$ and $\alpha_3 = \alpha_4 = 0$ in the corresponding expressions in Table II provides

$$I_1(t) = I_2(t) = U(t) = 0, \qquad I_3(t) = \frac{t}{R_{n_B}}.$$

With $V_0^{\downarrow} = V_{out}^{T_-^{\downarrow}}(0)$ as our initial condition, (2) yields

$$V_{out}^{T_-^{\downarrow}}(t) = V_{out}^{T_-^{\downarrow}}(0) e^{\frac{-t}{CR_{n_B}}}. \tag{8}$$

Now, turning our attention to $V_{out}^{T_+^{\downarrow\downarrow}}(t)$ confronts us with a more intricate case: Whereas $I_2(t) = I_3(t) = 0$ again, evaluating $I_1(t)$ requires us to study the function

$$f(s) = \frac{1}{\frac{\alpha_1}{s+\Delta} + \frac{\alpha_2}{s} + 2R}, \tag{9}$$

as

$$I_1(t) = \int_0^t f(s)ds, \qquad G(t) = I_1(t)/C, \tag{10}$$

$$\int_0^t e^{G(s)}U(s)ds = \frac{V_{DD}}{C}\int_0^t e^{\frac{I_1(s)}{C}}f(s)ds. \tag{11}$$

Computing the above integrals is complicated by the fact that $f(s)$ also involves the parameter Δ, which prohibits uniform closed-form solutions of both (10) and (11). We hence need to determine accurate approximations for $f(s)$, which will be different for different ranges of s and Δ, which will in turn depend on the unknown parameters α_1, α_2 and R. Fortunately, a closer look at (9) reveals that piecewise approximations can be built on the basis of (1) distinguishing different ranges (cases) for Δ w.r.t. the parameters α_1, α_2 and R, and (2) splitting the integration interval $[0, t]$ into certain subintervals, which may depend on the particular case. Note that this splitting typically also involves an additional free parameter $\epsilon = \eta\Delta$ (for some $\eta \in \mathbb{R}$), which takes care of values of s close to Δ. More specifically, we came up with the following four cases and the corresponding approximations for $f(s)$ in the appropriate subintervals:

- **Case 1:** ($0 \le \Delta < \frac{\alpha_2}{2R}$)

$$f(s) \approx \begin{cases} \frac{s}{\alpha_2} & 0 \le s < \Delta - \epsilon \\ \frac{2s}{\alpha_1+2\alpha_2} & \Delta - \epsilon \le s < \Delta + \epsilon \\ \frac{s}{\alpha_1+\alpha_2} & \Delta + \epsilon \le s < \frac{\alpha_1+\alpha_2}{2R} \\ \frac{1}{2R} & \frac{\alpha_1+\alpha_2}{2R} \le s \le t \end{cases}$$

- **Case 2:** ($\frac{\alpha_2}{2R} \le \Delta < \frac{\alpha_1+2\alpha_2}{4R}$)

$$f(s) \approx \begin{cases} \frac{s}{\alpha_2} & 0 \le s < \frac{\alpha_2}{2R} \\ \frac{1}{2R} & \frac{\alpha_2}{2R} \le s < \Delta - \epsilon \\ \frac{2s}{\alpha_1+2\alpha_2} & \Delta - \epsilon \le s < \Delta + \epsilon \\ \frac{s}{\alpha_1+\alpha_2} & \Delta + \epsilon \le s < \frac{\alpha_1+\alpha_2}{2R} \\ \frac{1}{2R} & \frac{\alpha_1+\alpha_2}{2R} \le s \le t \end{cases}$$

- **Case 3:** ($\frac{\alpha_1+2\alpha_2}{4R} \le \Delta < \frac{\alpha_1+\alpha_2}{2R}$)

$$f(s) \approx \begin{cases} \frac{s}{\alpha_2} & 0 \le s < \frac{\alpha_2}{2R} \\ \frac{1}{2R} & \frac{\alpha_2}{2R} \le s < \Delta + \epsilon \\ \frac{s}{\alpha_1+\alpha_2} & \Delta + \epsilon \le s < \frac{\alpha_1+\alpha_2}{2R} \\ \frac{1}{2R} & \frac{\alpha_1+\alpha_2}{2R} \le s \le t \end{cases}$$

- **Case 4:** ($\frac{\alpha_1+\alpha_2}{2R} \le \Delta < t$)

$$f(s) \approx \begin{cases} \frac{s}{\alpha_2} & 0 \le s < \frac{\alpha_2}{2R} \\ \frac{1}{2R} & \frac{\alpha_2}{2R} \le s \le t \end{cases}$$

To explain how to determine these approximations, we elaborate on how this is done for Case 1; similar arguments can be used to justify the remaining Cases 2–4. (i) For the range $0 \le s < \Delta - \epsilon$, which expresses the situation where the integration variable s is fairly smaller than Δ, we observe $s + \Delta \approx \Delta$, and therefore, $\frac{\alpha_1}{s+\Delta} \approx \frac{\alpha_1}{\Delta}$. Assuming that the slope parameters α_1 and α_2 are approximately the same, this leads us to $\frac{\alpha_1}{\Delta} + \frac{\alpha_2}{s} \approx \frac{\alpha_2}{s}$. Furthermore, since $s < \Delta < \frac{\alpha_2}{2R}$ here, we get $\frac{\alpha_2}{s} + 2R \approx \frac{\alpha_2}{s}$ and hence $f(s) \approx \frac{s}{\alpha_2}$ as asserted. (ii) For the range $\Delta - \epsilon \le s < \Delta + \epsilon$, where s is relatively close to Δ, we can substitute Δ by s, which leads to $\frac{\alpha_1}{s+\Delta} + \frac{\alpha_2}{s} \approx \frac{2\alpha_1+\alpha_2}{2s}$. Besides, since $s \approx \Delta < \frac{\alpha_2}{2R} < \frac{\alpha_1+2\alpha_2}{4R}$, we get $\frac{2\alpha_1+\alpha_2}{2s} + 2R \approx \frac{2\alpha_1+\alpha_2}{2s}$, which justifies the asserted approximation for $f(s)$. (iii) Turning our attention to the range $\Delta + \epsilon \le s < \frac{\alpha_1+\alpha_2}{2R}$, we obtain $\frac{\alpha_1}{s+\Delta} \approx \frac{\alpha_1}{s}$ and hence $\frac{\alpha_1}{s+\Delta} + \frac{\alpha_2}{s} \approx \frac{\alpha_1+\alpha_2}{s}$. Since $s < \frac{\alpha_1+\alpha_2}{2R}$, this leads to $\frac{\alpha_1+\alpha_2}{s} + 2R \approx \frac{\alpha_1+\alpha_2}{s}$ and hence to the asserted approximation. (iv) Finally, for the remaining range $\frac{\alpha_1+\alpha_2}{2R} \le s < t$, the term $2R$ is dominant in the denominator of $f(s)$, hence $f(s) \approx \frac{1}{2R}$.

It is worth pointing out that we had to split the range $0 \le s < \Delta - \epsilon$ into two parts for Case 2 to improve the approximation accuracy, whereas we could merge some ranges for Case 4. Whereas we did not bother to determine analytic bounds for the error of the above first-order Taylor approximations, given the very small absolut values of Δ and s, it is clear that it should be very small. This is also confirmed by the model validation simulations in Section V.

The above approximations allow us to easily compute accurate approximations for (10) and (11), and thus to determine the output trajectory $V_{out}^{T\downarrow\downarrow}(t)$, for every choice of Δ. More specifically, we get $I_k(t) \approx \frac{t}{2R} + i_k$ for Case $k \in \{1, 2, 3, 4\}$, with

$$i_1 = \frac{(\Delta-\epsilon)^2}{2\alpha_2} - \frac{(\Delta+\epsilon)^2}{2(\alpha_1+\alpha_2)} + \frac{4\epsilon\Delta}{\alpha_1+2\alpha_2} - \frac{\alpha_1+\alpha_2}{8R^2}, \tag{12}$$

$$i_2 = \frac{4R(\Delta-\epsilon)-(\alpha_1+2\alpha_2)}{8R^2} - \frac{(\Delta+\epsilon)^2}{2(\alpha_1+\alpha_2)} + \frac{4\epsilon\Delta}{\alpha_1+2\alpha_2}, \tag{13}$$

$$i_3 = \frac{4R(\Delta+\epsilon)-(\alpha_1+2\alpha_2)}{8R^2} - \frac{(\Delta+\epsilon)^2}{2(\alpha_1+\alpha_2)}, \tag{14}$$

$$i_4 = -\frac{\alpha_2}{8R^2}, \tag{15}$$

which do not depend on t. Furthermore,

$$\int_0^t e^{G(s)}U(s)ds \approx \frac{V_{DD}}{C}\cdot$$

$$\begin{cases} e^{\frac{i_1}{C}}\big(\frac{\int_0^{\Delta-\epsilon} s\cdot e^{\frac{s}{2RC}}ds}{\alpha_2} + \frac{2(\int_{\Delta-\epsilon}^{\Delta+\epsilon} s\cdot e^{\frac{s}{2RC}}ds)}{\alpha_1+2\alpha_2} + \\ \frac{\int_{\Delta+\epsilon}^{\frac{\alpha_1+\alpha_2}{2R}} s\cdot e^{\frac{s}{2RC}}ds}{\alpha_1+\alpha_2} + \frac{\int_{\frac{\alpha_1+\alpha_2}{2R}}^{t} e^{\frac{s}{2RC}}ds}{2R}\big) & \text{Case 1} \\ e^{\frac{i_2}{C}}\big(\frac{\int_0^{\frac{\alpha_2}{2R}} s\cdot e^{\frac{s}{2RC}}ds}{\alpha_2} + \frac{\int_{\frac{\alpha_2}{2R}}^{\Delta-\epsilon} e^{\frac{s}{2RC}}ds}{2R} + \frac{2(\int_{\Delta-\epsilon}^{\Delta+\epsilon} s\cdot e^{\frac{s}{2RC}}ds)}{\alpha_1+2\alpha_2} + \\ \frac{\int_{\Delta+\epsilon}^{\frac{\alpha_1+\alpha_2}{2R}} s\cdot e^{\frac{s}{2RC}}ds}{\alpha_1+\alpha_2} + \frac{\int_{\frac{\alpha_1+\alpha_2}{2R}}^{t} e^{\frac{s}{2RC}}ds}{2R}\big) & \text{Case 2} \\ e^{\frac{i_3}{C}}\big(\frac{\int_0^{\frac{\alpha_2}{2R}} s\cdot e^{\frac{s}{2RC}}ds}{\alpha_2} + \frac{\int_{\frac{\alpha_2}{2R}}^{\Delta+\epsilon} e^{\frac{s}{2RC}}ds}{2R} + \frac{\int_{\Delta+\epsilon}^{\frac{\alpha_1+\alpha_2}{2R}} s\cdot e^{\frac{s}{2RC}}ds}{\alpha_1+\alpha_2} \\ + \frac{\int_{\frac{\alpha_1+\alpha_2}{2R}}^{t} e^{\frac{s}{2RC}}ds}{2R}\big) & \text{Case 3} \\ e^{\frac{i_4}{C}}\big(\frac{\int_0^{\frac{\alpha_2}{2R}} s\cdot e^{\frac{s}{2RC}}ds}{\alpha_2} + \frac{\int_{\frac{\alpha_2}{2R}}^{t} e^{\frac{s}{2RC}}ds}{2R}\big) & \text{Case 4} \end{cases}$$

$$= V_{DD}\cdot e^{\frac{i_k}{C}}\big(e^{\frac{t}{2RC}} - \gamma_k\big)$$

for Case $k \in \{1, 2, 3, 4\}$, where

$$\gamma_1 = \frac{4R^2C}{\alpha_1+\alpha_2}e^{\frac{\alpha_1+\alpha_2}{4R^2C}} - \frac{4R^2C}{\alpha_2} - \big(\frac{2R(\Delta-\epsilon)-4R^2C}{\alpha_2} - \frac{4R(\Delta-\epsilon)-8R^2C}{\alpha_1+2\alpha_2}\big)e^{\frac{\Delta-\epsilon}{2RC}} - \big(\frac{4R(\Delta+\epsilon)-8R^2C}{\alpha_1+2\alpha_2} - \frac{2R(\Delta+\epsilon)-4R^2C}{\alpha_1+\alpha_2}\big)e^{\frac{\Delta+\epsilon}{2RC}}, \tag{16}$$

$$\gamma_2 = \frac{4R^2C}{\alpha_1+\alpha_2}e^{\frac{\alpha_1+\alpha_2}{4R^2C}} - \big(1 - \frac{4R(\Delta-\epsilon)-8R^2C}{\alpha_1+2\alpha_2}\big)e^{\frac{\Delta-\epsilon}{2RC}} - \big(\frac{4R(\Delta+\epsilon)-8R^2C}{\alpha_1+2\alpha_2} - \frac{2R(\Delta+\epsilon)-4R^2C}{\alpha_1+\alpha_2}\big)e^{\frac{\Delta+\epsilon}{2RC}} + \frac{4R^2C}{\alpha_2}\big(e^{\frac{\alpha_2}{4R^2C}} - 1\big), \tag{17}$$

$$\gamma_3 = \frac{4R^2C}{\alpha_1+\alpha_2} e^{\frac{\alpha_1+\alpha_2}{4R^2C}} - \big(1 - \frac{2R(\Delta+\epsilon) - 4R^2C}{\alpha_1+\alpha_2}\big) e^{\frac{\Delta+\epsilon}{2RC}} + \frac{4R^2C}{\alpha_2}\big(e^{\frac{\alpha_2}{4R^2C}} - 1\big), \tag{18}$$

$$\gamma_4 = \frac{4R^2C}{\alpha_2}\big(e^{\frac{\alpha_2}{4R^2C}} - 1\big). \tag{19}$$

Combining the above solutions for (10) and (11) according to (2) provides an accurate expression for the output trajectory $V_{out}^{T_+^{\downarrow\downarrow}}(t)$, namely,

$$V_{out}^{T_+^{\downarrow\downarrow}}(t) \approx V_{out}^{T_-^{\downarrow}}(\Delta) e^{\frac{-(2i_k R+t)}{2RC}} + V_{DD}(1-\gamma_k e^{-\frac{t}{2RC}})$$

for Case $k \in \{1,2,3,4\}$, where $V_{out}^{T_-^{\downarrow}}(\Delta) = V_{out}^{T_-^{\downarrow}}(0) e^{\frac{-\Delta}{CR_{n_B}}}$, $i_1, \ldots, i_4$, $\gamma_1, \ldots, \gamma_4$, are defined in (8), (12), ..., (19), respectively. In addition, our symmetry immediately provides us with a trajectory corresponding to the case of negative Δ:

$$V_{out}^{T_-^{\downarrow\downarrow}}(t) \approx V_{out}^{T_+^{\downarrow}}(\Delta) e^{\frac{-(2i'_k R+t)}{2RC}} + V_{DD}(1-\gamma'_k e^{-\frac{t}{2RC}}),$$

for Case $k \in \{1,2,3,4\}$, where $V_{out}^{T_+^{\downarrow}}(\Delta) = V_{out}^{T_+^{\downarrow}}(0) e^{-\frac{|\Delta|}{CR_{n_A}}}$ and $i'_1, \ldots, i'_4$, $\gamma'_1, \ldots, \gamma'_4$ are obtained by substituting α_1 by α_2 and Δ by $|\Delta|$ in (12), ..., (19), respectively.

C. Single Input Switching

Whereas the main focus of our model is the proper modeling of MIS effects, it of course also needs to handle the "simple" single input switching case. This is done exactly as in the IDM [2], by continuously switching between the trajectories of the modes $(0,0)$ and $(1,0)$ (for input A) resp. $(0,0)$ and $(0,1)$ (for input B). Fortunately, we can adapt the above trajectory formulas for the MIS cases to also obtain the ones for the SIS cases. In fact, in the SIS cases, Δ just represents the time difference between the current and the previous transition of the *same* input. For switching from $(1,0)$ and $(0,0)$, for example, it suffices to replace the initial value $V_{out}^{T_-^{\downarrow}}(\Delta)$ in $V_{out}^{T_+^{\downarrow\downarrow}}(t)$ by the initial value $V_{out}^{T_-^{\uparrow}}(\Delta)$, which gives the output voltage at the time the previous switch from $(0,0)$ to $(1,0)$.

V. Parameterization and Modeling MIS Effects

Since the ultimate goal of our hybrid model is to develop a basis for dynamic digital timing simulations, our main goal are explicit analytic formulas for the input-to-output delay functions $\delta_M^{\downarrow}(\Delta)$ and $\delta_M^{\uparrow}(\Delta)$ for both rising and falling output transitions. Moreover, we have to answer the question how to determine the parameters α_1, α_2, C, R, R_{n_A}, R_{n_B}, and η for a given technology. Such a parametrization is of course necessary for checking whether and how well our new model is capable of faithfully reproducing the MIS effects described in Section II.

To accomplish the first task, by inverting the explicit formulas obtained for the trajectories $V_{out}^{T_+^{\uparrow\uparrow}}(t)$ and $V_{out}^{T_-^{\uparrow\uparrow}}(t)$ resp. for $V_{out}^{T_+^{\downarrow\downarrow}}(t)$ and $V_{out}^{T_-^{\downarrow\downarrow}}(t)$ obtained in the previous section, we obtain the exact resp. approximate analytic expressions for $\delta_{M,+}^{\downarrow}(\Delta)$ (for $\Delta \geq 0$), $\delta_{M,-}^{\downarrow}(\Delta)$ (for $\Delta < 0$) and $\delta_{M,+}^{\uparrow}(\Delta)$ (for $\Delta \geq 0$), $\delta_{M,-}^{\uparrow}(\Delta)$ (for $\Delta < 0$) in terms of the model's parameters given in Theorem 1.

Theorem 1 (MIS Delay functions)**.** *For any $-\infty \leq \Delta \leq \infty$, the MIS delay functions for falling and rising output transitions of our model are given by:*

$$\delta_{M,+}^{\downarrow}(\Delta) = \begin{cases} -\frac{\ln(0.5)CR_{n_A}R_{n_B}+\Delta R_{n_B}}{R_{n_A}+R_{n_B}} + \Delta & 0 \leq \Delta < -\ln(0.5)CR_{n_A} \\ -\ln(0.5)CR_{n_A} & \Delta \geq -\ln(0.5)CR_{n_A} \end{cases}$$

$$\delta_{M,-}^{\downarrow}(\Delta) = \begin{cases} -\frac{\ln(0.5)CR_{n_A}R_{n_B}+|\Delta|R_{n_A}}{R_{n_A}+R_{n_B}} + |\Delta| & |\Delta| < -\ln(0.5)CR_{n_B} \\ -\ln(0.5)CR_{n_B} & |\Delta| \geq -\ln(0.5)CR_{n_B} \end{cases}$$

and, for Case $k \in \{1,2,3,4\}$,

$$\delta_{M,+}^{\uparrow}(\Delta) \approx 2RC\big(\ln(\gamma_k) - \ln(0.5)\big),$$
$$\delta_{M,-}^{\uparrow}(\Delta) \approx 2RC\big(\ln(\gamma'_k) - \ln(0.5)\big).$$

Proof. We sketch how $\delta_{M,+}^{\downarrow}(\Delta)$ is computed; the expression for $\delta_{M,-}^{\downarrow}(\Delta)$ is obtained analogously. Recall that falling output transitions imply rising input transitions and vice versa. Given the trajectory $V_{out}^{T_+^{\uparrow\uparrow}}(t)$ in (6), we start out from $V_{out}^{T_-^{\uparrow}}(0) = V_{DD}$ and need to compute the time $\delta_{M,+}^{\downarrow}(\Delta)$ when either (i) already the preceding trajectory $V_{out}^{T_-^{\uparrow}}(t)$ or else (ii) $V_{out}^{T_+^{\uparrow\uparrow}}(t)$ itself (which is started at time Δ) hits $V_{DD}/2$. Note that this reflects the fact that already the first rising input (at time 0) causes the output to switch to 0. Since all these trajectories only involve a single exponential, they are easy to invert. It turns out that case (i) occurs for values $\Delta \geq -\ln(0.5)CR_{n_A}$, case (ii) for smaller Δ.

Similarly, for computing $\delta_{M,+}^{\uparrow}(\Delta)$, we need to consider the trajectory $V_{out}^{T_+^{\downarrow\downarrow}}(t)$. Here, we have to start out from the initial value $V_{out}^{T_-^{\downarrow}}(\Delta) = 0$ and just need to compute the time $\delta_{M,+}^{\uparrow}(\Delta)$ until $V_{out}^{T_+^{\downarrow\downarrow}}(t)$ hits $V_{DD}/2$. This reflects the fact that it is only the second falling input (at time Δ) that causes the output to switch to 1. □

Whereas the specific initial values $V_{out}^{T_-^{\uparrow}}(0) = V_{DD}$ and $V_{out}^{T_-^{\downarrow}}(0) = 0$ used in Theorem 1 are sufficient for evaluating the MIS behavior of our model, we need generalized expressions for the dynamic digital timing simulation experiments in Section VI. In Theorem 2, we therefore provide the delay functions for arbitrary initial values.

Theorem 2. *[MIS and SIS delay functions for arbitray initial values] For any $-\infty \leq \Delta \leq \infty$, the extended delay functions for falling and rising output transitions of our model, when starting from a given initial value $V_{out}(0)$, are*

$$\delta_{EM,+}^{\downarrow}(\Delta) = \begin{cases} -\frac{\ell CR_{n_A}R_{n_B}+\Delta R_{n_B}}{R_{n_A}+R_{n_B}} + \Delta & 0 \leq \Delta < -\ell CR_{n_A} \\ -\ell CR_{n_A} & \Delta \geq -\ell CR_{n_A} \end{cases}$$

$$\delta_{EM,+}^{\uparrow}(\Delta) \approx 2RC\big(\ln(\frac{2V_{DD}\gamma_k - 2V_{out}(0)e^{-\frac{\Delta}{CR_{n_B}}}e^{-\frac{i_k}{C}}}{V_{DD}})\big),$$

where $\ell = \ln\big(V_{DD}/2V_{out}(0)\big)$ and Case $k \in \{1,2,3,4\}$. $\delta_{EM,-}^{\downarrow}(\Delta)$ and $\delta_{EM,-}^{\uparrow}(\Delta)$ can be easily obtained by our symmetry.

The delay functions given in Theorem 1 (that is, in Theorem 2) not only facilitate fast dynamic timing analysis, but are also instrumental for understanding which parameters affect the which delay value. In fact, the formulas could even be used for explicit parametrization of a given circuit when certain delay values are known.

A. Parameterization for 15 nm

Like in [13], in this subsection, we will fit our model to the characteristic MIS delay values $\delta_S^{\downarrow}(-\infty)$, $\delta_S^{\downarrow}(0)$, $\delta_S^{\downarrow}(\infty)$ according to Fig. 1a and the corresponding values $\delta_S^{\uparrow}(-\infty)$, $\delta_S^{\uparrow}(0)$, $\delta_S^{\uparrow}(\infty)$ in Fig. 1b in Section II.

Interestingly, our first attempt to simultaneously fit all six delay values for determining all parameters at once turned out to be naive since impossible. To understand why this is the case, note that the on-resistors of the two nMOS transistors R_{n_A} and R_{n_B} should roughly be the same. Consequently, we obtain

$$\frac{\delta_{M,+}^{\downarrow}(\infty)}{\delta_{M,+}^{\downarrow}(0)} = \frac{R_{n_A} + R_{n_B}}{R_{n_B}} \approx 2.$$

Unfortunately, however, the desired ratio is $\frac{\delta_S^{\downarrow}(-\infty)}{\delta_S^{\downarrow}(0)} \approx \frac{38\,\text{ps}}{28\,\text{ps}}$, which cannot fit these two values with reasonable choices for R_{n_A} and R_{n_B}. As in [13], we fixed this problem by adding (that is, subtracting) a suitably chosen pure delay $\delta_{\min} = 18\,\text{ps}$ (foreseen in the original IDM [2]), which just defers the switching to the new state upon an input transition. This results in an effective ratio of $\frac{20\,\text{ps}}{10\,\text{ps}} = 2$, which could finally be matched by least squares fitting. Of course, when using our model in digital timing analysis, $\delta_{\min}$ must be added to the computed delay values.

More specifically, starting out from some desired load capacitance C,[2] we first determined R_{n_A} and R_{n_B} by fitting $\delta_{M,-}^{\downarrow}(-\infty)$ (or $\delta_{M,+}^{\downarrow}(\infty)$) and $\delta_{M,+}^{\downarrow}(0)$ to match $\delta_S^{\downarrow}(-\infty) - \delta_{\min}$ (or $\delta_S^{\downarrow}(\infty) - \delta_{\min}$) and $\delta_S^{\downarrow}(0) - \delta_{\min}$. Once these parameter values had been obtained, we fixed those and determined the remaining parameters R, α_1, α_2 and η by fitting $\delta_{M,-}^{\uparrow}(-\infty)$, $\delta_{M,+}^{\uparrow}(0)$ and $\delta_{M,+}^{\uparrow}(\infty)$ to match $\delta_S^{\uparrow}(-\infty) - \delta_{\min}$, $\delta_S^{\uparrow}(0) - \delta_{\min}$ and $\delta_S^{\uparrow}(\infty) - \delta_{\min}$. Note that the same $\delta_{\min} = 18\,\text{ps}$ is used for both rising and falling output transitions. The result of our parametrization for the circuit in Section II is shown in Table III.

TABLE III: Model parameter values for the 15 nm CMOS NOR gate used for producing Fig. 1a and Fig. 1b.

Parameters found by fitting the falling output transition case		
$R_{n_A} = 8.360562682200\ k\Omega$	$R_{n_B} = 8.255562682200\ k\Omega$	$C = 3.6331599443276\ fF$

Parameters found by fitting the rising output transition case			
$R = 6.6999626822002\ k\Omega$	$\alpha_1 = 0.859 \cdot 10^{-7}\ \Omega s$	$\alpha_2 = 0.268 \cdot 10^{-7}\ \Omega s$	$\eta = 0.01$

Utilizing the parameters in Table III, we can finally visualize the delay predictions of our model. Fig. 4 shows the very good fit of $\delta_M^{\downarrow}(\Delta)$ for a falling output transition compared to the analog simulation result presented in Fig. 1a. Similarly, Fig. 5 also shows a good coverage of the MIS effect for rising output transitions ($\delta_{M,\pm}^{\uparrow}(\Delta)$) compared to the actual delay observed in Fig. 1b. We therefore conclude that our new hybrid model fully captures all the MIS effects introduced in Section II, including the case of $\Delta < 0$ for rising output transitions where the model proposed in [13] fails.

As a final remark, we note that the modeling inaccuracies visible in Fig. 5 are primarily caused by the errors introduced by our approximate solutions of (10) and (11). The first one is the lack of perfectly fitting the actual delay with the computed one for $\Delta = 0$, which looks quite bad in the figure but actually causes a relative error of about 0.95 % only. The other two are the small dent shapes around $\Delta \approx \pm 7.64$ ps, which are caused by the approximation error at the ending boundary of the range in Case 3. It could easily be circumvented by slightly moving the border between Case 3 and Case 4 to a smaller value, i.e., to $\frac{\alpha_1+\alpha_2}{2R} - \varepsilon$ for some $\varepsilon > 0$. Since the induced error is marginal, we did not bother with further complicating our analysis.

[2]Note that we observed that our model scales well with C, i.e., a parametrization starting out from a different C yields different parameters but essentially the same delays.

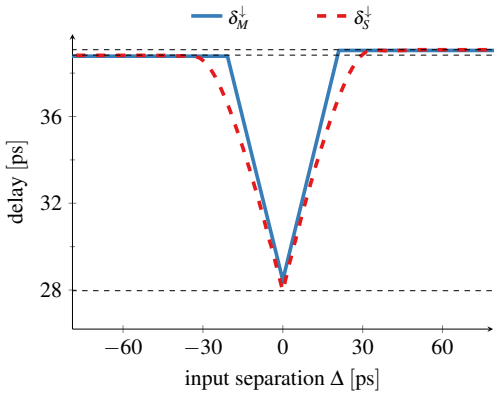


Fig. 4: Computed ($\delta_M^{\downarrow}(\Delta)$) and measured ($\delta_S^{\downarrow}(\Delta)$) MIS delays for falling output transitions.

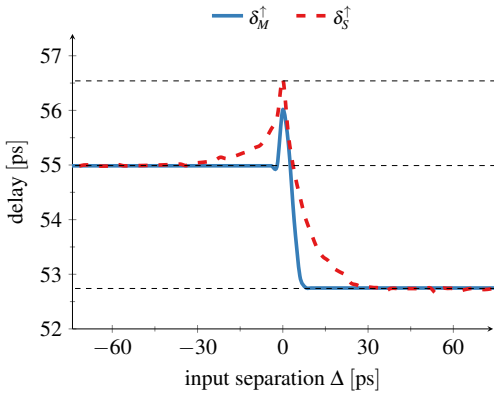
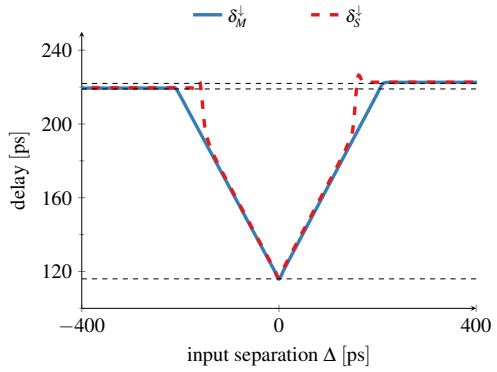


Fig. 5: Computed ($\delta_M^{\uparrow}(\Delta)$) and measured ($\delta_S^{\uparrow}(\Delta)$) MIS delays for rising output transitions.

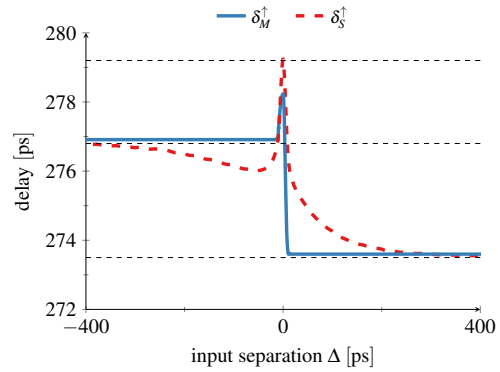
B. Other parameterizations

A crucial feature of any model is wide applicability. Ideally, our hybrid delay model should be applicable to any CMOS technology, for any supply voltage, temperature, age etc., in the sense that it is possible to determine a parametrization that allows our model to match the MIS delays of any given CMOS NOR implementation. Overall, it is reasonable to conjecture that our model is applicable whenever the Shichman-Hodges transistor model [14], i.e., (3) and (4), reasonably applies. Whereas it is of course impossible for us to prove such a claim, we can demonstrate that this is the case for some quite different technology and operation conditions, namely the UMC 65 nm technology with $V_{DD} = 1.2\,\text{V}$ supply voltage and a larger load capacitance C.[3] The red dashed curves in Fig. 6a (falling output) resp. Fig. 6b (rising output), which correspond to Fig. 4 resp. Fig. 5, show the gate delays depending on the input separation time Δ obtained via *SPICE* simulations.

[3]We note that we played with several different conditions and configurations, which all confirmed that the parameters of our model can be easily matched.



(a) Falling output delay



(b) Rising output delay

Fig. 6: Computed ($\delta_M^{\uparrow}(\Delta)$) and measured ($\delta_S^{\uparrow}(\Delta)$) MIS delays for 65 nm technology.

Parametrizing our model for these 65 nm MIS delays turned out to be remarkable easy: Exactly as for our 15 nm technology, by initially fixing some value for the load capacitance C and trying to match $\delta_{M,-}^{\downarrow}(-\infty)$ (or $\delta_{M,+}^{\downarrow}(\infty)$) and $\delta_{M,+}^{\downarrow}(0)$ with $\delta_S^{\downarrow}(-\infty) - \delta_{\min}$ (or $\delta_S^{\downarrow}(\infty) - \delta_{\min}$) and $\delta_S^{\downarrow}(0) - \delta_{\min}$, we determined accurate values for R_{n_A} and R_{n_B}. After fixing the latter, we obtained the values of the remaining parameters by fitting the remaining delay values. Note carefully, however, that we had to use a different pure delay $\delta_{\min} = 10.8\,\text{ps}$ here, since $\frac{\delta_S^{\downarrow}(-\infty)}{\delta_S^{\downarrow}(0)} \approx \frac{222\,\text{ps}}{116\,\text{ps}}$. Table IV provides the resulting list of parameters. Not surprisingly, since the gate delay is considerably higher than for our 15 nm data, we indeed face a significantly larger value for the load capacitance C also in our model.

TABLE IV: Model parameter values for the 65 nm CMOS NOR gate used for producing the red curves in Fig. 6.

Parameters found by fitting the falling output transition case		
$R_{n_A} = 8.409562682200\ k\Omega$	$R_{n_B} = 8.285562682200\ k\Omega$	$C = 30.6331599443276\ fF$

Parameters found by fitting the rising output transition case			
$R = 5.1916426822002\ k\Omega$	$\alpha_1 = 0.959 \cdot 10^{-7}\ \Omega s$	$\alpha_2 = 0.273 \cdot 10^{-7}\ \Omega s$	$\eta = 0.01$

Utilizing the parameters in Table IV, we finally obtained the blue curves in Fig. 6a and Fig. 6b, which illustrate the computed delays of our model for the 65 nm technology. It is apparent that they match the real delays very well: We see an ideal fitting for the case of falling output transition as well as for the case of rising output transitions at marginal Δ values. Moreover, considerably promising is the negligible absolute error value of 0.3% for the mismatch associated with $\Delta = 0$.

VI. Modeling Accuracy Experiments

In this section, we experimentally compare the modeling accuracy of our new model to the ideal switch hybrid model [13], the IDM and to classic inertial delays, using the publicly available Involution Tool [5]. Albeit it performs dynamic digital timing simulation in VHDL, it also supports delay models implemented in Python. We hence implemented our model, that is, the delay functions given in Theorem 2, in Python.

For our experimental evaluation, we used the same setup as in [13], namely the 15 nm Nangate Open Cell Library featuring FreePDK15™ FinFET models [15] ($V_{DD} = 0.8\,\text{V}$) that has also been used in Section II. Based on a Verilog description of our NOR gate, we performed optimization, placement and routing by utilizing the Cadence tools Genus and Innovus (version 19.11). We also extracted the parasitic networks from the final layout to obtain *SPICE* models. These models allowed us to perform simulations with Spectre (version 19.1), which are used to produce golden reference digital signal traces, by recording their $V_{DD}/2$ crossing times. All simulations in our delay model used the final parameters from Table III. For the ideal switch model and the IDM Exp-channel, the parameters reported in [13] were used.

In order to quantify and compare the typical (average) modeling accuracy of our competing delay models, we stimulated our NOR circuit with randomly generated waveforms. Our experiments targeted both fast (100/50) and slow (200/100) pulse trains on every input, with (LOCAL) and without (GLOBAL) concurrent transitions. For example, the configuration *100/50 - LOCAL* of the Involution Tool generates transitions independently on both input A and B, according to a normal distribution with $\mu = 100\,\text{ps}$ and $\sigma = 50\,\text{ps}$. The configuration *200/100 - GLOBAL* generates transitions either on input A or on B, with $\mu = 200\,\text{ps}$ and $\sigma = 100\,\text{ps}$. Each simulation run consisted of 500 transitions and has been repeated 200 times.

We note that the simulation times (all in the few second-range) for all our models turned out to be similar. More specifically, the inertial delay model (which is natively implemented in ModelSim) is around 33% faster than both the IDM model and the two hybrid models in all our configurations. Simulation times in *SPICE* match the ones of ModelSim in configurations like *100/50 - LOCAL*, but are as slow as our hybrid models in *5000/5 - GLOBAL*, for example. Since the Python implementations of our models are of course not at all optimized for simulation performance, we assume that fully engineered implementations would be substantially faster.

As our modeling accurracy comparison metric, we used the area under the deviation trace, which is the absolute value of the difference between the *SPICE* trace and the trace generated by the respective delay model integrated over time. However, since the absolute values largely depend on the number of transitions, and are therefore meaningless per se, we normalized them with respect to the inertial delays, which is our baseline model. Consequently, lower bars indicate less area under the deviation trace and hence better results. Fig. 7 shows the results of our experiments. It is apparent that our model outperforms the ideal switch hybrid model [13] in the case of fast pulse trains, where the increased modelling accuracy of the falling output MIS delay comes into effect: Around 5% is gained in the case of *100/50 - LOCAL*.

VII. Conclusions

We provided a novel hybrid ODE model for a NOR gate, which faithfully covers all MIS effects. As a first-order model of dimension 1, it is even simpler than the immediate switch hybrid ODE model proposed in [13], albeit it does not share its failure to model the MIS

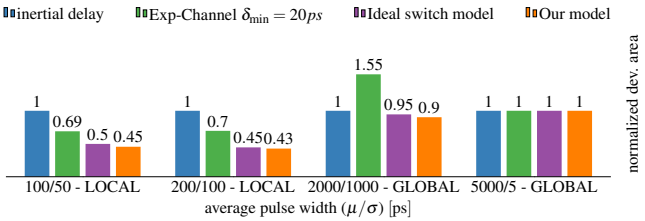


Fig. 7: Accuracy of inertial delay, Exp-Channel, ideal switch model and our new model, compared to *SPICE* simulations of a NOR gate.

effect for rising output transitions. Thanks to its simplicity, it allows to compute accurate approximation formulas for the gate delays, which makes the model readily applicable in dynamic digital timing analysis and facilitates easy model parametrization. An experimental comparison of the modeling accuracy against alternative approaches confirmed its superior performance. Part of our current/future work is devoted to applying our modeling approach to other (more complex) gates, like Muller C-gates and XOR gates.

REFERENCES

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

[2] 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.

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

[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] V. Chandramouli and K. A. Sakallah, "Modeling the effects of temporal proximity of input transitions on gate propagation delay and transition time," in *Proc. DAC'96*, 1996, p. 617–622. [Online]. Available: https://doi.org/10.1145/240518.240635

[10] 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.–.

[11] 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.

[12] J. C. Ebergen, S. Fairbanks, and I. E. Sutherland, "Predicting performance of micropipelines using charlie diagrams," in *Proceedings Fourth International Symposium on Advanced Research in Asynchronous Circuits and Systems (ASYNC'98)*, 1998, pp. 238–246.

[13] A. Ferdowsi, J. Maier, D. Öhlinger, and U. Schmid, "A simple hybrid model for accurate delay modeling of a multi-inpu t gate," in *Proceedings of the 2022 Design, Automation & Test in Europe Confer ence & Exhibition*, ser. DATE '22, 2022.

[14] H. Shichman and D. A. Hodges, "Modeling and simulation of insulated-gate field-effect transistor switching circuits," *IEEE Journal of Solid-State Circuits*, vol. 3, no. 3, pp. 285–289, Sep. 1968.

[15] 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