

Ultra-Low Power DLMS Adaptive Filter For Hearing Aid Applications

Hyung-il Kim
Purdue University

Department of Electrical and Computer Engineering
West Lafayette, IN 47907, USA
+1-765-494-3372

hyungil@ecn.purdue.edu

Kaushik Roy
Purdue University

Department of Electrical and Computer Engineering
West Lafayette, IN 47907, USA
+1-765-494-2361

kaushik@ecn.purdue.edu

ABSTRACT

We present an ultra-low power DLMS (delayed least mean square) adaptive filter working in the sub-threshold region for hearing aid applications. Sub-threshold operation was accomplished by using a parallel architecture with pseudo NMOS logic style. The parallel architecture enabled us to run the system at a lower clock rate with a reduced supply voltage, while maintaining the same throughput. Pseudo NMOS logic operating in the sub-threshold region (Sub-Pseudo NMOS) provided better power-delay product than sub-threshold CMOS (Sub-CMOS) logic. Simulation results show that the system can process voice signals at a throughput of 22kHz with a supply voltage of 400mV and achieve 91% improvement in energy compared to the non-parallel architecture using standard CMOS logic.

Keywords

DLMS adaptive filter, sub-threshold operation, parallel architecture, Sub-Pseudo NMOS, Sub-CMOS

1. INTRODUCTION

Modern hearing aid devices are compact enough to fit in the ear canal, even with sophisticated signal processing algorithms. Up to 10 independently programmable channels with respective automatic gain control are contained in current-day hearing aid DSP chips. Also other advanced signal processing techniques such as adaptive filtering for interference cancellations are becoming reality in the next generation hearing aid devices. But due to the miniaturized battery size, completely-in-the-canal hearing aids have a battery life of only around 100 hours. This is troublesome because the patient has to change the battery every several days. Thus, obtaining the required performance within a limited power budget is the most challenging goal in custom hearing aid device designs.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ISLPED'01, August 6-7, 2001, Huntington Beach, California, USA.
Copyright 2001 ACM 1-58113-371-5/01/0008.

Advanced signal processing techniques are used in hearing aid devices to resolve the acoustic feedback problem. This defect occurs when the amplified signal from the speaker leaks back to the microphone. This originates from the leakage sound that is propagated through the human body or through the clearance between the hearing aid device and the ear canal. Annoying defects of acoustic feedback, such as whistling, screeching or howling deteriorate the sound quality and limit the effective gain for stable operation. Several different techniques for reducing acoustic feedback in hearing aids were investigated [1,2]. The various methods fundamentally utilize adaptive noise cancellation to eliminate acoustic feedback.

Hearing aid devices are clearly one of the most suitable application areas for sub-threshold logic since ultra-low power consumption takes first priority, while the clock rate is merely in the kHz range. Digital sub-threshold logic has successfully achieved ultra-low power consumption in areas where performance is of secondary importance [3,4].

By simply reducing the supply voltage below the threshold voltage, we can operate the circuits using only the minute leakage current. Although the delay increases rapidly, ultra-low power can be achieved without major alteration of the circuit. In the sub-threshold region, the current through a transistor has an exponential dependency upon gate voltage, threshold voltage and temperature [5]. However the W/L ratio has a linear dependency upon the transistor current, so sizing has much less effect on the transistor current than it had in the normal strong inversion region. Thus the disadvantages of ratioed logic in the strong inversion region such as degradation of noise margin and VTC (voltage transfer characteristic) are diminished in the sub-threshold region. Simulations showed that Sub-Pseudo NMOS can have 46% less power-delay product than the Sub-CMOS logic.

In our prototype implementation, we have demonstrated an ultra-low power adaptive filter for hearing aid devices, operating in the sub-threshold region. For the system to run at a supply voltage lower than the threshold voltage, a non-folded parallel architecture of the DLMS algorithm was realized. Using this parallelism, we could run the system at a lower clock rate, reduce the supply voltage, and thus achieve ultra-low power dissipation, maintaining the same throughput. Pseudo NMOS logic was used instead of standard CMOS logic to utilize the better power-delay product. As a result, we were able to scale down the supply voltage to 400 mV

and achieve a 91% improvement in power compared to the folded architecture of an LMS adaptive filter using standard CMOS logic.

The remainder of this paper is organized as follows: In section 2, characteristics of the MOS transistors and advantages of Sub-Pseudo NMOS is discussed. Section 3 deals with various architectures of adaptive filter implementation, focussing on the advantages of the DLMS algorithm for non-folded pipelined adaptive filter architectures. Implementation of the DLMS adaptive filter for hearing aid devices using different architecture and different logic families is explored in section 4. Comparisons showing the superiority of a parallel architecture using Sub-Pseudo NMOS logic are presented in section 5. And conclusions are made in section 6.

2. SUB-THRESHOLD LOGIC

2.1 Sub-Threshold Characteristics of a MOS Transistor

Characteristics of a MOS transistor operating in the sub-threshold region are significantly different from that in the normal strong inversion region. The current through the MOS transistor, which is a quadratic (or linear for short channel transistors) function of the gate voltage in the strong inversion region, becomes an exponential function in the sub-threshold region. (The transistor current characteristic is shown in Fig. 1.) As described in the following equation, it is also an exponential function of threshold voltage (V_{th}) and temperature (T).

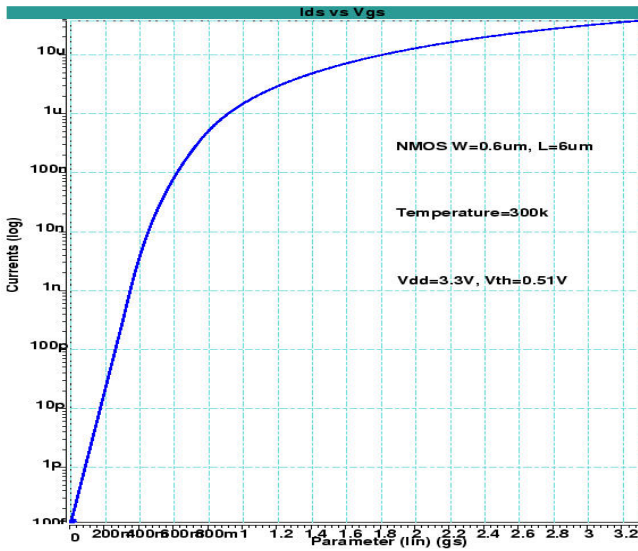


Figure 1. Relationship between transistor current and gate voltage. figure.

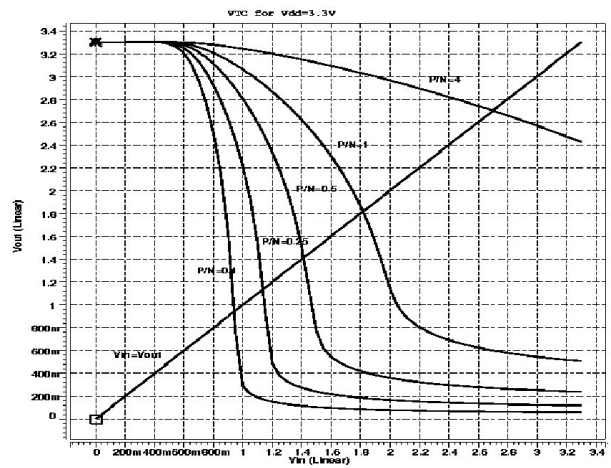
$$I_{ds} = \mu_{eff} C_{ox} \frac{W}{L} (m-1) \left(\frac{kT}{q} \right)^2 e^{\frac{(V_g - V_{th})}{m k T / q}} \left(1 - e^{-\frac{V_{ds}}{k T / q}} \right)$$

eq. (1)

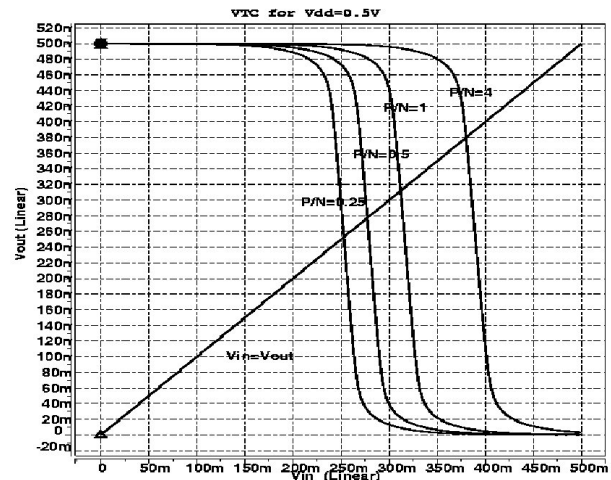
Variations of threshold voltage and temperature have an exponential effect on the sub-threshold current and without any regulating scheme, the transistor will experience severe current variations due to these dominating parameters. In the sub-threshold region, circuits are operated using the minute leakage current, thus resulting in ultra-low power consumption. However, since the driving current decreases exponentially, the delay of the circuit will increase sharply. Hence, sub-threshold operation of logics can only be applied to limited areas where performance is of secondary importance

2.2 Sub-Pseudo NMOS Logic

Pseudo NMOS logic operating in the strong inversion region has the advantage of reduced load capacitance, less interconnection and smaller area over standard CMOS. For the improved performance, we have to pay the cost of higher leakage current,



(a)



(b)

Figure 2. VTC of a pseudo NMOS inverter in the (a) strong inversion region and (b) sub-threshold region for different W/L ratios.

less noise margin and degradation of the VTC (voltage-transfer-characteristic) compared to the standard CMOS. This is described in Fig 2-(a).

Attractively, Sub-Pseudo NMOS logic inherit all the advantages mentioned above without going through degradation in noise margin or VTC. From equation (1), the current is a linear function of the W/L ratio while current exponentially changes with gate voltage, threshold voltage and temperature. The effect of the W/L ratio on the transistor current is much less than that in the strong inversion, so the VTC of the Sub-Pseudo NMOS does not suffer much due to variations of the transistor sizes. As we can compare from figures 2-(a) and 2-(b), the VTC of the Sub-Pseudo NMOS inverter becomes similar to that of an ideal standard CMOS inverter. The exponential relationship between the current and gate voltage lowers the Vol (output low voltage) providing robustness. For Sub-Pseudo NMOS logic designs, more aggressive sizing is possible for better performance since more noise margin is guaranteed by the exponential nature of the device. Whereas in the strong inversion, careful transistor sizing is required.

Sub-Pseudo NMOS is also more efficient than Sub-CMOS in terms of power-delay product. Simulation results of Sub-Pseudo NMOS logic gates and Sub-CMOS logic gates are compared in table 1. We observe that for nand, nor, xor gates and inverters, the power-delay product is always better in pseudo NMOS logic by at least a factor of 1.9.

3. DELAYED LMS ALGORITHM FOR PIPELINED ARCHITECTURE

Adaptive filters have been successfully applied in a wide variety of areas including channel equalization and noise cancellation. LMS algorithm is generally the most popularly used adaptation technique because of its simplicity and ease of computation [6,7]. Architectures of LMS adaptive filters can be implemented in several different ways depending on whether the primary design constraint is area, power or performance.

First, we consider a folded architecture as shown in Fig. 3, similar to a DSP processor core, where computation is performed by a single functional unit. Since computation must be done by a single functional unit, additional memory is required to store the data and weight coefficients. Also a control logic provides the appropriate control signals for each state. This folded architecture has the advantage of small area but it requires multiple clock cycles and has a total delay of $(2N + 1)t_m + 2Nt_a$ to process 1

data sample, where t_m , t_a , N are the delay of the multiplier, delay of the adder and the filter length, respectively [8]. As the tap length of the LMS filter increases, this architecture suffers from a large delay.

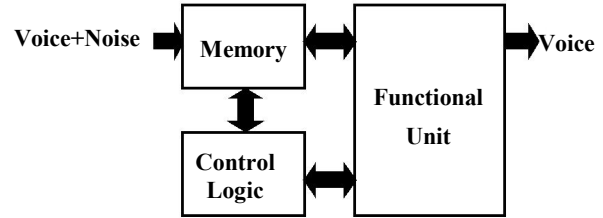


Figure 3. A folded architecture of an LMS adaptive filter.

A different architecture for parallel processing is to have a filter with multiple functional units cascaded to each another as shown in Fig. 6. Computations can be completed in a single clock cycle, but still a delay proportional to the filter length is present. This is due to the computation of the feedback error, which must be completed before the weights can be updated [8]. Within the confines of the conventional LMS algorithm, pipelining is unrealizable even with this non-folded architecture, thus the throughput of the LMS filter is limited.

Recent studies show that by inserting a fixed delay in the conventional LMS algorithm, a pipelining architecture can be realized with a desirable adaptation characteristic [9]. Following equations describe the filtering and weight updating procedures of the resulting DLMS algorithm which is nothing other than the conventional LMS algorithm using the delayed version of the feedback error and input data for weight update.

$$W(n+1) = W(n) + \mu e(n-N)U(n-N)$$

$$e(n-N) = d(n-N) - W^T(n-N)U(n-N)$$

where μ , N , $e(n)$, $d(n)$ are the gain constant, filter length, error and desired signal, respectively. The filter coefficients ($W(n)$) and input data ($U(n)$) are expressed in vector notations, respectively as

Table 1. Power and delay comparisons of Sub-CMOS logic and Sub-Pseudo NMOS logic (Vdd=500mV).

	CMOS			Pseudo NMOS		
	Power(W)	Delay(sec)	PDP(W*sec)	Power(W)	Delay(sec)	PDP(W*sec)
INV	2.90e-08	6.84e-08	1.98e-15	3.10e-08	4.52e-08	1.40e-15
NAND	3.33e-08	1.24e-07	4.13e-15	2.57e-08	7.67e-08	1.97e-15
NOR	3.60e-08	1.33e-07	4.78e-15	4.58e-08	4.39e-08	2.01e-15
XOR	8.36e-08	4.30e-07	3.60e-14	7.63e-08	2.48e-07	1.89e-14

$$W(n) = [\omega_0(n) \ \omega_1(n) \ \dots \ \omega_{N-1}(n)]^T$$

$$U(n) = [u(n) \ u(n-1) \ \dots \ u(n-N+1)]^T$$

In this particular equation, the inserted delay is equal to the length of the filter N . By distributing this delay throughout the systolic architecture using retiming techniques, the critical delay can be reduced to $2t_m + t_a$ which is independent of the filter length [8].

Although the DLMS algorithm has some drawbacks such as longer convergence time, larger minimum mean square error and larger output latency compared to the conventional LMS filter, the DLMS filter architecture provides a significant improvement in performance by utilizing pipelining. For better performance, we can use the parallelism of the non-folded architecture to achieve less power dissipation. By using a non-folded architecture instead of a folded architecture, we can have the same throughput with a reduced clock frequency, and at the same time, a reduced supply voltage. Of course the area will increase because the parallel architecture requires multiple functional units, but we can achieve a significant improvement in power consumption through this trade off. By utilizing this parallel architecture in our hearing aid application, we can scale the supply voltage down to the sub-threshold region and achieve ultra-low power consumption while maintaining the same throughput.

4. SYSTEM IMPLEMENTATION

The prototype adaptive filter for hearing aid applications in Fig. 4 has a few features, which differ from the traditional adaptive filters. First, a gain normally up to 20 dB is inserted in the loop to amplify the acoustic signal, secondly instead of having a separate reference signal input, the delayed error output of the filter is fed back as the reference signal, and third, a delay Δ_x , is inserted to compensate for the acoustic feedback delay. For our prototype design, the filter length was 12 and Δ_x for the delay compensation was 22 samples. Also the word length of each signal was 8 bits and the gain was set as unity.

The implementation of the adaptive filter was completed using two different architectures: the folded architecture and the non-folded architecture. For the folded architecture shown in Fig. 3, each block was described using VHDL. For our design, where the filter length was 12, it turned out that 34 clock cycles were required for a single data sample to be processed. Logic synthesis was done using the CMU standard cell library and the HP 0.35 μ m bulk CMOS technology was used when extracting the layout.

The non-folded DLMS architecture was realized by cascading the functional unit modules as in Fig. 6. Similar to the implementation of the folded architecture, the functional unit used in the non-folded design was described in VHDL. From the signal flow graph of the module shown in Fig. 5, the maximum throughput of $1/(2t_m + t_a)$ could be achieved.

To verify the merits of Sub-Pseudo NMOS, both pseudo NMOS and standard CMOS logic styles were implemented separately. For the pseudo NMOS implementation, we made a new pseudo NMOS cell library by substituting a grounded PMOS device for the whole pull-up network in the layouts of the CMU standard library cells. Transistor sizing of the pseudo NMOS was done in

an aggressive fashion geared towards better performance. For example, the W/L ratio of the pseudo NMOS inverter was designed as 1.0, which from Fig. 2 is inappropriate in the normal strong inversion region. This strategy is possible for Sub-Pseudo NMOS, since they are less sensitive to sizing than pseudo NMOS in the strong inversion region.

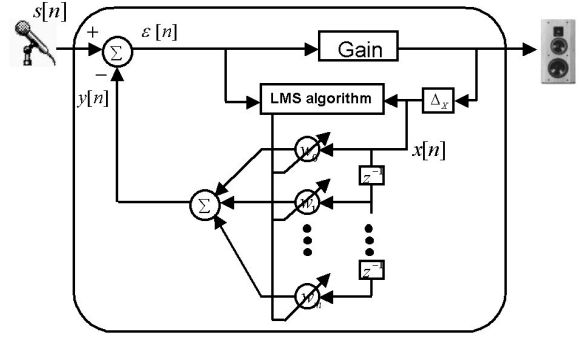


Figure 4. Block diagram of the prototype adaptive filter for hearing aid applications.

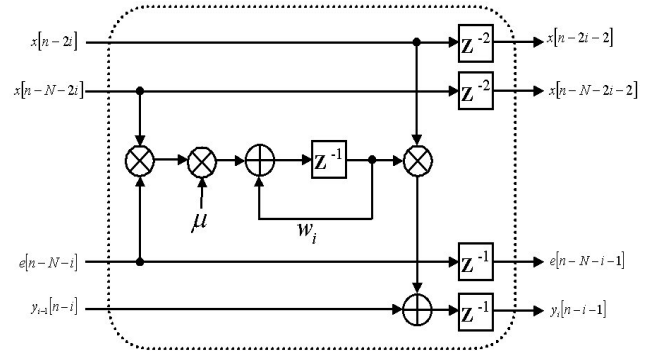


Figure 5. Signal flow graph of the functional unit module for the non-folded pipeline DLMS filter.

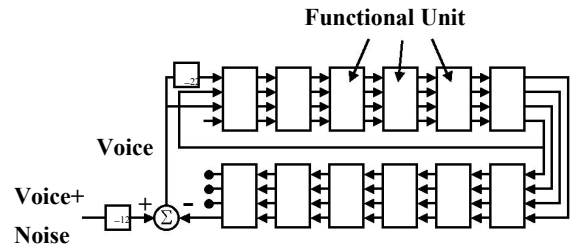


Figure 6. A non-folded architecture of a DLMS adaptive filter. This was implemented by cascading the functional unit module shown in Fig. 5.

5. SIMULATION RESULTS

The simulated filter outputs of the folded LMS architecture and non-folded DLMS architecture are shown in Fig. 7. The input data was a 1.0 kHz sinusoidal signal and the sampling frequency was 22 kHz. The LMS filter shows a faster convergence time and smaller convergence error compared to the DLMS filter. This can also be noticed in Fig. 8, where the minimum mean square error of the DLMS filter converges to a higher value, with a longer convergence time than the LMS filter.

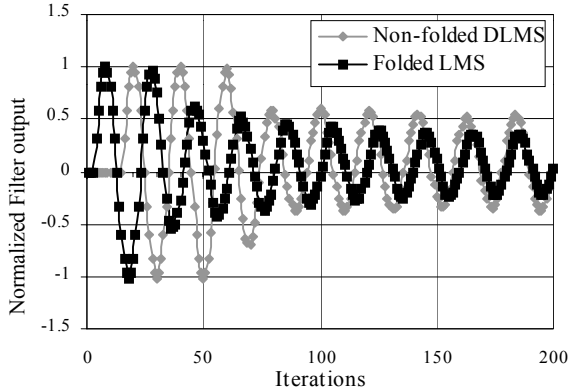


Figure 7. Convergence characteristics of the folded and non-folded adaptive filter architectures.

Table 2 shows the system attributes of the 3 different

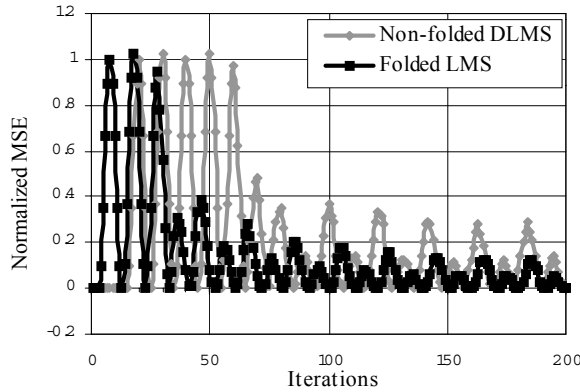


Figure 8. Mean square error of the folded and non-folded adaptive filter architectures.

implementations at a same throughput of 22 kHz. Since it takes 1 clock cycle to process 1 input data, the clock frequency of the non-folded architectures is same as the throughput. Whereas the folded architecture requires a clock frequency of 748 kHz since 34 clock cycles are needed for one operation. Since the folded architecture had to run at a higher clock rate than the non-folded architecture, its supply voltage had to be higher than the non-folded architecture, and hence the energy consumption per operation was almost 8 times larger. To calculate this energy efficiency, we derived energy consumption per operation, by multiplying the power, the critical path delay and the clock cycles per operation. Using a non-folded architecture with Sub-CMOS logic style, we were able to reduce the supply voltage down to 450 mV and achieve an energy efficiency of 2.47 nJ/operation. With the Sub-Pseudo NMOS logic, we could further lower the supply voltage down to 400 mV and achieve a 28% improvement in energy consumption per operation compared to Sub-CMOS logic. The number of transistors increases approximately 3.6 times if we use a non-folded instead of a folded architecture but we could get significant improvement in energy efficiency.

Power-delay product of pseudo NMOS logic and standard CMOS logic are shown in Fig. 9 for different supply voltage. In this prototype design, the power-delay product of the Sub-Pseudo NMOS logic was 46 % less than the Sub-CMOS logic for a supply voltage of 400 mV. Consequently from the power and delay relationships, less energy is required for the same delay by using Sub-Pseudo NMOS instead of Sub-CMOS.

Since the sub-threshold transistor current depends exponentially on the threshold voltage and temperature, in order to guarantee a robust operation, a negative feedback principle can be applied to suppress the variations due to process and temperature changes. In our previous research, an 8-by-8 carry save array multiplier was fabricated which had some interesting features such as a leakage current monitor and a substrate bias circuit [10]. These circuit blocks were used to compensate for the variations in the circuit and provide robust operation. The test chip showed out to have stable operation down to a supply voltage of only 300mV when the threshold voltage was 450mV. We should further consider these sophisticated circuit techniques in our DLMS filter implementation for the sake of robust operation.

6. CONCLUSIONS

As modern hearing aid devices are miniaturized, acoustic signals must be processed with a much smaller power budget due to the reduced battery size. At the same time, sophisticated signal processing algorithms such as sub-band filtering or adaptive noise

Table 2. Simulation results for the 3 different adaptive filter implementation.

	Clock frequency	V_{dd}	Energy per Operation	# of Transistors
Folded Standard CMOS (LMS)	748 kHz	650 mV	19.1 nJ	31121
Non-folded Sub-CMOS(DLMS)	22 kHz	450 mV	2.47 nJ	110916
Non-folded Sub-Pseudo NMOS (DLMS)	22 kHz	400 mV	1.77 nJ	85764

cancellation are required for better quality of sound. In this paper we introduced different architectures and different logic styles of adaptive filter designs for hearing aid applications. To design an ultra-low power system by operating in the sub-threshold region, a non-folded architecture with multiple functional units was proposed. We were able to maintain the same throughput with less power dissipation by reducing the clock rate and the supply voltage. Comparisons between the non-folded and folded architectures with same technology show, that we can save power up to 87% by trading off area for power. Though the number of transistors increased around 3.6 times, this unrolling strategy could enable sub-threshold operation of the chip for ultra-low power dissipation.

We also explored the suitability of Sub-Pseudo NMOS. Due to the exponential relationship between transistor current and gate voltage, Sub-Pseudo NMOS proved to be comparable to CMOS in robustness, noise margin and power consumption. At the same time, pseudo NMOS logic inherits all the advantages it had in the normal strong inversion region, such as better performance, less area and reduced routing. As a result, Sub-Pseudo NMOS showed 28% higher energy efficiency than Sub-CMOS for our prototype implementation.

Consequently, by applying both the architecture and circuit level optimization techniques, the non-folded DLMS filter using Sub-Pseudo NMOS logic was capable of operating in the sub-threshold region consuming ultra-low power with a desired performance.

7. ACKNOWLEDGEMENTS

This research has been funded in part by SRC under contract 98-HJ-638 and by NSF under contract CCR-9901152.

8. REFERENCES

- [1] J. M. Kates, "Feedback Cancellation in Hearing Aids: Results from a Computer Simulation", *IEEE Trans. on Signal Processing*, vol. 39, no. 3, March 1991
- [2] J. A. Maxwell and P. M. Zurek, "Reducing Acoustic Feedback in Hearing Aids", *IEEE Trans. on Speech and Audio Processing*, col. 3, no. 4, July 1993
- [3] H. Soeleman, K. Roy and B. Paul, "Robust Ultra-Low Power Sub-threshold DTMOS Logic", *International Symposium on Low Power Electronics and Design*, pp. 25-30, 2000
- [4] H. Soeleman and K. Roy, "Ultra-Low Power Digital Sub-threshold Logic Circuits", *International Symposium on Low Power Electronics and Design*, pp. 94-96, 1999
- [5] J. Rabaey, *Digital Integrated Circuits*, Prentice Hall, 1996
- [6] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 3rd Edition, 1996
- [7] B. Widrow and S. D. Starns, *Adaptive Signal Processing*, Prentice Hall, 1st, 1998
- [8] A. Harada, K. Nishikawa and H. Kiya, "Pipeline Architecture of the LMS Adaptive Digital Filter with the Minimum Output Latency", *IEICE Trans. Fundamentals*, vol. E81-A, no. 8, August 1998
- [9] M. D. Meyer and D. P. Agrawal, "A High Sampling Rate Delayed LMS Filter Architecture", *IEEE Trans. on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 40, no. 11, Nov. 1993
- [10] H. Soeleman, *Ultra-Low Power Digital Sub-threshold Logic Design*, Ph.D thesis, Dept. of ECE, Purdue University, W. Lafayette, IN, 2000