# Time Resolution of NMOS Sampling Switches Used on Low-Swing Signals

Henrik O. Johansson and Christer Svensson, *Member, IEEE*

***Abstract*— A number of recently reported CMOS line receivers and downconversion mixers are based on sampling. A key component in these designs is the NMOS sampling switch. It can sample a very high bandwidth signal, several GHz for a 0.8-$\mu$m transistor. We present an expression for the aperture time for an NMOS switch when the input has low swing. The switch can, under this condition, be modeled as a device that determines a weighted average over time of the input signal. The weight function is derived. The aperture time function shows that the maximum theoretical time resolution for a switch in 0.8-$\mu$m standard CMOS is 21 ps (~48 Gb/s). SPICE simulations agree with the theory. Transient two-dimensional (2-D) device simulations do not contradict the predicted results. Experiments on a switch made in a 0.8-$\mu$m standard CMOS process show successful sampling of every thirty-second bit of a 5-Gb/s data stream.**

***Index Terms*—Aperture time, CMOS integrated circuits, high-speed integrated circuits, sample and hold circuits.**

## I. Introduction

THE maximum tracking speed of an NMOS transistor is high for low swing signals, since both the channel-resistance and the capacitive load can be made small. If the transistor is cut off fast, either high-speed data can be sampled and demultiplexed, or fast analog, e.g., RF, signals can be sampled. Some publications show measured results on receivers/demultiplexers based on parallel sampling which receives low-swing data at speeds up to 3.0 Gb/s in CMOS [1], [2] and NMOS [3]. And measurements on CMOS down-conversion mixers up to 1.5 GHz have been reported in [4] and [5].

One of the speed limiting factors of the sampling receiver and downconversion mixer is the time resolution of the sampling switch. In this paper we present an expression of the aperture time of an NMOS switch, Fig. 1. The aperture time is the width of the sampling operation. The expression shows the time resolution limit and the tradeoff possibilities between the design variables when designing for a specific time resolution.

We use the following terminology for some of the properties of a sample-and-hold circuit, Fig. 2. We show in Section II that a sample is a weighted average of the input signal for an NMOS switch. The weight function is the *sampling function* (SF). The width of the most significant part of the SF is the *aperture time* (or *aperture window width*). The *aperture delay* is the time elapsed from the sampling edge of the clock to the actual moment when the sample is taken, i.e., the middle of the aperture window. (The aperture delay is called aperture time in some literature.) We relate the aperture time to the maximum bit rate under the condition of no *aperture jitter* or *aperture uncertainty* which is the uncertainty of the sampling moment. These definitions is based on the terminology in [6]–[10].

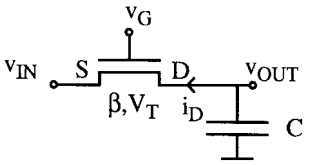


Fig. 1. NMOS sampling switch circuit, with NMOS transistor and load capacitance.

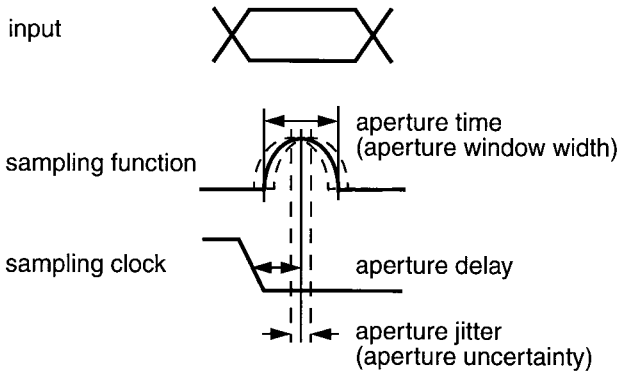


Fig. 2. Definition of some properties of a sample-and-hold circuit.

The organization of this paper is as follows. In Section II the derivation of the aperture time for the NMOS sampling switch is presented. In Section III the expression is compared to circuit simulations. The expression is used to predict the maximum time resolution for a standard CMOS process in Section IV, and the result is checked with transient two-dimensional (2-D) device simulations. In Section V, experiments on a previously reported CMOS line receiver [2] are presented.

## II. Theoretical Study of Time Resolution

Two characteristics of the sampling switch and the operation of it are of importance for the time resolution. First, the

Manuscript received July 3, 1996; revised August 21, 1997.

The authors are with Electronic Devices, Department of Physics and Measurement Technology, Linköping University, S-58183 Linköping, Sweden.

Publisher Item Identifier S 0018-9200(98)00720-3.

tracking speed of the switch must be fast in order to let the output follow the input, second, the transition-time from tracking to hold must be fast enough to utilize the tracking speed. The aim of this derivation is to get an expression of the *aperture time* that includes both these characteristics. The derivations shown are for NMOS transistors but are easily adapted to PMOS transistors.

### A. RC Time

The tracking speed of an NMOS transistor depends on the RC time. When the NMOS transistor is in nonsaturation, the large signal drain current is given by [11]

$$i_D = \frac{\beta}{2}\left[2(v_{\rm GS} - V_T)v_{\rm DS} - v_{\rm DS}^2\right] \tag{1}$$

where $\beta$ is the device transconductance, $v_{\rm GS}$ is the large-signal gate-source voltage, $V_T$ is the threshold voltage, and $v_{\rm DS}$ is the large-signal drain-source voltage. From the derivative of (1) we get the small-signal channel resistance

$$R = \frac{\partial v_{\rm DS}}{\partial i_D} = \frac{1}{\dfrac{k'(W-2\Delta W)}{L-2\Delta L}(v_{\rm GS} - V_T - v_{\rm DS})}. \tag{2}$$

The transistor size is given by the width $W$ and length $L$, and $k'$ is the process transconductance. The parameters $\Delta W$ and $\Delta L$ are the width and length lateral diffusion constants. The relationship between device and process transconductance is $\beta = k'(W-2\Delta W)/(L-2\Delta L)$. To attain a small resistance $v_{\rm GS}$ and $(W-2\Delta W)/(L-2\Delta L)$ should be large and $v_{\rm DS}$ should be low (low-swing signals). For the expressions of the capacitance $C$, see Appendix A.

### B. Sampling Function

Intuitively, one can assume that a MOS switch can be modeled as a device that determines a weighted-average over time of the input signal and adds some dc offset. The weight-function is the SF. The Fourier transform of the SF gives the frequency response of the switch [10]. The dc offset is caused by feedthrough of gate-voltage slopes to the drain.

We will now derive an expression of the output of an NMOS transistor used as a switch. From this expression we identify the SF. In the derivation we assume that the drain-source voltage is small and that there is no clock feedthrough. Small drain-source voltage is no major limitation since low swing is used. No clock feedthrough is harder to motivate, but we will see in Section III-B that the associated errors are reasonable.

The circuit of interest is shown in Fig. 1. The gate voltage is assumed to have a linear slope as shown in Fig. 3. The load capacitor $C$ is given by (A1) in Appendix A. Since we have assumed low swing and no clock feedthrough, the load capacitance will be constant. The device transconductance is calculated with effective transistor length and width. The threshold voltage is given by

$$V_T = V_{T0} + \gamma\left(\sqrt{2|\phi_F| + v_{\rm SB}} - \sqrt{2|\phi_F|}\right) \tag{3}$$

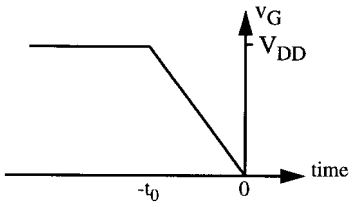


Fig. 3. Gate voltage as function of time.

where $V_{T0}$ is the zero bias threshold voltage, $\gamma$ is the body factor, $2|\phi_F|$ is the surface inversion potential, and $v_{\rm SB}$ is the large signal source-bulk voltage.

The dc-bias of the input and output voltages $v_{\rm IN}$ and $v_{\rm OUT}$, respectively, is $V_{\rm DC}$. Since $v_{\rm DS}$ is assumed small, the transistor is assumed to operate in nonsaturation until it is turned off when $v_G$ is less than $V_{\rm DC} + V_T$, and hence the large signal drain current is [11]

$$i_D = \begin{cases} \dfrac{\beta}{2}\left[2(v_{\rm GS} - V_T)v_{\rm DS} - v_{\rm DS}^2\right], & v_{\rm GS} \geq V_T \\ 0, & v_{\rm GS} < V_T \end{cases}. \tag{4}$$

Since $v_{\rm DS}$ is small compared to $(v_{\rm GS} - V_T)$, the $-v_{\rm DS}^2$ term is omitted (this is not true when $v_{\rm GS}$ is close to $V_T$) and (4) is approximated by

$$i_D \approx \begin{cases} \beta(v_{\rm GS} - V_T)v_{\rm DS}, & v_{\rm GS} \geq V_T \\ 0, & v_{\rm GS} < V_T \end{cases}. \tag{5}$$

Note that this expression of the drain current is also applicable for negative values of $v_{\rm DS}$ with the drain and source terminals defined as in Fig. 1. The gate voltage, Fig. 3, is

$$v_G = \begin{cases} V_{\rm DD}, & t < -t_0 \\ -V_{\rm DD}\dfrac{t}{t_0}, & -t_0 \leq t \leq 0 \\ 0, & 0 < t \end{cases}. \tag{6}$$

The input and output voltages are expressed as

$$v_{\rm IN} = V_{\rm DC} + v_{\rm in} \qquad v_{\rm OUT} = V_{\rm DC} + v_{\rm out} \tag{7}$$

where $v_{\rm in}$ and $v_{\rm out}$ are small-signal voltages. The output voltage and drain current relation is

$$\frac{dv_{\rm OUT}}{dt} = -\frac{i_D}{C}. \tag{8}$$

The derivation needed in order to obtain an expression of $v_{\rm OUT}$ which approximately satisfies (5)–(8) is found in Appendix B. The solution for $v_{\rm OUT}$ after sampling is given by (B2c). We see that the output voltage after the transistor is turned off is constant and is an integral over all previous values of the input multiplied by a weight function

$$v_{\rm sample} = v_{\rm OUT}\left(-t_0\frac{V_{\rm DC} + V_T}{V_{\rm DD}}\right) = \int_{-\infty}^{\infty} v_{\rm IN}(\tau)h(\tau)\,d\tau. \tag{9}$$

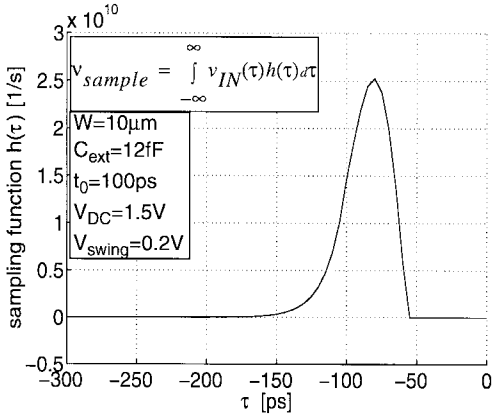


Fig. 4. Plot of a sampling function $h(\tau)$. The output of the NMOS sampling switch consists of a weighted average where the sampling function is the weight function.

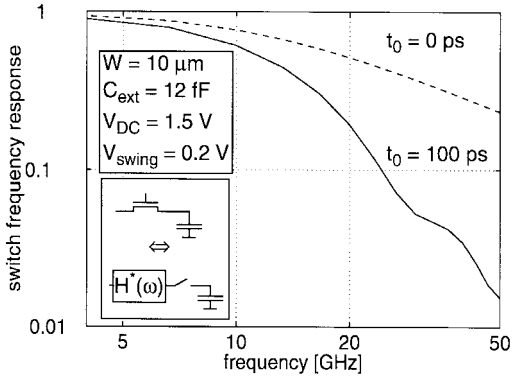


Fig. 5. Frequency response for an NMOS switch for two different gate voltage falltimes, 100 ps (solid) and 0 ps (dashed). The $-3$-dB bandwidth is 6.6 and 10.0 GHz, respectively. The corresponding sampling function for the 100 ps falltime case is shown in Fig. 4.

The weight-function $h(\tau)$ is the SF. By identification, we get (10), shown at the bottom of the page.

A plot of an SF shows a peak with a limited distribution in time, Fig. 4. The integral of the SF over the interval $(-\infty, \infty)$ is one, which is expected since the output should match the input if the input is constant. The (conjugate of the) Fourier transform of the SF gives the frequency response of the switch. The frequency response related to the SF in Fig. 4 is shown in Fig. 5 together with the case when the falltime is 0 ps. In tracking mode, the switch is a low-pass RC filter and the transfer function is the conjugate of the transform of the SF for 0 ps gate falltime. The bandwidth is larger for the tracking operation than for a sampling operation with nonzero falltime. For an ideal switch, the SF is Dirac's delta function, $\delta(\tau)$, which has infinite bandwidth.

### C. Aperture Time

We define the aperture time as the width of the peak of the SF where 80% of the sensitivity is confined

$$w_{80} = t_{90} - t_{10} \tag{11}$$

where the times $t_{10}$ and $t_{90}$ are defined as

$$0.1 = \int_{-\infty}^{t_{10}} h(\tau)\, d\tau \qquad 0.9 = \int_{-\infty}^{t_{90}} h(\tau)\, d\tau. \tag{12}$$

Hence, we allow a maximum attenuation of the signal of 20% through the sampling switch due to aperture effects.

Depending on the values of the input variables to the SF, the $t_{10}$ and $t_{90}$ times will be in different domains of $h(\tau)$, therefore we will get solutions of $t_{10}$ and $t_{90}$ which have different expressions for different domains. When the solutions to (12), (C1a), and (C1b) in Appendix C, are put into (11) and the integrals in the conditions are calculated, we get the following expressions of the aperture time for three different domains

$$w_{80} = \frac{C}{\beta}\frac{\ln(0.9) - \ln(0.1)}{(V_{\rm DD} - V_{\rm DC} - V_T)}, \qquad t_0 \le \frac{2V_{\rm DD}C\ln(1/0.9)}{\beta(V_{\rm DD} - V_{\rm DC} - V_T)^2}, \tag{13a}$$

$$w_{80} = \frac{t_0}{2} - \frac{t_0(V_{\rm DC} + V_T)}{2V_{\rm DD}} - \sqrt{\frac{2t_0C}{\beta V_{\rm DD}}\ln(0.9)} - \frac{C}{\beta}\frac{\ln(0.1)}{(V_{\rm DD} - V_{\rm DC} - V_T)}, \qquad \frac{2V_{\rm DD}C\ln(1/0.9)}{\beta(V_{\rm DD} - V_{\rm DC} - V_T)^2} \le t_0 \le \frac{2V_{\rm DD}C\ln(1/0.1)}{\beta(V_{\rm DD} - V_{\rm DC} - V_T)^2}, \tag{13b}$$

and

$$w_{80} = \sqrt{\frac{2t_0C}{\beta V_{\rm DD}}}\left[\sqrt{-\ln(0.1)} - \sqrt{-\ln(0.9)}\right], \qquad \frac{2V_{\rm DD}C\ln(1/0.1)}{\beta(V_{\rm DD} - V_{\rm DC} - V_T)^2} \le t_0. \tag{13c}$$

$$h(\tau) = \frac{\beta}{C}\exp\left\{-\frac{\beta}{C}\left[\frac{t_0(V_{\rm DC} + V_T)^2}{2V_{\rm DD}}\right]\right\} \cdot \begin{cases} (V_{\rm DD} - V_{\rm DC} - V_T)\exp\left\{\dfrac{\beta}{C}\left[\dfrac{V_{\rm DD}t_0}{2} + \tau(V_{\rm DD} - V_{\rm DC} - V_T)\right]\right\}, & \tau < -t_0 \\ \left(-V_{\rm DD}\dfrac{\tau}{t_0} - V_{\rm DC} - V_T\right)\exp\left\{-\dfrac{\beta}{C}\left[\dfrac{V_{\rm DD}\tau^2}{2t_0} + \tau(V_{\rm DC} + V_T)\right]\right\}, & -t_0 \le \tau < -t_0\dfrac{V_{\rm DC} + V_T}{V_{\rm DD}} \\ 0, & -t_0\dfrac{V_{\rm DC} + V_T}{V_{\rm DD}} < \tau \end{cases} \tag{10}$$

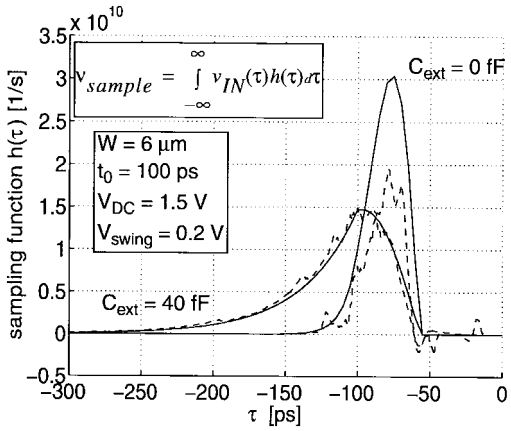


Fig. 6. Plot of the sampling function for two different values of the external capacitive load, $C_{\rm ext} = 0$ and 40 fF. The solid lines are the theoretical functions and the dashed are from simulations.

The aperture time in (13a), which is valid for short fall times of the gate voltage, is the same as for an ideal switch with an R and a C. Note that $1/[\beta(V_{\rm DD} - V_{\rm DC} - V_T)]$ is the channel resistance of the NMOS transistor when $v_{\rm DS} = 0$ V, compare with (2). The expression in (13b) is valid when the aperture time depends on both the gate-voltage slope and the input dc-bias. In (13c), the aperture time is not dependent on $V_{\rm DC}$ but on the slope $V_{\rm DD}/t_0$. Hence, we have the opposite situation compared to (13a), the fall time is slower than the tracking speed.

## III. Comparison of Theory with Circuit Simulations

We compare the derived expressions for the sampling function, SF (10), and aperture time, $w_{80}$ (13a)–(13c), with results from SPICE level 2 simulations. These comparisons are made for a switch in a 0.8-$\mu$m standard CMOS process [12]. The simulations include effects that have been neglected in the derivations: finite input swing, clock feedthrough, the transistor may operate in saturation, and the parasitic capacitances, threshold voltage, and mobility are modeled with higher accuracy. Simulations are made for five cases with varied transistor-width, external load capacitance, gate-voltage-falltime, input dc-bias, and input swing.

### A. Scheme to Extract Sampling Function from Circuit Simulations

The SF cannot be obtained directly from circuit simulations. Instead, the SF is extracted out of data generated by the circuit simulator. In the theoretical derivation we got the result that the output of the switch depends on the input as shown in (9). We assume that this expression is valid also for the circuit simulations and generalize it to represent samples where the sampling is done at different times $t_S$

$$v_{\rm sample}(t_S) = \int_{-\infty}^{\infty} v_{\rm IN}(\tau) h(\tau - t_S)\, d\tau. \tag{14}$$

The SF, $h(t_S)$, is obtained by first taking the derivative of the step response and then reversing the input-variable axis by changing the sign of the input variable. The input to the switch is then $v_{\rm IN}(t) = u(t)$ where $u(t)$ is the unit step function (Heaviside's step function). The derivative of $v_{\rm sample}(t_S)$ with respect to $t_S$ yields

$$\begin{aligned} v'_{\rm sample}(t_S) &= \frac{d}{dt_S} \int_{-\infty}^{\infty} u(t) h(\tau - t_S)\, d\tau \\ &= \frac{d}{dt_S} \int_{0}^{\infty} h(\tau - t_S)\, d\tau \\ &= \frac{d}{dt_S} \int_{-t_S}^{\infty} h(t)\, dt = h(-t_S) \end{aligned} \tag{15}$$

where the integration variable $\tau$ is substituted with $t + t_S$. The result, $h(-t_S)$, has the wrong sign of the input variable so it is changed by a substitution of $-t_S$ with $t_S$

$$h(t_S) = v'_{\rm sample}(-t_S). \tag{16}$$

Since we are interested in the SF when the swing is low, the input is modified to $v_{\rm IN}(t) = V_{\rm SWING} u(t)$ and the SF is then

$$h(t_S) = \frac{1}{V_{\rm SWING}} \cdot v'_{\rm sample}(-t_S). \tag{17}$$

Circuit simulations give numerical data for $v_{\rm sample}(t_S)$ and the derivative is obtained with numerical processing by using an approximation of the derivative operator. The following approximation which removes noise was developed and used:

$$y'(x) \approx \frac{1}{16k} \cdot \sum_{n=0}^{3} \{y(x + n \cdot k) - y[x - (n+1)k]\}. \tag{18}$$

### B. Comparison

SF's, from both theory and simulations, for a number of transistor widths (2.0 and 10.0 $\mu$m) and gate-voltage falltimes (100, 200, 300, and 400 ps) have been compared and the agreement is good. All other parameters are constant. Theoretical and simulated SF's for two different capacitive loads are compared in Fig. 6 and here the agreement is good in the case when $C_{\rm ext} = 40$ fF but only fair when $C_{\rm ext} = 0$ fF. The nonsimilarity depends on an attenuation of the swing. This attenuation is due to the large clock feedthrough when $C_{\rm ext}$ is low in combination with the voltage dependent diffusion capacitance. The input swing is represented by a charge difference after sampling. The diffusion capacitance increases with the clock feedthrough, and hence the charge difference gives a smaller voltage difference.

In Fig. 7 the SF's for two different values of the input dc bias are shown. The agreement is good for the low dc bias, $V_{\rm DC} = 0.5$ V, but is bad for the high dc bias, $V_{\rm DC} = 3.1$ V, due to the assumption of $v_{\rm DS}$ being small compared to $(v_{\rm GS} - V_T)$ is no longer valid when $V_{\rm DC}$ is close to $(V_{\rm DD} - V_T)$.

The aperture time expression is compared to simulations when the design and operation parameters, transistor width, external load capacitance, gate-voltage falltime, and input dc bias are varied. For example, a plot for varied transistor width is shown in Fig. 8. In order to compare only the time resolution, the attenuation of the input swing due to clock

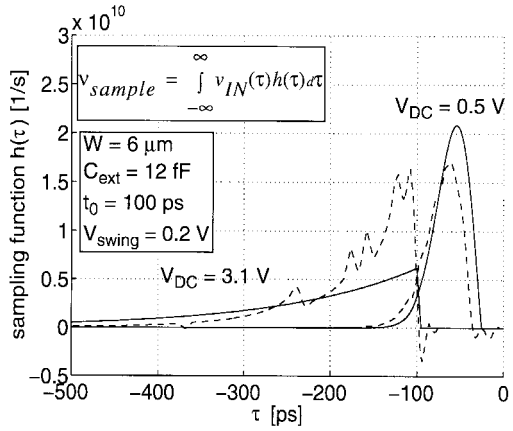


Fig. 7. Plot of the sampling function for two different values of the input dc bias, $V_{\rm DC} = 0.5$ and 3.1 V. The solid lines are the theoretical functions and the dashed are from simulations.

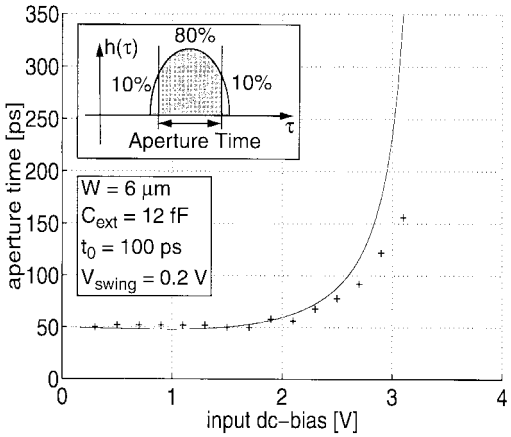


Fig. 9. Aperture time as function of input dc bias. Simulated values are marked with "+" and theory is drawn with a solid line. The definition of aperture time is shown in the upper inset and circuit data in the lower.

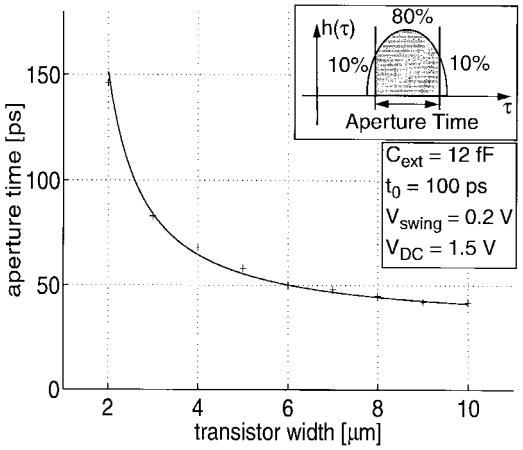


Fig. 8. Aperture time as function of the transistor width. Simulated values are marked with "+" and theory is drawn with a solid line. The definition of aperture time is shown in the upper inset and circuit data in the lower.

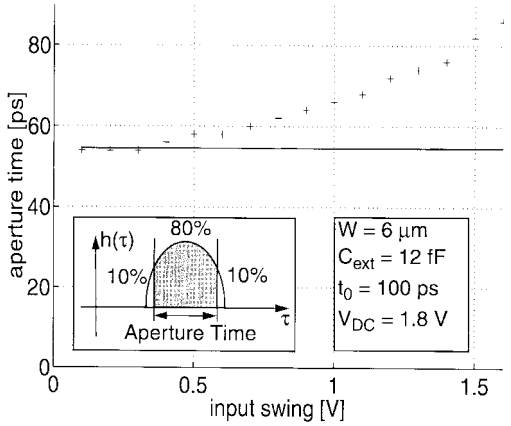


Fig. 10. Aperture time as function of input signal swing. Solid line is theory and "+" are circuit simulation results. The definition of aperture time is shown in the left inset and circuit data in the right.

feedthrough has been compensated. The resemblance is good for all cases except when the input dc bias is varied, Fig. 9. The reason for the deviation is, as explained earlier, that the assumption that $v_{\rm DS}$ is small compared to $(v_{\rm GS} - V_T)$ is not true for large biases.

In the derivation of the aperture time, it is assumed that the input swing is small. Therefore, the theory predicts the same aperture time for all input swings. In reality, the aperture time increases with signal swing. In Fig. 10, aperture times from simulations and theory are plotted and they agree for swings up to ~0.8 V, which is enough for the applications that the derivation is intended for [1]–[5].

## IV. Maximum Time Resolution

We will now apply the aperture time formula on an NMOS switch in a 0.8-$\mu$m standard CMOS process [12]. The switch is realized in common-drain layout style and the switch transistor has a width of 2 × 10 $\mu$m, a layout is shown as an inset in Fig. 11. The external capacitor is assumed small compared to the ones associated with the transistors and is neglected. The capacitive load is then the sum of the capacitances from the diffusion area, $C_{{\rm diff}A}$, diffusion perimeter, $C_{{\rm diff}P}$, gate-drain overlap, $C_{\rm GDoverlap}$, and half the channel, $C_{\rm channel}$

$$\begin{aligned} C &= C_{{\rm diff}A} + C_{{\rm diff}P} + C_{\rm GDoverlap} + \tfrac{1}{2}C_{\rm channel} \\ &= WL_{{\rm diff}D}C_J + (2W + 2L_{{\rm diff}D})C_{\rm JSW} \\ &\quad + 2(W - 2\Delta W)C_{\rm GDO} + (W - 2\Delta W)(L - \Delta L)C_{ox} \end{aligned} \tag{19}$$

where $W$ is the width of one of the two transistors and $L_{{\rm diff}D}$ is the length of the drain. The device transconductance of the switch is

$$\beta = 2\,\frac{W - 2\Delta W}{L - 2\Delta L}\,\mu_0 C_{ox}. \tag{20}$$

We set $V_{\rm DD}$ to 5 V and $V_{\rm DC}$ to 0 V and plot the aperture time as a function of the gate-voltage falltime, Fig. 11. For an ideal gate-voltage transition ($t_{\rm fall} = 0$ ps), the aperture time is below 10 ps. The fastest falling edge attainable in this process is an inverter which uses common-drain layout-style of the NMOS transistors and the capacitances from an external load and pull-up device (e.g., a PMOS transistor) are negligible. The output falltime depends on the input rise time. Hence, we

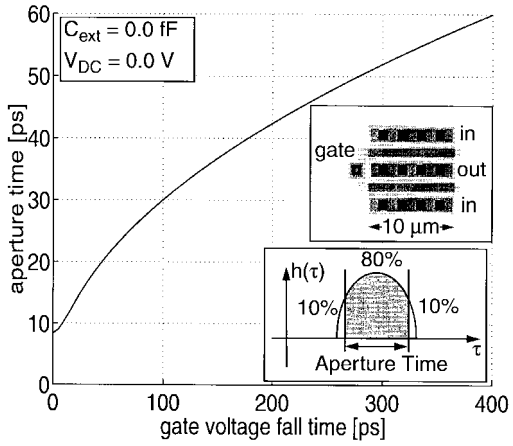


Fig. 11. Aperture time as a function of gate voltage falltime for an NMOS sampling switch with common-drain layout-style. The switch is shown in the upper inset and the definition of aperture time is shown in the lower inset.

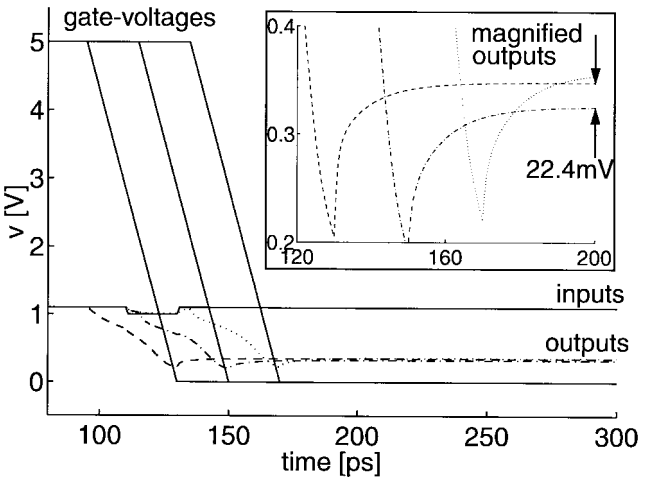


Fig. 12. Input, output, and gate voltages from three device simulations of a sampling switch where the gate-voltage phase has been changed in 20 ps steps. The switch samples on the bit before, on, and after the single zero bit. A magnified part of the output is shown in the upper right corner. The input voltage is the (horizontal) solid line. The dashed, dot–dashed, and dotted line are the output voltages, and the solid (vertical) lines are the gate-voltages.

let the input be driven by a four times larger PMOS transistor. SPICE simulations of extracted layout of this circuit show a falltime (90–10%) of 31.5 ps, which is equivalent to a linear falling edge with a falltime (100–0%) of 35 ps. This yields an aperture time of 18 ps in Fig. 11. This indicates the possibility to sample single bits of a 55-Gb/s input if zero aperture-jitter is assumed.

A word of caution is needed here since we are outside the region where the basis of the derivations, the ordinary current equations, are used normally. To check this result we need to build circuits. Unfortunately, other effects such as aperture-jitter and limited bandwidth of the bondwire-inductance/pad-capacitance filter may limit the performance. Therefore, great care must be taken when designing such a circuit and its package.

### A. Device Simulations

To check if there are some effects that are not shown by SPICE level 2 simulations which degrade the performance, transient 2-D device simulations are done with the MEDICI simulator. The device model used in the simulations is designed by the authors to be fairly similar to the devices in the process which is used for the SPICE simulations [12]. The simulated circuit does not use common-drain layout-style for the switch. The width lateral diffusion constant is not included, since the simulator is two-dimensional.

Three sampling simulations are shown in Fig. 12 with the outputs magnified to the upper right. The input data has 0.1-V swing, a dc bias of 1.05 V, and consists of a single negative pulse (20 ps long). Sampling is done 20 ps before, on, and 20 ps after the single pulse. The gate voltage has a falltime (100–0%) of 35 ps. Note that the actual sampling is done $\sim$20 ps before the moment when the gate voltage intersect the input voltage. As shown in Fig. 12, the 20-ps pulse is detected by the sampling.

Settled output values after sampling are listed in Table I from six simulations with both positive and negative pulses. Assuming we have a differential data input and two samplers, the differential output voltage is also listed. As seen, the swing

TABLE I
OUTPUT VOLTAGES AFTER SAMPLING FROM DEVICE SIMULATIONS

| sampling time (relative to pulse) [ps] | Output voltage for positive pulse [mV] | Output voltage for negative pulse [mV] | Differential output voltage [mV] |
|---|---|---|---|
| -20 | 317.3 | 347.6 | -30.3 |
| 0 | 339.4 | 325.2 | +14.2 |
| +20 | 309.1 | 357.2 | -48.1 |

is reduced from 100 to 14.2 mV for the samples on the pulse. The reduction is due to clock feedthrough and the fact that the sampling function is not close to zero for the adjacent bits to the pulse, i.e., the data rate is too high for the bandwidth of the sampling. Still data can be recovered of the 50 Gb/s data stream with a low-offset comparator with enough time for a decision.

There is no sign of RC transmission-line delay through the channel [13] in the device simulations. Therefore, the partial differential wave equations for an RC transmission line, with R and C values for the channel and also for the poly gate, are solved by numerical methods to check if these delays have any importance. The step response at the output of an NMOS sampling switch, when an edge with 1 ps rise time is applied on the input and the initial gate-source voltage is 4 V, has a delay of 0.9 ps which is so small that it is neglected, Fig. 13.

The propagation of an edge through a 10-$\mu$m wide poly-gate has a delay of 2.7 ps. This is one-sixth of the minimum aperture time, which is on the limit to be neglected. For safety, this skew is added to the minimum aperture time of 18 ps, which gives 21 ps (48 Gb/s). By using a process with silicided poly, this delay could be reduced.

## V. EXPERIMENTS

We have not compared the maximum time resolution predictions above with measurements, but we have measured indications on high time resolution possibilities. We have

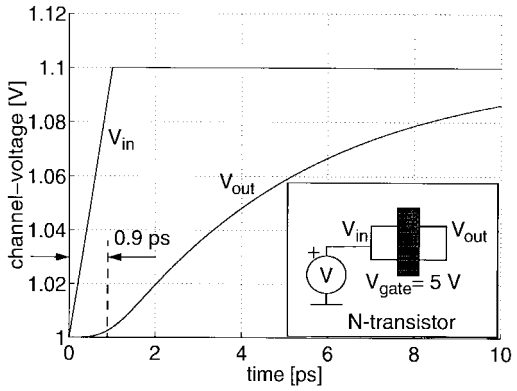


Fig. 13. Waveforms from numerical solution of a lumped RC model of the NMOS transistor channel, indicating a transport delay of 0.9 ps through the channel.

TABLE II
MEASUREMENTS ON A SINGLE 0.8-$\mu$m NMOS SAMPLER OPERATED AT 1/32 OF THE BIT RATE

| Bit-rate [Gbit/s] | Oscilloscope check | BER |
|---|---|---|
| 5.0 | OK | 3.0e-6 |
| 6.0 | OK | no synch |
| 7.0 | OK | no synch |
| 8.0 | NOT OK | no synch |

succeeded to sample out single bits, every thirty-second bit, with one of the sampling units of the circuit described in [2], from a 7-Gb/s data stream. The bit-error rate (BER) was, however, too high for the error detector to synchronize. At 5.0 Gb/s the BER was measured to be 3.0e−6, Table II. No special care was taken in the design of the input pads of the measured circuit and there is an estimated bandwidth (−3 dB) of 3.5 GHz of the LP filter that consists of the bond wire inductance, terminating resistor, and pad capacitance. Therefore, the theoretical results are not contradicted by these measurements.

## VI. CONCLUSIONS

We have derived expressions for the aperture time and sampling function of NMOS sampling switches when used on low swing signals. The PMOS expressions can be obtained by a similar derivation. The derivation is based on the ordinary MOS current equations. We have also shown that the switch under this condition can be modeled as a device which determines a weighted average over time of the input. The derivations match SPICE level 2 simulations well except when the input swing is close to the initial gate source voltage.

When the aperture-time formula is applied on an aggressively designed and operated sampling switch in a 0.8-$\mu$m CMOS process and we also consider delays in the gate, we get a minimum aperture time of 21 ps. This makes it possible to sample out single bits out of a 48-Gb/s data stream. Transient 2-D device simulations do not contradict the results. Measurements on a 0.8-$\mu$m CMOS circuit show successful sampling of every thirty-second bit of a 5.0-Gb/s input.

## APPENDIX A
## EXPRESSIONS FOR THE OUTPUT LOAD CAPACITANCE

The output load capacitance $C$ is the sum of the external load capacitance $C_{\text{ext}}$ and all parasitic capacitances

$$\begin{aligned} C &= C_{\text{ext}} + C_{\text{diff}A} + C_{\text{diff}P} + C_{\text{GDoverlap}} + \tfrac{1}{2}C_{\text{channel}} \\ &= C_{\text{ext}} + WL_{\text{diff}}C_J + (2W + 2L_{\text{diff}})C_{\text{JSW}} \\ &\quad + (W - 2\Delta W)C_{\text{GDO}} + \tfrac{1}{2}(W - 2\Delta W) \\ &\quad \cdot (L - 2\Delta L)C_{ox}. \end{aligned} \tag{A1}$$

Several sources for parasitic capacitance are considered: first, the drain diffusion junction capacitance, where there are contributions from the bottom area, $C_{\text{diff}A}$, and perimeter, $C_{\text{diff}P}$; further, the gate drain overlap capacitance $C_{\text{GDoverlap}}$ and half the gate-channel capacitance $C_{\text{channel}}$. The diffusion capacitances are voltage dependent

$$C_{\text{diff}A} = \text{Area} \cdot C_{\text{JA}} \cdot \left(1 + \frac{v_{\text{DB}}}{V_{\text{PB}}}\right)^{-\text{MJ}} \tag{A2a}$$

and

$$C_{\text{diff}P} = \text{Perimeter} \cdot C_{\text{JP}} \cdot \left(1 + \frac{v_{\text{DB}}}{V_{\text{PHP}}}\right)^{-\text{MJSW}} \tag{A2b}$$

where $C_{\text{JA}}$ and $C_{\text{JP}}$ are the zero bias junction capacitances for area and perimeter, respectively, $v_{\text{DB}}$ is the large signal drain-bulk voltage which we approximate with the dc component $V_{\text{DB}}$. The potentials $V_{\text{PB}}$ and $V_{\text{PHP}}$ are the bulk junction contact potentials for area and perimeter, respectively. The parameters MJ and MJSW are the bulk junction grading coefficients for area and perimeter, respectively.

## APPENDIX B
## SOLUTION OF THE ORDINARY DIFFERENTIAL EQUATION

If we combine (4)–(8) and assume that the gate source voltage $v_{\text{GS}}$ is independent of the voltages $v_{\text{in}}$ and $v_{\text{out}}$ (we have assumed low swing), we get (B1), shown at the bottom of the page. The solution for $v_{\text{OUT}}(t)$ from this equation when

$$-i_D = C\frac{dv_{\text{OUT}}}{dt} \approx \begin{cases} \beta(V_{\text{DD}} - V_{\text{DC}} - V_T)(v_{\text{IN}} - v_{\text{OUT}}), & t < -t_0 \\ \beta\left(-V_{\text{DD}}\dfrac{t}{t_0} - V_{\text{DC}} - V_T\right)(v_{\text{IN}} - v_{\text{OUT}}), & -t_0 \le t < -t_0\dfrac{V_{\text{DC}} + V_T}{V_{\text{DD}}} \\ 0, & -t_0\dfrac{V_{\text{DC}} + V_T}{V_{\text{DD}}} \le t \end{cases} \tag{B1}$$

$$t_{10} = \begin{cases} -\dfrac{t_0}{2} - \dfrac{t_0(V_{\rm DC}+V_T)}{2V_{\rm DD}} + \dfrac{C}{\beta}\dfrac{\ln(0.1)}{(V_{\rm DD}-V_{\rm DC}-V_T)}, & \displaystyle\int_{-\infty}^{-t_0} h(\tau)\,d\tau \geq 0.1 \\ -\dfrac{t_0(V_{\rm DC}+V_T)}{V_{\rm DD}} - \sqrt{-\dfrac{2t_0C}{\beta V_{\rm DD}}\ln(0.1)}, & \displaystyle\int_{-\infty}^{-t_0} h(\tau)\,d\tau < 0.1 \end{cases} \tag{C1a}$$

$$t_{90} = \begin{cases} -\dfrac{t_0}{2} - \dfrac{t_0(V_{\rm DC}+V_T)}{2V_{\rm DD}} + \dfrac{C}{\beta}\dfrac{\ln(0.9)}{(V_{\rm DD}-V_{\rm DC}-V_T)}, & \displaystyle\int_{-\infty}^{-t_0} h(\tau)\,d\tau \geq 0.9 \\ -\dfrac{t_0(V_{\rm DC}+V_T)}{V_{\rm DD}} - \sqrt{-\dfrac{2t_0C}{\beta V_{\rm DD}}\ln(0.9)}, & \displaystyle\int_{-\infty}^{-t_0} h(\tau)\,d\tau < 0.9 \end{cases} \tag{C1b}$$

$V_T$ is assumed constant is

$$\begin{aligned} v_{\rm OUT}(t) &= \exp\left[-t\,\frac{\beta}{C}(V_{\rm DD}-V_{\rm DC}-V_T)\right] \\ &\quad\cdot\int_{-\infty}^{t} v_{\rm IN}(\tau)\,\frac{\beta}{C}\,(V_{\rm DD}-V_{\rm DC}-V_T) \\ &\quad\cdot\exp\left[\tau\,\frac{\beta}{C}\,(V_{\rm DD}-V_{\rm DC}-V_T)\right]d\tau, \qquad t<-t_0 \end{aligned} \tag{B2a}$$

$$\begin{aligned} v_{\rm OUT}(t) &= \exp\left\{\frac{\beta}{C}\left[\frac{V_{\rm DD}t^2}{2t_0}+t(V_{\rm DC}+V_T)\right]\right\} \\ &\quad\cdot\left(\int_{-\infty}^{-t_0} v_{\rm IN}(\tau)\,\frac{\beta}{C}\,(V_{\rm DD}-V_{\rm DC}-V_T)\right. \\ &\quad\cdot\exp\left\{\frac{\beta}{C}\left[\frac{V_{\rm DD}t_0}{2}+\tau(V_{\rm DD}-V_{\rm DC}-V_T)\right]\right\}d\tau \\ &\quad+\int_{-t_0}^{t} v_{\rm IN}(\tau)\,\frac{\beta}{C}\left(-V_{\rm DD}\,\frac{\tau}{t_0}-V_{\rm DC}-V_T\right) \\ &\quad\cdot\left.\exp\left\{-\frac{\beta}{C}\left[\frac{V_{\rm DD}\tau^2}{2t_0}+\tau(V_{\rm DC}+V_T)\right]\right\}d\tau\right), \\ &\quad -t_0 \leq t < -t_0\,\frac{V_{\rm DC}+V_T}{V_{\rm DD}} \end{aligned} \tag{B2b}$$

$$\begin{aligned} v_{\rm OUT}(t) &= v_{\rm OUT}\left(-t_0\,\frac{V_{\rm DC}+V_T}{V_{\rm DD}}\right) \\ &= \frac{\beta}{C}\,\exp\left\{-\frac{\beta}{C}\left[\frac{t_0(V_{\rm DC}+V_T)^2}{2V_{\rm DD}}\right]\right\} \\ &\quad\cdot\left(\int_{-\infty}^{-t_0} v_{\rm IN}(\tau)(V_{\rm DD}-V_{\rm DC}-V_T)\right. \\ &\quad\cdot\exp\left\{\frac{\beta}{C}\left[\frac{V_{\rm DD}t_0}{2}+\tau(V_{\rm DD}-V_{\rm DC}-V_T)\right]\right\}d\tau \\ &\quad+\int_{-t_0}^{-t_0[(V_{\rm DC}+V_T)/V_{\rm DD}]} v_{\rm IN}(\tau) \\ &\quad\cdot\left(-V_{\rm DD}\,\frac{\tau}{t_0}-V_{\rm DC}-V_T\right) \\ &\quad\cdot\left.\exp\left\{-\frac{\beta}{C}\left[\frac{V_{\rm DD}\tau^2}{2t_0}+\tau(V_{\rm DC}+V_T)\right]\right\}d\tau\right), \\ &\quad -t_0\,\frac{V_{\rm DC}+V_T}{V_{\rm DD}} < t. \end{aligned} \tag{B2c}$$

## APPENDIX C
## SOLUTIONS OF INTEGRAL EQUATIONS

The solutions to the integral equations in (12) are given in (C1a) and (C1b) at the top of the page.

## ACKNOWLEDGMENT

The authors would like to thank the following people in their group: J. Yuan for the introduction to high-speed CMOS receivers and J.-E. Eklund and A. Edman for fruitful discussions. Further, they would like to thank A. Djupsjöbacka at Ericsson Components AB for the use of their BER equipment and for assistance during the measurements and M. Östling at the Royal Institute of Technology for the use of their computers and device simulation software.

## REFERENCES

[1] C. K. Yang and M. Horowitz, "A 0.8 $\mu$m CMOS 2.5Gbps oversampling receiver for serial links," in *Proc. ISSCC'96 IEEE Solid-State Circuit Conf.*, San Francisco, CA, Feb. 8–10, 1996, pp. 200–201.

[2] H. O. Johansson, J. Yuan, and C. Svensson, "A 4 Gsamples/s line-receiver in 0.8 $\mu$m CMOS," in *Proc. Symp. VLSI Circuits*, Honolulu, HI, June 13–15, 1996, pp. 116–117.

[3] R. J. Bayruns, E. A. Hofstatter, and H. T. Weston, "A fine line NMOS 3-Gbit/s 12-channel time-division multiplexer-demultiplexer chip set," *IEEE J. Solid-State Circuits*, vol. 24, pp. 814–821, June 1989.

[4] P. Y. Chan, A. Rofougaran, K. A. Ahmed, and A. A. Abidi, "A highly linear 1-GHz CMOS downconversion mixer," in *Proc. ESSCIRC'93 European Solid-State Circuit Conf.*, Sevilla, Spain, Sept. 22–24, 1993, pp. 210–213.

[5] J. Crols and M. S. J. Steyaert, "A 1.5 GHz highly linear CMOS downconversion mixer," *IEEE J. Solid-State Circuits*, vol. 30, pp. 736–742, July 1995.

[6] J. Sylvan, "Using sample and hold circuits in multichannel environments," *Powerconversion and Intell. Motion*, no. 15, pp. 12–15, Apr. 1989.

[7] M. Sauerwald, "Effects of aperture time and jitter in a sampled data system," National Semiconductor Corporation, 1994, Application Note AD-03.

[8] B. Razavi, *Principles of Data Conversion System Design.* Piscataway, NJ: IEEE Press, 1995.

[9] D. F. Hoeschele, *Analog-to-Digital and Digital-to-Analog Conversion Techniques.* New York, NY: Wiley, 1994.

[10] C. Schiller and P. Byrne, "A 4-GHz 8-b ADC system," *IEEE J. Solid-State Circuits*, vol. 26, pp. 1781–1789, Dec. 1991.

[11] J. P. Uyemura, *Fundamentals of MOS Digital Integrated Circuits.* Reading, MA: Addison-Wesley, 1988.

[12] AMS, "0.8 $\mu$m CMOS process parameters," Austria Mikro Systeme International AG, Document Number 9933006 Revision N/C, 1993.

[13] M. Bagheri and Y. Tsividis, "A small signal DC-to-high-frequency nonquasistatic model for the four-terminal MOSFET valid in all regions of operation," *IEEE Trans. Electron Devices*, vol. ED-32, pp. 2383–2391, Nov. 1985.

**Henrik O. Johansson** was born in Dalhem, Gotland, Sweden, in 1966. He received the M.S. degrees in electrical engineering and applied physics in 1993 from Linköping University, Sweden, and Case Western Reserve University, Cleveland, OH, and the Licentiate of Engineering degree in microelectronics in 1997 from Linköping University. He is currently a Ph.D. candidate at the Electronic Devices Group, Department of Physics, Linköping University.

His research interests are in high-speed CMOS off-chip communication and phase-locked loops.

**Christer Svensson** (M'96) was born in Borås, Sweden, in 1941. He received the M.S. and Ph.D. degrees from Chalmers University of Technology, Sweden, in 1965 and 1970, respectively.

He was with Chalmers University from 1965 to 1978, where he performed research on MOS transistors, nonvolatile memories, and gas sensors. He joined Linköping University in 1978 and in 1983 became a Professor in Electronic Devices. He initiated a new research group on integrated circuit design. His present interests are high-performance and low-power analog and digital CMOS circuit techniques for computing, signal processing, and sensors. He has published more than 80 papers in international journals and conferences and holds five patents.

Dr. Svensson was awarded the Solid-State Circuits Council 1988–1989 Best Paper Award. He is a member of the Royal Swedish Academy of Engineering Sciences