Timing macromodels for CMOS static set/reset latches and their applications

C.-Y. Wu
C. Li
J.S. Hwang

Indexing terms: Modelling, Logic

Abstract: Efficient timing macromodels for CMOS static NAND-type and NOR-type latches are developed, to compute analytically their signal timing under different input state transitions. The timing equations in the macromodels are derived from the effective dominant pole of the linearised large-signal equivalent circuit of a latch under the characteristic-waveform consideration. Through extensive comparisons with SPICE simulations, it is found that the macromodels have a maximum error of 22% for the total propagation delay times of the latches, with different device sizes, capacitive loads, device parameter variations, noncharacteristic-waveform input excitations and input-state transitions. When incorporated with the timing models of CMOS combinational logic gates, the macromodels can also be applied to characterise the signal timing of static sequential integrated circuits. Application examples on two CMOS clocked flip-flops and experimental verifications on a fabricated CMOS master-slave T flip-flop are successfully made to confirm the accuracy and applicability of the developed macromodels. Reasonable accuracy, wide applicable ranges and CPU-time, and memory efficiency have made the macromodels very attractive in many CAD applications.

List of principal symbols

$C_{bd(bs)}$ = bulk-drain (bulk-source) *pn* junction capacitance of a MOSFET
$C_{BD(BS)}$ = linearised bulk-drain (bulk-source) *pn* junction capacitance of a MOSFET
C_{gb} = gate-bulk capacitance of a MOSFET
$C_{gd(gs)}$ = gate-drain (gate-source) capacitance of a MOSFET
C_{GBOV} = gate-bulk overlap capacitance per unit channel width (SPICE device parameter)
$C_{GDOV(GSOV)}$ = gate-drain (gate-source) overlap capacitance per unit channel width (SPICE device parameter)
C_L = fixed load capacitance of a logic gate
DELTA = channel width factor (SPICE device parameter)
GAMMA = bulk threshold parameter in SPICE, which represents the proportionality factor relating the change in threshold voltage to backgate bias
I_d = drain current of a MOSFET
L = effective or electrical channel length of a MOSFET
L_{mask} = mask channel length of a MOSFET
N = fan-out number
$NSUB$ = substrate doping concentration (SPICE device parameter)
$P_{f(r)}$ = effective dominant pole in the fall (rise) characteristic waveform case
q = magnitude of electronic charge
$t_{df(df2)}$ = initial fall delay times of the voltage waveform $V_{21}(V_{23})$
T_{ox} = channel oxide thickness
UCRIT = critical field for mobility degradation (SPICE device parameter)
UEXP = critical field exponent in mobility degradation (SPICE device parameter)
UO = surface mobility (SPICE device parameter)
UTRA = horizontal field factor in mobility degradation (SPICE device parameter)
V_{BS} = bulk-source reverse bias of a MOSFET
$V_{DS(GS)}$ = drain-source (gate-source) voltage of a MOSFET
V_{FB} = flat-band voltage
V_{TO} = zero-bias threshold voltage of a MOSFET (SPICE device parameter)
W = effective or electrical channel width of a MOSFET
XJ = metallurgical junction depth of a MOSFET (SPICE device parameter)
$\varepsilon_{SI(SIO_2)}$ = permittivity of Si semiconductor (silicon dioxide)
μ'_S = linearised carrier mobility
ϕ_F = Fermi potential

Paper 6007E (C2) first received 18th September 1987, and in revised form 2nd February 1988

The authors are with the Department of Electronics Engineering, Institute of Electronics, National Chiao Tung University, 75 Po-Ai Street, Hsin-Chu, Taiwan, Republic of China

1 Introduction

It is known that the set/reset (S/R) latch is one of the commonly-used building blocks in static sequential circuits; it serves as a basic core in a static flip-flop. Generally an S/R latch can be formed by cross-coupling two NOR gates or NAND gates. Both types of S/R latches have complex regenerative feedback paths which may cause difficulty in numerical convergence, or lead to too much CPU-time consumption in transient simulations using SPICE or other circuit simulators. This problem becomes worse for complex VLSI circuits which may contain more latches.

It is the aim of this paper to solve this problem by developing a general timing macromodel for CMOS S/R latches. In this modelling approach, the large-signal equivalent circuit of a CMOS latch is first constructed according to the characteristic waveform [1, 2] consideration. Then the circuit is linearised by using a similar technique as in the case of CMOS combinational logic gates [2]. From the linearised circuit, the effective dominant pole can be found by using the dominant-pole-dominant-zero (DPDZ) technique [2]. Then the signal timing of the latch can be explicitly expressed in terms of various device and circuit parameters. These expressions form the timing macromodels of CMOS S/R latches.

Applying the general timing macromodels, the signal timing of various CMOS S/R latches with different MOS channel dimensions, capacitive loads, device parameters and input excitation waveforms can be quickly calculated with satisfactory accuracy. Moreover, the developed macromodels can be applied to analyse the speed characteristics of the latches, calculate the signal timing of various static CMOS flip-flops, and determine suitable device channel dimensions from a given set of timing specifications.

2 Macromodel construction

Since the actual chip internal voltage waveforms are some sorts of characteristic waveforms [1–2], the timing macromodels to be developed for CMOS static S/R latches are based on the characteristic-waveform considerations [2]. The resultant macromodels, however, can be applied to the noncharacteristic waveform case. This makes the macromodels more practical and versatile in analysing the actual chip timing.

As an illustrative example, a CMOS NAND-type S/R latch will be modelled in this Section. The characteristic waveform of the latch can be generated from SPICE transient simulations on a chain of identical latches with the same capacitive loads, as shown in Fig. 1. Generally, the desired characteristic waveform, which is independent of any input excitations and nearly the same in each intermediate stage, can be obtained after three or four stages from the excited input port. Typical characteristic waveforms are shown in Fig. 2, where the voltage waveform at each node of the driving stage or its load stage is denoted by the corresponding node number indicated in the circuit of Fig. 1.

As may be seen from Figs. 1 and 2, the rising voltage V_{17}, which is the input voltage to the right NAND gate of the driving stage, has negligible effect on the output voltage V_{21} because the voltage V_{20} at that time is kept at a low voltage, to turn off the NMOS M_{N3} and to maintain the voltage V_{21} at V_{DD}. The only effective triggering input voltage to the driving stage, therefore, is the falling voltage V_{16}. Due to the excitation of the voltage V_{16} at the left NAND gate, its output voltage V_{20} has a characteristic rising waveform, with the characteristic rise time T_r defined as the true interval from $V_{20} = 0.1\ V_{DD}$ to $V_{20} = 0.9\ V_{DD}$. This rise time T_r will be characterised. Since the characteristic waveforms appear among those stages, the waveform of the voltage V_{17} is the same as that of the voltage V_{20}.

In the load stage shown in Fig. 1, the voltage V_{20} turns on the NMOS M_{N4L} and thus discharges the voltage V_{26} to 0 V. However, it has negligible effect on the voltage V_{24}, just as the voltage V_{17} does on the voltage V_{21}. In the driving stage, the voltage V_{20} drives the right NAND gate to lower the voltage V_{21}. This falling voltage V_{21} is then fed back to the gate of the NMOS M_{N1}. Since the voltage V_{16} has already been lowered toward 0 V, when the voltage V_{21} starts to decrease from V_{DD}, the voltage

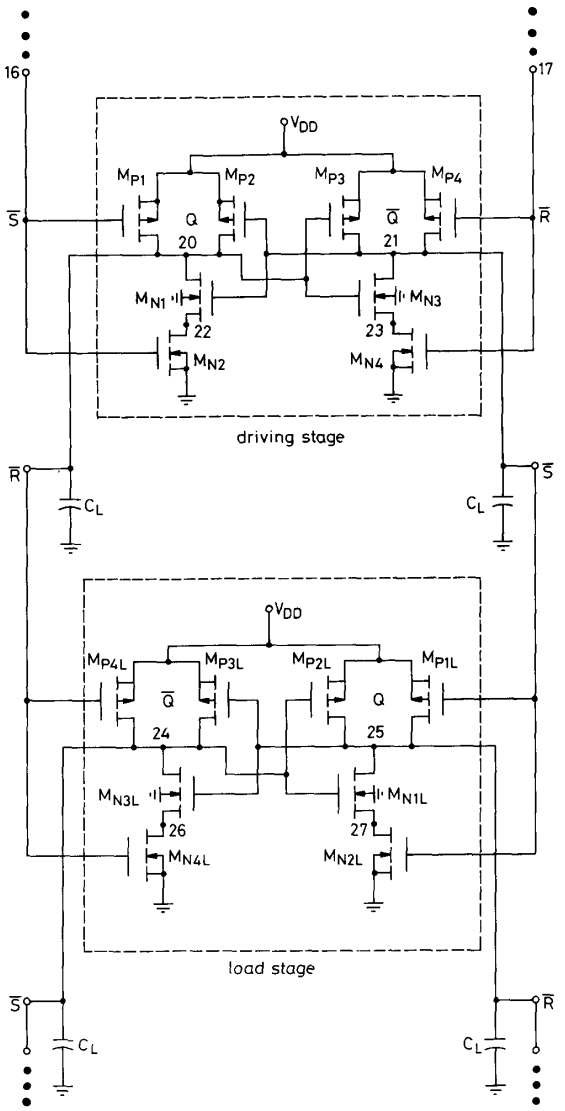

Fig. 1 *A driving stage and a load stage within a string of identical CMOS static NAND-type latches*

V_{20} is nearly independent of the feedback signal V_{21}. The corresponding rise time T_r, therefore, can be modelled by considering the left NAND gate, with the driving signal V_{16} but with the feedback signal V_{21} set to a constant level.

Based on similar considerations, the characteristics fall time T_f, defined as the time interval from $V_{21} = 0.9\ V_{DD}$ to $V_{21} = 0.1\ V_{DD}$, can be characterised by considering the right NAND gate with the input voltage V_{20} and with another input voltage V_{17} set to V_{DD}. Note that both V_{21} and V_{16} have the same characteristic fall waveforms.

To find the equivalent circuit for the rise-time calculation, the transient behaviour of each node voltage during the rise time must first be investigated. It is found that during this interval the voltages V_{17}, V_{23} and V_{24} are either kept at constant levels or changed slowly. The voltage V_{21} is set to a constant voltage as mentioned

above. Since these slowly-changing or constant voltages, together with the power supply voltage, have negligible effect or no effect on the transient behaviour [2], they are

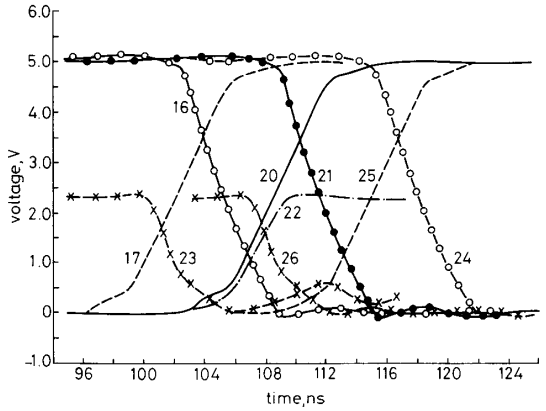

Fig. 2 *Typical characteristic waveforms of a CMOS NAND-type latch*

shorted to ground. The resultant circuit is shown in Fig. 3, which will be used to generate the equivalent circuit.

The operating regions of all the MOSFETs in Fig. 3 can be determined by comparing the drain-source voltage

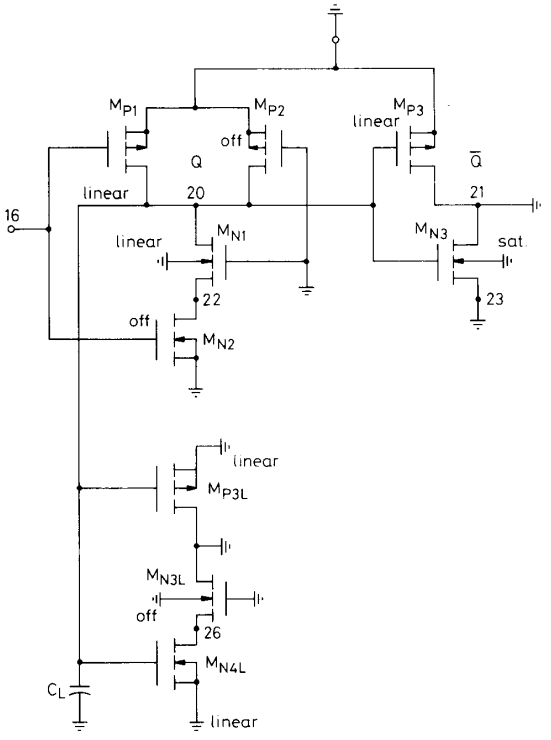

Fig. 3 *The MOS circuit used to characterise the characteristic rise time of a NAND-type latch*

V_{DS} with the simulated drain-source saturation voltage V_{DSAT} during the whole interval T_r. In some MOSFETs two operating regions are involved during the interval T_r. To simplify the calculation, only one region is considered. It is found that such a simplification is a good compromise between calculation complexity and calculation error. The determined operating regions are indicated in the circuit in Fig. 3.

By using the large-signal equivalent circuit of a

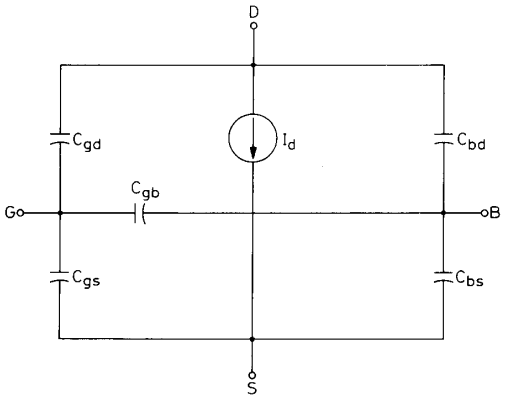

Fig. 4 *Large-signal equivalent circuit of a MOSFET*

MOSFET in different operation regions (Fig. 4 and Table 1) [4], the overall equivalent circuit of the circuit in Fig. 3 is generated, and is shown in Fig. 5*a*. In this circuit, the capacitances C_1, C_2, C_3 and C_4 can be expressed in terms of device capacitances and load capacitance. The expressions are given in Table 2.

Table 1: Expressions of various gate capacitances in different operating regions

Capacitances	Linear region	Saturation region	Off region
C_{gs}	$C_{GSOV}W + C_0WL/2$	$C_{GSOV}W + 2C_0WL/3$	$C_{GSOV}W$
C_{gd}	$C_{GDOV}W + C_0WL/2$	$C_{GDOV}W$	$C_{GDOV}W$
C_{gb}	$C_{GBOV}W$	$C_{GBOV}W$	$C_{GBOV}W + C_0WL$

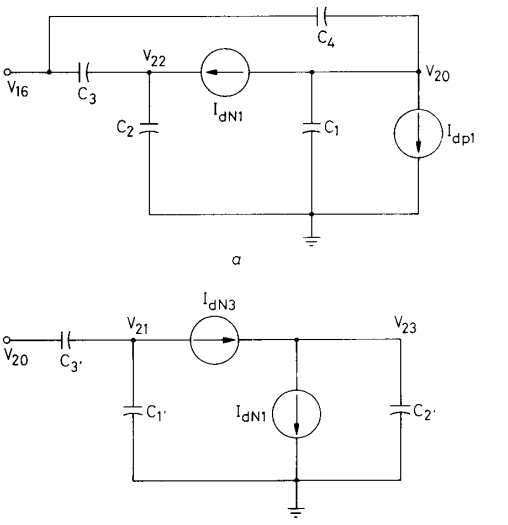

Fig. 5 *Linearised large-signal equivalent circuit used to characterise characteristic times*

a Rise time
b Fall time

That the characteristic waveforms are nearly independent of the input excitations implies that the output voltage V_{20} strongly depends on the poles or zeros of the

Table 2: Expressions of capacitances and conductance factors in the rise time case and the fall time case

Rise time case:

$$C_1 = C_{BDP1} + C_{BDP2} + C_{GDOVP}W_{P2} + C_{BDN1} + C_{ON}W_{N1}L_{N1}/2 + C_{GDOVN}W_{N1} + C_{GSOVP}W_{P3} + C_{GDOVP}W_{P3} + C_{OP}W_{P3}L_{P3} + C_{GSOVN}W_{N3} + C_{GDOVN}W_{N3} + 2C_{ON}W_{N3}L_{N3}/2 + C_L + NC_{1L}$$

$$C_{1L} = C_{GSOVP}W_{P3L} + C_{GDOVP}W_{P3L} + C_{OP}W_{P3L}L_{P3L} + C_{GSOVN}W_{N4L} + C_{GDOVN}W_{N4L} + C_{ON}W_{N4L}L_{N4L}$$

$$C_2 = C_{BSN1} + C_{BDN2} + C_{GSOVN}W_{N1} + C_{ON}W_{N1}L_{N1}/2$$

$$C_3 = C_{GDOVN}W_{N2}$$

$$C_4 = C_{GDOVP}W_{P1} + C_{OP}W_{P1}L_{P1}/2$$

$$\alpha_1 \equiv \beta_{P1}[7\eta_{P1}V_{DD}/8 - V_{BINSP1} - (\eta_{P1} - 1)V_{DD} - 2\gamma_{SP1}(2\phi_{FP1} + V_{DD}/4)^{1/2}/3 - V_{DD}4^{-P_f/P_r}\exp(-P_f t_{dr})]$$

$$\alpha_2 \equiv \beta_{N1}[V_{DD} - V_{BINSN1} - 7\eta_{N1}V_{DD}/8 - 2\gamma_{SN1}(2\phi_{FN1} + 3V_{DD}/4)^{1/2}/3]$$

$$\alpha_3 \equiv \beta_{N1}\{V_{DD} - V_{BINSN1} - 7(V_{DD} - V_{TNF})\eta_{N1}/8 - 2\gamma_{SN1}[2\phi_{FN1} + 3(V_{DD} - V_{TNF})/4]^{1/2}/3\}$$

Fall time case:

$$C'_1 = C_{BDP3} + C_{BDP4} + C_{GDOVP}W_{P4} + C_{BDN3} + C_{ON}W_{N1}L_{N1} + C_{GDOVN}W_{N1} + C_{OP}W_{P2}L_{P2}/2 + C_{GSOVP}W_{P2} + C_L + NC'_{2L}$$

$$C'_{2L} = C_{GSOVN}W_{N2L} + C_{GDOVN}W_{N2L} + C_{ON}W_{N2L}L_{N2L} + C_{GSOVP}W_{P2L} + C_{GDOVP}W_{P2L} + 2C_{OP}W_{P2L}L_{P2L}/3$$

$$C'_2 = C_{BDN4} + C_{GDOVN}W_{N4} + C_{ON}W_{N4}L_{N4}/2 + C_{BSN3} + C_{GSOVN}W_{N3} + C_{ON}W_{N3}L_{N3}/2$$

$$C'_3 = C_{GDOVP}W_{P3} + C_{GDOVN}W_{N3} + C_{ON}W_{N3}L_{N3}/2$$

$$\alpha'_1 \equiv \beta_{N3}[V_{DD} - V_{DD}4^{-P_f/P_r}\exp(-P_r t_{df}) - V_{BINSN3} - \eta_{N3}V_{DD}/8 - 2\gamma_{SN3}(2\phi_{FN3} + V_{DD}/4)^{1/2}/3]$$

$$\alpha'_2 \equiv \beta_{N3}[V_{DD} - V_{DD}4^{-P_f/P_r}\exp(-P_r t_{df}) - V_{BINSN3} - \eta_{N3}V_{DD}/20 - 2\gamma_{SN3}(2\phi_{FN3} + V_{DD}/10)^{1/2}/3]$$

$$\alpha'_3 \equiv \beta_{N4}[V_{DD} - V_{BINSN4} - \eta_{N4}V_{DD}/20 - 2\gamma_{SN4}(2\phi_{FN4} + V_{DD}/10)^{1/2}/3]$$

circuit in Fig. 5*a*. To characterise analytically the signal timing of the output signal V_{20}, through the poles and zeros, the nonlinear circuit in Fig. 5*a* must be linearised. A linearisation technique previously proposed [2] is adopted here to linearise the *pn* junction capacitance and the drain current.

After the linearisation point at $t = t_e$ is chosen, the corresponding gate-source, bulk-source and drain-source voltages V'_{GS}, V'_{BS} and V'_{DS} for each MOSFET can be determined. The *pn* junction capacitances C_{BD} and C_{BS} at the linearisation point can be calculated by using the formula in SPICE2 [3]. Setting those voltage-dependent capacitances in C_1 and C_2 expressions to their corresponding calculated constant values, all the nonlinear capacitances become linear and have fixed values.

In the drain current linearisation, the linear-region drain-current equation in SPICE2 [3] is modified by linearising the square-root terms [2], discarding constant terms and replacing the voltage-dependent mobility by its fixed value at the linearisation point. The resultant expressions are given in Table 3. By applying the equations in Table 3, I_{dP1} and I_{dN1} in Fig. 5*a* can be written as

$$I_{dP1} = \beta_{P1}[-V_{BINSP1} - (\eta_{P1} - 1)V_{DD} - 2\gamma_{SP1}(2\phi_{FP1} + V_{DD} - V'_{20})^{1/2}/3]V_{20} - \beta_{P1}V_{16}V_{20} + \beta_{P1}\eta_{P1}V_{20}^2/2 + \beta_{P1}V_{DD}V_{16} \quad (1)$$

$$I_{dN1} = \beta_{N1}[V_{DD} - V_{BINSN1} - 2\gamma_{SN1}(2\phi_{FN1} + V'_{20})^{1/2}/3]V_{20} - \beta_{N1}\eta_{N1}V_{20}^2/2 - \beta_{N1}[V_{DD} - V_{BINSN1} - 2\gamma_{SN1}(2\phi_{FN1} + V'_{22})^{1/2}/3]V_{22} - \beta_{N1}\eta_{N1}V_{22}^2/2 \quad (2)$$

In the I_{dN1} expression, V_{21} is set to V_{DD} and $V_{GSN1} = V_{21} - V_{22} = V_{DD} - V_{22}$.

To further linearise the product terms and the square terms in I_{dP1} and I_{dN1}, the functions of V_{16} and V_{20} in the time domain must be determined. According to the dominant-pole approximation, the output voltage in each stage is a single-pole response. Thus V_{16}, V_{20} and V_{22} can be written as

$$V_{16}(t) = V_{DD}\exp(-P_f t)u(t) \quad (3)$$

$$V_{20}(t) = V_{DD}\{1 - \exp[-P_r(t - t_{dr})]\}u(t - t_{dr}) \quad (4)$$

$$V_{22}(t) = (V_{DD} - V_{TNF})\{1 - \exp[-P_r(t - t_{dr2})]\} \times u(t - t_{dr2}) \quad (5)$$

where $t_{dr}(t_{dr2})$ is the initial rise time [2] between the voltage waveforms V_{16} and V_{20} (V_{22}). The voltage V_{TNF} is the threshold voltage of M_{N1} with substrate bias $V_{DD} - V_{TNF}$. It can be expressed as

$$V_{TNF} = -\{2[\gamma_{SN}(2\phi_{FN})^{1/2} - V_{TN}] + \gamma_{SN}^2\}/2 + \{[2\gamma_{SN}(2\phi_{FN})^{1/2} - 2V_{TN} + \gamma_{SN}^2]^2 + 4[2\gamma_{SN}V_{TN}(2\phi_{FN})^{1/2} + \gamma_{SN}V_{DD} - V_{TN}^2]\}^{1/2}/2 \quad (6a)$$

where

$$V_{TN} \equiv V_{BINN} + \gamma_{SN}(2\phi_{FN})^{1/2} \quad (6b)$$

Generally, the linearisation point can be adjusted to minimise the calculation error. It is found that the optimal position for the linearisation point is the centre point of the linear region of the MOSFET under consideration. In this case, the linearisation point is chosen to be at $V_{20} = 3V_{DD}/4$. Thus the time $t = t_e$ can be expressed as

$$t_e = t_{dr} + (\ln 4)/P_r \quad (7)$$

Once t_e is determined, the V'_{GS}, V'_{BS} and V'_{DS} in each MOSFET can be determined accordingly to calculate the *pn* junction capacitances, the mobilities and the drain currents.

By using the same technique [2], the terms V_{20}^2, $V_{16}V_{20}$ and V_{22}^2 in eqns. 1 and 2 can be linearised. The resultant linearised currents I_{dP1} and I_{dN1} are

$$I_{dP1} = \alpha_1 V_{20} \quad (8)$$

$$I_{dN1} = \alpha_2 V_{20} - \alpha_3 V_{22} \quad (9)$$

Table 3: Linearised MOSFET drain current equation

$$I_d = \beta\left[\left(V_{GS} - V_{BIN} - \frac{\eta}{2}V_{DS}\right)\cdot V_{DS} - \tfrac{2}{3}\gamma_S(2\phi_F + V'_{DS} - V'_{BS})^{1/2}(V_{DS} - V_{BS}) + \tfrac{2}{3}\gamma_S(2\phi_F - V'_{BS})^{1/2}V_{BS}\right]$$

where

$$\beta = \frac{W}{L}\mu'_s C_o$$

$$C_o = \varepsilon_{SIO_2}/T_{ox}$$

$$\mu'_s = U0\{UCRIT \cdot \varepsilon_{SI}/\{C_o[V'_{GS} - V_{TH} - UTRA \cdot \min(V'_{DS}, 2\phi_F)]\}\}^{UEXP}$$

$$V_{BIN} = V_{BI} + (\eta - 1)(2\phi_F - V_{BS})$$

$$V_{BINS} = V_{BI} + (\eta - 1)2\phi_F$$

$$V_{BI} = V_{FB} + 2\phi_F = V_{TO} - GAMMA \cdot \sqrt{2\phi_F}$$

$$\eta = 1 + (\pi \cdot DELTA \cdot \varepsilon_{SI})/(4C_o W)$$

$$\gamma_s = GAMMA \cdot (1 - \alpha_S - \alpha_D)$$

$$\alpha_S = \tfrac{1}{2}(XJ/L)[(1 + 2W_S/XJ)^{1/2} - 1]$$

$$\alpha_D = \tfrac{1}{2}(XJ/L)[(1 + 2W_D/XJ)^{1/2} - 1]$$

$$W_S = XD \cdot (2\phi_F - V'_{BS})^{1/2}$$

$$W_D = XD \cdot (2\phi_F - V'_{BS} + V'_{DS})^{1/2}$$

$$XD = (2\varepsilon_{SI}/q \cdot NSUB)^{1/2}$$

$$V_{TH} = V_{BIN} + \gamma_s \cdot (2\phi_F - V'_{BS})^{1/2}$$

$$V'_{DS} = V_{DS}(t)\Big|_{t=t_e}$$

$$V'_{BS} = V_{BS}(t)\Big|_{t=t_e}$$

$$V'_{GS} = V_{GS}(t)\Big|_{t=t_e}$$

t_e = time at the linearisation point

where the conductance factors α_1, α_2 and α_3 are expressed in Table 2. The factor $P_f t_{dr}$ in α_1 is nearly constant in different cases. It is, therefore, set to a fixed value of 0.7.

The linearisations of both capacitances and currents make the large-signal equivalent circuit in Fig. 5 a linear one. Its dominant pole P_D and dominant zero Z_D can thus be analytically expressed as [2]

$$1/P_D = C_1/\alpha_1 + [C_2/\alpha_3 + \alpha_2 C_2/\alpha_1\alpha_3]/2 \quad (10)$$

$$1/Z_D = C_2 C_4/2[\alpha_3(C_3 + C_4)] \quad (11)$$

According to the dominant-pole-dominant-zero (DPDZ) method [2], one can determine the effective characteristic rise pole P_r. Its expression is

$$1/P_r = 1/P_D - 1/Z_D \quad (12)$$

To calculate the characteristic fall pole of the voltage V_{21}, the right NAND gate in the driving stage with the input voltage V_{20} and the suitable loading is considered. The large-signal equivalent circuit can be similarly obtained as shown in Fig. 5*b*. Although the waveform of the voltage V_{23} is different for the circuits with different device dimensions, it is a falling waveform around the linearisation point at $V_{21} = V_{DD}/4$ and can be approximately characterised by the fall pole P_f and the initial delay t_{df2}. The expression is

$$V_{23}(t) = (V_{DD} - V_{TNF})[u(t) - u(t - t_{df2})] + (V_{DD} - V_{TNF})\exp[-P_f(t - t_{df2})] \times u(t - t_{df2}) \quad (13)$$

Since the voltage V_{23} is nearly equal to $V_{DD}/10$ at the linearisation point, the linearisation of V_{23}^2 can be done as

$$V_{23}^2(t) \rightarrow V_{DD}V_{23}(t)/10$$

Based upon the DPDZ method, the characteristic fall pole P_f can be expressed as

$$P_f = \{C'_1/\alpha'_1 + \alpha'_2 C'_1/(\alpha'_1\alpha'_3) + [C'_3/\alpha'_1 + \alpha'_2 C'_3/(\alpha'_1\alpha'_3) + C'_2/\alpha'_3 - C'_3/(\alpha'_2 + \alpha'_3)]/2\}^{-1} \quad (14)$$

where the expressions of the capacitances C'_1, C'_2 and C'_3 and the conductance factors α'_1, α'_2 and α'_3 are listed in Table 2. The factor $P_r t_{df}$ in α'_1 and α'_2 is set to a constant of 1.1. Because α_1 in the expression of P_r is a function of P_f whereas α'_1 and α'_2 in P_f is a function P_r, eqns. 12 and 14 must be solved together by using the numerical iteration.

The rise time T_r and the fall time T_f can be computed by using the solved P_r and P_f in the formula

$$T_r = (\ln 9)/P_r \quad (15)$$

$$T_f = (\ln 9)/P_f \quad (16)$$

The rise delay time T_{PLH}, defined as the time interval between $V_{16} = V_{DD}/2$ and $V_{20} = V_{DD}/2$ can be empirically determined as

$$T_{PLH} = X_1 T_r + X_2 T_f = 0.66T_r - 0.11T_f \quad (17)$$

where X_1 and X_2 are universal constants for different CMOS NAND-type S/R latches. They were determined to be 0.66 and -0.11, respectively. Note that T_{PLH} is the delay time, between $\bar{S}$ and Q, of the latch.

Similarly, T_{PHL}, the delay time between Q and $\bar{Q}$, can be expressed as

$$T_{PHL} = 0.73T_f - 0.05T_r \quad (18)$$

The pair delay T_p defined as the sum of T_{PLH} and T_{PHL} can be written as

$$T_p = 0.61T_r + 0.62T_f \quad (19)$$

The pair delay T_p is the delay between the input $\bar{S}$ and the output $\bar{Q}$, or equivalently between $\bar{R}$ and Q. It is the propagation delay of the latch.

In the NAND-type S/R latch, the ambiguous input state is 00 and the effective input excitation is a falling voltage. Thus only the following three input state transitions must be considered in characterising the delay time of the latches:

$$\bar{S}\bar{R}: 10 \leftrightarrow 01 \quad (20a)$$

$$11 \rightarrow 01 \quad (20b)$$

$$11 \rightarrow 10 \quad (20c)$$

For the transition in eqn. 20*a*, both the two inputs $\bar{S}$ and $\bar{R}$ are excited. This case was modelled as described above. For the other two transitions which are identical to each other because of the symmetric structure of the latch, the signal timing was similarly modelled with one input of the latch kept in the logic 1 state and the other excited by a falling voltage.

Based on the derived timing equations, a complete timing macromodel for the CMOS NAND-type S/R latch can be formed. Given the device dimensions, the rise/fall time of the input waveforms and the output loads of the latch, the rise, fall and delay times of the latch under all the possible excitations can be computed. On the other hand, the desired device sizes can be synthesised through the macromodel with the given timing specifications.

For the CMOS NOR-type S/R latch, the effective input excitation is a rising voltage, and the ambiguous input state is 11. The timing macromodel for the latch was similarly developed. Generally, the developed macromodels for both NAND-type and NOR-type latches have a reasonable accuracy and a wide applicable range, as will be verified in the following Section.

3 Macromodel verification

To check the accuracy and the generality of the timing macromodels, comparisons with SPICE simulation results were extensively made for the latches, with different device sizes, device parameters, capacitive loads and input excitations. Fig. 6*a* shows the comparisons on the

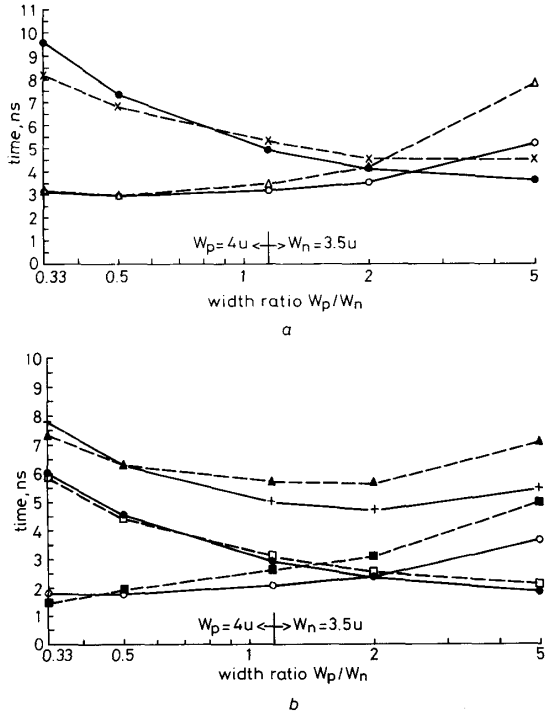

Fig. 6 *Calculated and simulated times and delays of characteristic waveforms in a CMOS NAND-type latch with $C_L = 0$ and under the two-input excitation*

a Rise and fall times
- rise (theory)
- time (ESPICE)
- fall (theory)
- time (ESPICE)

b Rise and fall delays, pair delays
- rise (theory)
- delay (ESPICE)
- fall (theory)
- delay (ESPICE)
- pair (theory)
- delay (ESPICE)

rise/fall time of the NAND-type latches under the two-input excitation of eqn. 20*a* and with $L_{mask} = 3.5\ \mu m$, $C_L = 0$ pF and different width ratios whereas Fig. 6*b* shows the corresponding comparisons on the rise/fall delay and the pair delay. The maximum error is 30% in the rise/fall time and 22% in the pair delay. Better accuracy is shown in the timing of the latches with commonly used device dimensions. For a large fixed capacitive load C_L up to 5 pF, the error decreases as shown in Figs. 7*a* and *b*.

For the NOR-type latches under the two-input excitation, the calculated timing has a similar error characteristic when compared with SPICE simulation results. Part of the comparisons are shown in Figs. 8*a* and *b* for $C_L =$ 0 pF.

The calculated signal timing of various NAND-type latches under the single-input excitation of eqns. 20*b* and *c* was compared with SPICE simulation results. Gener-

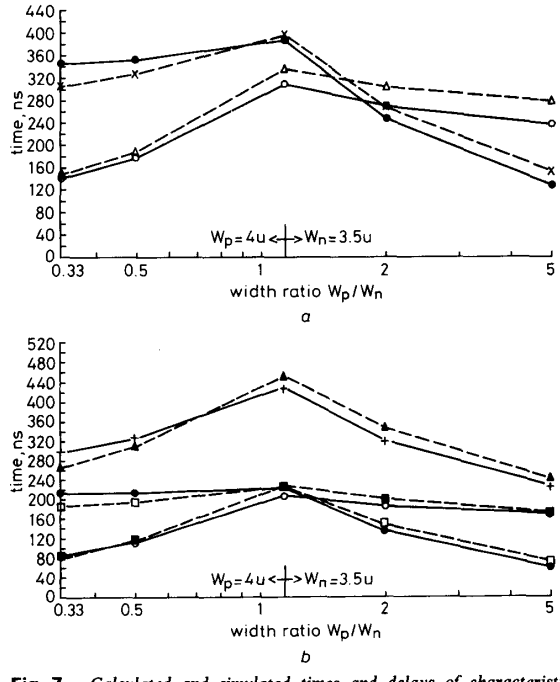

Fig. 7 *Calculated and simulated times and delays of characteristic waveforms for a CMOS NAND-type latch with $C_L = 5$ pF and under the two-input excitation*

a Rise and fall times
- rise (theory)
- time (ESPICE)
- fall (theory)
- time (ESPICE)

b Rise and fall delays, pair delays
- rise (theory)
- delay (ESPICE)
- fall (theory)
- delay (ESPICE)
- pair (theory)
- delay (ESPICE)

ally, the signal timing under the single-input excitation is close to that under the two-input excitation. Their error characteristics are also similar. Part of the comparison is listed in Table 4 where the comparisons on NOR-table latches are also made. All the latches considered in Table 4 have a minimum load of only one fanout gate and no C_L. This case generally shows a maximum error in the timing calculation.

To investigate the accuracy of the macromodels under device parameter variations, comparisons for the latches with different values of the zero-bias long-channel threshold voltage V_{TO} and mobility parameter UO were made. It is found that the error characteristics remain the same under large parameter variations. Part of the comparisons are shown in Fig. 9*a* for the NAND-type latches with V_{TO} down to 0.3 V. The corresponding comparisons for NOR-type latches are shown in Fig. 9*b*.

Although the macromodels are developed from the characteristic-waveform consideration, it can also be applied to the noncharacteristic-waveform case. Extensive comparisons between the calculated and the simulated timing data were performed for the NAND-type (NOR-type) latches under the input excitations, with the fall times (rise times) from 1 to 100 ns. For the CMOS latches with commonly used device dimensions the error of the timing macromodels is similar to that in the characteristic-waveform case, even when the input excitation waveforms greatly deviate from the characteristic waveforms. For the latches with $W_p/W_n = 0.33$ or 5, the

same error can be kept for the input waveforms not deviating much from the characteristic waveforms. Part of the comparisons are shown in Fig. 10*a*, (Fig. 10*b*) for the

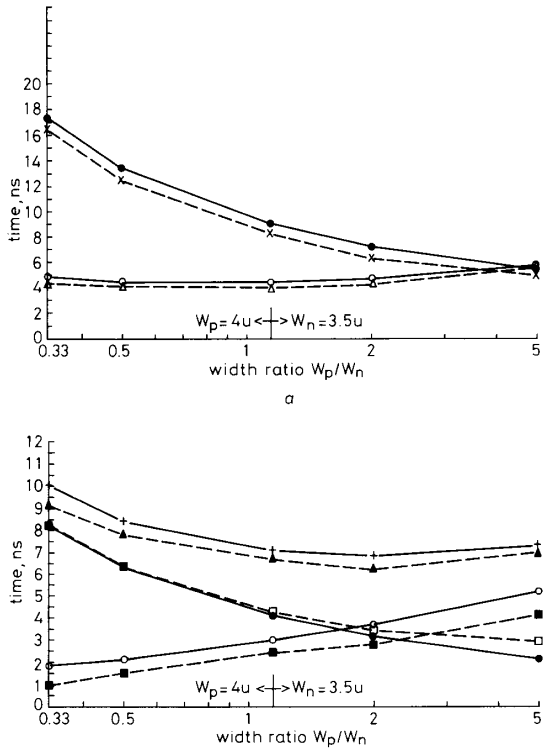

Fig. 8 *Calculated and simulated times and delays of characteristic waveforms for a CMOS NOR-type latch with $C_L = 0$ and under the two-input excitation*

a Rise and fall times
- rise (theory)
- time (ESPICE)
- fall (theory)
- time (ESPICE)

b Rise and fall delays, pair delays
- rise (theory)
- delay (ESPICE)
- fall (theory)
- delay (ESPICE)
- pair (theory)
- delay (ESPICE)

NAND-type (NOR-type) latches under the input excitations with fall times (rise times) from 1 to 20 ns.

Through extensive verifications, it is seen that the developed macromodels can be applied to compute the total propagation delay times of different static CMOS latches with a maximum error of 22%. The same accuracy can be maintained for the CMOS latches with the effective channel length down to 1.5 μm, width ratios W_p/W_n from 0.3 to 5 and the capacitive lead C_L up to 5 pF. It also can be maintained under large device parameter variations and noncharacteristic-waveform input excitations. As to the CPU-time consumption, the macromodel calculation is about 100 times as fast as the SPICE simulation. Reasonable accuracy, wide applicable range and little computation time make the developed macromodels practical, useful and efficient in computing the CMOS latch delay.

4 Application and experimental verification

By incorporating the timing models for CMOS combinational logic gates [2] into the developed timing macromodels of CMOS latches, the signal timing of CMOS static sequential logic circuits can be efficiently computed. Two different types of CMOS static flip-flops were characterised to demonstrate such an application.

The first flip-flop is a CMOS clocked S/R flip-flop shown in Fig. 11A where the device dimensions are given.

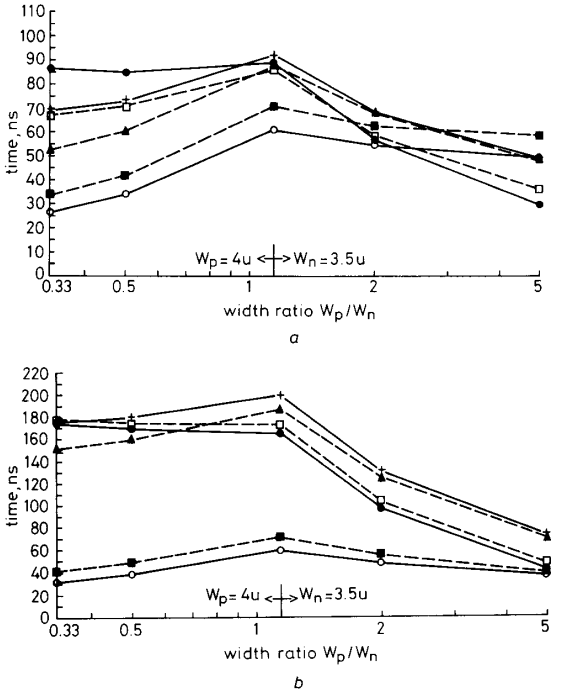

Fig. 9 *Calculated and simulated rise time, fall time and pair delay of characteristic waveforms for a CMOS latch*

$V_{TOP} = V_{TON} = 0.3$ V; $C_L = 0$
a NAND-type latch
b NOR-type latch
- rise (theory)
- delay (ESPICE)
- fall (theory)
- delay (ESPICE)
- pair (theory)
- delay (ESPICE)

To show the worst-case error, the load stage connected to the output nodes Q or $\bar{Q}$ is a CMOS inverter which represents a minimum load to the flip-flop. The flip-flop is driven by a rising clock with a rise time of 2 ns, while the inputs $\bar{S}$ and $\bar{R}$ are kept in the logic 1 and 0 states, respectively. The calculated and the simulated propagation delay of the flip-flop, defined as the total delay from the clock input to the output $\bar{Q}$, and the propagation delay of the latch are listed in Table 5. The maximum error is 22%.

The second flip-flop is a CMOS clocked master-slave JK flip-flop as shown in Fig. 11B. The computed and the simulated latch delay times and total delay times in both master and slave stages, with $J = 1$ and $K = 0$, are listed in Table 5. In the master stage the error in the total delay is higher than that in the latch delay, owing to the higher error in the calculated delay of the three-input NAND gate. Except the total delay of the master stage, all other delay times have a maximum error below 22%.

To verify experimentally part of the developed macromodels, the signal timing of the CMOS clocked master-slave T flip-flop, designed and implemented through a CMOS 5 μm gate array, was measured and calculated. The logic diagram of the fabricated flip-flop is shown in Fig. 12. Consider the slave stage of the flip-flop with its input node A in the logic 1 state. The negative edge of the

Table 4: Timing data obtained from macromodels and SPICE for the NAND-type and NOR-type latches under different input excitations

Latch type	W_p/W_N μm/μm	Input excitation	Data type	Rise time, ns	Fall time, ns	Pair delay, ns
NAND	4.0/3.5	($\bar{S}\bar{R}$)	SPICE	4.39	3.29	5.05
		(11) → (01)	macromodel	4.31	3.80	4.46
		(11) → (10)	error, %	−1.9	15.5	11.8
		($\bar{S}\bar{R}$)	SPICE	5.33	3.47	5.71
		(01) ↔ (10)	macromodel	5.05	3.93	4.95
			error, %	−5.3	13.4	−13.3
	7.0/3.5	($\bar{S}\bar{R}$)	SPICE	3.74	4.14	4.97
		(11) → (01)	macromodel	3.75	4.76	4.62
		(11) → (10)	error, %	0.3	15.1	−7.0
		($\bar{S}\bar{R}$)	SPICE	4.56	4.19	5.64
		(01) ↔ (10)	macromodel	4.44	4.93	5.11
			error, %	−2.7	17.8	−9.4
NOR	4.0/3.5	(SR)	SPICE	8.17	3.53	6.22
		(00) ↔ (01)	macromodel	9.05	3.92	6.60
		(00) → (10)	error, %	10.8	11.1	6.10
		(SR)	SPICE	8.25	4.00	6.66
		(10) ↔ (01)	macromodel	9.05	4.39	7.07
			error, %	9.80	9.80	6.10
	17.5/3.5	(SR)	SPICE	4.67	4.36	6.24
		(00) → (01)	macromodel	5.37	4.80	6.37
		(00) → (10)	error, %	15.1	10.2	2.10
		(SR)	SPICE	4.92	5.51	7.00
		(10) ↔ (01)	macromodel	5.38	5.75	7.31
			error, %	9.30	4.40	4.50

clock CLK drives the NAND 2 gate to generate a falling voltage at the node B, which triggers the NAND-type latch. The total delay time for the output Q to reach the logic threshold point of the NAND 3 gate is nearly equal to the sum of the delay times of the inverter 1, the NAND 2 gate and the latch, i.e. $T_{d1} + T_{d2} + T_p$. At that time, if the positive edge of the clock CLK has reached the node B after the total delay time $T_{d1} + T_{d2}$ of the inverter 1 and the NAND 2 gate, the output state of the latch will become ambiguous. The minimum required negative (CLK = 0) clock width T_L, therefore, can be related to the various delay times by

$$T_L + T_{d1} + T_{d2} = T_{d1} + T_{d2} + T_P \tag{21}$$

Eqn. 21 can be reduced to

$$T_L = T_P \tag{22}$$

This means that the minimum negative clock width in the flip-flop must be equal to the pair delay of the slave latch. If T_L is smaller than T_P, ambiguous states can be detected at the output nodes OUT1 and OUT2.

The minimum required negative clock width T_L can be experimentally determined by keeping the input T at V_{DD}, and applying a voltage pulse with a large positive width and a short adjustable negative pulse width to the clock input. The negative pulse width is then reduced until ambiguous voltage states are observed. This pulse width is the measured T_L.

Fig. 13*a* shows the measured waveforms of the applied voltage pulse and the output voltage at the output node OUT2. The negative pulse width is about 32 ns, and a normal output rising waveform is detected. Reducing the negative pulse width to 19.3 ns, the ambiguous output voltage state can be observed as shown in Fig. 13*b*. Thus the measured slave latch delay is 19.3 ns. The calculated pair delay of the slave latch is 17.31 ns, which has an error of 10.3%. This reasonable accuracy is consistent with that obtained from the comparison between theoretical calculations and SPICE simulations in Section 3.

5 Discussion and conclusion

The timing macromodels for CMOS NAND- and NOR-type S/R latches have been developed to compute analytically their signal timing. The timing equations in the macromodels are derived from the effective dominant pole of the linearised large-signal equivalent circuit of the latch under the characteristic waveform consideration.

Table 5: Timing data obtained from SPICE and the combined models for two clocked CMOS flip-flops

Flip-flop type		Clock timing	Output load at Q and $\bar{Q}$	Delay type	Model calculation, ns	SPICE simulation, ns	Error, %
Clocked S/R		2 ns (rise)	1 CMOS inverter	latch delay	4.34	5.53	−21.6
				total delay	6.28	7.45	15.7
Clocked	master stage	2 ns (rise)	slave stage	latch delay	4.51	5.74	−21.4
				total delay	7.0	9.21	−24.0
JK	slave stage	2 ns (fall)	1 CMOS inverter	latch delay	6.91	7.79	−11.3
				total delay	10.36	11.46	−9.6

Through extensive comparisons with SPICE simulation results, it is found that the developed macromodels have a maximum error of 22% in the total propagation delay of the CMOS NAND- and NOR-type latches, with different device effective channel length down to 1.5 μm, different channel width ratios and different capacitive loads.

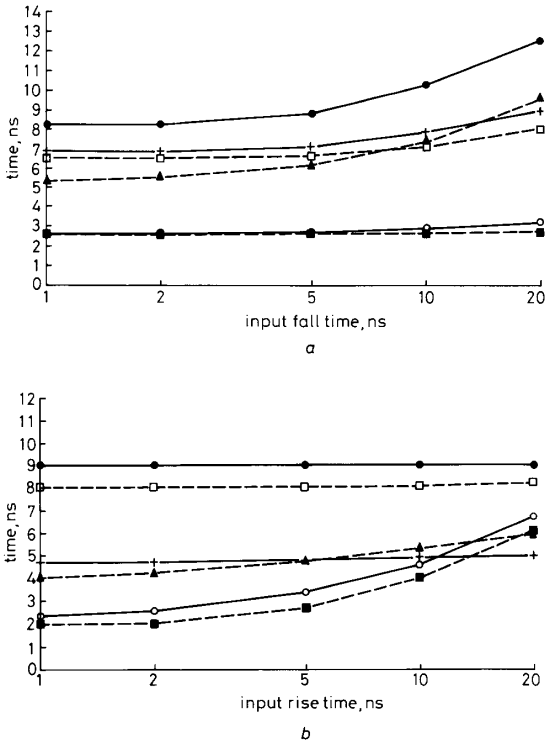

Fig. 10 *Calculated and simulated rise time, fall time and pair delay for a CMOS latch driven by different input voltages with different fall or rise times*

a NAND-type: $W_p = 4.0$ μm; $W_n = 12.0$ μm
b NOR-type: $W_p = 4.0$ μm; $W_n = 3.5$ μm
●——● rise (theory)
□- - - -□ delay (ESPICE)
○——○ fall (theory)
■- - - -■ delay (ESPICE)
+——+ pair (theory)
▲- - - -▲ delay (ESPICE)

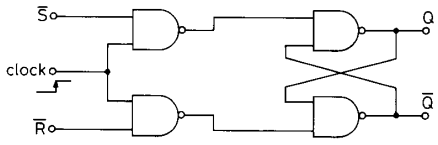

Fig. 11A *Logic diagram of a CMOS clocked S/R flip-flop*

$L_{mask} = 3.5$ μm; $W_p = 4.0$ μm; $W_n = 3.5$ μm
$L_{mask} = 5$ μm
NAND: $W_p = 13$ μm; $W_n = 40$ μm
INV: $W_p = 40$ μm; $W_n = 15$ μm

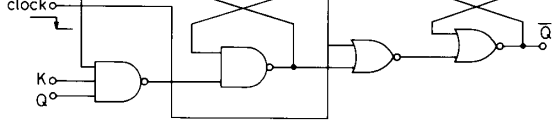

Fig. 11B *Logic diagram of a CMOS clocked master-slave JK flip-flop*

$L_{mask} = 3.5$ μm; $W_p = 4.0$ μm; $W_n = 3.5$ μm
$L_{mask} = 5$ μm
NAND: $W_p = 13$ μm; $W_n = 40$ μm
INV: $W_p = 40$ μm; $W_n = 15$ μm

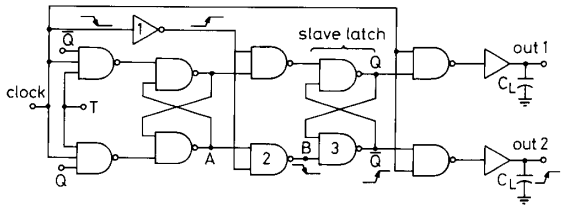

Fig. 12 *Logic diagram of a CMOS clocked master-slave T flip-flop fabricated in a 5 μm CMOS gate array*

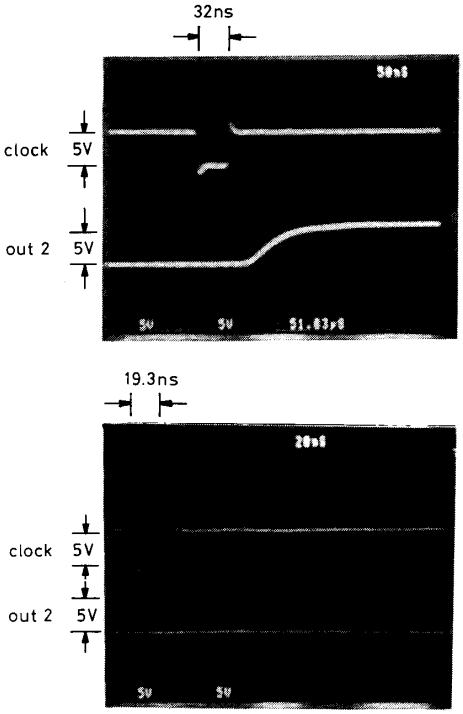

Fig. 13 *Measured clock and output waveforms in a fabricated CMOS T flip-flop driven by a clock with a negative pulse width*

a Negative pulse width = 32 ns
b Negative pulse width = 19.3 ns

The same accuracy is kept for the latches under device parameter variations, noncharacteristic-waveform input excitations and single- and two-input excitation. Moreover, the computation time of the macromodels is about 100 times as fast as that of the SPICE simulations. Thus the difficulties of intolerably long CPU time and possible numerical divergence in the full transient simulations can be avoided. The features of reasonable accuracy, wide applicable ranges and less CPU-time and memory consumption make the developed macromodels practical and efficient in timing analysis of CMOS latches.

By combining the timing models of CMOS combination logic gates [2] with the developed latch macromodels, the signal timing of CMOS static sequential logic circuits can be efficiently computed. As an application example, the signal timing of a CMOS clocked S/R flip-flop and a CMOS clocked master-slave JK flip-flop were characterised. It is shown that the accuracy of the combined model is quite satisfactory. Experimental verification of the macromodels on the delay of the NAND-type

latch is also successfully performed, through the fabricated CMOS clocked master-slave T flip-flop.

Besides timing analysis, timing synthesis [2] and speed optimisation are the expected applications of the developed macromodels. They will be investigated in detail when the macromodels are incorporated into the CAD program TISA [2].

6 Acknowledgment

The research was supported by the Microelectronics and Information Science and Technology and Electronics Research and Service Organisation, Industrial Technology and Research Institution, Republic of China.

7 References

1 BURNS, J.R.: 'Switching response of complementary-symmetry MOS transistor logic circuits', *RCA Review*, 1964, pp. 627–661
2 WU, C.Y., HWANG, J.S., CHANG, C., and CHANG, C.C.: 'An efficient timing model for CMOS combinational logic gates', *IEEE Trans.*, 1985, **CAD–4**, pp. 636–650
3 VLADIMIRESCU, A., and LIU, S.: 'The simulation of MOS integrated circuits using SPICE2', UCB/ERL M8017, Electronics Research Laboratory, College of Engineering, Univ. of California, Berkeley, California, Feb. 1980
4 ELMASRY, M.I.: 'Digital MOS integrated circuits: A tutorial' *in* 'Digital MOS Integrated Circuits' (IEEE Press, New York, 1981) pp. 4–27