A 1-V Fully Differential Sample-and-Hold Circuit using Hybrid Cascode 
Compensated DTMOS-based Folded OTA

Christian J. B. FAYOMI,
Hervé F. ACHIGUI
Wireless Smart Devices Labs
Computer Science Department
Université du Québec à Montréal /President Kennedy Hall
201 President Kennedy Avenue, Montreal, CANADA H2X 3Y7

Gilson I. WIRTH
Federal University of Rio Grande do Sul, Departamento de Engenharia Elétrica, Porto Alegre (RS), BRAZIL

Tadeusz KWASNIEWSKI
Carleton University, Department of Electronics
Ottawa, Ontario, CANADA

ABSTRACT
This paper presents the design and preliminary results of a sample-and-hold circuit based on a novel implementation of a dynamic threshold MOS (DTMOS) hybrid compensated folded OTA. The heart of this circuit is a new low-voltage fully-differential hybrid cascode compensated DTMOS folded OTA. The use of DTMOS reduces the input/output common mode requirement on the OTA input while hybrid cascode compensation yields to a higher amplification bandwidth compared to the standard Miller and compensation techniques. To overcome input sampling switch limitations imposed by the low supply voltage we make use of a low-voltage low stress and reliable clock signal doubler. Preliminary post-layout simulation results in a 0.18 μm digital CMOS process show that a resolution greater than 8 bits can be obtained with a 1.0 V supply voltage using a 2 MHz clock signal. Further investigations on the performance limit of the proposed method as well as reliability concerns will be performed on the final experimental test chip.

1. INTRODUCTION
The performance and the accuracy of advanced digital signal processing systems for measurement and signal analysis purposes are strongly determined by the properties of the implemented A/D and D/A conversion functions. Sample-and-hold (S/H) amplifiers used as front-ends in A/D converters are key building blocks of such systems because they make it possible to achieve high conversion speed and high linearity at the same time. On the other hand, the parasitic behavior of the S/H deteriorates the analog input signal. Most of the signal quality impairments caused thereby can not be compensated by computation functions in the downstream digital signal processing units. When designing S/H circuits for lower voltage, quite quickly we encounter severe difficulties: the switch-driving and the operational transconductance amplifier (OTA) problems. Previous low-voltage S/H circuits have either used the switched-opamp technique [1] or utilised the technique of [2]. The switched-opamp technique operates at a low (10 kHz) sampling frequency, which is limited by the opamp turn-on time. With respect to the switched-opamp techniques, the technique of [2] does not turn off the opamp. However, the available output swing is reduced. Other methods have either limited input bandwidth or introduced high design complexity.

Although the above techniques are useful and commercially utilized, it should be understood that they represent a tradeoff with respect to speed, power consumption and design flexibility. Alternatively, bootstrapped analog switches have been extensively used for rail-to-rail switching functions in low-voltage SC circuits [3]. They show a constant charge injection through the whole range of operating supply voltage at the expense of an input dependent clock feedthrough. Reliability problems introduced by this technique have been extensively addressed in [3], in particular, for devices with reduced oxide thickness, tOx. The utilization of the bootstrapped low-voltage analog switch for advanced VLSI processes will be limited in the next few years by the need for a low-stress clock voltage doubler circuit.

In this paper we present a new design strategy to develop a low-voltage CMOS fully differential S/H circuit based on a two-stage hybrid cascode compensated DTMOS folded OTA. We will start with a detailed description of the proposed sample-and-hold circuit in Section 2. Section 3 is related to the design and reliability considerations of the proposed circuit. Section 4 will focus on the preliminary post-layout simulation results and will conclude in Section 5.

2. PROPOSED SAMPLE-AND-HOLD CIRCUIT
The sample-and-hold circuit often used in CMOS A/D converters is depicted in Figure 1. Here, in the acquisition mode (during Φ1) switches S1–S4 are on, S5 and S6 are off. Thus, the OTA is reset and the voltage across sampling capacitors Cs track the analog input. In the transition to the hold mode (during Φ2), S1 and S4, and subsequently, S1 and S2 turn off, and S5 and S6 turn on.

![Figure 1: CMOS sample-and-hold circuit architecture.](image)

While this switching sequence suppresses input-dependent charge injection, fundamentally the circuit suffers from a long hold settling time because the differential output always starts from zero at the beginning of the hold mode. Furthermore, the use of an OTA (or opamp) to establish virtual ground at nodes P and Q makes the operation from low-supply voltages difficult. Especially, (a) the opamp suffers from various trade-offs in dynamic range, linearity, and speed as its supply voltage is reduced, and (b) the OTA (or opamp) requires an input/output common mode level approximately equal to half the supply voltage, thereby limiting the gate-source overdrive.
voltage of \( S_5-S_6 \) and degrading the settling behavior. Two design strategies have been proposed to resolve those limitations.

To overcome the input sampling switch limitations imposed by the low supply voltage we make use of a low-voltage low stress and reliable clock signal doubler depicted in Figure 2 [3]. When the input clock signal \( \phi \) is high, the output signal \( \phi_{b2} \) is low while nodes \( b \) and \( b' \) are precharged to \( V_{DD} \). When clock signal \( \phi \) becomes low, node \( b \) is boosted to \( 2 \times V_{DD} \), while node \( b' \) is discharged to \( V_{DD} - |V_{th,p}| \) and the output signal becomes \( 2 \times V_{DD} \). Feedback transistor \( P_4 \) keeps the gate-drain voltage of transistor \( P_3 \) to a reasonably low level and is subjected to a maximum source-gate voltage of \( V_{DD} + |V_{th,p}| \), which is acceptable in most processes. The purpose of transistor \( N_2 \) is to keep the gate-drain voltage of \( N_1 \) far below \( 2 \times V_{DD} \) in the off state; thus preventing its breakdown. However, \( N_2 \) is subject to a slightly higher drain-source voltage \( V_{DD} + V_{th,n} \). Special care is needed in the layout and sizing of this device.

Such opamps are employed in the switched-capacitor circuits and other applications which do not require the rail-to-rail input stage. It is worth mentioning that the OTA structure has been chosen for its simplicity and the hybrid cascode compensation method can be applied to both rail-to-rail and non rail-to-rail input stage two-stage OTAs. As shown in Figure 3, two separate capacitors, \( C_a \) and \( C_m \), have been used for compensation of the OTA where \( C_a \) is used in the signal path and \( C_m \) in a non-signal path. The use of DTMOS-based first stage reduces the input/output common mode requirement on the OTA [4]. A detailed small signal analysis of the hybrid cascade compensation two-stage non DTMOS OTAs is fully described in [5]. The dominant pole is the same as that of the conventional Miller and cascade compensation schemes with a small difference. The sum of the two capacitors is replaced with one that is used in the Miller and/or cascade compensations. The main advantages of the hybrid cascade compensation scheme can be summarized as follows:

a. It has one extra zero and pole compared to the conventional cascade compensation. However considering \( g_m2 = g_m3 \) and \( C_a = C_m \) the first zero is cancelled with the second pole and the order of the system is reduced to three.

b. The value of its non-dominant poles and zeros are larger by a factor of about 1.4 compared to the conventional cascade compensation. This results in enhancement of the unity-gain bandwidth of the OTA with the same phase margin.

c. The first zero of the hybrid cascade compensation is about two times that of the improved cascade compensation. This requires a higher transconductance for the cascode transistor in the improved cascade compensation where its realization can be a main problem to achieve a good settling behavior. In the case of \( g_m2 = g_m3 \) and \( C_a = C_m \) the zero of the hybrid cascade compensation is much greater than that of the improved cascade compensation since the first zero of the hybrid cascade compensation is cancelled with its second pole.

d. The hybrid cascade compensation also achieves about 3.7-DB DC gain greater than the conventional cascade compensation if both circuits are designed for the same phase margin. The DC gain of the improved cascade compensation is about 6.5 dB less the hybrid cascade compensation due to its small signal transistors’ transconductance. The slew rate and input referred thermal noise of all cascode compensation schemes are approximately the same.

Design and reliability considerations and layout guidelines will be addressed in the following section.

3. **Design and Reliability Considerations**

Although it is difficult to derive an analytical set of design equations for the proposed circuit, some general design and reliability guidelines are given.

The capacitor value \( C_{c2} \) of Figure 2 should be chosen as small as possible for area considerations but large enough to sufficiently charge the load to the desired voltage levels. As process technologies continue to scale down, lateral fringing becomes more important. The lateral spacing of the metal layers shrinks with scaling, while the thickness of the metal
layers and the vertical spacing of the metal layers stay relatively constant. This means that structures utilizing lateral flux enjoy a significant improvement with process scaling, unlike conventional structures that depend on vertical flux [3]. A higher capacitance density can be achieved by using a lateral flux capacitor in addition to vertical flux. These capacitors are called fringing-effects-based capacitors and have been used to implement the capacitors in the design.

The device sizes (P1, N2, and N1 of Figure 2) should be chosen to create sufficiently fast rise and falling times at the load. The load consists of the gate capacitances of the switching devices S1 through Sn (Figure 1) and any parasitic capacitance. Capacitors Cg1 (Figure 2) must be sufficiently large to supply charge to the gate of the switching device in addition to all parasitic capacitances in the charging.

The reliability of the circuit (clock signal doubler) can be further improved by carefully laying out some of the critical devices. Although the relative voltages between gate, source, and drain do not exceed VDD, the drain-to-substrate and source-to-substrate voltages of some devices exceed VDD (assuming n-well process). In particular, device N2 of Figure 2 is subjected to this large voltage. Typically a CMOS technology is designed such that the reverse breakdown of a stand-alone n+/p- junction is approximately 3×VDD. This voltage, called BVSS, is tested under the condition that the gate and source are grounded. In an nMOS transistor, however, a n+/p+ junction is formed between the n+ drain (or source) and the p+ channel-stop implant. Using a circular drain or "doughnut" layout, the p+ channel stop can be removed around the drain to add another 1-2 V to the breakdown voltage. Thus, for improved reliability, the drain of devices N2 should be laid out circularly. The "doughnut" transistor structure also has the advantage of being area-efficient with less parasitic capacitance at the drain, making this structure attractive for high-speed applications [3]. Finally, the "off" drain-source voltage of N2 can exceed VDD introducing a potential punch-through problem. If, however, the channel length of this device is increased (typical 1.5×Lmin) this punch-through voltage can be significantly beyond the supply voltage.

In the technical sense, the use of DTMOS in the input differential pair for the two-stage OTA indeed helps in reducing the requirement on the input common mode, VICM. This however is particularly true for continuous time systems, not sampled systems. As presented in the S/H schematic (Figure 1), in one phase VC, is sampled on Cg, which can be different than VDD/2 and can be as low as VSS if need be, hence the use of DTMOS is not a necessity. Moreover, for rail to rail output stages where the output is allowed to swing close to VDD and VSS, the use of DTMOS can be problematic. In the sample phase, most S/H have a reset phase to minimize output common mode output drift. This leads to large transitions in the output once the hold phase is engaged, which in turn are fed back to the input. Using a DTMOS input stage then will completely forward bias the bulk-drain junction for swings close to VSS, leading to reliability issues. For this reason the two-stage OTA common-mode input VC is set to VDD/2. The minimum and maximum input signals are respectively kept in the range of 1/4×VDD and 3/4×VDD.

The following section will focus on the preliminary post-layout simulation results.

4. Preliminary Results

The CMOS fully differential S/H circuit based on the two-stage hybrid cascode compensated DTMOS folded OTA circuit of Figure 1 has been implemented in a standard 0.18 μm CMOS process to show the application of the above procedure. The process parameters are summarized in Table I.

<table>
<thead>
<tr>
<th>Table I: 0.18 μm digital CMOS average parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>Vth0 (V)</td>
</tr>
<tr>
<td>μCox (μA/V^2)</td>
</tr>
<tr>
<td>gamma, γ (V^-0.5)</td>
</tr>
<tr>
<td>K1 (V^-0.5)</td>
</tr>
<tr>
<td>K2</td>
</tr>
<tr>
<td>Tox (nm)</td>
</tr>
</tbody>
</table>

Figure 4 shows the post-layout simulation results of the two-stage hybrid cascode compensated DTMOS folded OTA using a 1-V supply voltage. An open-loop gain of 78 dB, a 67.0 MHz unity gain frequency and a 67 degree phase margin under a load capacitor of 0.25 pF while dissipating 3 mW including bias circuit.
simplifying the spectral analysis resulting from the FFT. Thus, the dynamic performance such as signal-to-noise ratio (SNR) and total harmonic distortion (THD) can be tested, [6], [7].

To further evaluate the non-linearity, we generated the power spectral density (PSD) of the sample-and-hold output signal via Matlab. A spectrum analyzer connected to the output node can also be used. A total of \( N = 256 \) data samples were used with 21 signal cycles to ensure the coherency sampling requirement.

The simulated spectrum of Figure 5 shows a sample-and-hold operation of an input signal amplitude of 0.5 \( V_{pp} \) under a 1-V supply voltage using a 1 MHz sampling clock frequency. Figure 6 shows the spectrum of the sample-and-hold waveform under sampling frequency of 2 MHz based on coherency sampling method. The input frequency of 164.0625 kHz and a signal amplitude of 0.5 \( V_{pp} \) under a supply voltage \( V_{DD} = 1 \) V is used. The measured SINAD obtained is 80.7099 dB a in the Nyquist bandwidth. This corresponds to an effective resolution of 13.11 bits.

4. Conclusion

This paper presents the design and preliminary results of a sample-and-hold circuit based on a novel implementation of a dynamic threshold MOS (DTMOS) hybrid compensated folded OTA. The heart of this circuit is a low-voltage fully-differential hybrid cascode compensated DTMOS-based folded OTA. The use of DTMOS reduces the input/output common mode requirement on the OTA input while hybrid cascode compensation yields to a higher amplification bandwidth compared to the standard Miller and compensation techniques.

Preliminary post-layout simulation results in a 0.18 \( \mu m \) digital CMOS process show that a resolution greater than 8 bits can be obtained with a 1.0 V supply voltage using a 2 MHz clock signal. Further investigations on the performance limit of the proposed method as well as reliability concerns will be performed on the final experimental test chip. The method used for post-layout simulation could be applied to the test chip characterization.

The proposed circuit can be used in low-voltage high-speed and high-resolution applications such as flash, successive approximation and pipelined ADCs, to name just a few examples.

Acknowledgments

The authors would like to acknowledge the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Foundation for Innovation (CFI), the Fonds du Québec pour la Recherche sur la Nature et les Technologies (FQRNT), and the Canadian Microelectronics Corporation (CMC) and the Microelectronics Strategic Alliance of Quebec (ReSMIQ). They also wish to thank the reviewers and Mary-Rose Morrison for their comments, suggestions and inputs to this work.

References