A Low-Offset Double-Tail Latch-Type Voltage Sense Amplifier

Daniël Schinkel, Eisse Mensink, Eric Klumperink, Ed van Tuijl, Bram Nauta

Abstract— A latch-type voltage sense amplifier in 90nm CMOS is designed with a separated input and cross-coupled stage. This separation enables fast operation over a wide common-mode and supply voltage range. With a one-sigma offset of 8mV, the circuit consumes 92fJ/decision at 1.2V supply. It has an input equivalent noise of 1.5mV and requires only 18ps setup plus hold time.

Index Terms— Sense Amplifier, Clocked Comparator, Offset, Common-mode range

I. INTRODUCTION

LATCH-TYPE sense amplifiers, or sense amplifier based flip-flops, are very effective comparators. They achieve fast decisions due to a strong positive feedback and their differential input enables a low offset. Sense amplifiers (SA) are hence widely applied in e.g. memories, A/D converters, data receivers, and lately also in on-chip transceivers [1-3]. Especially voltage-mode SA’s, as shown in Figure 1, have become quite popular [4-8] due to their high input impedance, full swing output and absence of static power consumption. However, the stack of transistors in a conventional voltage-mode SA requires quite a large voltage headroom, which is problematic in low-voltage deep-submicron CMOS technologies. Furthermore, the speed and offset of such a circuit are very dependent on the common-mode voltage of the input \( V_{cm} \) [7], which is a problem in applications with wide common-mode ranges, for example A/D converters.

To circumvent these drawbacks, a latch-type voltage sense amplifier with a separated input and cross-coupled stage was introduced in [9], the ‘double-tail’ sense amplifier. There already exist many types of sense amplifiers with a separated differential input stage [10, 11], but those are circuits that do consume static power (such as current-mode logic latches).

This work is supported by the Technology Foundation STW, an applied science division of NWO, and the technology program of the Ministry of Economic Affairs, under project TCS.5791.

This work was carried out at the IC-Design Group, University of Twente, Enschede, 7500 AE, The Netherlands.

D. Schinkel is now with Axiom-IC, Colosseum 28, 7521PT Enschede, The Netherlands (e-mail: daniel.schinkel@axiom-ic.com).

E. Mensink is now with Bruco, Oostermaat 2, 7623CS Borne, The Netherlands (e-mail: eisse.mensink@bruco.nl).

E.A.M. Klumperink, Ed van Tuijl and B. Nauta are with the University of Twente (e-mail: e.a.m.klumperink@utwente.nl, b.nauta@utwente.nl).

A.J.M. van Tuijl is also with Axiom-IC (e-mail: ed.van.tuijl@axiom-ic.com).

Figure 1: Conventional latch-type voltage sense amplifier. The dotted transistors are examples of common variations.

The double-tail sense amplifier is a fully dynamic circuit without any static power consumption. It can operate at a lower supply and has a more stable offset than its conventional counterpart, with a significantly better offset per power ratio at high common-mode voltages. This makes it a very suitable sense-amplifier for offset-critical applications.

This paper discusses the double-tail sense amplifier in more detail. The next section starts with a short discussion of the conventional sense amplifier and its drawbacks. Section three subsequently discusses the operation of the double-tail sense amplifier and what its advantages are. In section four the two are compared. Section five discusses the measurements and section six rounds up with the conclusions.

II. CONVENTIONAL SENSE AMPLIFIER AND ITS DRAWBACKS

An extensive (numerical) analysis of the operation of the conventional sense amplifier from Figure 1 is given in [7]. Here we will suffice with a short description of its operation (also see Figure 2b for signal-graphs of a functionally similar circuit).

The circuit operates similar to other dynamic circuits with a reset or pre-charge phase and an evaluation phase. During the first part of the clock cycle, when the \( Clk \) is low (0V), the output nodes of the cross-coupled inverters (M1-M4) are reset to \( Vdd \), using the reset transistors M7 and M8. The second part of the clock cycle is the actual sense and evaluation
phase. When the Clk signal starts to rise, the tail (M9) of the differential pair (M5, M6) is turned on. The differential pair will discharge the Di and later the output (Out) nodes and an input-dependent voltage difference will build-up on these nodes. When the Di nodes have dropped about a threshold voltage (Vt) under Vdd, then the NMOS transistors of the cross-coupled inverters (M1, M3) turn on, marking the start of the positive feedback. When the Di nodes are about 2Vt lower than the supply, the PMOS transistors of the inverters (M2, M4) also turn on; further enhancing the positive feedback and enabling the regeneration of a small differential voltage at Vin to a full swing differential output.

The circuit has a large number of variations. First of all, it is of course possible to make a complementary version of this circuit, with all the NMOS and PMOS transistors swapped [12]. Second, an often found addition is a transistor between Di+ and Di-, as shown with the dotted transistor in Figure 1. This transistor prevents that the output of the SA becomes floating when the input signal changes polarity after a decision has already been made [6]. The other two dotted transistors in Figure 1 are additional reset transistors that reset the Di nodes [1, 8, 13]. They improve operation at high common-mode input voltages (as discussed later) and also significantly reduce the hysteresis (or memory effect) of the comparator. Other variations include the use of special clocks, for example with non 50% duty cycles, or with slightly different timings for the tail and reset transistors [12] to optimize the timing of the various phases in the operation cycle.

The one thing that many variants of the dynamic sense amplifiers have in common is that the cross-coupled (latching) inverters are placed in series with the differential pair. This series (or cascode) configuration has several drawbacks in applications with limited voltage headroom. The first drawback is the fact that there is only a very short time in which the differential pair actually has gain. This is especially a problem when the input has a common-mode voltage Vcm close to the Vdd (which is often the case for example with memories, and also in transceivers for low-swing data-communication, as in [3]). In that case, the differential pair will enter triode region when the Di node voltages drop below Vdd-Vt and the short period between the end of the reset-phase and this moment is the (sampling or sensing) interval in which the input is amplified and integrated onto the capacitances at the Di nodes. Low amplification (due to a short integration time) of the input signal means a high sensitivity to offset from stages further in the signal chain, in this case offset originating from M1 and M3. In the conventional circuit the Di nodes are not reset to Vdd, but to about one Vt below Vdd (through transistors M1, M3), which further reduces the integration time.

The integration time can be lengthened by also resetting Di to Vdd with the additional reset transistors, as shown with the two dotted transistors in Figure 1. Simulations with the circuit in a 0.13μm CMOS process showed that the additional reset transistors can reduce the input-equivalent offset of M1, M3 by a factor of three (given Vin-com.mode = 1.1V and Vdd = 1.2V).

However, the additional reset transistors solve only one of the drawbacks. The remaining drawback is the fact that there is only one current path, via tail transistor M9, which defines the current for both the differential amplifier and the latch (the cross-coupled inverters). On the one hand, one would like a small tail current to keep the differential pair in weak inversion and obtain a long integration interval and a better Gm/I ratio. On the other hand, a large tail current is desirable to enable fast discharge and regeneration in the latch. For regeneration, it is also not favorable that this tail current depends on the common-mode voltage of the input, which it
A solution to circumvent these drawbacks is to decouple the available current for the latch from the available current for the differential pair. This is accomplished with the double-tail circuit as discussed next.

III. DOUBLE-TAIL CIRCUIT

The schematic of the double-tail sense amplifier is shown in Figure 2a. This topology has less stacking and can therefore operate at lower supply voltages. The double tail enables both a large current in the latching stage (wide M12), for fast latching independent of the Vcm, and a small current in the input stage (small M9), for low offset.

The signal behavior of the double-tail SA is shown in Figure 2b. During the reset phase (Clk=0V), transistors M7 and M8 pre-charge the Di nodes to VDD, which in turn causes M10 and M11 to discharge the output nodes to ground (so there is no need for dedicated reset transistors at the output nodes). After the reset phase the tail transistors M9 and M12 turn on (Clk=VDD). At the Di nodes, the common-mode voltage then drops monotonically with a rate defined by \( Gm_9/C_Di \) and on top of this, an input dependent differential voltage \( \Delta V_{Di} \) will build up. The intermediate stage formed by M10 and M11 passes \( \Delta V_{Di} \) to the cross-coupled inverters and also provides additional shielding between in- and output, with less kickback noise [10] as a result. The cross-coupled inverters start to regenerate the voltage difference as soon as the common-mode voltage at the Di nodes is no longer high enough for M10 and M11 to clamp the outputs to ground. The ideal operating point (Vcm) and the timing of the various phases can be tuned with the transistor sizes.

Compared to the conventional sense amplifier, this circuit requires a few additional transistors, but as the total area can be comparable, as will be shown in the next section. It also requires the availability of both a clock and a clock-not signal. Often, both a clock and a clock-not are already available in a system. If not, then a simple inverter can generate the clock-not from the clock, as the clock-not is allowed to trail the clock signal without a significant impact on performance.

To be able to optimize the design of the sense-amplifier, linear time-variant (LTV) models were developed. For the double-tail sense amplifier, such a model is shown in Figure 3. The signals in the model represent the differential signals in the actual circuit and the time-variance is controlled by the common-mode signals. The input stage acts as an integrator that is reset when the input transistors enter deep-triode. Gm2 is the intermediate stage (M10 and M11) and the last four blocks represent the actual latch (the cross-coupled inverters). As mentioned above, the latch becomes active when the intermediate stage is no longer able to clamp the outputs to ground. The time-constant of the positive feedback of the latch is \( \tau_{latch} = Gm_3/C_2 \).

Although the stages of the actual circuit do not become active or inactive instantaneously, the model is still able to predict the behavior of the actual sense amplifier circuits quite accurately.

The delay of the sense amplifier for example, basically consists of two parts, similar as described in [7]. First, there is the fixed delay for the sampling part in which the differential pair integrates the input onto the Di nodes, without the latch being active. The second part of the delay is from the latch and this part is logarithmically dependent on the sampled input voltage, as the positive feedback creates an exponentially increasing signal.

Due to this exponential increase, it is not necessary to keep the input stage active for more than about three to four times \( \tau_{latch} \) after the latch is turned on, as the contribution of the input of the latch quickly becomes insignificant compared to the internally build-up signal.

To maximize the gain in the input-stage of the sense amplifier, and hence minimize offset and noise contributions from later stages, it would be ideal to first turn on only the input integrator and leave the latch inactive until the integrator has amplified the input with a suitable factor. However, that would come at a cost of an increased delay. In this double-tail circuit it is also not a possible approach, as the timing of the different phases are all linked to the common-mode behavior of the Di nodes. In this circuit, a way to maximize the gain of the input integrator is to keep the input differential pair (M5, M6) operating in (or at the edge of) weak inversion. In weak inversion the Gm/Iq is highest. This is important because the effective amplification factor equals the integrator gain (Gm/Ci) times the integration time (proportional to C/Iq).

A more detailed numerical analysis that also takes the timing of the latch into account can be carried out with the help of the LVT models. These models can be used both for analysis of the amount of input-equivalent offset as for the amount of input-equivalent sampling noise.

Although the rms value for the noise is high in sense amplifiers due to the very wide sampling bandwidths, it is still significantly lower than the offset (as will be confirmed in the measurements section). We therefore used the models to determine how the transistor parameters should be tuned to get the lowest offset at a certain total area, while maintaining high speed. The dominant cause for offset in this circuit is \( Vt \) mismatch, which is inversely related to the square root of the area of a transistor, as are most types of offset sources [14]. The total area is distributed over the transistors in such a way that the (area) derivative of the contribution to the input-equivalent offset are equal for all transistors. The result is a minimum offset for a given area.

The transistors that contribute to the offset are in order of importance: the input transistors (M5, M6), the PMOSTs from
the latch (M2, M4) and the intermediate stage (M10, M11). The NMOSTs from the latch have only a very minor contribution to the offset, as the signal is already strongly amplified when these transistors become active. These NMOSTs are still important for the speed of the latter part of the regeneration phase and are hence optimized for this criterion. The reset transistors (M7, M8) also have a very low contribution to offset. They do however have an impact on the amount of hysteresis. Their effective impedance determines the amount of signal residue at the Di nodes at the end of the reset phase. We dimensioned these reset transistors such that the input equivalent hysteresis is significantly lower than the offset, specified at the maximum clock frequency (lower than 0.5mV with a 3GHz clock).

As all the nodes in the circuit are dynamic – with their capacitances charged and discharged in every clock-cycle – area translates directly to power and an optimal offset/area is roughly equivalent to an optimal offset/power.

Area-scaling (impedance scaling) for the complete design can subsequently be used to match the total input-equivalent offset to any desired value. Note that with this procedure, the total input-equivalent noise will scale with the same factor as offset to any desired value. Note that with this procedure, the additional reset transistors at the Di nodes are a must in the conventional topology, to avoid unrealistically high offsets.

At the nominal conditions, equally high performance figures are obtained for both the double-tail as the conventional variant, with only 100ps Clk-to-output delay (including a clock-buffer) and with only 90fJ/bit consumption for 10mV offset. Note that this 100ps includes about 50ps of clock-to-sample delay (negative setup time of ~50ps).

### IV. Comparison with Conventional

To compare the conventional and the double-tail sense amplifier, both circuits were simulated in a 90nm CMOS technology with \( V_{DD}=1.2V \). Both circuits were optimized with the help of the linear time-variant models and the transistor dimensions were scaled to get an equal offset standard deviation of \( \sigma_{os}=10mV \) at the nominal input common-mode voltage of \( V_{cm}=1.1V \) (the same conditions that are found in [3]). At this high \( V_{cm} \) the additional reset transistors at the Di nodes are a must in the conventional topology, to avoid unrealistically high offsets.

At the nominal conditions, equally high performance figures are obtained for both the double-tail as the conventional variant, with only 100ps Clk-to-output delay (including a clock-buffer) and with only 90fJ/bit consumption for 10mV offset. Note that this 100ps includes about 50ps of clock-to-sample delay (negative setup time of ~50ps).

When the operating conditions are changed, then the two circuits start to behave differently. Figure 4 shows the simulated performance – in terms of (clk to output) delay and energy/cycle – as a function of the supply voltage (with \( V_{cm} \) being 0.1V lower than the supply). It is clear that the double-tail topology is faster and can operate at lower supply voltages, while it consumes approximately the same power as the conventional topology. The double-tail topology could for example operate at a supply of 0.5V at a cost of only 10fJ/cycle with 1000ps delay, versus 2350ps for the conventional circuit. Note that this is for a design that is optimized to operate with \( V_{DD}=1.2V \). Optimization for \( V_{DD}=0.5V \) would give smaller delay, as a wider tail would be used for the input section.

Figure 5 shows the simulated performance as a function of the \( V_{cm} \). Again, the double-tail topology is faster and has a wider common-mode range. The power consumption is nearly equal, except at low input common-mode voltage, where the double-tail topology is able to make faster decisions at the cost of power.

The most interesting difference at high \( V_{cm} \) is not shown in the figure, but is the difference in offset standard deviation. At \( V_{cm}=1.4V \), the offset for the conventional topology increases to \( \sigma_{os}=30mV \), while the double tail offset becomes only \( \sigma_{os}=15mV \), a factor two difference. At common-mode levels lower than the nominal value of 1.1V, the offset of both types of converters remain roughly equal (10% lower for the conventional circuit at \( V_{cm}=0.5V \)).

The standard deviation for the offset was extracted by monte-carlo simulations with 1000 trials. The differential input voltage \( \Delta V_{in} \) of the sense-amplifier was set at a value around the expected standard deviation and the percentage of the trials with the correct positive decision (\( p \)) was subsequently used to calculate the actual offset standard deviation. The effect of hysteresis was excluded from this analysis by first using a large negative \( V_{in} \) during one clock-cycle, followed by the actual decision test. In that way all sense amplifiers start in the same negative state. Assuming a Gaussian distribution with cumulative distribution \( Q \), the offset standard deviation then becomes:

\[
\sigma_{os} = \frac{\Delta V_{in} - V_{hysteresis}}{Q^{-1}(p)}.
\]
The double-tail sense amplifier was implemented in a 1.2V 90nm CMOS technology, as part of a low-swing on-chip data transceiver which operates around a V_{cm} of 1.1V [3]. The V_{cm} can have large variations due to e.g. crosstalk effects. A double-tail SA with dedicated input and output pads (for probe station measurement) was placed on the same die. The layout of the double-tail SA is shown in the inset of the chip micrograph in Figure 6. An SR-latch (made from two NOR gates) is connected to the output of the SA to create static output signals without loss of timing information from the core of the SA. When required, more advanced ‘slave’ stages could be used [6]. A simple SR-latch is for example not ideal when the sense amplifier is used at very high speeds, as it has a ‘non-overlapping’ behavior –the falling edge always comes first– which creates a significant state dependent delay. An SR-latch furthermore also has a state-dependent input capacitance which increases the hysteresis of the total sense amplifier to about 1.5mV (simulated). But, for application in the low-swing data transceiver, the SR-latch sufficed.

Figure 7 shows the measured relative delay under different conditions (the absolute delay is not measurable due to additional delay from the output drivers). As intended, the minimal delay is found at V_{cm}=1.1V. At a V_{cm} of 0.6V, there is still only 20ps increase in delay. The delay versus ΔVin is 44ps/decade under nominal conditions. In comparison, measurements in [7] on a conventional topology in CMOS 0.13μm with V_{DD}=1.5V show a delay versus ΔVin of 100 to 170ps/dec and a 250ps increase in delay when V_{cm} is lowered to 0.6V.

The offset in [7] is also very dependent on the V_{cm} and rises from 8.5mV to 19mV when the V_{cm} changes from 1.05V to 1.5V. For our design, measurements on 20 samples gave an offset of σ_{ox}=8mV, at a V_{cm} of both 1.1V and 0.75V. If desired, area upscaling could further reduce the offset at the expense of power (P ∝ 1/σ_{ox}^2). Offset compensation schemes [8] are a good alternative if the application allows for the added complexity. The power consumed by the SA is 113fJ/decision when ΔVin has 50mV amplitude (f_{clk}=1GHz, V_{DD} = 1.2V, P = 113μW @ 1GHz or 225μW @ 2 GHz), which drops to 92fJ/decision for full-swing inputs.

Setup & hold times are extracted from BER measurements around the zero crossings of full-swing input patterns, as shown in Figure 9. No bit errors are measured outside an interval of 18ps, so the required setup+hold time is smaller...
than 18ps (as input jitter is part of the 18ps). A conventional circuit in 0.18μm CMOS [6] achieves 80ps, which would still be 40ps in 90nm CMOS according to scaling theory. In the double-tail topology, the setup+hold time could be further reduced with a wider tail transistor M9, but at the expense of increased offset and noise due to a shortening of the time that M5/M6 operate in saturation. Simulations predict that the current aperture time is already fast enough to sample data patterns of 40Gb/s, provided that interleaving is used to enable a suitable long regeneration phase.

VI. CONCLUSIONS

In conclusion, the double-tail topology has an added degree of freedom that enables better optimization of the balance between speed, offset, power and common-mode voltage when compared to conventional dynamic sense amplifiers.

This claim is supported by comparing the performance figures with other sense amplifiers, as shown in Table I. For a fair comparison, the published data from the various sense amplifier publications has been scaled to its equivalent value in a 90nm CMOS process, assuming standard scaling rules.

The double-tail sense amplifier also has a better isolation between input and output (lower kickback noise) and it can operate at lower supply voltages than its conventional counterpart.

ACKNOWLEDGMENT

The authors would like to thank Philips Research for chip fabrication, Pascal Wolkotte, Gerard Smit and the STW user committee for helpful discussions and Gerard Wienk and Henk de Vries for assistance.

REFERENCES


Table I. Sense Amplifier Comparison

<table>
<thead>
<tr>
<th>This work (90nm)</th>
<th>[7] (scaled to 90nm)</th>
<th>[8] (scaled to 90nm)</th>
<th>[6] (scaled to 90nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Setup+Hold time</td>
<td>18ps</td>
<td>--</td>
<td>40ps</td>
</tr>
<tr>
<td>Delay/log(Vin)</td>
<td>44ps/dec</td>
<td>&gt;70ps/dec</td>
<td>--</td>
</tr>
<tr>
<td>Input eq. noise</td>
<td>1.5mV</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Offset σ</td>
<td>8mV, 9.5-19mV</td>
<td>15mV</td>
<td>15mV (native)</td>
</tr>
<tr>
<td>Energy/decision</td>
<td>92fJ</td>
<td>--</td>
<td>110fJ</td>
</tr>
</tbody>
</table>


