

Body–Voltage Estimation in Digital PD-SOI Circuits and Its Application to Static Timing Analysis

Kenneth L. Shepard, *Member, IEEE*, and Dae-Jin Kim, *Student Member, IEEE*

Abstract—Partially depleted silicon-on-insulator (PD-SOI) has emerged as a technology of choice for high-performance low-power deep-submicrometer digital integrated circuits. An important challenge to the successful use of this technology involves successfully managing and predicting the large “uncertainties” in the body potential and consequently the threshold voltages that can result from unknown past switching activity. In this paper, we present a unique state-diagram abstraction of the PD-SOI field-effect transistor that can capture all of the past switching activity determining the body voltage. Based on this picture, four different estimation schemes are discussed that increasingly bound floating body uncertainty based on more detailed knowledge of switching activity. Using these estimation techniques within a prototype transistor-level static timing analysis engine, we demonstrate both the accuracy of the estimation and the reduction in delay uncertainty possible with these techniques.

Index Terms—Silicon-on-insulator, static timing analysis.

I. INTRODUCTION

SILICON-ON-INSULATOR (SOI) technology has long found niche applications for radiation-hardened or high-voltage integrated circuits. Recently, however, SOI has emerged as a technology for high-performance low-power deep-submicrometer digital integrated circuits [1]–[4]. For digital applications, fully depleted devices have been largely abandoned in favor of partially depleted technology because of the difficulty in controlling the threshold voltage of fully depleted thin-film transistors. Partially depleted SOI (PD-SOI) has two main advantages for digital applications: the reduction of parasitic source-drain depletion capacitances and the reduction of the body effect in stack structures and pass-transistor logic.

At the device and circuit level, however, the floating body effect in PD-SOI poses major challenges in the successful use of this technology. There is a parasitic bipolar effect that can result in noise failures if not correctly considered [5]. In addition, there can be large “uncertainties” in the body potential and, consequently, the threshold voltage of devices due to unknown past switching activity. For many circuits, the design margining required to protect against this uncertainty erodes all of the potential performance advantage under nominal operation. In addition, for many circuit styles in which noise margin is strongly

determined by threshold voltage (e.g., dynamic circuits), considerable overdesign for noise can also result from conservative body–voltage margining. An array of circuit design techniques already exist to attempt to contain the noise impact of the parasitic bipolar current and the delay and noise-margin variation due to the floating body [3], [4]. Some of these are quite counter to design practice in bulk; for example, precharging internal nodes of the nFET pulldown stack in domino logic to avoid parasitic bipolar currents. Previous circuit-level modeling work on PD-SOI has focused on device issues [6]–[10] or delay and noise effects due to the floating-body effect evident for particular circuits under periodic stimulus [11]–[13] (pulse stretching, frequency-dependent delay time).

In this paper, we present the first techniques to quantify floating-body effects over tens of millions of transistors through static analysis [14]. In our approach, we analyze each field-effect transistor (FET) of each circuit, determining the minimum and maximum possible body voltage (the body–voltage uncertainty) that could be achieved based on different “static” characterizations of past switching activity. These values are then used as “initial conditions” for the constituent simulations of channel-connected components (CCC) that are used in transistor-level static timing analysis [15]–[17]. The same techniques can be applied to static noise analysis [18], but this will not be considered in this paper.

We work with BSIM3SOI [19] models for an IBM PD-SOI technology described elsewhere [20]. Devices have a 0.25- μm effective channel length, 5-nm gate oxide, 350-nm back oxide, and 140-nm thin silicon film. Two supply voltages are considered—1.0 V and 2.5 V. The former might be used in low-power applications. While the detailed results we present here apply to this technology, the techniques are generally applicable to any PD-SOI technology.

In Section II, we describe a state-diagram model that can be used to abstract all of the past switching history of a PD-SOI FET. We then describe a simplified device physics that can be used to accurately predict the body voltages in PD-SOI circuits under various switching histories as abstracted in the state-diagram model. Section III describes a prototype transistor-level static timing analysis engine that incorporates these body voltage characterizations. Some results with example circuits are presented in Section IV. Section V concludes and offers direction for future work.

II. PD-SOI DEVICE PHYSICS

A. State Diagram and Concept of Reference State

The body potential of a PD-SOI FET is determined by capacitive coupling of the body to the gate, source, and drain by

Manuscript received July 23, 1999; revised January 3, 2001. This work was supported in part by the National Science Foundation under Grant CCR-97-34216 and in part by the IBM Corporation. This paper was recommended by Associate Editor Z. Yu.

The authors are with the Columbia Integrated Systems Laboratory, Department of Electrical Engineering, Columbia University, New York, NY 10027 USA.

Publisher Item Identifier S 0278-0070(01)04828-X.

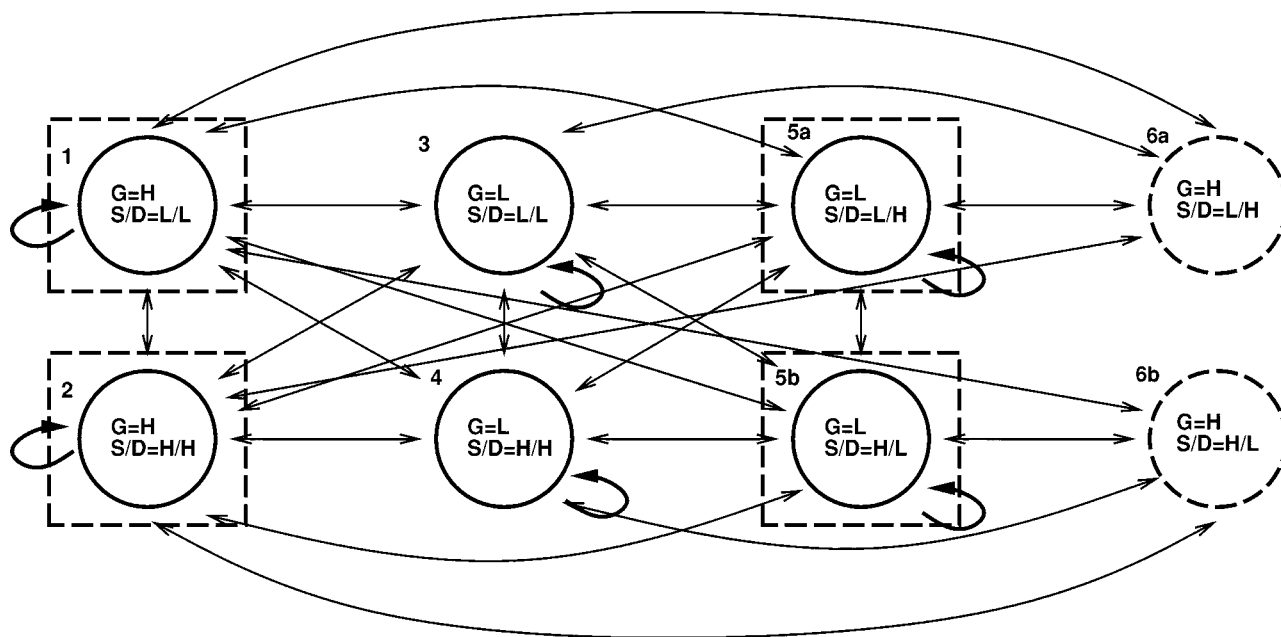


Fig. 1. State diagram for a PD-SOI nFET.

diode currents at the source–body and drain–body junctions [including gate-induced drain leakage (GIDL) [21]] and by impact ionization currents produced by current flow through the device (sometimes referred to as the on-state impact ionization current).¹ The impact ionization currents have a strong supply-voltage dependence, decreasing with decreasing supply voltages and becoming quite negligible for voltages below the silicon bandgap (1.1 V). Moreover, it is convenient to distinguish “fast” and “slow” processes. Fast processes can change the body potential on time scales on the order of or less than the cycle time, while slow processes require time scales much longer than the cycle time (up to milliseconds) to affect the body voltage. There are two fast mechanisms at work: switching transitions on the gate, source, or drain that are capacitively coupled to the body (which we call *coupling displacements*) and forward-bias diode currents across source–body and drain–body junctions with voltages exceeding the diode turn-on voltage (which we call *body discharge*). The slow processes involve charging or discharging the body through reverse-biased or very weakly forward-biased diode junctions and through impact ionization.

As a (usually) dynamic circuit node, the floating body has “memory.” To model the switching history determining the body voltage of a particular device, we use the state diagram abstraction shown in Fig. 1. (This diagram applies to the nFET. The state diagram of the pFET is the “dual” of this in which the gate is high rather than low in states 3–5 and low rather than high in states 1, 2, and 6.) The states denoted with solid circles represent “static” states, states in which the FET can be stable, in contrast with the “dynamic” states 6a and 6b, which are only present transiently during switching events. For example, state 1 corresponds to the case in which the gate is high and both the source and drain are low. Arrows indicate possible state transitions produced by switching events in the circuits

containing these FETs. These switching events can represent transitions from the logic state at the end of the previous cycle to the logic state at the end of the current cycle or can represent hazards that occur transiently within a cycle. States 5a and 5b can usually be treated equivalently as state 5; similarly, states 6a and 6b can usually be treated equivalently as state 6.

If the device is allowed to remain in one state for a very long time, the body voltage in each state will achieve a direct current (dc) value, denoted as s_i . The dc voltages in states 1 and 3 (s_1 and s_3) are zero, while the dc voltages in stages 2 and 4 (s_2 and s_4) are given by the supply voltage. s_5 is determined by the steady-state balance between a weakly forward-biased junction drawing current from the body and a reverse-biased junction, leaking current to the body, enhanced by GIDL currents. Similarly, s_6 is determined by the steady-state balance between a weakly forward-biased junction drawing current from the body and charging current due to reverse leakage of the other diode junction and on-state impact ionization. These values of s_i are shown for our example technology in Table I at both characterized supply voltages.

In the absence of body discharge, the coupling displacements that occur with each transition in Fig. 1 are completely reversible on “fast” time scales; that is, if one begins in state 1 and traverses the state diagram, returning to state 1 on a time scale much faster than any of the “slow” leakage mechanisms, the body voltage on return will be the same as the initial body voltage, a simple result of charge conservation. Because of this, we can represent the charge stored on the body as the value of the body voltage in one particular state of Fig. 1, the *reference state*, which we choose to be state 2 for the nFET and state 1 for the pFET. From this reference body voltage (V_B^{ref}), we can then determine the corresponding body voltage in each state i (V_B^i) according to

¹Gate leakage is also an emerging factor in determining the body voltage that we do not consider in this paper.

$$V_B^i = V_B^{\text{ref}} + d_i(V_B^{\text{ref}}).$$

TABLE I
VALUES OF s_i , V_i^{zero} , AND $V_i^{forward}$ FOR THE nFET AND pFET OF OUR EXAMPLE TECHNOLOGY

i	nfet			pfet			nfet			pfet		
	s_i	V_i^{zero}	$V_i^{forward}$	s_i	V_i^{zero}	$V_i^{forward}$	s_i	V_i^{zero}	$V_i^{forward}$	s_i	V_i^{zero}	$V_i^{forward}$
1	0	1.77	2.41	0.0	0.0	-0.6	0	0.53	1.13	0	0	-0.6
2	2.5	2.50	3.1	2.5	0.73	0.12	1.0	1.0	1.6	1.0	0.469	-0.127
3	0	2.15	2.85	0	-0.92	-0.98	0	0.87	1.5	0	-0.75	-0.907
4	2.5	3.43	3.5	2.5	0.35	-0.30	1.0	1.75	1.906	1.0	0.14	-0.5
5	0.437	2.58	2.79	2.068	-0.045	-0.23	0.278	1.12	1.47	0.722	-0.118	-0.47
6	0.711	1.60	1.60	1.976	1.24	1.24	0.357	0.61	0.85	0.722	0.467	0.17

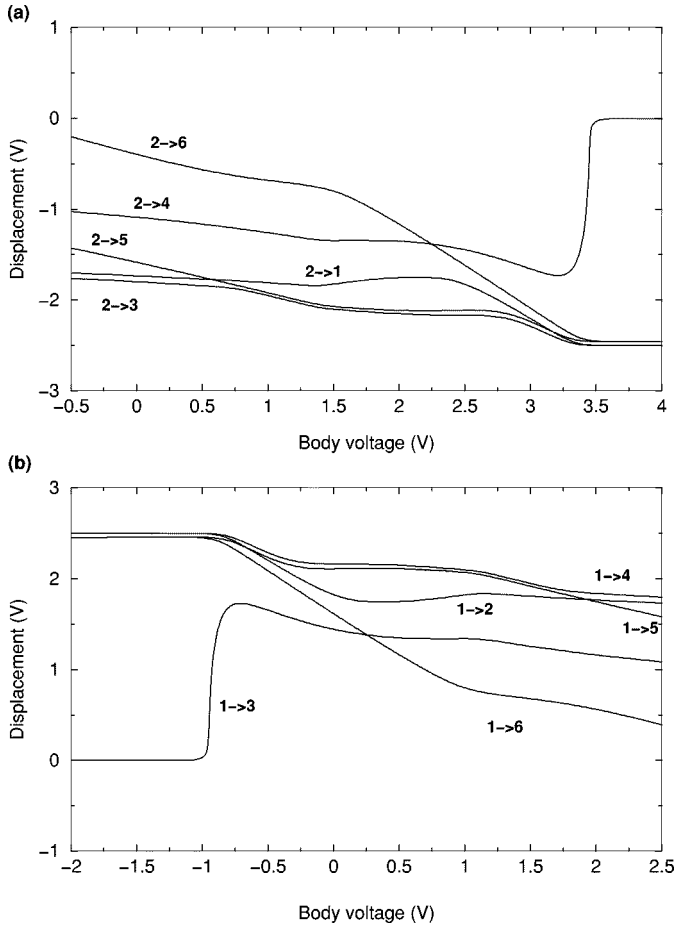


Fig. 2. Displacements as a function of reference state body voltage at a 2.5-V supply for both the (a) nFET for which the reference state is state 2 and (b) the pFET for which the reference state is state 1.

The displacements $d_i(V_B^{ref})$ are explicitly shown to be dependent on the reference body voltage because of the strong voltage-dependence of the capacitances of the source–body, drain–body, and gate–body. d_i is independent of device width W because the dominant components of capacitance scale proportionately with W . Fig. 2 shows these displacements as a function of V_B^{ref} for our example technology at the 2.5-V supply. In many cases, these values are not easily determined with direct simulation of a transition from the reference state. For example, reference body voltages for the nFET greater than about 3.1 V would result in strongly forward-biased body–drain and body–source junctions making it difficult to distinguish the displacement due to the transition from state 2 from the body discharge. As a result, these curves are instead determined by

isolated evaluation of the metal–oxide–semiconductor (MOS) capacitance–voltage (C – V) model.

B. $V_i^{forward}$ and V_i^{zero} for Each State

With the reference body voltage as a “state-independent” way of representing the charge trapped on the body, we proceed to characterize each state i in Fig. 1 by two values of this reference voltage V_i^{zero} and $V_i^{forward}$, shown in Table I for our example technology. V_i^{zero} represents the steady-state value of the reference body voltage achieved by remaining in state i for a long time. This follows immediately from the s_i in each state

$$s_i = V_i^{zero} + d_i(V_i^{zero}). \quad (1)$$

$V_i^{forward}$ represents the value of the reference body voltage for the nFET (pFET) to which the body would be very quickly pulled down (up) as a result of body discharge (charge) if state i were accessed with a higher (lower) reference body voltage than $V_i^{forward}$. The values shown in Table I for our example technology presume that the fast body discharge will bring the forward-biased-junction bias down to a turn-on voltage of 0.6 V. (It is important to note that fast body discharge can trigger parasitic bipolar leakage between source and drain for FETs in state 5.) This means, for example, that if a FET that reached a dc steady-state in state 4 (with a V_B^{ref} of 3.3 V) switches into state 2, the reference body voltage will quickly discharge to $V_2^{forward} = 3.1$ V. If the FET subsequently remains in state 2 for a long time, V_B^{ref} will eventually decrease to $V_2^{zero} = 2.5$ V.

The most important qualitative difference in the $V_i^{forward}$ and V_i^{zero} values between the 2.5 V and 1.0 V cases is the 6 state. At a supply voltage of 2.5 V, $V_6^{forward} = V_6^{zero}$ (which is not the case at the 1.0-V supply) because of the dominant effect of the on-state impact ionization current. This current is so large that a strongly forward-biased junction is required to balance it in steady-state.

C. “Simple” Uncertainty Modes—Full-Uncertainty, Initial-Condition, and Accessibility

Armed only with the d_i curves and the $V_i^{forward}$ and V_i^{zero} values for each state, we can already offer three possible static estimation modes.

1) *Full-Uncertainty Estimation*: In this case, we assume that we have no knowledge of the switching activity of the circuit. We must choose maximum and minimum possible values of the body voltage that cover all possible stimulus and history. We say that a state is *accessible* if the circuit topology allows the state to be visited. (For example, for the nFET of an inverter,

those states with the source high would not be accessible because the source of the nFET is tied to ground.) We let \mathcal{A} represent the set of all such accessible states, including possibly the dynamic state 6. In this case, the minimum and maximum body voltages are given by

$$(V_B^{\text{ref}})_{\text{max}} = \max_{j \in \mathcal{A}} V_j^{\text{zero}} \quad (2)$$

$$(V_B^{\text{ref}})_{\text{min}} = \min_{j \in \mathcal{A}} V_j^{\text{zero}}. \quad (3)$$

2) *Initial-Condition Body-Voltage Estimation:* In this case, we assume that the circuit has been “quiet” for a long period of time but we do not know the specific quiescent state of the circuit. In this case, the minimum and maximum body voltage are given by

$$(V_B^{\text{ref}})_{\text{max}} = \max_{j \in \mathcal{A}_{\text{static}}} V_j^{\text{zero}} \quad (4)$$

$$(V_B^{\text{ref}})_{\text{min}} = \min_{j \in \mathcal{A}_{\text{static}}} V_j^{\text{zero}} \quad (5)$$

where $\mathcal{A}_{\text{static}}$ is the set of all accessible *static* states (i.e., states 1 through 5). Note that for the technology of Table I, for example, the minimum nFET full-uncertainty body-voltage value is less than initial-condition value. This means that it is possible for a switching nFET to have a lower body voltage than a quiescent FET because of the effect of state 6.

3) *Accessibility Analysis:* If one is assured that the circuit is under steady switching activity such that every accessible state is visited with reasonable frequency (i.e., on a time scale that is faster than the “slow” body-voltage mechanisms), then the V_j^{forward} values for the nFET (pFET) will cap the maximum (minimum) possible value of the body voltage. For the nFET

$$(V_B^{\text{ref}})_{\text{min}} = \min_{j \in \mathcal{A}} V_j^{\text{zero}} \quad (6)$$

$$(V_B^{\text{ref}})_{\text{max}} = \min \left(\max_{j \in \mathcal{A}} V_j^{\text{zero}}, \min_{j \in \mathcal{A}_{\text{static}}} V_j^{\text{forward}} \right) \quad (7)$$

while for the pFET

$$(V_B^{\text{ref}})_{\text{min}} = \max \left(\min_{j \in \mathcal{A}} V_j^{\text{zero}}, \max_{j \in \mathcal{A}_{\text{static}}} V_j^{\text{forward}} \right) \quad (8)$$

$$(V_B^{\text{ref}})_{\text{max}} = \max_{j \in \mathcal{A}} V_j^{\text{zero}}. \quad (9)$$

The assumption here is that it does not matter how long an accessible state is visited; it will be long enough to discharge the body down to the forward-bias turn-on voltage of the source-body or drain-body diodes. In general, we do not include state 6 in this “accessibility” analysis (we use $\mathcal{A}_{\text{static}}$ in the above equations) because state 6 is a switching state and, therefore, cannot be assured to meet this criterion. In those cases in which the relaxation in state 6 is extraordinarily “fast” because of a large on-state impact ionization current, state 6 can be safely included in the accessibility analysis.

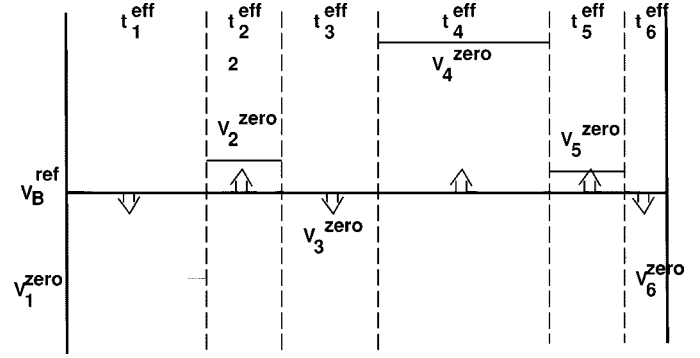


Fig. 3. Average cycle for an nFET, which, when repeated over and over, models the behavior of the body over a long period of time. Body voltage in state 2 determines the reference body voltage V_B^{ref} . Each state has a target reference voltage V_i^{zero} to which the body voltage is relaxing.

References [22] and [12] report on minimizing floating body effects in the complementary MOS (CMOS) inverter through the use of “charge-balanced” devices. Within the context of the formalism developed in this section, the charge-balanced nFET simply satisfies the condition $V_1^{\text{zero}} = V_5^{\text{zero}}$. If only states 1 and 5 were accessible, then the dc and steady-state voltages would be equal. Similarly, the charge balanced pFET satisfies $V_2^{\text{zero}} = V_3^{\text{zero}}$. However, the additional accessibility of 6 can upset this balance, dependent to a large extent on the magnitude of the on-state impact ionization current.

D. Detailed Body-Voltage Model

It is possible to tighten the estimates provided by the simple “uncertainty” modes with stochastic techniques (combined with timing information from static timing analysis) in which we consider the behavior of the body over a long period of time to be determined by an “average” cycle repeated over and over; in fact, we will characterize two average cycles: one to minimize and one to maximize the body voltage. Such an average cycle is shown in Fig. 3 divided into a series of time slices t_i^{eff} , which characterize the amount of time per cycle on the average that the FET spends in state i . Of course, the time slices sum to the cycle time

$$\sum_{i=1}^6 t_i^{\text{eff}} = t_{\text{cycle}}. \quad (10)$$

In each time slice i , the reference body voltage is relaxing to the target value V_i^{zero} with a characteristic relaxation time denoted by τ_i . The body-voltage-dependent time constants for this relaxation are denoted by the τ_i and can be captured from the device models as part of a technology precharacterization (usually with a piecewise-linear (PWL) representation of the logarithm of τ_i as a function of V_B^{ref} , since τ_i varies over several orders of magnitude). Moreover, τ_i is independent of device width W because both the body currents and body capacitance scale proportionately with W . Except for the fast discharge associated with a source-body or drain-body junction that becomes strongly forward biased, these time constants are much larger than t_{cycle} and any voltage change during a single time slice would be imperceptible in Fig. 3. From this simple picture, one can relate the

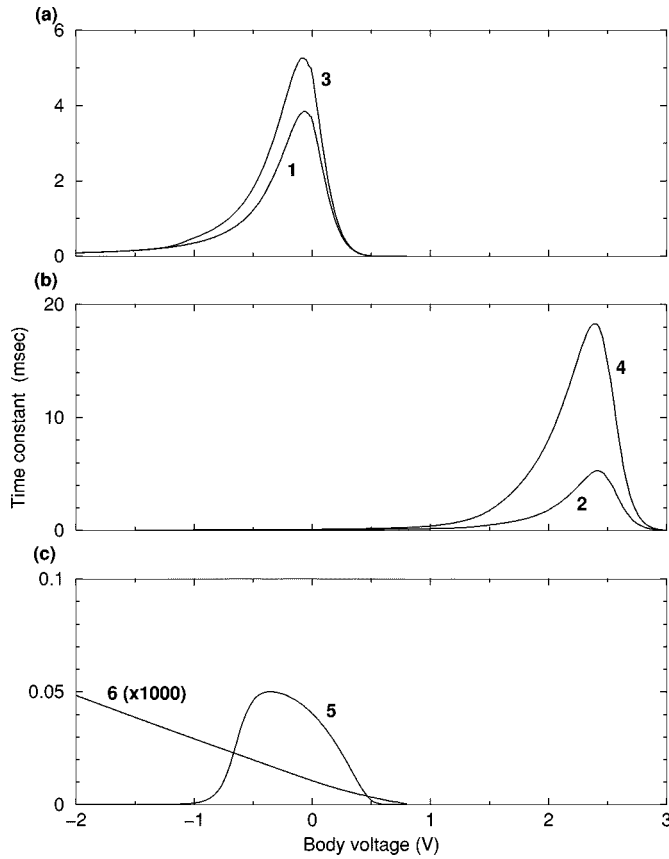


Fig. 4. nFET relaxation times τ_i as a function of “local” body voltage for a supply voltage of 2.5 V. (a) τ_1 and τ_3 . (b) τ_2 and τ_4 . (c) τ_5 and τ_6 .

reference body voltage at the end of the cycle V_{n+1} to the reference body voltage at the beginning of the cycle V_n by

$$V_{n+1} = V_n \exp \sum_{i=1}^6 -t_i^{\text{eff}}/\tau_i + \sum_{i=1}^6 V_i^{\text{zero}} (1 - e^{-t_i^{\text{eff}}/\tau_i}) \exp \sum_{j=i+1}^6 t_j^{\text{eff}}/\tau_j. \quad (11)$$

The steady-state solution of this difference equation (in the approximation that the τ_i are much greater than t_{cycle}) is given by

$$V_B^{\text{ref}} = \frac{\sum_{i=1}^6 (t_i^{\text{eff}}/\tau_i) V_i^{\text{zero}}}{\sum_{i=1}^6 t_i^{\text{eff}}/\tau_i}. \quad (12)$$

Because of the V_B^{ref} -dependence of the τ_i , this equations must be solved self-consistently.

Fig. 4 plots the relaxation times τ_i for the nFET at the 2.5-V supply as a function of the “local” body voltage (that is, the body voltage in state i rather than the reference body voltage). The relaxation times τ_1 and τ_3 peak around a zero body voltage and then decrease for negative body voltage as the source–body and drain–body leakage currents increase. The decrease for positive body voltage is even more substantial as the junctions become

forward biased and draw significant current. The same trends are observed for the relaxation times τ_2 and τ_4 . In this case, the relaxation times peak around the supply voltage (the source and drain are both at the supply voltage for states 2 and 4). τ_5 peaks around zero. For voltages above the peak, τ_5 decreases strongly as the body–source junction becomes forward biased. For voltages below the peak, τ_5 also decreases as both the body–source and body–drain junctions become more reverse-biased. τ_5 peaks at a much smaller time-constant value than τ_1 through τ_4 because even at a body voltage near the peak, one of the junctions (the drain–body junction) is strongly reversed biased and, therefore, leaking considerable current. For state 6, the time constants are considerably smaller (factor of 1000) than they are in state 5. This is due to the contribution of the on-state impact ionization current to the leakage currents from the floating body. (This very low relaxation time will mean that devices that have any time in state 6 in their switching transients will have their body voltages effectively “pinned” at s_6 .) At 1.0 V, by contrast, τ_5 and τ_6 are comparable in magnitude.

In detailed steady-state analysis, we seek to determine the maximum and minimum possible V_B^{ref} values from (12) with consideration of *all* of the allowable t_i^{eff} . If the τ_i in (12) were constant and equal, then to maximize V_B^{ref} , one would simply maximize the time in those states with higher V_i^{zero} (subject to constraints on the maximum and minimum possible values of the t_i^{eff} imposed by the stochastic analysis); similarly, to minimize V_B^{ref} , one must maximize the time in those states with lower V_i^{zero} . The τ_i values, however, are not constant, but because they have the property that they decrease rapidly for V_B^{ref} different from V_i^{zero} , states with the largest $|V_B^{\text{ref}} - V_i^{\text{zero}}|$ continue to have the biggest “pull” and the V_i^{zero} “ordering rule” continues to hold. To formalize this, we define the rank of the state i as an integer r_i indicating the priority of states to *maximize* the body voltage. Higher r_i means higher priority. For example, from Table I, for the nFET at 2.5 V, $r_4 = 6$, $r_5 = 5$, $r_2 = 4$, $r_3 = 3$, $r_1 = 2$, and $r_6 = 1$. We can also define a complementary rank \bar{r}_i for *minimizing* V_B^{ref} ; these priorities are simply the reverse of the maximum case. This ordering [which, quite expectedly, favors states with high (low) source or drain voltages to maximize (minimize) the body voltages] is technology-independent except for the relative positions of state 5 and 6, which depend on both technology and supply voltage.

As a specific example of how the steady-state body voltage depends on the t_i^{eff} , we consider in Fig. 5 an nFET at the 1.0-V supply. The cycle time t_{cycle} is 10 ns. In Fig. 5(a), we assume a hypothetical case in which only states 1 and 4 are accessible and plot the reference body voltage as a function of t_4^{eff} given that $t_1^{\text{eff}} = t_{\text{cycle}} - t_4^{\text{eff}}$. The values of V_1^{zero} , V_1^{forward} , and V_4^{zero} are also noted. Given the accessibility of states 1 and 4, the range from V_1^{zero} to V_1^{forward} (denoted with the arrow) would be given by accessibility body–voltage estimation. The detailed body voltage estimate becomes V_1^{zero} as $t_4^{\text{eff}} \rightarrow 0$ and monotonically increases with increasing t_4^{eff} ($r_4 > r_1$). As $t_4^{\text{eff}} \rightarrow t_{\text{cycle}}$ (and $t_1^{\text{eff}} \rightarrow 0$), V_B^{ref} rapidly increases to V_4^{zero} . V_B^{ref} can only increase above V_1^{forward} with $t_1^{\text{eff}} = 0$ on the scale of the “fast” forward-biased discharge of the body. In Fig. 5(b), we show the case in which only states 1 and 2 are accessible. V_B^{ref} increases monotonically as t_2^{eff} increases ($r_2 > r_1$). In this case,

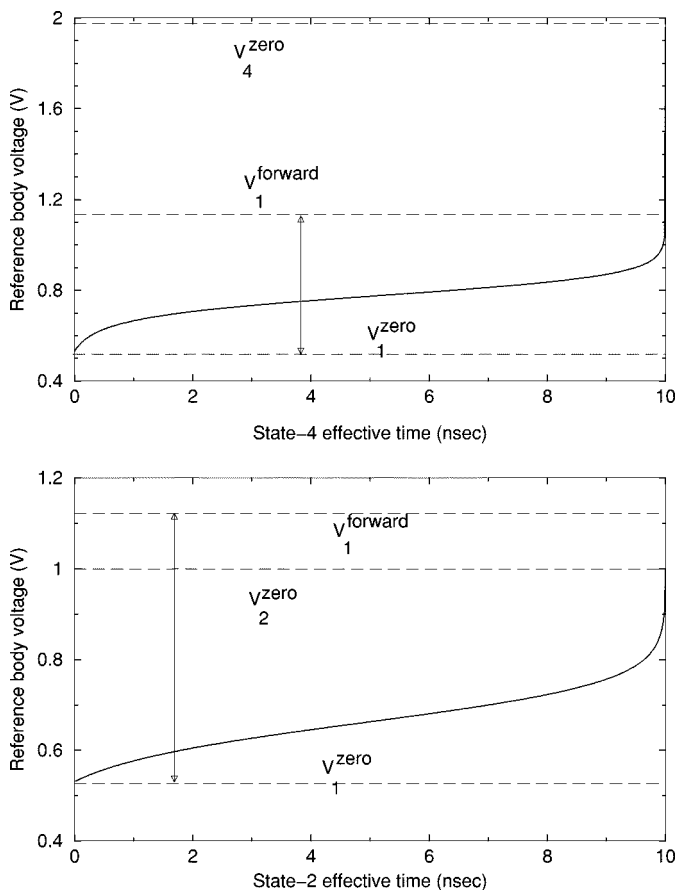


Fig. 5. Two examples for an nFET at 1.0-V supply to demonstrate how the steady-state reference body voltage depends on the t_i^{eff} . (a) Only states 1 and 4 are accessible. (b) Only states 1 and 2 are accessible.

$V_2^{\text{zero}} < V_1^{\text{forward}}$ so that V_B^{ref} as $t_2^{\text{eff}} \rightarrow 0$ does not actually reach the upper bound of accessibility analysis.

To determine allowable values of t_i^{eff} for detailed steady-state analysis, we need information about the circuit environment of each transistor, both logical and temporal. We characterize the logical environment of each FET by a set of signal probabilities that determine the possible states of the source, gate, and drain of the transistor at the end of a cycle. For the nFET, these are:

$P(G)$	Probability that at the end of a cycle the gate is driven high.
$P(D \bar{G})$	Probability that at the end of a cycle the drain is driven high given that the gate is low.
$P(\bar{D} \bar{G})$	Probability that at the end of a cycle the drain is driven low given that the gate is low.
$P(S \bar{G})$	Probability that at the end of a cycle the source is driven high given that the gate is low.
$P(\bar{S} \bar{G})$	Probability that at the end of a cycle the source is driven low given that the gate is low.
$P(S G) = P(D G)$	Probability that at the end of a cycle the drain (and source) are driven high given that the gate is high.

$P(\bar{S}|G) = P(\bar{D}|G)$ Probability that at the end of a cycle that the drain (and source) are driven low given that the gate is high.

The corresponding probabilities for the pFET are $P(G)$, $P(S|G)$, $P(\bar{S}|G)$, $P(D|G)$, $P(\bar{D}|G)$, $P(S|\bar{G}) = P(D|\bar{G})$, $P(\bar{S}|\bar{G}) = P(\bar{D}|\bar{G})$. By driven low (high), we mean that there is a path to ground (supply). We note that the following conditions on the conditional signal probabilities must hold:

$$\begin{aligned} P(D|\bar{G}) + P(\bar{D}|\bar{G}) &\leq 1 \\ P(S|\bar{G}) + P(\bar{S}|\bar{G}) &\leq 1 \\ P(D|G) + P(\bar{D}|G) &\leq 1 \\ P(S|G) + P(\bar{S}|G) &\leq 1. \end{aligned}$$

None of these have to sum to precisely one because of the possibility of floating nodes.

With these signal probabilities, one can view the state diagram in Fig. 1 as representing a Markov process with six-by-six transition matrix \mathbf{A} . Assigning different transition matrices to the minimum and maximum cases yields the difference equations

$$\begin{aligned} (P_i^{\text{max}})_{k+1} &= \mathbf{A}_{\text{max}}(P_i^{\text{max}})_k \\ (P_i^{\text{min}})_{k+1} &= \mathbf{A}_{\text{min}}(P_i^{\text{min}})_k \end{aligned}$$

where the six-by-six matrices \mathbf{A}_{max} and \mathbf{A}_{min} are given by

$$\begin{aligned} \mathbf{A}_{\text{max}} &= (p_1^{\text{max}} \ p_2^{\text{max}} \ p_3^{\text{max}} \ p_4^{\text{max}} \ p_{5a}^{\text{max}} \ p_{5b}^{\text{max}})^T \\ &\quad \times (1 \ 1 \ 1 \ 1 \ 1 \ 1) \\ \mathbf{A}_{\text{min}} &= (p_1^{\text{min}} \ p_2^{\text{min}} \ p_3^{\text{min}} \ p_4^{\text{min}} \ p_{5a}^{\text{min}} \ p_{5b}^{\text{min}})^T \\ &\quad \times (1 \ 1 \ 1 \ 1 \ 1 \ 1). \end{aligned}$$

$(P_i)_k$ is the probability of being in state i at the end of cycle k . p_i is the probability of making a transition to state i and follows directly from the source, gate, and drain signal probabilities. For example, for the nFET, $p_3^{\text{max}} = (1 - P(G))P(\bar{S}|\bar{G})P(\bar{D}|\bar{G})$ while $p_3^{\text{min}} = (1 - P(G))(1 - P(S|\bar{G}))(1 - P(D|\bar{G}))$. The maximum (minimum) case assumes that the floating node condition on the source or drain takes a high (low) voltage value. Diagonalizing \mathbf{A} (trivially) and finding the eigenvector associated with eigenvalue one (normalized so that the sum of the elements of the vector is one) gives the steady-state values of the P_i . From these probabilities, one can calculate a set of 36 transition probabilities (that is, the probability that at given FET in a given cycle is transitioning from state i to state j) for both the minimum or maximum cases: $P_{i \rightarrow j}^{\text{max}} = P_i^{\text{max}} p_j^{\text{max}}$ and $P_{i \rightarrow j}^{\text{min}} = P_i^{\text{min}} p_j^{\text{min}}$.

We must next determine the fraction of the cycle time ($t_i^{j \rightarrow k}$) that can be spent in each state i as part of the transition $j \rightarrow k$ to maximize (or minimize) the body voltage among the set of possible waveforms. To do this, we require early and late arrival times (rising and falling) for the source, gate, and drain of the FET under consideration (the target FET). We denote these arrival times for the early case as the following.

$S_{\text{rise}}^{\text{early}}$	Earliest time the source of the FET can be driven high.
$S_{\text{fall}}^{\text{early}}$	Earliest time the source of the FET can be driven low.
$D_{\text{rise}}^{\text{early}}$	Earliest time the drain of the FET can be driven high.
$D_{\text{fall}}^{\text{early}}$	Earliest time the drain of the FET can be driven low.

$G_{\text{fall}}^{\text{early}}$ Earliest time the gate of the FET can be driven low.

$G_{\text{rise}}^{\text{early}}$ Earliest time the gate of the FET can be driven high.

There are comparable arrival times associated with the late state:

$S_{\text{rise}}^{\text{late}}$, $S_{\text{fall}}^{\text{late}}$, $D_{\text{rise}}^{\text{late}}$, $D_{\text{fall}}^{\text{late}}$, $C_{\text{fall}}^{\text{late}}$, and $C_{\text{rise}}^{\text{late}}$. For the source and drain arrival time, we are assuming that the target FET is off.

The details of how these are obtained in the context of static timing analysis is described in Appendix II. Each transition has associated with it a set of arrival times necessary to make that transition, a *transition set*, denoted as $\mathcal{T}_{i \rightarrow j}$. For example, for the $1 \rightarrow 2$ transition for the nFET, the associated arrival time set is $\mathcal{T}_{1 \rightarrow 2} = \{S_{\text{rise}}, D_{\text{rise}}\}$. For $1 \rightarrow 4$ for the nFET, the associated arrival time set is $\mathcal{T}_{1 \rightarrow 4} = \{G_{\text{fall}}, S_{\text{rise}}, D_{\text{rise}}\}$. We can define max and min operators that act on the transition sets. $\max(\mathcal{T}_{i \rightarrow j})$ returns the largest of the early arrival times in the transition set, while $\min(\mathcal{T}_{i \rightarrow j})$ returns the smallest of the late arrival times in the transition set.

To indicate the states of Fig. 1 involved in a cycle and to handle the possibility of hazards, we can denote the waveform in a cycle (in this case involving a transition from i to j) using a *transition notation* as follows:

$$i \xrightarrow{\max} k \xrightarrow{\min} j.$$

In this cycle, a hazard to state k occurs as part of the transition. The transition notation must involve only static states and indicates the amount of time spent in each of these static states as part of the transition. Specifically for this example, $t_{i \rightarrow j}^{\max} = \max(\mathcal{T}_{i \rightarrow k})$, $t_{i \rightarrow j}^{\min} = \min(\mathcal{T}_{k \rightarrow j}) - \max(\mathcal{T}_{i \rightarrow k})$, and $t_j^{\max} = t_{\text{cycle}} - \min(\mathcal{T}_{k \rightarrow j})$. Hazards are introduced when they act to increase (in the case that we are seeking the maximum body voltage) or decrease (in the case that we are seeking the minimum body voltage) the steady-state body voltage that would result from the particular waveform being repeated indefinitely. Appendix I discusses an algorithm for determining which hazards to include in a given transition.

This transition analysis determines the $t_i^{j \rightarrow k}$ for the static states i . $t_i^{j \rightarrow k}$ depends, in principle, on the number of switching events in a cycle that occur as a result of current flow through the target device. Each such switching event contributes an amount t_{switch} to $t_i^{j \rightarrow k}$, where t_{switch} is the approximate switching time of the FET. For the purposes of our detailed steady-state body-voltage estimation, we assume that if state 6 is accessible, exactly one switching event occurs per cycle. We find that, in practice, the detailed results are not very sensitive to the number of assumed switching events or the exact value of t_{switch} . For cases in which the on-state impact ionization current is large, any $t_i^{j \rightarrow k} > 0$ produces pinning at $V_B^{\text{ref}} = V_6^{\text{zero}} = V_6^{\text{forward}}$. For cases in which the on-state impact ionization does not dominate the current to the body in state 6, t_i^{eff} is usually much smaller than the t_i^{eff} associated with the static states. As a result, the detailed body voltage has very little sensitivity to the exact value of t_i^{eff} .

From the $t_i^{j \rightarrow k}$ determined above, we can calculate an effective amount of time (t_i^{eff}) on the average per cycle that the FET is in the state i

$$\begin{aligned} (t_i^{\text{eff}})^{\max} &= \sum_{j,k} P_{j \rightarrow k}^{\max} (t_i^{j \rightarrow k})^{\max} \\ (t_i^{\text{eff}})^{\min} &= \sum_{j,k} P_{j \rightarrow k}^{\min} (t_i^{j \rightarrow k})^{\min}. \end{aligned}$$

Obviously, there are two sets of t_i^{eff} values, one to maximize and one to minimize the body voltage.

III. STATIC TIMING ANALYSIS

Now consider the application of these body-voltage estimates [both the “simple” uncertainty modes and the more complex detailed analysis] to a static transistor-level timing analysis engine: SOI static timing analyzer (SOI-STA). The design is partitioned into CCCs for analysis, as is traditionally done in static transistor-level tools [16]. SOI-STA utilizes a breadth-first search (BFS) of the resulting timing graph, which ensures that the arrival times on all of the inputs are known when the delay through the CCC must be calculated. This enables detailed body-voltage estimates on all of the FETs of the CCC to be used in delay simulation. Delay propagation through a CCC occurs as a result of a single switching event on an input (i.e., simultaneously switching inputs are not considered).

Depending on the degree of knowledge we have of past switching activity (no knowledge, quiescent, steady switching, steady switching with known signal probabilities and arrival times), we can use one of the techniques discussed in Section II (full-uncertainty, initial-condition, accessibility, detailed) to determine the minimum and maximum possible value the reference body voltage can have for each FET of the CCC under analysis. For the given sensitization of the CCC for delay calculation, each FET is in a known state, the minimum or maximum body voltage of which can be determined by a displacement from the reference voltage. The body voltage values are then used as the “initial conditions” for the required delay simulation. Early-mode calculation for rising transitions sensitizes the nFETs (pFETs) of the pull-up path to be maximum (minimum) and (to reduce the “fight” during switching) the nFETs (pFETs) of the pull-down path to be minimum (maximum). This same sensitization applies to late-mode fall transitions. Early-mode calculation for falling transitions sensitizes the nFETs (pFETs) of the pull-down path to be maximum (minimum) and the nFETs (pFETs) of the pull-up path to be minimum (maximum). The same sensitization applies to late-mode rising transitions. The fan-out CCCs are included in the delay calculation to improve the delay accuracy, an improvement over the grounded-cap-load approach first reported in [14]. (This, along with some bug fixes, accounts for some of the differences between the detailed results presented here and those in [14].) To maximize (minimize) device capacitance for late-mode (early-mode), the maximum (minimum) nFET and minimum (maximum) pFET body voltages are used.

SOI-STA propagates full PWL waveforms. One of the complexities of BFS timing analysis is determining the late- and early-mode waveforms at the output of each CCC. It can sometimes be the case that the waveform with the maximum (minimum) delay (as measured by the 50% point) is not the waveform with the slowest (fastest) slew (as measured by the 20%–80% rise–fall time). In these cases, we propagate a “hybrid” waveform. We choose the waveform with the largest (smallest) slew as the late-mode (early-mode) waveform and translate it in time so that it has the largest (smallest) delay. SOI-STA also propagates signal probabilities using

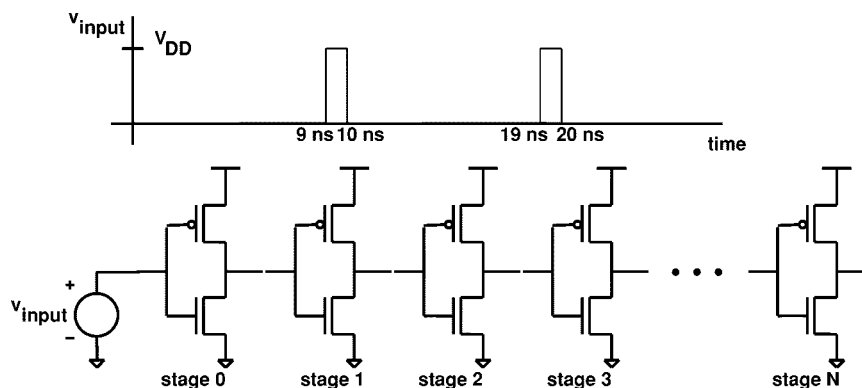


Fig. 6. Chain of identical inverters stimulated with the waveform shown in the inset. Slew times on the input waveforms are 100 ps.

assumptions of spatial and temporal independence, borrowing from similar techniques in static power analysis. If detailed body-voltage estimation is used, once the signal probabilities and arrival times are known at the inputs of a CCC, these probabilities are translated into FET signal probabilities and arrival time values as discussed in Appendix II.

IV. RESULTS AND DISCUSSION

We present static timing results from SOI-STA for three examples (of increasing complexity) and compare these with the results of circuit simulations in which vectors are chosen both to correctly sensitize the delay path in question and to match the assumed switching behavior.

A. Inverter Chain

The first circuit is a chain of identical inverters as shown in Fig. 6 stimulated with the periodic waveform shown in the inset. This is the simplest possible circuit example and has been well-studied from a dynamic point-of-view in previous work [12], [13]. The input waveform repeats every 10 ns and is equivalent to a signal probability of 0.1 on the input of even-stage inverters and 0.9 on the input of odd-stage inverters. Before $t = 0$, even (odd) stages have a zero (one) on their input. We consider results at two supplies, 2.5 V in Fig. 7 and 1.0 V in Fig. 8. In Fig. 7(a), we present the body voltages for the FETs as determined by circuit simulation for both even (in solid lines) and odd (in dashed lines) inverter stages. For both the even and odd cases, there are four curves: two curves for the pFET and two curves for the nFET. The pFET curves labeled 5 and the nFET curves labeled 1 correspond to the case in which the inverter input is high, so that the body voltage is measured with the pFET in state 5 and the nFET in state 1. The pFET curves labeled 2 and the nFET curves labeled 5 correspond to the case in which the inverter input is low so that the body voltage is measured with the pFET in state 2 and the nFET in state 5. We notice that initially the even-stage FETs have higher body voltages than the odd-stage FETs. This is because the even-stage FETs have their gates held low before $t = 0$ with the pFET (nFET) in state 2 (5), while the odd-stage FETs have their gates held high before $t = 0$ with the pFET (nFET) in state 5 (1). $r_5 > r_1$ for the nFET and $r_2 > r_5$ for the pFET. As shown in the stage delay results in Fig. 7(b) (even stage delay in solid, odd-stage delay in dashed), this gives the even stages initially a longer rise

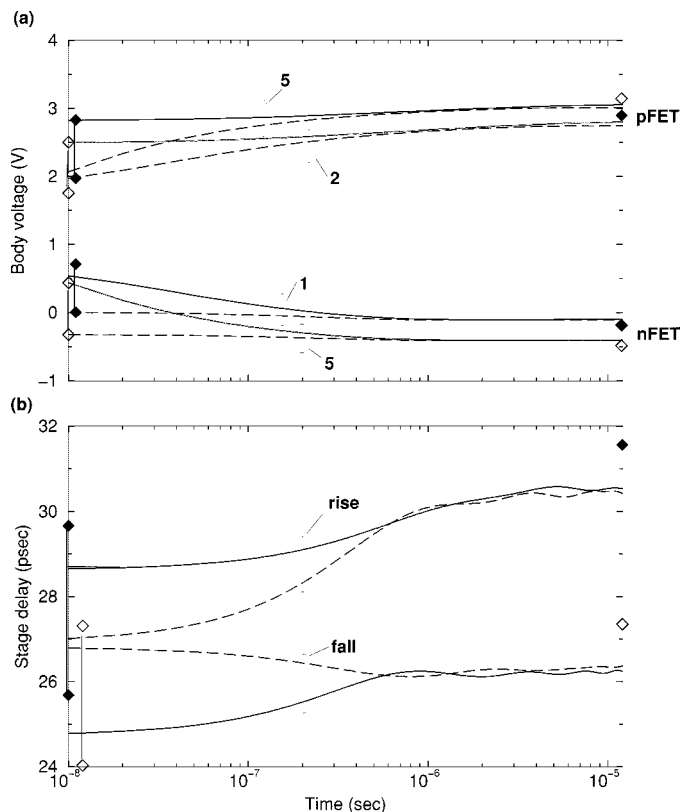


Fig. 7. Inverter chain results for a supply voltage of 2.5 V. (a) Body voltage. (b) Inverter delay.

time, but smaller fall time than the odd stages. Initially, then the input pulses are stretched [12] as they move down the inverter chain. As switching begins, state 6 becomes accessible and the body voltages become pinned in about a microsecond (the time scale of τ_6) to their steady-state values. We also note that in this case, the steady-state nFET (pFET) body voltage is slightly less (more) than the initial-condition minimum (maximum) value. Six-state pinning means that there is no difference between the steady-state body voltages of the FETs in even and odd stages. In steady-state, then, the rise and fall delays become the same in the even and odd stages and there is no pulse stretching.

We now consider how these results compare with SOI-STA. Early and late arrival times at the input of the chain, both rise and fall, are set to 100 ps. The rise and fall times at the input of the chain are also set to 100 ps. The diamonds on the right

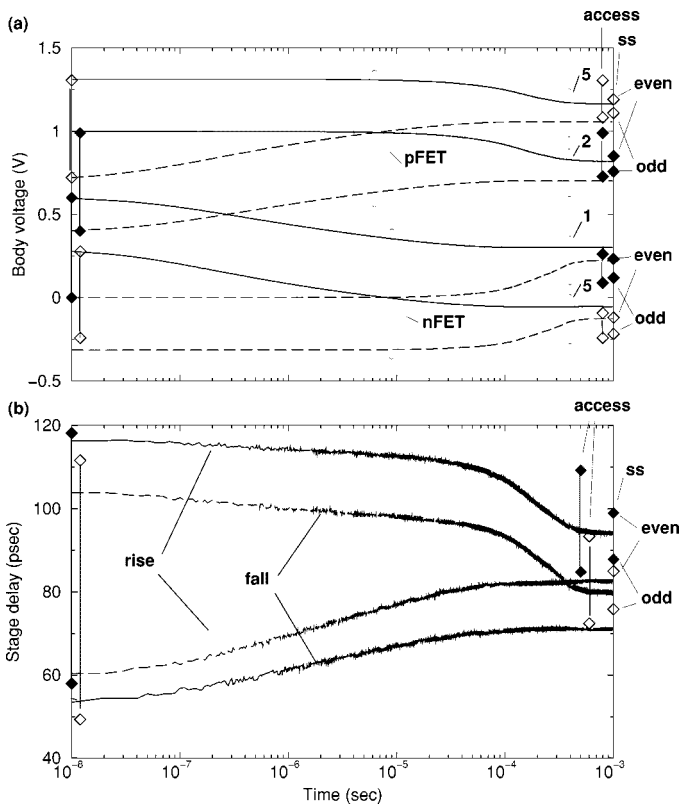


Fig. 8. Inverter chain results for a supply voltage of 1.0 V. (a) Body voltage. (b) Inverter delay.

vertical axis in Fig. 7(a) correspond to the *detailed* steady-state body voltages estimated by SOI-STA for an input signal probability of 0.1, propagated as 0.9 to the inputs of the odd stages. State 1 and state 2 for the pFET and nFET, respectively, are shown as solid diamonds. State 5 for the pFET and nFET are shown as hollow diamonds. Because of the six-state pinning in the steady state, the detailed body voltages are the same for both the even and odd stages with minimum and maximum values indistinguishably close. Including state 6, the *accessibility* analysis produces the same body-voltage estimates as the steady-state results. On the left vertical axis, we show the minimum and maximum *initial-condition* body voltage values estimated by SOI-STA. These bound very closely the initial condition body voltages observed in the dynamic simulation, since in this case, the dynamic simulation covers all the accessible states. In Fig. 7(b), we show the rising (solid diamond) and falling (hollow diamond) delays predicted by SOI-STA for the initial-condition (left axis) and steady-state (right axis) cases.²

The SOI-STA steady-state delays slightly overestimate those determined by dynamic simulation because of the slight overestimation (underestimation) of pFET (nFET) body voltages.

We can contrast these results with Fig. 8 in which the same simulation and analysis is done on the inverter chain at a supply voltage of 1.0 V. Here, we do not have six-state pinning because of the considerably reduced on-state impact ionization current at the lower supply. Dashed lines once again correspond to odd inverter stages and solid lines to even inverter stages.

²The “wiggles” in the delay curve beyond 1 μ s are rounding-error artifacts of the circuit simulation.

Disparate steady-state body voltages occur for inverters in the even and odd stages of the chain as shown in Fig. 8(a). The diamonds on the right vertical axis denote the SOI-STA-determined detailed steady-state body voltages, which agree with the circuit simulation results within a few percent. Minimum and maximum values are indistinguishably close³, as they were in the 2.5-V case, but in this case, there are different steady-state body voltage values for the even and odd stages of the chain. The diamonds slightly to the left of the right vertical axis are the minimum and maximum body voltages determined by SOI-STA from the simple accessibility analysis. These numbers bound the full body voltage variation of both odd and even inverter stages. The diamonds on the left vertical axis give the initial condition values that bound nearly precisely the variation observed in circuit simulation across even and odd stages. Fig. 8(b) shows the inverter delay as a function of time for both even (solid) and odd (dashed) stages. (The “noise” in these curves is due to numerical round-off error as increasingly large numbers are subtracted to give the stage delay.) Because of the steady-state body-voltage differences between even and odd stages at this supply, the steady-state delays for falling (rising) outputs is slower (faster) for even stages than for odd stages. Pulse-stretching occurs in this case in the steady state. The diamonds on the right vertical axis are the SOI-STA-determined detailed steady-state delays. Solid diamonds correspond to rising outputs, while hollow diamonds correspond to falling outputs. The delays determined by SOI-STA using the accessibility analysis, which bound the delay variation for both even and odd stages, are also noted. The solid diamonds on the left vertical axis give the initial condition delay variation for rising outputs, while the hollow diamonds correspond to falling outputs. With known switching activity, one could significantly reduce the delay uncertainty with which one would have to ensure functionality of the design. Comparing Figs. 7 and 8, the impact of the hysteric V_T variation is far greater at the reduced supply because V_T is a larger portion of the supply voltage. Also, comparing the time scales of the 2.5- and 1.0-V cases, we find that the time scale for reaching steady-state is far longer for the 1.0-V case than the 2.5-V case because of the reduced on-state impact ionization current.

B. Ripple-Carry Adder

The next example we consider is a static ripple-carry adder as shown in Fig. 9(b). Fig. 9(a) shows the component full-adder circuit. In the circuit simulation, we use the input waveforms shown in the inset of Fig. 9(a), which sensitizes the critical path of this circuit, the carry chain. The “A” waveform is applied to each A input and the “B” waveform is applied to each B input. The “C” waveform is applied to the C_{in} input of the ripple-carry [see Fig. 9(b)]. These waveforms correspond to signal probabilities of 0.5 on the A and B inputs of each full-adder cell and 0.5 on the C input of each full-adder cell. For these input signal probabilities, the signal probability of C_{out} is 0.5, so that each cell sees identical switching activity. Before $t = 0$, A is high and B is low for each cell. For even stages, C is high, while for odd

³There is no path delay variation to produce any significant differences between the early and late arrival times.

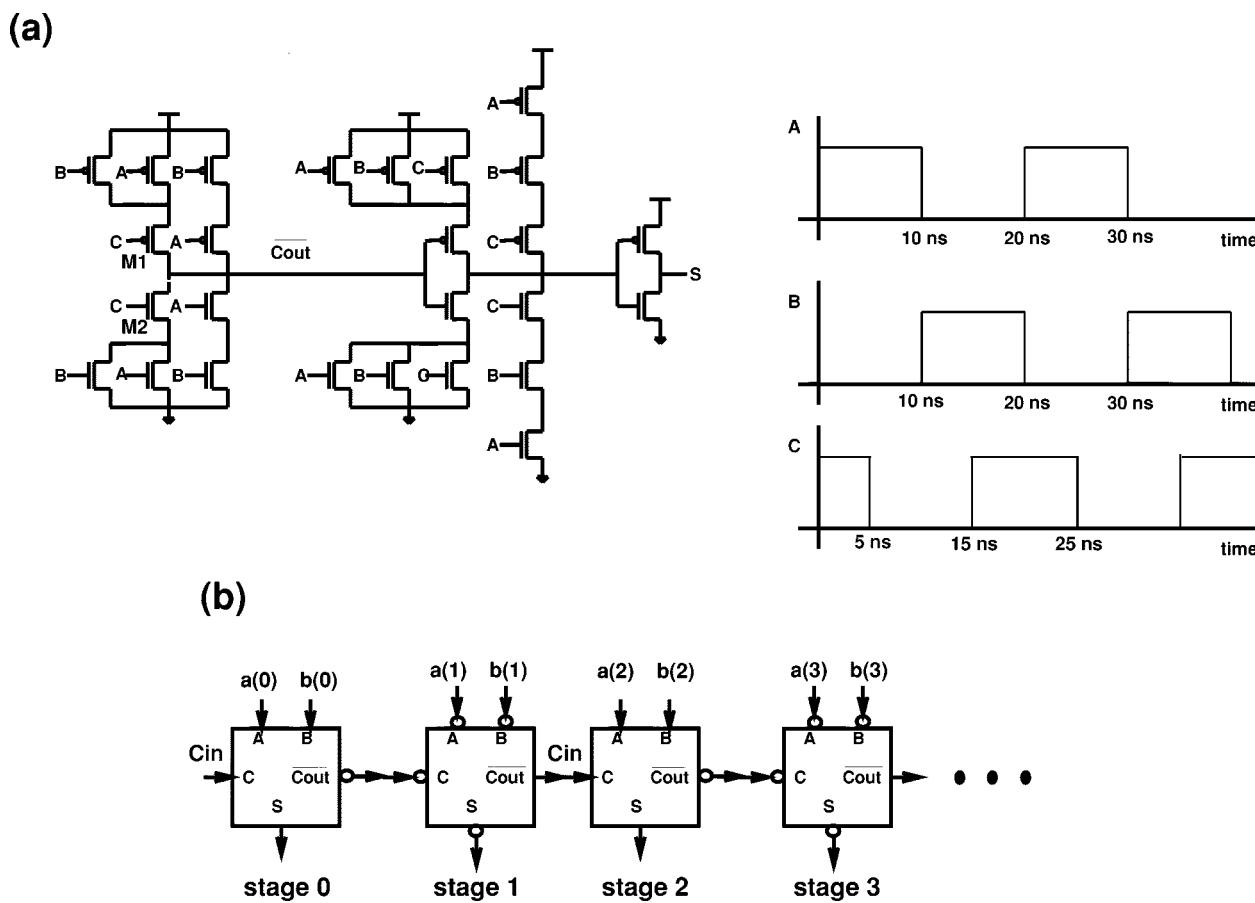


Fig. 9. Ripple-carry adder circuit. (a) Associated full-adder circuit. (b) connection of the full-adder cells. Inset of (a) gives the waveform applied for dynamic simulation.

stages C is low. Fig. 10 shows the results for a supply voltage of 1.0 V. In Fig. 10(a), we compare the body voltages of transistors $M1$ and $M2$ of Fig. 9(a) with the SOI-STA initial-condition, accessibility, and detailed steady-state results. The $M1$ curves labeled 5 (2) and $M2$ curve labeled 1 (5) correspond to the case in which C is high (low) for a given cell. The solid lines correspond to even stages and the dashed lines to odd stages. The steady-state values match almost exactly the values determined from circuit simulation, a circuit simulation run that can take *days* to complete. The initial-condition body-voltage values almost precisely bound the circuit simulation results. Fig. 10(b) shows the complete stage delay of the full-adder cell from C to \overline{Cout} for both rising and falling \overline{Cout} for both even (solid lines) and odd (dashed lines) stages. The detailed, accessibility, and initial-condition delays from SOI-STA bound the simulation values. The uncertainty of the detailed steady-state values derives entirely from the difference in the delay depending on whether A is zero and B is one or A is one and B is zero. These cases present different loads on \overline{Cout} (the “sidebranch” loading).

C. 4-2 Compressor

The last and most complex example presented is a 4-2 compressor circuit from a tree multiplier design [24] (see Fig. 11). We specifically consider the long delay path I4-Ap-C-D-E-F-S

and the short delay path I2-C-E-S. In the dynamic analysis, we use input waveforms of “50% duty cycle” to sensitize these paths, corresponding to signal probabilities of 0.5 on all of the inputs in SOI-STA detailed steady-state analysis. In Table II, we compare the initial-condition delays determined by circuit simulation with the SOI-STA-determined full-uncertainty delays. (Here, we do full-uncertainty analysis, as opposed to initial-condition analysis as done previously. The only difference is that stage 6 can be included in the set of accessible states.) Table III does a similar comparison between the steady-state delays determined by circuit simulation (after more than 50 000 cycles of simulation) and the SOI-STA-determined detailed and accessibility steady-state delays. In all cases, the SOI-STA-determined delays bound the SPICE delays. One should also notice the considerable reduction in uncertainty between the full-uncertainty and detailed steady-state delays as the component of this uncertainty due to body voltage variation is noticeably reduced.

V. CONCLUSION AND FUTURE WORK

In this paper, we have presented a circuit-focused model of the floating-body potential of PD-SOI FETs. This model allows one to determine the body voltage and its associated uncertainty, depending on knowledge of the switching activity of the FETs in question. Four types of estimation are possible

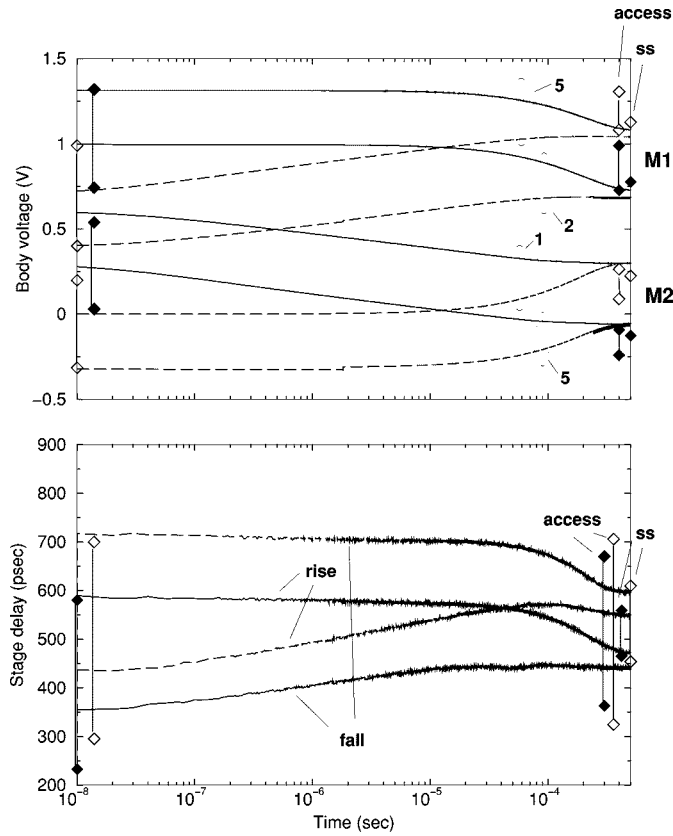


Fig. 10. Carry chain results for a supply voltage of 1.0 V. (a) Body voltage. (b) Stage delay.

depending on switching assumptions and the amount of information known about the logical and temporal environment of the circuit under analysis. We have incorporated this model into a prototype transistor-level static timing analysis engine to demonstrate the impact reduced body-voltage uncertainty can have on performance evaluation. We find that the body-voltage uncertainty can be significantly reduced with fairly conservative assumptions about switching behavior.

PD-SOI technology delivers the most potential performance benefit for circuits with high stack structures and pass-transistor logic. It is precisely the FETs of these circuit structures that show the greatest potential body-voltage uncertainty because of the large number of accessible states. It is also these FETs that show the most dramatic reduction in uncertainty with the knowledge that they are under relatively consistent switching activity.

Future work will include incorporating these body voltage estimates into transistor-level static noise analysis. In addition, we intend to consider design techniques whereby a normally inactive block could be periodically stimulated to keep it "primed" so that when it is eventually exercised, it has more tightly predictable body voltage variation. This is similar to some of the circuit techniques that attempt to force discharge of the body during "noncritical" periods of circuit operation (e.g., precharge in dynamic logic) to reduce parasitic bipolar leakage. In many ways, this could also be viewed as analogous to dynamic random access memory (DRAM) refresh. More work would be required to determine the necessary frequency and nature of this pattern.

APPENDIX I ADDING HAZARDS TO A GIVEN TRANSITION IN DETAILED ANALYSIS

Specifically, a state k can be inserted between states i and j according to one of the following cases.

Case 1) If $(r_k > r_i) \wedge (r_k > r_j) \wedge (\max(\mathcal{T}_{i \rightarrow k}) < \min(\mathcal{T}_{k \rightarrow j}))$, the state k can be inserted between i and j as a hazard to increase the steady-state body voltage. The cycle after this insertion is

$$i \xrightarrow{\max} k \xrightarrow{\min} j.$$

In this case, increasing the amount of time in state k at the expense of states i and j increases the body voltage. If $(\bar{r}_k > \bar{r}_i) \wedge (\bar{r}_k > \bar{r}_j) \wedge (\max(\mathcal{T}_{i \rightarrow k}) < \min(\mathcal{T}_{k \rightarrow j}))$, the state k can be inserted between i and j as a hazard to decrease the body voltage. The cycle after this insertion is the same as in the maximum case. However, in this case, increasing the amount of time in state k at the expense of states i and j decreases the body voltage.

Case 2) If $(r_k > r_j) \wedge (r_k < r_i) \wedge (\min(\mathcal{T}_{i \rightarrow k}) \geq \min(\mathcal{T}_{i \rightarrow j})) \wedge (\min(\mathcal{T}_{k \rightarrow j}) > \min(\mathcal{T}_{i \rightarrow k}))$, the state k can be inserted between i and j as a hazard to increase the body voltage. In this case, the initial cycle is

$$i \xrightarrow{\min} j.$$

After insertion, it is

$$i \xrightarrow{\min} k \xrightarrow{\min} j.$$

State k is inserted only if adding it does not decrease the time in state i . If $(\bar{r}_k > \bar{r}_j) \wedge (\bar{r}_k < \bar{r}_i) \wedge (\min(\mathcal{T}_{i \rightarrow k}) \geq \min(\mathcal{T}_{i \rightarrow j})) \wedge (\min(\mathcal{T}_{k \rightarrow j}) > \min(\mathcal{T}_{i \rightarrow k}))$, the state k can be inserted between i and j as a hazard to decrease the body voltage. The cycle before and after insertion is the same as in the maximum case.

Case 3) If $(r_k > r_i) \wedge (r_k < r_j) \wedge (\max(\mathcal{T}_{k \rightarrow j}) \leq \max(\mathcal{T}_{i \rightarrow j})) \wedge (\max(\mathcal{T}_{k \rightarrow j}) > \max(\mathcal{T}_{i \rightarrow k}))$, the state k can be inserted between i and j as a hazard to increase the body voltage. In this case, the initial cycle is

$$i \xrightarrow{\max} j.$$

After insertion, it is

$$i \xrightarrow{\max} k \xrightarrow{\max} j.$$

State k is inserted only if adding it does not decrease the time in state j . If $(\bar{r}_k > \bar{r}_i) \wedge (\bar{r}_k < \bar{r}_j) \wedge (\max(\mathcal{T}_{k \rightarrow j}) \leq \max(\mathcal{T}_{i \rightarrow j})) \wedge (\max(\mathcal{T}_{k \rightarrow j}) > \max(\mathcal{T}_{i \rightarrow k}))$, the state k can be inserted between i and j as a hazard to decrease the body voltage. The cycle before and after insertion is the same as in the maximum case.

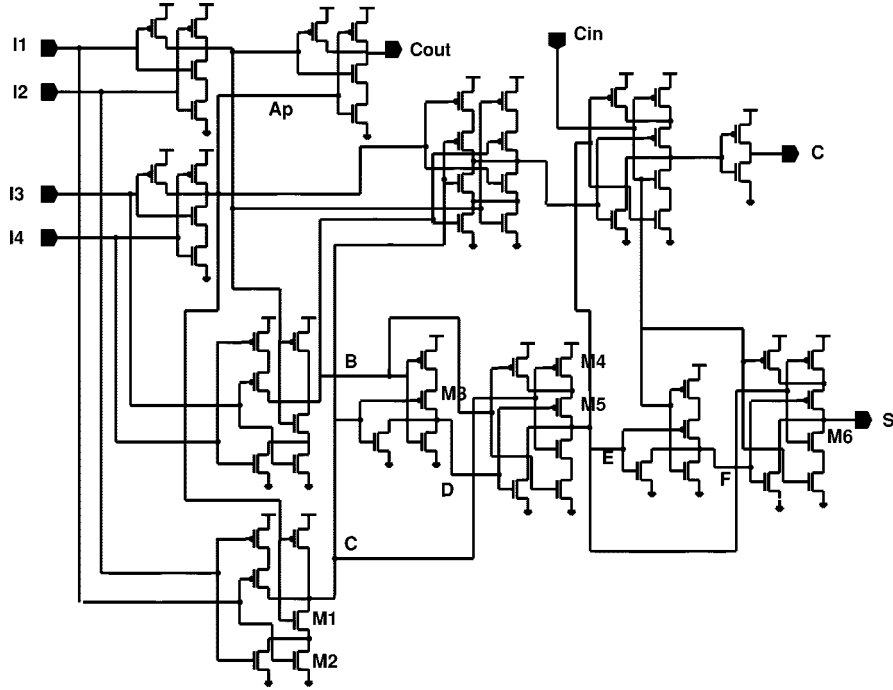


Fig. 11. 4-2 compressor circuit.

TABLE II

INITIAL-CONDITION DELAYS FOR THE COMPRESSOR FROM CIRCUIT SIMULATION ARE COMPARED WITH THE SOI-STA-DETERMINED FULL-UNCERTAINTY AND INITIAL-CONDITION DELAYS

	SOI-STA full-uncertainty				SPICE initial			
	short path		long path		short path		long path	
	rise	fall	rise	fall	rise	fall	rise	fall
I2	100	100	100	100	100	100	100	100
I4	100	100	100	100	100	100	100	100
Ap			302	491			235	315
C	302	317	866	712	326	379	476	432
D			1057	1079			723	553
E	526	533	1421	1359	743	762	785	943
F			1620	1697			1112	936
S	672	685	1926	1829	927	837	1104	1174

These preliminaries lead to a straightforward algorithm for determining the $t_k^{i \rightarrow j}$ to maximize or minimize the body voltage for a given transition. For the maximum case (the minimum case is the same except the complementary ranks \bar{r}_i are used), the following applies.

- 1) If $r_j > r_k$, then the starting cycle is

$$j \xrightarrow{\min} k$$

else the starting cycle is

$$j \xrightarrow{\max} k.$$

- 2) Find the accessible state k' with the largest rank (different from j and k and not previously inserted or attempted) that can be inserted between states j and k . If no such state exists, then exit with the current cycle. However, if such a state exists then the new cycle is either $j \xrightarrow{\max} k' \xrightarrow{\min} k$, $j \xrightarrow{\max} k' \xrightarrow{\max} k$ or $j \xrightarrow{\min} k' \xrightarrow{\min} k$, depending on which case led to the insertion.
- 3) Repeat step 2 for each transition in the current average cycle. Repeat this until no further refinement is possible.

Example: We calculate $(t_i^{1 \rightarrow 5a})^{\max}$ for an nFET at the 2.5-V supply, assuming that all states are accessible; that is, $\forall i, a_i = 1$. Let us further assume that $t_{\text{cycle}} = 10$ ns, the arrival times in this specific example are as shown in Table IV, and the r_i are those derived from our example technology. In this case, we begin with the transition $1 \xrightarrow{\max} 5a$ since $r_5 > r_1$. We next try to insert the accessible state with the largest rank (4) between 1 and $5a$. This insertion corresponds to Case 1, since $r_4 > r_5 > r_1$ and is possible since $\max(\mathcal{T}_{1 \rightarrow 4}) = \max(G_{\text{fall}}^{\text{early}}, S_{\text{rise}}^{\text{early}}, D_{\text{rise}}^{\text{early}}) = 2$ is less than $\min(\mathcal{T}_{4 \rightarrow 5a}) = G_{\text{fall}}^{\text{late}} = 7$. At this point, the current cycle is

$$1 \xrightarrow{\max} 4 \xrightarrow{\min} 5a.$$

We next attempt to insert state 2 between 1 and 4 (Case 3, $r_1 < r_2 < r_4$). This is not possible since although $\max(\mathcal{T}_{1 \rightarrow 2}) = \max(S_{\text{rise}}^{\text{early}}, D_{\text{rise}}^{\text{early}}) = 2$ equals $\max(\mathcal{T}_{1 \rightarrow 4}) = 2$, $\max(\mathcal{T}_{2 \rightarrow 4}) = G_{\text{fall}}^{\text{early}} = 1$ is less than $\max(\mathcal{T}_{1 \rightarrow 2}) = 2$. Similarly, state 2 cannot be inserted between 4 and $5a$ (Case 2, $r_5 < r_2 < r_4$ because although $\min(\mathcal{T}_{4 \rightarrow 2}) = G_{\text{fall}}^{\text{late}} = 7$ equals $\min(\mathcal{T}_{4 \rightarrow 5a}) = 7$, $\min(\mathcal{T}_{2 \rightarrow 5a}) = \min(G_{\text{fall}}^{\text{late}}, S_{\text{fall}}^{\text{late}}) = 7$ is not greater than $\min(\mathcal{T}_{4 \rightarrow 2}) = 7$. We next try to insert 3 between 1 and 4 and between 4 and $5a$ in the same way and find that only the former is possible, yielding the cycle

$$1 \xrightarrow{\max} 3 \xrightarrow{\max} 4 \xrightarrow{\min} 5a.$$

Last, we attempt to insert $5a$ between 1 and 3 and find that this is not possible; therefore, the final cycle is

$$1 \xrightarrow{\max} 3 \xrightarrow{\max} 4 \xrightarrow{\min} 5a.$$

From this, we find $t_1^{1 \rightarrow 5a} = \max(\mathcal{T}_{1 \rightarrow 3}) = G_{\text{fall}}^{\text{early}} = 1$, $t_3^{1 \rightarrow 5a} = \max(\mathcal{T}_{3 \rightarrow 4}) - \max(\mathcal{T}_{1 \rightarrow 3}) = \max(S_{\text{rise}}^{\text{early}}, D_{\text{rise}}^{\text{early}}) - G_{\text{fall}}^{\text{early}} = 1$, $t_4^{1 \rightarrow 5a} = \min(\mathcal{T}_{4 \rightarrow 5a}) - \max(\mathcal{T}_{3 \rightarrow 4}) = S_{\text{fall}}^{\text{late}} - \max(S_{\text{rise}}^{\text{early}}, D_{\text{rise}}^{\text{early}}) = 7$, $t_{5a}^{1 \rightarrow 5a} = t_{\text{cycle}} - \min(\mathcal{T}_{4 \rightarrow 5a}) = t_{\text{cycle}} - S_{\text{fall}}^{\text{late}} = 1$, and $t_2^{1 \rightarrow 5a} = t_{5b}^{1 \rightarrow 5a} = 0$.

TABLE III
STEADY-STATE DELAYS FOR THE COMPRESSOR FROM CIRCUIT SIMULATION ARE COMPARED WITH THE SOI-STA-DETERMINED DETAILED AND ACCESSIBILITY STEADY-STATE DELAYS

	SOI-STA detailed				SOI-STA accessibility				SPICE steady-state			
	short path		long path		short path		long path		short path		long path	
	rise	fall	rise	fall	rise	fall	rise	fall	rise	fall	rise	fall
I2	100	100	100	100	100	100	100	100	100	100	100	100
I4	100	100	100	100	100	100	100	100	100	100	100	100
Ap			267	409			296	449			266	362
C	464	485	706	605	389	365	804	673	488	503	594	576
D			897	877			1007	999			784	794
E	970	985	1168	1146	733	746	1332	1287	1040	1017	1059	1099
F			1363	1409			1540	1590			1246	1352
S	1372	1347	1610	1534	1026	986	1812	1737	1384	1365	1508	1483

TABLE IV
ARRIVAL TIMES FOR THE nFET IN THE EXAMPLE CALCULATION OF $(t_{i \rightarrow 5a}^1)_{\max}$

Arrival time	Time (nsec)	Arrival time	Time (nsec)
G_{rise}^{early}	1	G_{rise}^{late}	3
G_{fall}^{early}	1	G_{fall}^{late}	7
D_{rise}^{early}	2	D_{rise}^{late}	9
S_{rise}^{early}	2	S_{rise}^{late}	5
D_{fall}^{early}	2	D_{fall}^{late}	9
S_{fall}^{early}	2	S_{fall}^{late}	9

APPENDIX II

DETERMINING FET SIGNAL PROBABILITIES AND ARRIVAL TIME VALUES

We first consider the calculation of the signal probabilities. If we let i and j denote two channel nodes of the CCC, then similar to [25], we can define the k th path $P_{i,j}^k$ as one connection of FETs between i and j . We can also define a path function $f_{P_{i,j}^k}$ as a Boolean function indicating whether the k th path is conducting. Let n_i denote a controlling nFET gate input function in the path and let p_i denote a controlling pFET gate input function in the path. Then, the path function is given by

$$f_{P_{i,j}^k} = \bigwedge_{n_i \in P_{i,j}^k} n_i \wedge \bigwedge_{p_i \in P_{i,j}^k} \overline{p_i}.$$

If there are N paths between i and j , then the total path function $f_{P_{i,j}}$ is given by

$$f_{P_{i,j}} = \bigvee_{\forall P_{i,j}^k} f_{P_{i,j}^k}.$$

The path probability $P(P_{i,j})$, the probability that at the end of the cycle, the path from i to j is conducting, follows from elementary probability theory or binary decision diagram analysis [23]. The source and drain conditional probabilities required for the detailed body voltage estimation are given by specific path probabilities. For example, $P(D|\overline{G})$ is the path probability between the drain node and V_{DD} with the gate of the target transistor low.

We next consider calculating the FET arrival times which determine the temporal circuit environment of each FET. G_{fall}^{early} , G_{rise}^{early} , G_{fall}^{late} , G_{rise}^{late} are determined directly from the CCC arrival times. To determine S_{rise}^{early} , D_{rise}^{early} , D_{rise}^{late} , S_{rise}^{late} , we trace

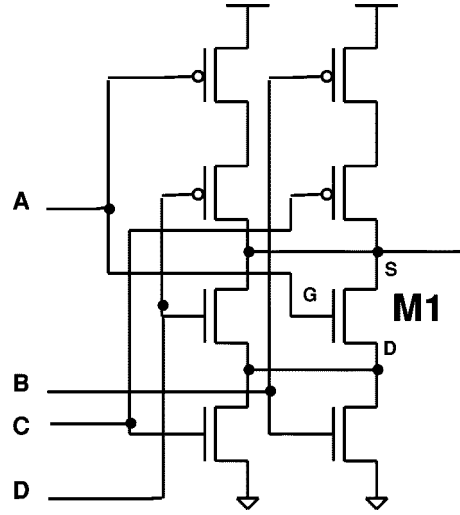


Fig. 12. Example for determining the channel signal probabilities and channel arrival times.

all the paths from the channel node (node i) to V_{DD} (node 1) with the target transistor off

$$\forall P_{i,1}: S_{rise}^{early} = \min \left(\forall n_i \in P_{i,1}, \forall p_i \in P_{i,1}: \left(G_{rise}^{early} \right)_{n_i}, \left(G_{fall}^{early} \right)_{p_i} \right)$$

$$\forall P_{i,1}: S_{rise}^{late} = \max \left(\forall n_i \in P_{i,1}, \forall p_i \in P_{i,1}: \left(G_{rise}^{late} \right)_{n_i}, \left(G_{fall}^{late} \right)_{p_i} \right)$$

with identical expressions for D_{rise}^{early} and D_{rise}^{late} . These relations determine the earliest or latest time a conducting path from node i to V_{DD} can be established if at the beginning of the cycle no such path exists. Similarly, to determine S_{fall}^{early} , D_{fall}^{early} , D_{fall}^{late} , S_{fall}^{late} , we trace all the paths from the channel node (node i) to ground (node 0) with the target transistor off

$$\forall P_{i,0}: S_{fall}^{early} = \min \left(\forall n_i \in P_{i,0}, p_i \in P_{i,0}: \left(G_{rise}^{early} \right)_{n_i}, \left(G_{fall}^{early} \right)_{p_i} \right)$$

$$\forall P_{i,0}: S_{fall}^{late} = \max \left(\forall n_i \in P_{i,0}, p_i \in P_{i,0}: \left(G_{rise}^{late} \right)_{n_i}, \left(G_{fall}^{late} \right)_{p_i} \right)$$

with identical expressions for $D_{\text{fall}}^{\text{early}}$ and $D_{\text{fall}}^{\text{late}}$. These relations determine the earliest or latest time a conducting path from node i to ground can be established if at the beginning of the cycle no such path exists. In cases in which there are no paths, we set early arrival times to t_{cycle} and late arrival times to zero.

Example: Consider transistor M1 in the CCC of Fig. 12. The channel conditional probabilities are

$$\begin{aligned} P(S|G) &= P(D|G) = (1 - P(B))(1 - P(C)) \\ P(S|\bar{G}) &= (1 - P(B))(1 - P(C)) + (1 - P(D)) \\ &\quad - (1 - P(D))(1 - P(B))(1 - P(C)) \\ P(D|\bar{G}) &= P(D)(1 - P(B))(1 - P(C)). \end{aligned}$$

The arrival times for the source, for example, are

$$\begin{aligned} S_{\text{rise}}^{\text{early}} &= \min(A_{\text{fall}}^{\text{early}}, B_{\text{fall}}^{\text{early}}, C_{\text{fall}}^{\text{early}}, D_{\text{fall}}^{\text{early}}) \\ S_{\text{rise}}^{\text{late}} &= \max(A_{\text{fall}}^{\text{late}}, B_{\text{fall}}^{\text{late}}, C_{\text{fall}}^{\text{late}}, D_{\text{fall}}^{\text{late}}) \\ S_{\text{fall}}^{\text{early}} &= \min(D_{\text{rise}}^{\text{early}}, C_{\text{rise}}^{\text{early}}, B_{\text{rise}}^{\text{early}}) \\ S_{\text{fall}}^{\text{late}} &= \max(D_{\text{rise}}^{\text{late}}, C_{\text{rise}}^{\text{late}}, B_{\text{rise}}^{\text{late}}). \end{aligned}$$

ACKNOWLEDGMENT

The authors would like to thank C. T. Chuang, R. Puri, G. Sai-Halasz, and M. R. Rosenfield of IBM Yorktown for encouragement, helpful discussions, and preprints of their work.

REFERENCES

- [1] C. T. Chuang, P.-F. Lu, and C. J. Anderson, "SOI for digital CMOS VLSI: Design considerations and advances," *Proc. IEEE*, vol. 86, pp. 689–720, Apr. 1998.
- [2] M. Canada *et al.*, "A 580MHz RISC microprocessor in SOI," in *IEEE Int. Solid-State Circuits Conf. Digest Tech. Papers*, Feb. 1999, pp. 430–431.
- [3] D. H. Allen *et al.*, "A 0.20 μm 1.8 V SOI 550 MHz 64 b PowerPC microprocessor with Cu interconnects," in *IEEE Int. Solid-State Circuits Conf. Digest Tech. Papers*, Feb. 1999, pp. 438–439.
- [4] C. T. Chuang and R. Puri, "SOI digital CMOS VLSI—A design perspective," in *36th ACM/IEEE Design Automation Conf.*, June 1999, pp. 709–714.
- [5] P.-F. Lu *et al.*, "Floating-body effects in partially depleted SOI CMOS circuits," *IEEE J. Solid-State Circuits*, vol. 23, pp. 1241–1253, Aug. 1997.
- [6] J. Gautier and J. Y.-C. Sun, "On the transient operation of partially depleted SOI NMOSFETs," *IEEE Electron Device Lett.*, vol. 16, pp. 497–499, Nov. 1995.
- [7] J. G. Fossum, R. Sundaresan, and M. Matloubian, "Anomalous sub-threshold current-voltage characteristics of n-channel SOI MOSFETs," *IEEE Electron Device Lett.*, vol. EDL-8, pp. 544–546, Nov. 1987.
- [8] H.-K. Lim and J. G. Fossum, "Current-voltage characteristics of thin-film SOI MOSFETs in strong inversion," *IEEE Trans. Electron Devices*, vol. ED-31, pp. 401–408, Apr. 1984.
- [9] L. T. Su, D. A. Antoniadis, M. I. Flik, and J. E. Chung, "Measurement and modeling of self-heating effects in SOI nMOSFETs," in *Proc. Int. Electron Devices Meeting*, Dec. 1992, pp. 357–360.
- [10] L. T. Su, J. B. Jacobs, J. Chung, and D. A. Antoniadis, "Deep-submicrometer channel design in silicon-on-insulator (SOI) MOSFETs," *IEEE Electron Device Lett.*, vol. 15, pp. 183–185, May 1994.
- [11] R. Puri and C. T. Chuang, "Hysteresis effect in pass-transistorbased partially-depleted SOI CMOS circuits," in *Proc. Int. SOI Conf.*, Oct. 1998, pp. 103–104.
- [12] A. Wei, D. A. Antoniadis, and L. A. Bair, "Minimizing floating-body-induced threshold voltage variation in partially depleted SOI CMOS," *IEEE Electron Device Lett.*, vol. 17, pp. 391–394, Aug. 1996.

- [13] M. M. Pelella, C. T. Chuang, J. G. Fossum, C. Tretz, B. W. Curran, and M. G. Rosenfield, "Hysteresis in floating-body PD/SOI circuits," in *Proc. Int. Symp. VLSI Technology, Systems, Applications*, Taipei, Taiwan, June 1999, pp. 278–281.
- [14] K. L. Shepard and D. Kim, "Body-voltage estimation in digital PD-SOI circuits and its application to static timing analysis," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, Nov. 1999, pp. 531–538.
- [15] R. B. Hitchcock, G. L. Smith, and D. D. Cheng, "Timing analysis for computer hardware," *IBM J. Res. Develop.*, vol. 26, no. 1, pp. 100–105, 1982.
- [16] N. P. Jouppi, "Timing analysis and performance improvement of MOS VLSI designs," *IEEE Trans. Computer-Aided Design*, vol. CAD-6, pp. 650–665, July 1987.
- [17] T. M. Burks, K. A. Sakallah, and T. N. Mudge, "Critical paths in circuits with level-sensitive latches," *IEEE Trans. VLSI Syst.*, vol. 3, pp. 273–291, June 1995.
- [18] K. L. Shepard, V. Narayanan, and R. Rose, "Harmony: Static noise analysis for deep-submicron digital integrated circuits," *IEEE Trans. Computer-Aided Design*, vol. 18, pp. 1132–1150, Aug. 1999.
- [19] *BSIM3SOI Manual*, Univ. California, Berkeley, CA, 1998, ver. 1.3.
- [20] G. G. Shahidi *et al.*, "SOI for 1-volt CMOS technology and application to a 512kb SRAM with 3.5 ns access time," in *Proc. Int. Electron Devices Meeting*, Dec. 1993, pp. 813–816.
- [21] S. A. Parke, J. E. Moon, H. C. Wann, P. K. Ko, and C. Hu, "Design for suppression of gate-induced drain leakage in LDD MOSFETs using a quasitwo-dimensional analytical model," *IEEE Trans. Electron Devices*, vol. 39, pp. 1697–1703, July 1992.
- [22] A. Wei and D. A. Antoniadis, "Design methodology for minimizing hysteretic VT-variation in partially-depleted SOI CMOS," in *Proc. Int. Electron Devices Meeting*, Dec. 1997, pp. 411–414.
- [23] F. N. Najm, "A survey of power estimation techniques in VLSI circuits," *IEEE Trans. VLSI Syst.*, vol. 2, pp. 446–455, Dec. 1994.
- [24] G. Goto, T. Sato, M. Nakajima, and T. Sukemura, "A 54-by-54-b regularly structured tree multiplier," *IEEE J. Solid-State Circuits*, vol. 27, p. 1229, Sept. 1992.
- [25] A. Kuehlmann, A. Srinivasan, and D. P. Lapotin, "Verity—A formal verification program for custom CMOS circuits," *IBM J. Res. Develop.*, vol. 39, no. 1/2, pp. 149–165, 1995.

Kenneth L. Shepard (S'85–M'91) received the B.S.E. degree from Princeton University, Princeton, NJ, in 1987 and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1988 and 1992, respectively.

From 1992 to 1997, he was a Research Staff Member and Manager in the VLSI Design Department at the IBM T. J. Watson Research Center, Yorktown Heights, NY, where he was responsible for the design methodology for IBM's G4 S/390 microprocessors. Since 1997, he has been at Columbia University, where he is now an Associate Professor. He also serves as Chief Technology Officer of CadMOS Design Technology, San Jose, CA. Current research interests include design tools for advanced CMOS technology, on-chip test and measurement circuitry, low-power design techniques for digital signal processing, low-power intrachip communications, and CMOS imaging applied to biological applications.

Dr. Shepard received the Fannie and John Hertz Foundation Doctoral Thesis Prize in 1992, IBM Research Division Awards in 1995 and 1998, a National Science Foundation CAREER Award in 1998, IBM University Partnership Awards in 1998, 1999, and 2000, and the 1999 Distinguished Faculty Teaching Award from the Columbia Engineering School Alumni Association. He is an Associate Editor of IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS, is a program track chair for the International Conference on Computer Design, and serves on the program committee for the International Conference on Computer-Aided Design.

Dae-Jin Kim (S'96) received the B.E. degree in electrical engineering from the Cooper Union, New York, NY, in 1997 and the M.S. degree in electrical engineering from Columbia University, New York, NY, in 1999. He is currently working toward the Ph.D. degree in electrical engineering at the same school.

He is currently working at Manhattan Routing, a design-services startup in New York, NY, during the 2000 to 2001 academic year. His current research interest are digital VLSI design and computer-aided design.