Yield and Speed Optimization of a Latch-Type Voltage Sense Amplifier

Bernhard Wicht, *Member, IEEE*, Thomas Nirschl, and Doris Schmitt-Landsiedel, *Member, IEEE*

***Abstract*—A quantitative yield analysis of a latch-type voltage sense amplifier with a high-impedance differential input stage is presented. It investigates the impact of supply voltage, input dc level, transistor sizing, and temperature on the input offset voltage. The input dc level turns out to be most significant. Also, an analytical expression for the sensing delay is derived which shows low sensitivity on the input dc bias voltage. A figure of merit indicates that an input dc level of 0.7 V_{DD} is optimal regarding speed and yield. Experimental results in 130-nm CMOS technology confirm that the yield can be significantly improved by lowering the input dc voltage to about 70% of the supply voltage. Thereby, the offset standard deviation decreases from 19 to 8.5 mV without affecting the delay.**

***Index Terms*—Current sensing, latch delay, latch-type sense amplifier, sense amplifier, SRAM circuits, SRAM yield, yield optimization.**

I. Introduction

LATCH-TYPE sense amplifiers are applied to read the contents of several types of memory since they achieve a fast decision due to strong positive feedback [1]. Fig. 1(a) shows the well-known conventional latch. Two cross-coupled inverters provide positive feedback. The enable signal EN turns on the amplifier and starts the sensing operation. Depending on the polarity of the voltage difference between the nodes V1 and V2, the sense amplifier will flip in one or the other direction. It is essential for latch-type sense amplifiers in general that, once the decision process is finished, the current flow stops automatically. Hence, there is no static power consumption. Switching EN to low logic level resets the latch before the next read can start. The nodes V1 and V2 are input and output terminals at the same time. Therefore, the circuit cannot be connected directly to the bitline since the circuit would attempt to discharge the bitline capacitance during the decision phase and would increase delay and power. A solution is either to separate the bitline by a multiplexer or to use passgates, forming a decoupling resistor. Both devices cause a voltage drop that deteriorates the available input voltage difference. This way the voltage swing at the bitlines can easily reduce by half, resulting in lower speed and noise margin.

This drawback does not occur for the latch circuit shown in Fig. 1(b) because of a high-impedance input differential stage.

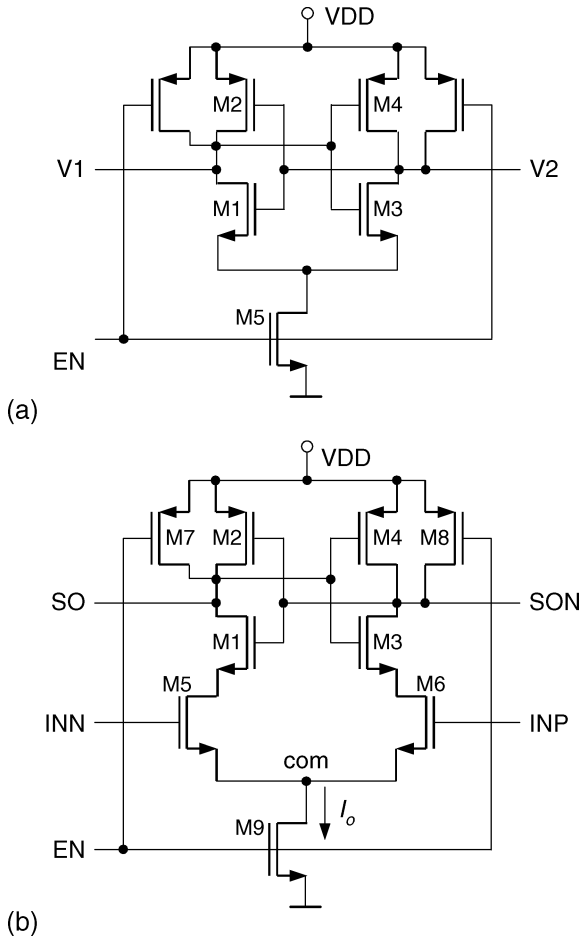


Fig. 1. Latch-type voltage sense amplifiers. (a) Conventional latch. (b) Sense amplifier according to [2].

It was introduced by Kobayashi *et al.* in 1993 [2]. This sense amplifier combines strong positive feedback with a high resistive input. The current flow of the differential input transistors M5 and M6 controls the serially connected latch circuit. A small difference between the currents through M5 and M6 converts to a large output voltage. The circuit has been used as an alternative to the conventional latch of Fig. 1(a) [3]. It also became popular in logic design to speed up signal propagation [4], [5]. As a sense-amplifier-based flip-flop, the circuit has been reported to be one of the fastest flip-flops in state-of-the-art processors/VLSI designs [6], and can significantly reduce power [7]. Due to its high-impedance input, it has recently been used as a second stage in a current-sensed SRAM operating down to 0.7 V [1], [8]. Also, a complementary transistor implementation is possible [4].

Manuscript received November 5, 2003; revised January 19, 2004.

B. Wicht is with Texas Instruments Deutschland GmbH, Mixed-Signal Power and Control, 85350 Freising, Germany (e-mail: b-wicht@ti.com).

T. Nirschl is with Infineon Technologies, 81609 Munich, Germany, and also with the Technical University Munich, 80333 Munich, Germany.

D. Schmitt-Landsiedel is with Technical University Munich, 80333 Munich, Germany.

Digital Object Identifier 10.1109/JSSC.2004.829399

The most critical point of latch-type sense amplifiers is that once the decision process has started, it does not recover unless the circuit is reset to the metastable point $(V(\mathrm{SO}) = V(\mathrm{SON}))$. In the presence of mismatch and noise, a wrong output signal can develop for small input voltage differences. On the other hand, the latch-type sense amplifier should be enabled at the smallest possible input voltage difference because of the delay until a voltage swing ΔV_{in} between the inputs has been developed. This refers to the capacitive discharge of the bitline or it represents the speed of a previous sensing stage. In general, this takes longer than the delay of the latch itself. Therefore, ΔV_{in} has a significant influence on the speed as well as on the yield of fabricated circuits.

This paper investigates the design issues for the latch-type sense amplifier of Fig. 1(b) to ensure a fast and correct decision even for small input signals [9]. With focus on current sensing, it will be shown why this latch makes an excellent second-stage comparator [8] after a current sensing input stage [1]. In such an application, the conventional circuit of Fig. 1(a) is not suitable as a second stage because it would draw a current out of a first current-sensing stage. The outline of this paper is as follows. Sections II and III cover the transient behavior and delay. The yield is analyzed in Sections IV and V. Section VI derives a figure of merit for optimum input biasing, while Section VII shows the influence of transistor sizing and temperature on yield and delay. A comparison to the conventional latch-type amplifier of Fig. 1(a) is presented in Section VIII. Section IX confirms the theory by experimental results, followed by the conclusions in Section X.

II. Transient Behavior

The transient behavior is shown in Fig. 2 (at a large value of V_{DD} for better visibility). During the off state, there is no major current flowing and the internal node voltages are defined by leakage currents or precharge transistors (not shown). The output terminals SO and SON are pulled to V_{DD} by the switch transistors M7 and M8. The sense-enable signal EN starts the sensing operation by turning on M9. Immediately an operation current I_o starts to flow and pulls down node com. Successively, the input transistor pair M5 and M6 turns on. Obviously, the delay up to this point is relatively small. Now three phases can be distinguished. During phase 1, the drain currents of M5 and M6 start to discharge the outputs SO and SON. With a voltage difference between the inputs, also the drain current of M5 and M6 will be different. This current imbalance causes different discharging speeds at SO and SON. So the difference increases. Since the output nodes are precharged to V_{DD} before reading, both p-channel transistors M2 and M4 remain turned off until the output voltage reaches $(V_{DD} - V_{thp})$. This marks the transition from phase 1 to phase 2. During phase 2, strong positive feedback enhances the output voltage difference starting from an initial difference V_o, as marked in Fig. 2. The latching completes during the third phase where one n-channel transistor is cut off. The operation current I_o flows during the transition of the output nodes. The current flow stops automatically after the transition so that the whole sense amplifier does not dissipate static power.

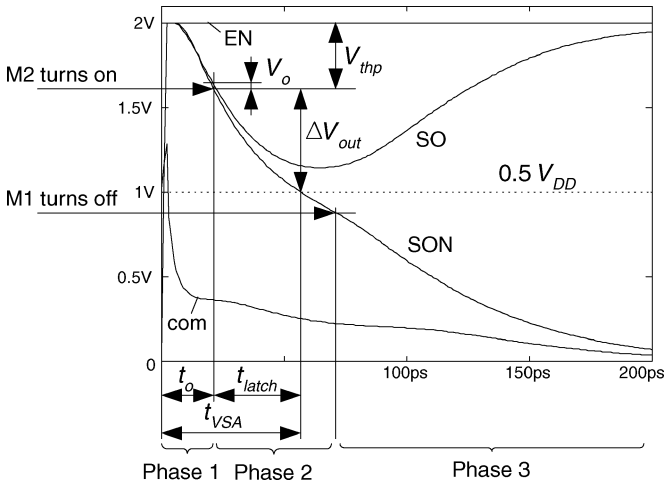


Fig. 2. Simulated transient behavior of the latch-type sense amplifier of Fig. 1(b).

III. Delay

For pure voltage sensing, the sense amplifier is usually connected to the bitline with a potential near V_{DD}. However, the power consumption during write recovery can significantly be reduced if the bitline voltage is set to a value far below V_{DD} [10], [11]. In an application as a second stage of a current sense amplifier, the input dc voltage can become nearly any value between ground and V_{DD}. In this paper, this input voltage level of the latch-type sense amplifier will be referred to as V_{INDC}. It is assumed that the voltage V_{INDC} is applied to input INP while INN receives a lower input voltage of $(V_{INDC} - \Delta V_{IN})$. This causes the output SO to approach V_{DD} during the decision and SON to go to zero.

Fig. 3 shows how the sensing delay t_{VSA} depends on V_{INDC}. Obviously, there is no speed penalty as long as V_{INDC} does not go below approximately 60% of V_{DD}. Below this value, the delay increases significantly due to reduced drain-source voltage of M9 which decreases the operation current I_o. However, for V_{INDC} between 60% of V_{DD} and V_{DD}, the delay remains almost constant. Indeed, it even slightly decreases and reaches a minimum. This is the major reason why this kind of latch-type sense amplifier is suitable as a second stage in current-sensed memories because the delay is not sensitive to the actual value of V_{INDC}. Fig. 3 also proves that the speed increases for higher supply voltage V_{DD} and larger input voltage difference ΔV_{IN}.

The influence of V_{INDC} but also of ΔV_{IN} and V_{DD} on the sense amplifier speed can be analyzed by estimating the total delay

$$t_{\mathrm{VSA}} = t_o + t_{\mathrm{latch}}. \tag{1}$$

According to Fig. 2, t_{VSA} is the sum of the delay in the first and second phase, t_o and t_{latch}, respectively. The delay t_o represents the capacitive discharge of a load capacitance C_L (at both outputs SO, SON) until the first p-channel transistor turns on. Since M6 has been assumed to have larger input potential, its drain current I_P causes faster discharge of output SON compared to SO which is driven by M5 with smaller current I_N.

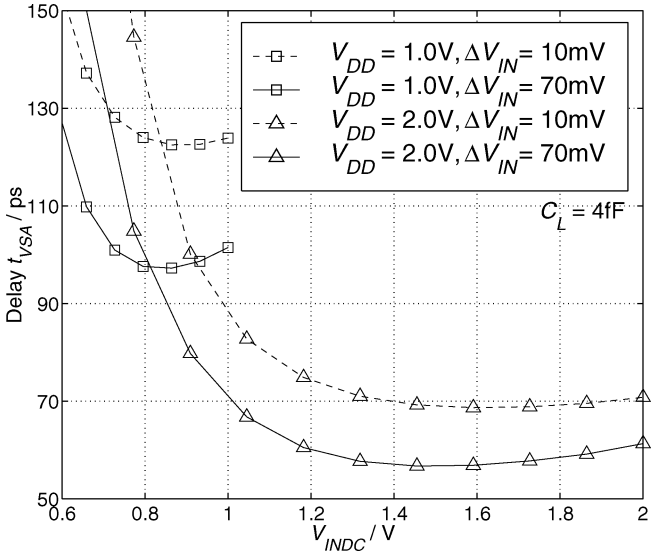


Fig. 3. Simulated delay of the sense amplifier of Fig. 1(b) versus dc input potential V_{INDC}.

Both currents can be assumed to be constant. Consequently, the discharge delay is given by

$$t_o = \frac{C_L V_{thp}}{I_P}. \tag{2}$$

The second term $t_{\rm latch}$ of (1) is the latching delay of two cross-coupled inverters. If a voltage swing of $\Delta V_{\rm out}$ at the falling output (e.g., SON) has to be obtained from an initial voltage difference V_o, the delay [12] is given by

$$t_{\rm latch} = \underbrace{\frac{C_L}{g_{\rm m,eff}}}_{\tau_{\rm latch}} \ln\left(2\frac{\Delta V_{\rm out}}{V_o}\right). \tag{3}$$

This delay depends (in a logarithmic manner) on the initial difference V_o between the outputs SO and SON at $t = t_o$. Based on (2), V_o can be calculated by

$$\begin{aligned} V_o &= |v_{\rm SO}(t = t_o) - v_{\rm SON}(t = t_o)| \\ &= V_{thp} - \frac{I_N t_o}{C_L} \\ &= V_{thp}\left(1 - \frac{I_N}{I_P}\right). \end{aligned} \tag{4}$$

The current difference $\Delta I_{IN} = |I_P - I_N|$ between the branches is much smaller than I_P, I_N. Therefore, I_P can be approximated by half of the drain current I_o of M9 and (4) can be rewritten as

$$V_o = 2V_{thp}\frac{\Delta I_{IN}}{I_o}. \tag{5}$$

The current I_o of transistor M9 is a function of V_{INDC} and V_{DD}. Assuming M9 in the triode region and M5 and M6 in saturation, I_o can be derived to be approximately

$$I_o \approx 2\beta(V_{INDC} - V_{th})^2\left(1 - \frac{0.75}{1 + \frac{V_{DD} - V_{th}}{V_{INDC} - V_{th}}}\right)^2. \tag{6}$$

In (6), all transconductance parameters β and all threshold voltages have been assumed to be equal. The equation basically reflects the output characteristics of M9. Short-channel effects are not so much of an influence since the transistor operates in triode region with a low drain-source voltage. Mobility degradation has some effect, but it depends on the gate voltage, which is constant in this case.

Since M5, M6, and M9 form a differential stage, the relation between ΔI_{IN} and ΔV_{IN} can be taken from basic textbooks. For small values of ΔV_{IN}, it can be expressed as

$$\Delta I_{IN} = \sqrt{2\beta I_o}\Delta V_{IN} \tag{7}$$

where β refers to M5 and M6.

With the current difference from (7), the initial voltage difference of (5) becomes

$$V_o = V_{thp}\sqrt{\frac{8\beta}{I_o}}\Delta V_{IN} \tag{8}$$

and it yields for the latch delay of (3)

$$t_{\rm latch} = \frac{C_L}{g_{\rm m,eff}} \ln\left\{\frac{1}{V_{thp}}\sqrt{\frac{I_o}{2\beta}}\frac{\Delta V_{\rm out}}{\Delta V_{IN}}\right\}. \tag{9}$$

With (2) and (9), the total delay of (1) becomes

$$t_{\rm VSA} = \underbrace{\frac{2C_L V_{thp}}{I_o}}_{t_o} + \underbrace{\frac{C_L}{g_{\rm m,eff}} \ln\left\{\frac{1}{V_{thp}}\sqrt{\frac{I_o}{2\beta}}\frac{\Delta V_{\rm out}}{\Delta V_{IN}}\right\}}_{t_{\rm latch}} \tag{10}$$

with the bias current I_o according to (6) provided by M9 after sensing is activated. Instead of the output swing $\Delta V_{\rm out}$, mostly an absolute threshold $V_{\rm out}$ is given. Then the output swing in (10) can be replaced by $\Delta V_{\rm out} = V_{DD} - V_{thp} - V_{\rm out}$.

Equation (10) explains the impact of various parameters. As expected, the total delay is directly proportional to the output load capacitance C_L. Only $t_{\rm latch}$ depends on the input difference ΔV_{IN}, resulting in higher total speed the larger ΔV_{IN} becomes. This agrees with the simulation results of Fig. 3.

Based on the derived delay equations, the influence of V_{INDC} can also be discussed. In the first order, I_o is proportional to the square of V_{INDC} [see (6)]. Therefore, t_o can be expected to increase if V_{INDC} is reduced while, on the other hand, $t_{\rm latch}$ decreases. The delay t_o of the first sensing phase increases because lower V_{INDC} causes smaller bias current I_o. This slows down the discharge of both output terminals SO and SON. On the other hand, (8) shows that a delayed discharge with smaller I_o results in an increased initial voltage difference V_o at the end of the first phase when the first p-channel transistor turns on. Hence, $t_{\rm latch}$ decreases due to the inverse dependence on V_o expressed by (3).

Based on this theoretical analysis, Fig. 4 shows the total delay $t_{\rm VSA}$ of the voltage sense amplifier and its components t_o and $t_{\rm latch}$ as a function of V_{INDC}. As discussed above, t_o decreases and $t_{\rm latch}$ increases with V_{INDC}. As a consequence, they cancel to some extent. Between 1.2 and 2 V, there remains only a slight influence of V_{INDC} on the total delay. This matches well with the simulated results of Fig. 3 even though they include short-channel effects, etc., which have been neglected in the above equations. Nevertheless, the relatively simple expression of (10), while suitable for estimating the sensing delay and the influence of its parameters, is not adequate for a quantitative delay calculation. Especially for low values of V_{INDC}, (10) is not precise enough. This is due to the fact that (10) is based

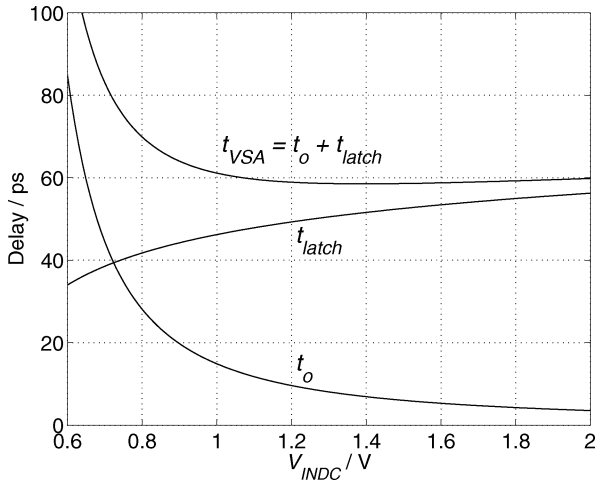


Fig. 4. Theoretical dependence of the sense amplifier delay on V_{INDC} according to (10). The total delay $t_{\rm VSA}$ is the sum of the output discharge delay t_o and the latching delay $t_{\rm latch}$ of the cross-coupled inverters ($V_{DD} = 2$ V, $\Delta V_{IN} = 70$ mV).

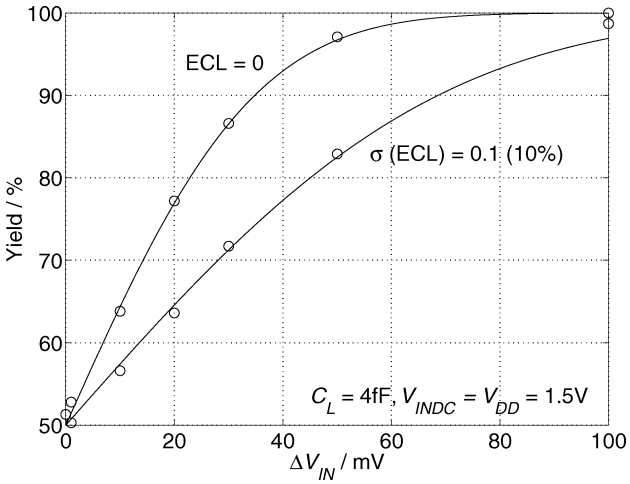


Fig. 5. Yield versus input voltage difference ΔV_{IN}. Being a measure for the number of correct decisions, the yield increases with ΔV_{IN}. $\mathrm{ECL} = 0$ corresponds to fully matched output capacitances. Increased capacitance mismatch causes lower yield as shown for a standard deviation of 10% of C_L.

on first-order transistor equations, partly on small-signal analysis and some approximations [see I_o in (6) and ΔI_{IN} in (7)]. Moreover, the effective transconductance $g_{m,\rm eff}$ is assumed to be constant with respect to V_{INDC}.

Obviously, for larger supply voltage V_{DD}, the operation current I_o increases. Therefore, the first phase becomes faster. Although I_o appears in the numerator of the latch delay $t_{\rm latch}$, it has to be considered that also the effective transconductance $g_{m,\rm eff}$ increases with V_{DD}. As a consequence, the total delay $t_{\rm VSA}$ can be reduced by increasing the supply voltage.

IV. Yield

As for all latch-type sense amplifiers, it is essential that the circuit is not activated until the input voltage difference ΔV_{IN} is large enough to ensure a correct decision. Due to local variations during the fabrication, mismatch occurs, causing the sense amplifier structure to become asymmetrical. The root causes may be many (geometry variations, random dopant fluctuations, oxide thickness variations, edge roughness, etc.) but in any case this imbalance can be represented by an input offset voltage V_{os}. This is the voltage difference that has to be applied to force the cross-coupled inverters to get into metastability, $V(\mathrm{SO}) = V(\mathrm{SON})$. Only for values of ΔV_{IN} larger than this offset the sensing circuit flips in the right direction. Otherwise, the sensing fails. Applied to a certain number of samples, this effect is referred to as parametric yield Y

$$Y = \frac{\text{number of correct decisions}}{\text{number of samples}} \cdot 100\% \qquad (11)$$

which represents the percentage of correct decisions.

Besides random mismatch, systematic offset occurs since matched devices can never be completely symmetric in layout. Technology factors such as mechanical strain can introduce further systematic offsets. The investigation presented here assumes a circuit without systematic offset to predict the limits in terms of yield with respect to random mismatch. Nevertheless, the presented results and strategy are also valid for systematic offset.

Besides the delay, the yield Y is an important characteristics of sense amplifiers. Its dependence on several parameters, especially on ΔV_{IN}, V_{INDC}, and V_{DD}, is analyzed below for the latch-type sense amplifier of Fig. 1(b). This statistical investigation is based on Monte Carlo simulation in addition to theoretical calculations. All Monte Carlo simulations incorporate local and global variations of transistor parameters, limited to $\pm 6\sigma$, where σ represents the standard deviation of each parameter. Also, random mismatch between the output load capacitances C_L at SO and SON has been considered by a statistical parameterECL (matching error between the load capacitances C_L) which is varied within a range of $\pm 5\sigma$. It is expected that the nominal layout is fully matched so that no systematic mismatch occurs, i.e., the mean value of ECL is zero. Within the simulation a nominal capacitance C_L is assigned to output SO, whereas for the complementary output SON the capacitance value is set to $(1 + \mathrm{ECL}) \cdot C_L$.

In general it can be expected that a correct decision becomes more likely the larger ΔV_{IN} and the smaller ECL is. Fig. 5 shows this relation for $V_{DD} = 1.5$ V and different degrees of output capacitance matching. $\mathrm{ECL} = 0$ means that only transistor mismatch but no C_L asymmetry affects the yield. Each marker represents the yield determined by a Monte Carlo simulation of 1000 samples. The same number of samples has been used consistently for all statistical simulations presented below.

From the theoretical point of view, the yield $Y(\Delta V_{IN})$ is identical to the probability P for the actual offset voltage V_{os} to lie below ΔV_{IN} where V_{os} follows a Gaussian probability distribution. For a fully symmetric circuit, the mean offset can be assumed to be zero. The yield is properly described by the normalized Gaussian distribution Φ in the form

$$\begin{aligned} Y(\Delta V_{IN}) &= P(V_{os} \le \Delta V_{IN}) \\ &= \Phi\left(x = \frac{\Delta V_{IN}}{\sigma_{os}}\right) \\ &= \frac{1}{2} + \int_0^x \varphi(y)dy \end{aligned} \qquad (12)$$

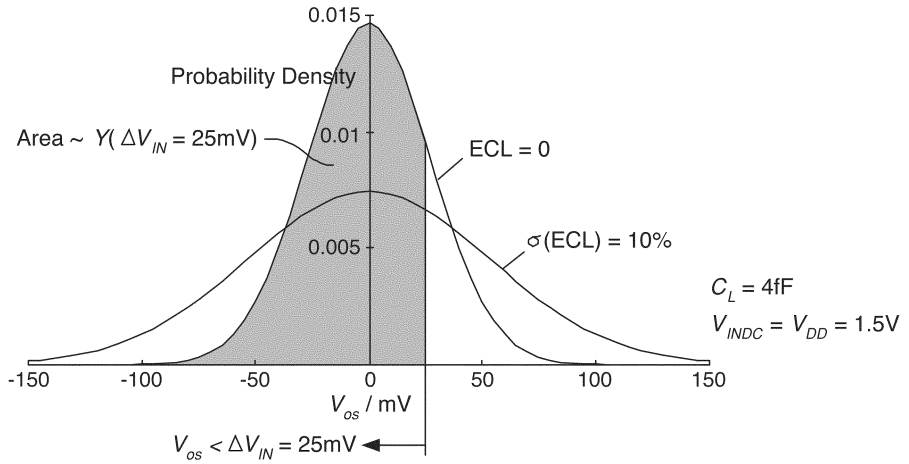


Fig. 6. Probability density function of the input offset voltage V_{os} according to Fig. 5.

with the normalized probability density function $\varphi(y)$ given by

$$\varphi(y) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{y^2}{2}}. \tag{13}$$

The distribution of V_{os}, and consequently Y, is well-defined by just *one* parameter, the standard deviation σ_{os} of the input offset voltage V_{os}.

The probability density function is plotted in Fig. 6. The variance increases if the output capacitance mismatch deteriorates (parameter ECL). The required standard deviation σ_{os} can be taken from Fig. 5. Equation (12) yields $Y(\Delta V_{IN} = \sigma_{os}) = 0.8413$. Hence, σ_{os} is equal to the value of ΔV_{IN} that corresponds to $Y = 84.13\%$. For the two curves of Fig. 5, a standard deviation of 27.2 and 53.5 mV can be read off. In general, an arbitrary value pair $(\Delta V_{IN};Y)$ is sufficient to calculate the standard deviation σ_{os} of V_{os}. Using the inverse function Φ^{-1} of the Gaussian distribution, σ_{os} can be extracted based on

$$\sigma_{os} = \frac{\Delta V_{IN}}{\Phi^{-1}(Y)}. \tag{14}$$

Fig. 6 also illustrates how the integral from $-\infty$ to ΔV_{IN}, i.e., the area under the curve, defines the yield for $\Delta V_{IN} = 25$ mV. In addition, it can be seen that $\Delta V_{IN} = 0$ results in 50% yield. Due to its symmetrical structure, it is evenly likely that the sense amplifier flips in one or the other direction. Therefore, half of the samples will fail [see also Fig. 5: $Y(\Delta V_{IN} = 0) = 50\%$].

The offset voltage V_{os} is determined by the mismatch between the differential input transistors and between the cross-coupled inverters. Fundamentals on transistor matching have been presented in [13], [14]. References [15]–[17] cover the offset of cross-coupled inverters and latches.

Since there is no analytical solution for the integral in the Gaussian distribution Φ, it has to be solved numerically. For this reason, numerical tools provide the error function

$$\mathrm{erf(x)} = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-y^2} dy \tag{15}$$

and its inverse erf^{-1}. Therefore, it is appropriate to rewrite (12) and (14) in the form

$$Y(\Delta V_{IN}) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{\Delta V_{IN}}{\sigma_{os}\sqrt{2}}\right)\right) \tag{16}$$

and

$$\sigma_{os} = \frac{\Delta V_{IN}}{\sqrt{2}\,\mathrm{erf}^{-1}(2Y-1)}. \tag{17}$$

It is possible to obtain the deviation of V_{os} directly by Monte Carlo simulation. However, since V_{os} corresponds to the metastable condition of the cross-coupled inverters, for a stable and accurate circuit simulation it is preferable to determine the yield instead of V_{os}. In addition, even though Fig. 6 is a measure for the possible offset voltage, for the design process curves like Fig. 5 are more appropriate since they directly provide a minimum input difference ΔV_{IN} for a given yield target.

V. Yield Improvement

In Fig. 5, the input potential V_{INDC} was set equal to V_{DD}. As pointed out before, in some applications (e.g., current sensing) V_{INDC} will be below V_{DD}. Its influence on the yield is investigated below. Fig. 7 shows the functional yield according to (11) versus ΔV_{IN} for values of V_{INDC} down to 40% of V_{DD}. Obviously, there is a significant influence of V_{INDC} on the yield. In this case, it improves almost by 30% if the input dc potential is reduced from 1.5 to 0.6 V at $\Delta V_{IN} = 30$ mV. To explain this effect, it has to be realized that lower V_{INDC} causes smaller bias current I_o. This, in turn, results in a larger initial voltage difference V_o between the cross-coupled inverter terminals SO and SON. Like the full circuit (V_{os}), the single block of cross-coupled inverters also suffers from a statistical offset voltage. If the initial voltage difference V_o is large, it is more likely that a correct decision occurs. Therefore, the yield increases inversely to V_{INDC}.

By means of a transient simulation, Fig. 8 shows how V_o increases if V_{INDC} is lowered to 60% of V_{DD}. The right-hand plots represent an enlarged section as marked by the gray box.

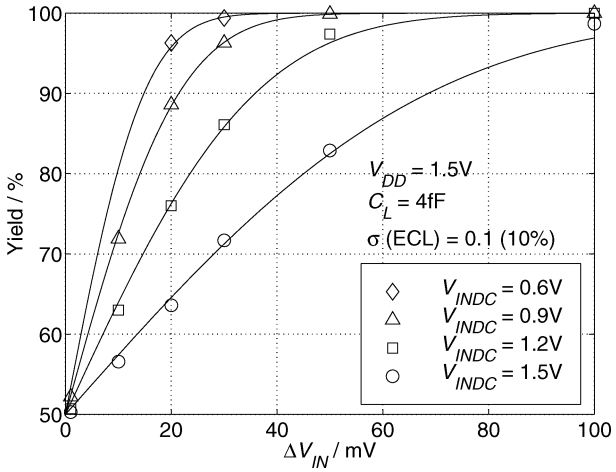


Fig. 7. Yield versus input voltage difference ΔV_{IN} for different input dc levels V_{INDC}.

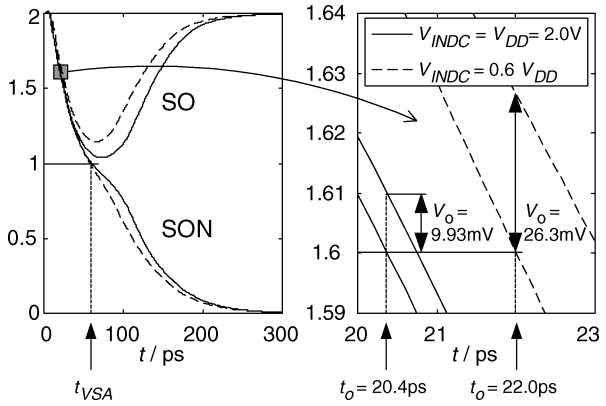
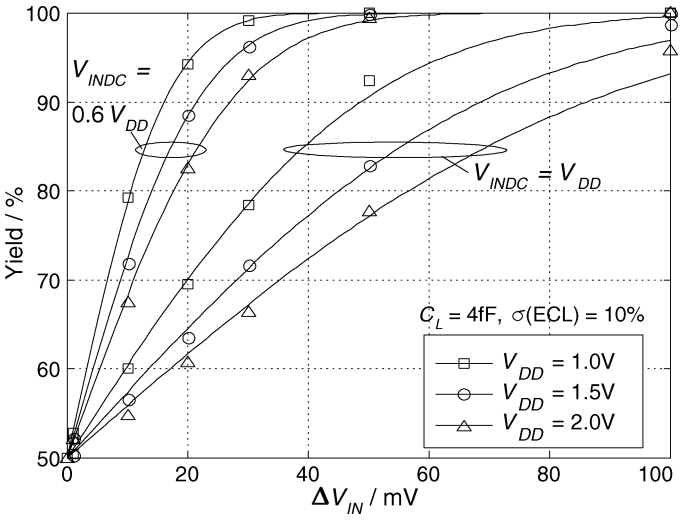
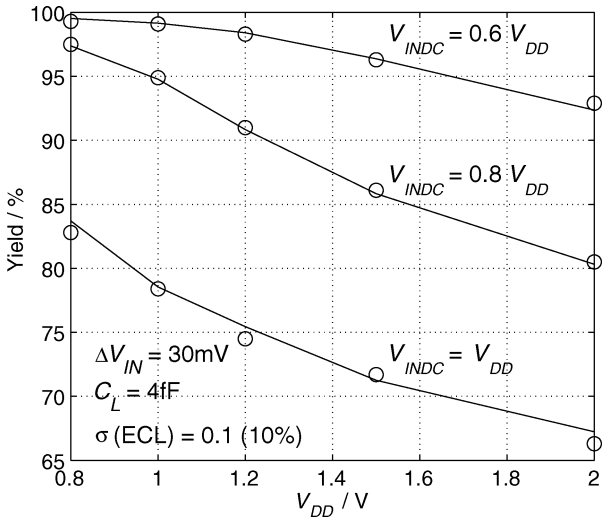


Fig. 8. Transient simulation of the sense amplifier output voltages showing the impact of V_{INDC}.



(a)



(b)

Fig. 9. Influence of V_{DD} on the functional yield Y in relation to V_{INDC}. (a) Yield versus input voltage difference ΔV_{IN} for different supply voltages. (b) Yield versus supply voltage V_{DD} at $\Delta V_{IN} = 30$ mV for different input potential V_{INDC}.

It is assumed that the p-channel transistors M2 and M4 are fully turned on when the output voltage reaches 1.6 V, i.e., $V_{SG} = 400$ mV $\approx V_{thp}$. Due to larger I_o, during the first phase the case $V_{INDC} = V_{DD}$ is faster ($t_o = 20.4$ ps and 22.0 ps, the right-hand plot of Fig. 8) but it results in an initial difference of only 9.93 mV compared to 26.3 mV for $V_{INDC} = 0.6V_{DD}$. This is the reason why the yield significantly increases by lowering V_{INDC}. Fig. 8 also illustrates how V_{INDC} influences the sensing speed. In spite of larger delay t_o during the first phase for $V_{INDC} = 0.6V_{DD}$, the larger value of V_o speeds up the second phase where the cross-coupled inverters amplify the output voltage difference. Hence, the total delay is almost equal in both cases. This is an important result, since it explains how the yield can be improved without decreasing the speed just by reducing the input dc level V_{INDC}.

The supply voltage itself also has an impact on V_o as can be seen from (6) and (8). As for V_{INDC}, for a given input difference ΔV_{IN} the initial output difference V_o increases if V_{DD} is reduced. However, V_{DD} is of less influence than V_{INDC}. In addition to the theoretical equations this is shown by Fig. 9. Even though there is a considerable increase in yield for lower supply voltages, an acceptable yield at a given value of V_{DD} can only be achieved by decreasing

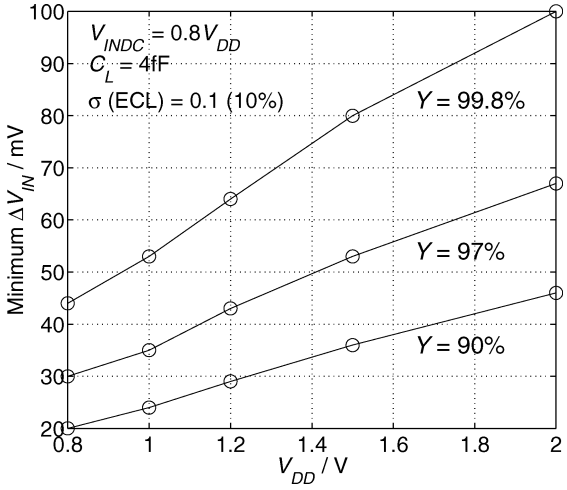


Fig. 10. Minimum input voltage difference ΔV_{IN} at a given yield Y versus V_{DD}.

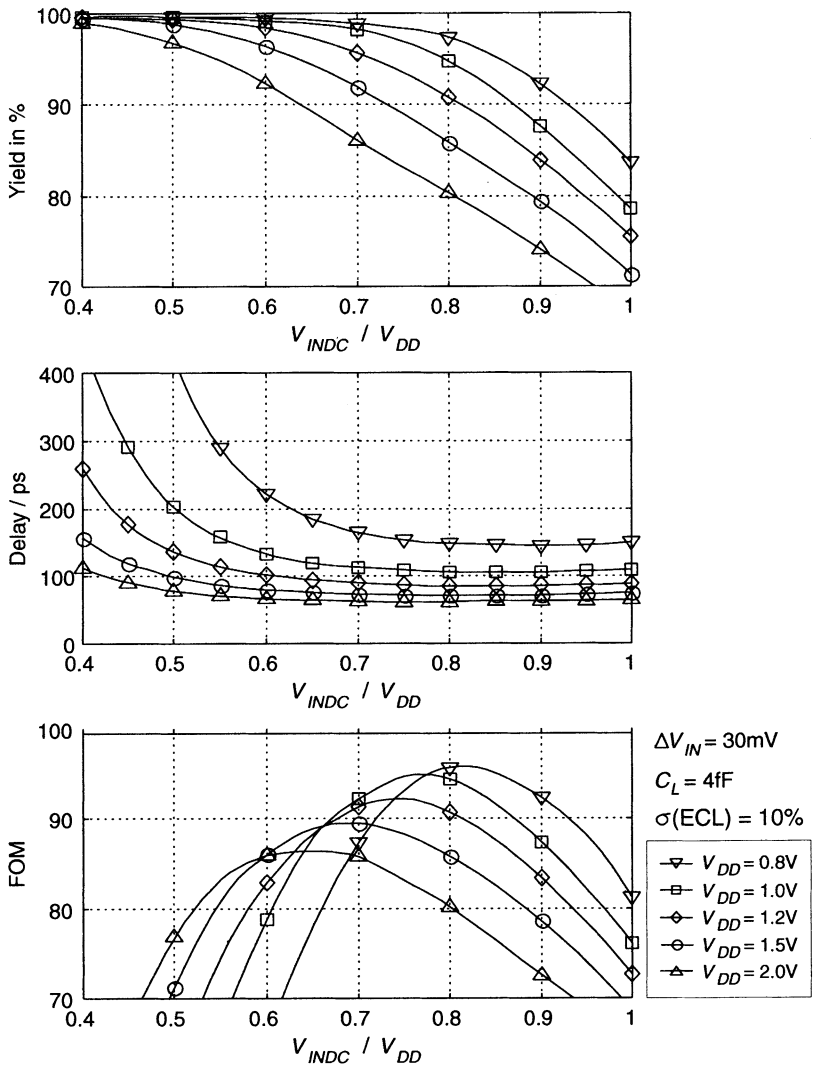


Fig. 11. Optimum V_{INDC} regarding yield (top) and speed (center) is indicated by the maximum FOM (bottom). FOM is defined as the ratio between yield and normalized delay.

V_{INDC} as the comparison between $V_{INDC} = 0.6V_{DD}$ and $V_{INDC} = V_{DD}$ shows. Another consequence is that for a certain yield specification, the minimum required ΔV_{IN} scales with the supply voltage. Fig. 10 shows that there is nearly a linear relation between ΔV_{IN} and V_{DD}. A yield of 99.8% requires $\Delta V_{IN} = 80$ mV at $V_{DD} = 1.5$ V and it reduces to only $\Delta V_{IN} = 53$ mV at $V_{DD} = 1.0$ V. The most stringent yield specification needs the largest value of ΔV_{IN}, e.g., at $V_{DD} = 1.5$ V the minimum input difference would decrease to 36 mV if the yield requirement would decline to 90%.

VI. Optimum Input DC Level V_{INDC}

Fig. 9 suggests that the smallest possible V_{INDC} gives best results regarding yield. On the other hand, the delay curve of Fig. 3 shows an opposite relation with respect to sensing delay. An optimum value of V_{INDC} can be derived by introducing a figure of merit (FOM)

$$\mathrm{FOM}(V_{\mathrm{INDC}}) = \frac{Y(V_{INDC})}{t_{\mathrm{VSA}}(V_{INDC})/\min(t_{\mathrm{VSA}})}. \qquad (18)$$

This expression incorporates the ratio between the functional yield Y and sense amplifier delay t_{VSA}. The optimum V_{INDC} corresponds to the maximum FOM. For better comparison t_{VSA} is normalized to the minimum delay. So the best FOM equals the upper limit of 100. The results are plotted in Fig. 11 (bottom) in relation to the yield (top) and delay diagram (center). To include V_{DD} as a parameter, V_{INDC} has been normalized by V_{DD}. A clear maximum FOM can be observed. It ranges from about $V_{INDC} = 0.8\ V_{DD}$ at 0.8-V supply to $V_{INDC} = 0.6\ V_{DD}$ at $V_{DD} = 2.0$ V. As a rule of thumb, $V_{INDC} = 0.7\ V_{DD}$ gives good results in any case.

VII. Transistor Sizing and Temperature

So far, a fixed transistor size has been assumed. Indeed, there is low influence of the W/L ratio on the performance except for the turn-on transistor M9. But, unlike for the voltage sense amplifier of Fig. 1(a), the n-channel switch transistor M9 of the

latch does not turn out to be critical due to the different principle of operation.[1] Fig. 12 shows that a sizing of $W/L = 1\ \mu\text{m}/300$ nm gives good results regarding yield and speed. Fig. 12(a) also includes the behavior at different temperatures with the highest speed at low temperature ($-55°\text{C}$).

Since the transistor matching improves with the size of the gate area [13] there is always a tradeoff with respect to the given area limitation. In most memory applications, the area is strictly limited so that no extra area can be spent for better matching properties. In this case, for example, the differential pair M5 and M6 has been set to $W/L = 1500$ nm$/240$ nm. The influence of the W/L ratio and temperature on the parametric yield is rather low. Fig. 12(b) shows for M9 that the impact can be neglected compared to the considerable yield improvement due to V_{INDC}. The low geometry influence is caused by the fact that the operation current I_o according to (6) depends in a quadratic manner on V_{INDC} but only linearly on β (which is proportional to W/L).

A consequence of Fig. 3 and Fig. 12 is that V_{INDC} has a large influence on the yield while the delay remains almost constant within a certain range of V_{INDC}. On the other hand, M9 definitely determines the total delay but there is rather a small impact on the yield. This means for the design of the sense amplifier that the delay can be adjusted by the sizing of M9 while a desired yield can be achieved by an appropriate value of V_{INDC}. So delay and yield can be adjusted independently.

VIII. Comparison to Conventional Latch

Compared to the sense amplifier of Fig. 1(a), the circuit structure of Fig. 1(b) can be considered as two stages. Besides the cross-coupled inverters, which both circuits have in common, the sense amplifier of Fig. 1(b) consists of a differential input transistor pair M5 and M6. Although this transistor pair is responsible for the desired high input resistance, it increases the input offset voltage V_{os}. In comparison to the conventional circuit of Fig. 1(b), a larger total offset voltage can be expected since it is determined by the superposition of the offset of both stages. This is confirmed by Fig. 13 where the yield is plotted versus ΔV_{IN}. The results of the Monte Carlo simulation for the conventional latch-type sense amplifier [Fig. 1(a)] are plotted in relation to the curves shown in Fig. 7 for the latch-type sense amplifier of Fig. 1(b). For fair comparison, the same geometry has been used for corresponding transistors. For the conventional latch, a similar influence of V_{INDC} can be observed but it is less intense. Since the circuit of Fig. 1(a) is not suitable for current sense amplifiers, an input dc voltage different from V_{DD} only makes sense for voltage sensing if the bitlines are precharged to a value below V_{DD}. If the latch of Fig. 1(b) operates at $V_{INDC} = 0.6\ V_{DD}$ it is quite competitive to the conventional latch. Moreover, due to the voltage drop across the passgates at the inputs of the conventional latch-type sense amplifier the real input voltage difference ΔV_{IN}, e.g., between the bitlines, has to be much larger than the values of Fig. 13.

[1]For the conventional latch it was found that the enable transistor [M5 in Fig. 1(a)] must not be too strong [18]. A medium W/L ratio gives best results since the delay between the turn on of M1 and M3 is desired to be as large as possible.

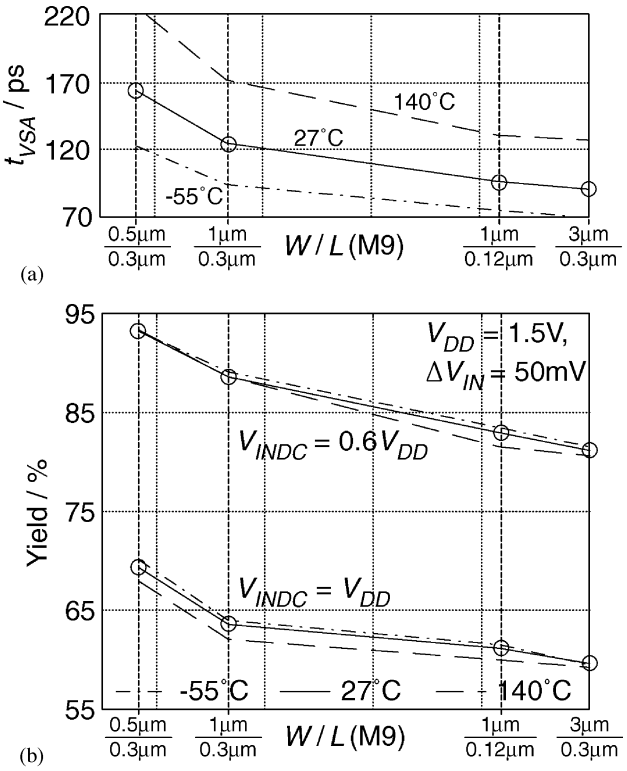


Fig. 12. Influence of temperature and turn on transistor (M9) sizing. (a) Delay ($V_{INDC} = 0.6\ V_{DD}$) and (b) yield versus W/L of M9.

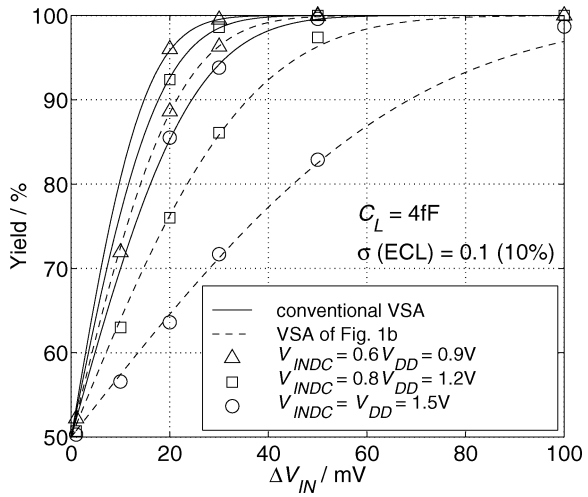


Fig. 13. Yield versus ΔV_{IN}. The yield of the conventional latch circuit [Fig. 1(a)] at $V_{INDC} = V_{DD}$ and of the sense amplifier of Fig. 1(b) at $V_{INDC} = 0.6\ V_{DD}$ is nearly equal.

Fig. 14 compares the delay of both latch types with respect to the input difference ΔV_{IN}. According to (3) and (10) the relation follows a logarithmic function. The speed increases with V_{DD} in either case. However, for the same sizing (W of turn-on transistor M9 set to 1 μm) the conventional sense amplifier is almost twice as fast. If the width of M9 is increased to 3 μm, the latch of Fig. 1(b) becomes as fast as the conventional latch-type sense amplifier. This is due to an increase of the operation current I_o [see (6)]. The delay dependence on ΔV_{IN} is defined by

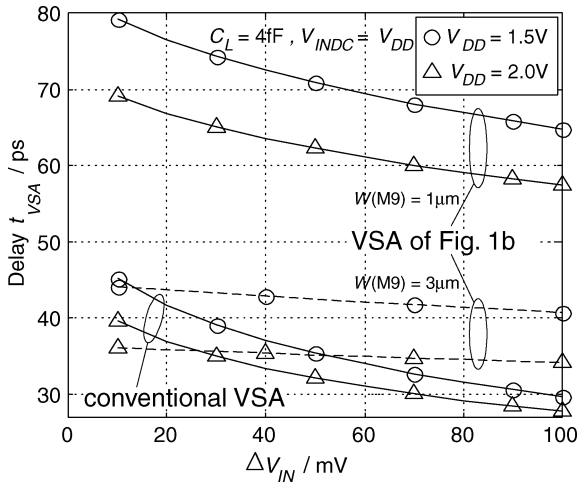


Fig. 14. Simulated delay versus ΔV_{IN}. If the width of transistor M9 is increased to 3 μm the same speed can be achieved for both sense amplifier types.

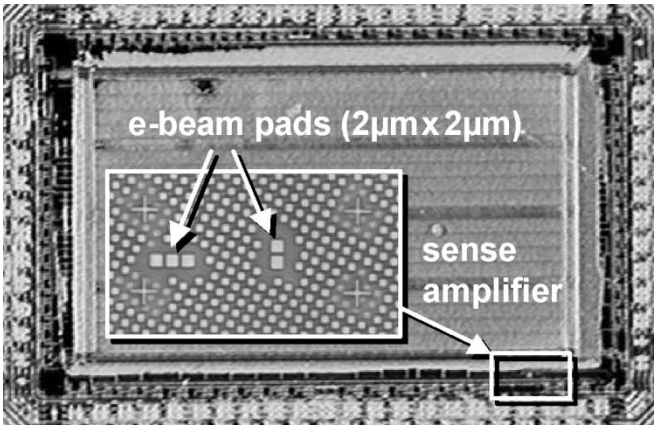


Fig. 15. Microphotograph of the 4 K $\times$ 32-bit SRAM in 130-nm CMOS technology with a separate sense amplifier according to Fig. 1(b) (enlargement).

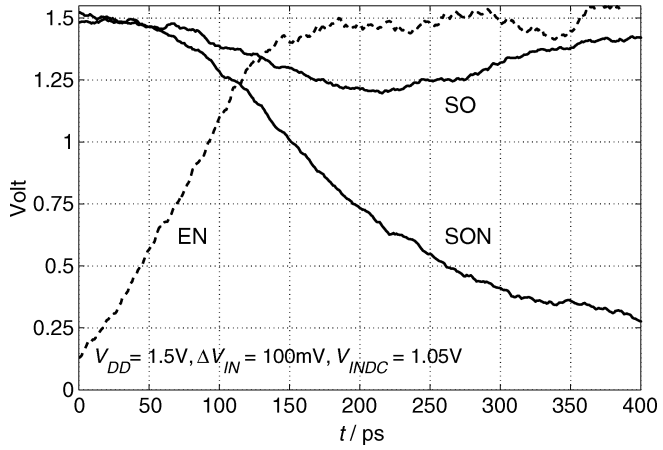


Fig. 16. Measured transient response.

the initial difference V_o. According to (8), V_o is a linear function of ΔV_{IN}. The slope factor as well as the delay itself decreases for larger operation current I_o given by (6). Therefore, the delay gets less sensitive to changes of ΔV_{IN} if V_{DD} and the W/L ratio of M9 are increased.

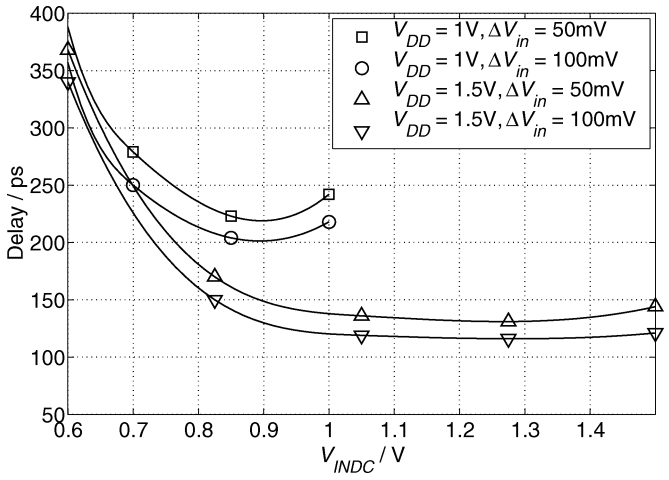


Fig. 17. Measured sensing delay versus input dc voltage.

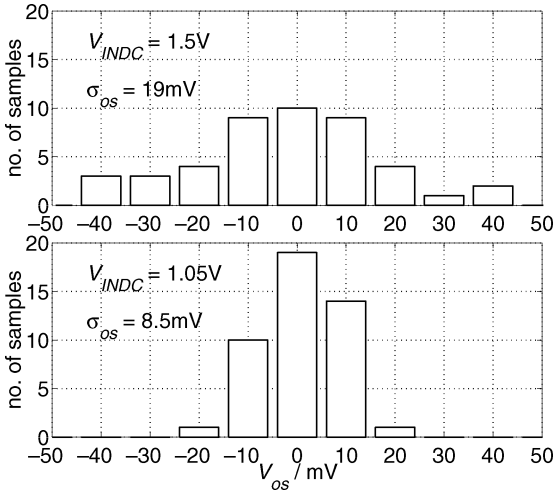


Fig. 18. Histogram of the input offset voltage for $V_{INDC} = V_{DD} = 1.5$ V (top) and $V_{INDC} = 0.7V_{DD}$ (bottom).

IX. Experimental Results

A 4 K $\times$ 32-bit SRAM with the considered voltage sense amplifier has been published previously [8]. In order to measure speed and yield, a separate sense amplifier has been placed beside the memory on the same chip as shown in Fig. 15. Small metal pads for contactless electron beam measurement have been placed at both output pins, at the enable signal and supply. Fig. 16 shows a typical transient response. In this case, a delay (from EN to SON, referred to 50% of V_{DD}) of 119 ps can be obtained. The delay for different supply voltages and input differences is plotted in Fig. 17 as a function of V_{INDC}. The same characteristic behavior as in Figs. 3 and 4 can be observed, even though the measured delay is larger due to a different load at the outputs. This confirms (10), derived in Section III.

Fig. 18 shows the histogram of the input offset voltage and Fig. 19 the corresponding yield obtained from 45 samples. If V_{INDC} is reduced from 1.5 to 1.05 V (70% of V_{DD}), the standard deviation of the offset voltage decreases from 19 to 8.5 mV. Consequently, the yield increases, e.g., from 85.2% to 99.2% at $\Delta V_{IN} = 20$ mV. This agrees well with the theory of Section V.

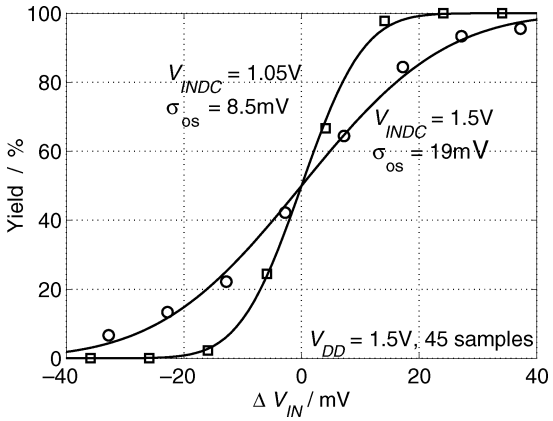


Fig. 19. Measured yield versus input voltage difference.

TABLE I
INFLUENCE OF DESIGN PARAMETERS AND TEMPERATURE ON YIELD AND SENSING DELAY

	ΔV_{IN}	V_{INDC}	V_{DD}	W/L(M9)	T
Yield $Y\uparrow$	$\uparrow$	$\Downarrow$	$\downarrow$	$(\downarrow)$	–
Delay $t_{VSA}\downarrow$	$\uparrow$	–	$\uparrow$	$\Uparrow$	$\downarrow$

Fig. 19 indicates a better matching of the transistors and load capacitances than assumed in Fig. 7. The presented measurement results could also be fitted on a log-normal distribution since the final distribution does not have to be Gaussian if multiple Gaussian variables affect the yield. In any case, the general behavior will be as shown in Figs. 18 and 19.

X. CONCLUSION

Due to strong positive feedback, latch-type sense amplifiers are fast, but in the presence of mismatch and noise an incorrect decision may occur if the input voltage difference is small. In this paper, a latch-type sense amplifier has been analyzed which is competitive with the conventional latch circuit but does not require any input decoupling that may deteriorate noise margin and speed.

An analytical expression for the sense amplifier delay has been derived which distinguishes two output transition phases. Based on this investigation and on Monte Carlo simulations the yield concerning the correct sensing decision has been analyzed. Yield and delay measurements from an implementation in a 130-nm CMOS technology confirm the theory.

Table I summarizes the influence of the most important parameters on yield and delay. A large input difference ΔV_{IN} is always of advantage, but it is limited due to its own generation delay (bitline discharge or current sensing stage). The input dc voltage V_{INDC} has most impact on yield while the geometry of the turn-on transistor M9 and the temperature T is of minor influence. On the other hand, the aspect ratio of M9 basically determines the sensing delay. As long as V_{INDC} is above approximately $0.6\ V_{DD}$, its influence on the delay can be neglected.

A figure of merit has been introduced that indicates an optimum input dc voltage. As a rule of thumb, the optimum performance in terms of speed and yield can be achieved for an input dc level of 70% of the supply voltage. Therefore, the latch is also suitable as a second-stage comparator in current sense amplifiers where the input bias will typically be around $0.7\ V_{DD}$. The required input voltage difference to achieve a correct decision is found to scale proportional to the supply voltage. For example, a yield of 99.8% is achieved for an input voltage swing of 5% of the supply voltage, e.g., 75 mV at 1.5-V supply.

ACKNOWLEDGMENT

The authors wish to thank R. Thanner and B. Krüger for their excellent measurement support.

REFERENCES

[1] B. Wicht, *Current Sense Amplifiers for Embedded SRAM in High-Performance System-on-a-Chip Designs*. Heidelberg, Germany: Springer Verlag, 2003.

[2] T. Kobayashi, K. Nogami, T. Shirotori, and Y. Fujimoto, "A current-contolled latch sense amplifier and a static power-saving input buffer for low-power architecture," *IEEE J. Solid-State Circuits*, vol. 28, pp. 523–527, Apr. 1993.

[3] J. M. Hill and J. Lachman, "A 900 MHz 2.25 MB cache with on-chip CPU—now in Cu SOI," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2001, pp. 176–177, 444.

[4] M. Matsui, H. Hara, Y. Uetani, L.-S. Kim, T. Nagamatsu, Y. Wanatabe, A. Chiba, K. Matsuda, and T. Sakurai, "A 200 MHz 13 mm^2 2-D DCT macrocell using sense-amplifying pipeline flip-flop scheme," *IEEE J. Solid-State Circuits*, vol. 29, pp. 1482–1490, Dec. 1994.

[5] B. Nikolić, V. Stojanović, V. G. Oklobdžija, W. Jia, J. Chiu, and M. Leung, "Sense amplifier-based flip-flop," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 1999, pp. 282–283, 468.

[6] B. Nikolić, V. G. Oklobdžija, V. Stojanović, W. Jia, J. K.-S. Chiu, and M. M.-T. Leung, "Improved sense-amplifier-based flip-flop: design and measurements," *IEEE J. Solid-State Circuits*, vol. 35, pp. 876–884, June 2000.

[7] H. Kawaguchi and T. Sakurai, "A reduced clock-swing flip-flop (RCSFF) for 63% power reduction," *IEEE J. Solid-State Circuits*, vol. 33, pp. 807–811, May 1998.

[8] B. Wicht, J.-Y. Larguier, and D. Schmitt-Landsiedel, "A 1.5 V 1.7 ns 4 k × 32 SRAM with a fully-differential auto-power-down current sense amplifier," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2003.

[9] B. Wicht, T. Nirschl, and D. Schmitt-Landsiedel, "Yield and speed optimization of a latch-type voltage sense amplifier," in *Proc. Eur. Solid-State Circuits Conf. (ESSCIRC)*, Sept. 2003, pp. 409–412.

[10] T. Nirschl, B. Wicht, and D. Schmitt-Landsiedel, "High speed, low power design rules for SRAM precharge and self-timing under technology variations," in *Proc. 11th Int. Workshop Power and Timing Modeling, Optimization and Simulation*, Yverdon-les-Bains, Switzerland, Sept. 2001, pp. 7.3.1–7.3.10.

[11] K. W. Mai, T. Mori, B. S. Amrutur, R. Ho, B. Wilburn, M. A. Horowitz, I. Fukushi, T. Izawa, and S. Mitarai, "Low-power SRAM design using half-swing pulse-mode techniques," *IEEE J. Solid-State Circuits*, vol. 33, pp. 1659–1671, Nov. 1998.

[12] D. Johns and K. Martin, *Analog Integrated Circuit Design*. New York: Wiley, 2000.

[13] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, "Matching properties of MOS transistors," *IEEE J. Solid-State Circuits*, vol. 24, pp. 1433–1440, Oct. 1989.

[14] R. Thewes, C. Linnenbank, U. Kollmer, S. Burges, M. DiLeo, M. Clincy, U. Schaper, R. Brederlow, R. Seibert, and W. Weber, "On the matching behavior of MOSFET small signal parameters," in *Proc. IEEE Int. Conf. Microelectronic Test Structures*, Mar. 2000, pp. 137–141.

[15] E. Seevinck, F. J. List, and J. Lohstroh, "Static-noise margin analysis of MOS SRAM cells," *IEEE J. Solid-State Circuits*, vol. SC-22, pp. 748–754, Oct. 1987.

[16] M. Eisele, J. Berthold, R. Thewes, E. Wohlrab, D. Schmitt-Landsiedel, and W. Weber, "Intra-die device parameter variations and their impact on digital CMOS gates at low supply voltages," in *IEDM Tech. Dig.*, 1995, pp. 67–69.

[17] S. J. Lovett, G. A. Gibbs, and A. Pancholy, "Yield and matching implications for static RAM memory array and sense-amplifier design," *IEEE J. Solid-State Circuits*, vol. 35, pp. 1200–1204, Aug. 2000.

[18] W. T. Lynch and H. J. Boll, "Optimization of the latching pulse for dynamic flip-flop sensors," *IEEE J. Solid-State Circuits*, vol. SC-9, pp. 49–55, Apr. 1974.

Bernhard Wicht (M'01) was born in Halle/Saale, Germany, in 1970. He received the Dipl.-Ing. degree from the Technical University Dresden, Germany, in 1996 and the Ph.D. degree from the Technical University Munich, Germany, in 2002. From 1994 to 1995, he studied at the University of Wisconsin–Madison supported by a scholarship of the German Academic Exchange Service (DAAD).

From 1996 to 1998, he was with MAZ Hamburg GmbH, Germany, as a Designer of analog ASICs. During that time, he was responsible for the design of integrated photo detectors, optical receivers, and analog postprocessing for measurement systems. In 1998, he joined the Institute for Technical Electronics of the Technical University Munich, Germany, as a Research Assistant, where he was working on memory sense amplifiers until 2002 in cooperation with Infineon Technologies. His results have been published in the book *Current Sense Amplifier for Embedded SRAM in High-Performance System-on-a-Chip Designs* (Springer Verlag, 2003). Since January 2003, he has been with the Mixed-Signal Power and Control division of Texas Instruments Deutschland GmbH, Freising, Germany, as a Designer for automotive sensor and power management ICs. He holds five patents and has several more pending.

For his outstanding Ph.D. thesis, Dr. Wicht received the 2003 Texas Instruments Prize at the Technical University Munich.

Thomas Nirschl was born in Dingolfing, Germany, in 1975. He received the Dipl.-Ing. degree from the Technical University Munich, Germany, in 2000. He did a part of his diploma thesis in Japan, supported by Siemens/Infineon Technologies. In 2001, he began working toward the Ph.D. degree at the Institute for Technical Electronics, Technical University Munich, Germany, in parallel to his work at Infineon Technologies. At the university, he is involved in designing circuitry with a novel silicon-based device, the tunneling field effect transistor.

In 2000, he joined Infineon Technologies, where he has been involved in the design and verification of the memory core cells for the standard logic applications. The design of test structures and statistical analysis are the main topics of his work. He holds two patents and has several more pending.

Doris Schmitt-Landsiedel (M'87) received the Dipl. Ing. degree in electrical engineering from the Technical University Karlsruhe, the diploma in physics from the University Freiburg, and the Dr. rer. nat. degree from the Technical University Munich.

Following some research projects on semiconductor lasers and nonlinear optics, she joined the Corporate Research and Development Department of Siemens AG, Munich, Germany, in 1981. There, she worked on scaling problems in MOS devices and on the design of high-speed logic and SRAM circuits. Since 1989, she has been Section Manager of a group of projects in future-generation DRAM design, analog and digital CMOS and BICMOS circuits and design-based yield analysis. Since 1996, she has been a Professor of electrical engineering and Director of the Institute for Technical Electronics at the Technical University Munich. Her research interests are in CMOS and sensor silicon technology, failure analysis and design for manufacturability of digital and mixed signal circuits.