Ultra Low Energy vs Throughput Design Exploration of 65 nm Sub-V_T CMOS Digital Filters

S.M. Yasser Sherazi, Joachim N. Rodrigues, Omer. C. Akgun, Henrik Sjöland, and Peter Nilsson

Department of Electrical and Information Technology, Lund University

Box 118, SE-221 00 Lund, Sweden

Email: {yasser.sherazi, joachim.rodrigues, omercan.akgun, henrik.sjoland, peter.nilsson}@eit.lth.se

Abstract—**This paper presents an analysis on energy dissipation of digital half band filter operated in the the sub-threshold (sub-V_T) region. The filter is implemented in a single 12-bit and various unfolded structures. The designs are synthesized in a 65 nm low-leakage high-threshold CMOS technology. The simulation results from application of an energy model shows that the unfolded by 2 implementation of the filter is the most energy efficient structure dissipating 22 % less energy compared to single implementation at energy minimum voltage. However, unfolded by 4 implementation, is the best suited for throughputs requirements of 100 Ksamp/sec to 1 Msamp/s, as it gives the least energy dissipation compared to any other implementation.**

I. INTRODUCTION

Presently miniaturized devices are getting more important in medicine, sensor networks, and many other applications. Engineers aim to develop ultra compact and low energy dissipating circuits that may be used in devices like hearing aids, medical implants, and remote sensors. We are working towards small wireless devices having ultra low energy dissipation targeting on-body or implanted devices. In these devices the minimal energy dissipation in active and standby mode, is of highest importance as it makes the battery last longer, which is important as it is non-trivial to change or charge a battery in an medical implant. Devices like hearing aids that communicate between the two ears to improve binaural hearing may benefit from this kind of wireless receiver. Another example is a neural sensor inside the body that communicates with a robotic arm or leg. If a radio is made sufficiently small and with minimal power consumption there will be vast possibilities for new applications.

In the conducted project the design constraints are, less than 1 mW and 1 µW power consumption in active and standby mode, respectively, capacity to handle data rates up to 250 kbits/s, realized on a single chip solution with an area cost of 1 mm² in 65 nm CMOS. A block digram shows the receiver system in Fig. 1, containing a RF front-end (2.5 GHz), an analog-to-digital converter, a digital baseband for demodulation and control, and finally, an analog decoder that processes the received data packets.

The main focus of this paper is on the digital baseband part of the receiver system. The first task of the digital baseband circuit is to re-sample data from 4 Msamp/s to 250 Ksamp/s. Therefore, a chain of decimation filters needs to be applied. To get lesser energy dissipation, we are employing voltage scaling techniques rigorously, hence making the designed
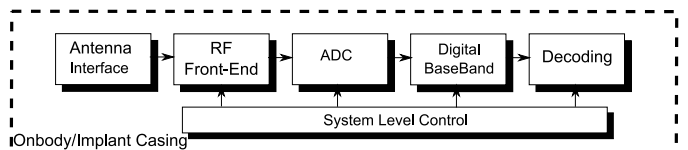

Fig. 1. Receiver system.

circuits run in the sub-threshold (sub-V_T) domain [1]. When operating in the sub-V_T domain, leakage currents are to be dealt with, which are the source of energy dissipation in an idle CMOS [2]. This current puts an important design constraint especially in implantable medical devices. Consequently, we need to optimize the circuits in terms of energy dissipation and throughput for sub-V_T operation.

In Sec. II we briefly present the applied sub-V_T energy model. In Sec. III we present a 12-bit architecture of a Half Band Digital (HBD) filter that is implemented as direct mapped and its various unfolded structures. In Sec. IV the results attained from the HBD filters are shown and discussed, and finally, the conclusions are presented in Sec. V.

II. SUB-V_T ENERGY MODEL

For a device it is shown in [3] that the current does not drop to *0* when gate to source voltage V_{GS} is equal to the threshold voltage V_T, $V_{GS} = V_T$, which is an indication for leakage currents, commonly referred to as the sub-V_T or weak inversion conduction [4]. The existent current is due to leakage and low in amperage, and in the sub-V_T domain used as the operating switching current. The drawback of sub-V_T circuits is speed penalty. However, circuits that operate at sub-V_T manage to satisfy the ultra low power requirements, since magnitudes less power is consumed compared to super-threshold circuits [4]. The total energy dissipation of static CMOS digital circuits is specified as

$$E_{total} = \underbrace{\alpha C_{tot} V_{DD}^2}_{E_{dyn}} + \underbrace{I_{leak} V_{DD} T_{clk}}_{E_{leak}} + \underbrace{I_{peak} t_{sc} V_{DD}}_{E_{sc}}, \quad (1)$$

where E_{dyn} is the average switching and E_{leak} is leakage energy dissipated during a clock cycle T_{clk}. As it is known that the energy dissipation due to short circuit energy (E_{sc}) in the sub-V_T domain is minor compared to the overall energy dissipation, which therefore is neglected [5]. In (1), E_{dyn} during one clock period is specified by the switching activity
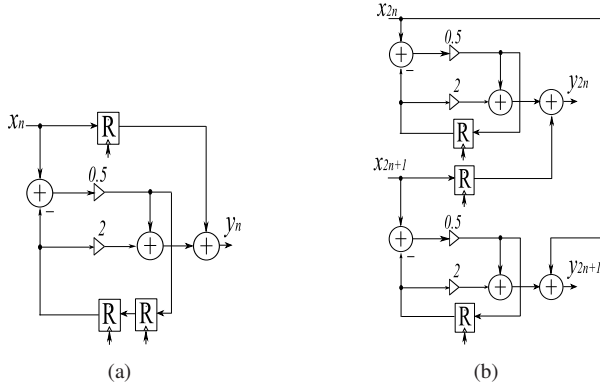
Fig. 2. Half Band Digital Filter. (a) single HBD filter (b) uf-2 HBD filter.

factor (α), and the maximum possible switched capacitance of the circuit (C_{tot}).

The model used to calculate energy dissipation was presented in [6], delivers SPICE-accurate results. This model acquires parameters such as k_{leak}, k_{cap}, k_{crit}, and μ_e, that are derived from high level simulations. The average leakage scaling factor of the circuit normalized to the average leakage current of a single inverter is k_{leak}. The scaling factor k_{cap} is normalized total capacitance of the circuit in terms of a single inverter capacitance. The k_{crit} is a coefficient that defines the critical path delay of the circuit in terms of a single inverter delay. Lastly, μ_e is the average switching activity of circuit per N samples operations, for detail reading refer to [6].

III. FILTER ARCHITECTURES

Minimum energy dissipation with medium to high throughput requirement puts stringent constraints on a design. Therefore, it is important to explore and analyse the architectures that best fulfill the requirements. This section presents the architectural differences in the single and unfolded versions of HBD filter after presenting the HBD Filter.

A. Half Band Digital Filter

An optimized third order filter structure is evaluated for minimum energy dissipation. The filter structure for the parallel implementation, see Fig. 2(a), is an optimized parallel third-order bi-reciprocal lattice wave digital filter, [7], considered as highly suitable as decimator or interpolator, for sample rate conversions with a factor of two. The benefit of using this type of filter is that all filtering may be performed at lower sample rates, with low arithmetic complexity, therefore, yielding low energy dissipation and a low chip area cost [8]. The transfer function of the proposed filter is given as

$$H_z = \frac{1 + 2z^{-1} + 2z^{-2} + z^{-3}}{2 + z^{-2}}, \tag{2}$$

showing that the filter has coefficients that do not need dedicated multipliers, simple shifting shifting will suffice, therefore saving the over all area of the circuit. Some initial analysis indicated that the required throughput will not be achieved by a single sample implementation of this filter. Therefore, unfolding was applied. Unfolding is a transformation technique that calculate j samples in a clock cycle where j is the
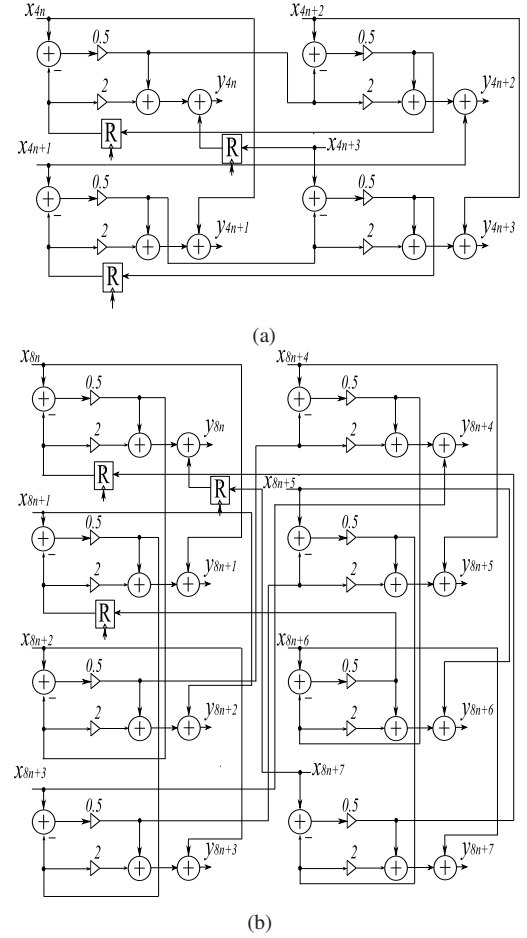

(a)

(b)

Fig. 3. Unfolded Architectures of the HBD filter. (a) uf-4 HBD filter (b) uf-8 HBD filter.

unfolding factor, unfolding has a property of preserving the number of delays in a Direct Form Graph (DFG) [9]. The basic HBD filter architecture is unfolded to get three more structures, i.e., unfolded by 2 (uf-2), unfolded by 4 (uf-4) and, unfolded by 8 (uf-8). Fig. 2(b), shows the *uf-2* version of the filter. In this architecture, the number of registers remain unchanged whereas the adders are doubled. Fortunately the critical path of this circuit is equal to the original HBD filter structure. Fig. 3(a) shows an architecture that was unfolded by a factor of 4. The number of adders has increased according to the unfolding factor. The critical path has increased, since two of the feed back paths do not contain a register. Similarly, Fig. 3(b), shows the architecture of *uf-8* HBD, the adders have increased by a factor of 8 when compared to the original HBD structure. The critical path increases, since six of the feed back paths do not contain any register. However, there are more samples processed, per clock cycle in the unfolded structures which wins over a limited increase in the critical path [10].

B. Hardware Mapping

All the cells used for implementation are from a low-leakage high-threshold (LL-HVT) standard cell library. For the parallel architecture, tight synthesis constraints were set to get minimum area and a short critical path. The parameters for the

TABLE I
EXTRACTED PARAMETER FOR CIRCUITS

Circuit	k_{leak}	k_{cap}	k_{crit}	μ_e	Area	t_p[nsec]
par	1113.6	835.4	127.4	0.727	1124	2.84
uf-2	1695.5	1375.7	127.4	0.708	1836	2.84
uf-4	3172.5	2797.9	164.2	0.703	3275	3.66
uf-8	5924.5	5422.3	232.2	0.890	6170	5.22

TABLE II
CIRCUIT CHARACTERIZATION AT EMV

Circuit	EMV [mV]	f [kHz]	E/Cyc [fJ]	E/smp [fJ]
par	241	14.6	45	45
uf-2	238	13.8	71	35
uf-4	247	13.0	150	38
uf-8	251	9.94	380	48

energy model were retrieved by a gate-level simulations with back annotated toggle and timing information, which includes occurring glitches. The parameters obtained are applied to the energy model to generate and characterize the designs in the sub-V_T domain.

IV. RESULT

In this section the architectures of the filter are evaluated with respect to energy and throughput. The parameters required for the energy model [6], extracted during synthesis and energy simulations, are presented in Table I. The values for k_{leak} follow the area cost, indicating reduced leakage with respect to area. The k parameters for the unfolded implementation does not scale with the unfolding factor j since the number of internal registers remain unchanged from the single sample implantation, although there is an increase in the number of input and output registers. Energy dissipation is calculated with the assumption that the designs operate at critical path speed, this gives an Energy Minimum Voltage (EMV) point [11]. The designs energy characteristics, over a scaled supply voltage V_{DD} per clock cycle operation in the sub-V_T domain, are presented in Fig. 4(a). It is shown that the single sample implementation denoted by *(par)* dissipates the minimum amount of energy per clock cycle when compared with the other three implementation. The reason being that the leakage for this circuit is lesser than the other circuits due to less area. The energy minima (per clock cycle) of 45.5 fJ for *par* implementation is achieved around 241 mV (indicated by the dot), which is lower than EMV of any other architecture, which confirms that lesser area contributes to less energy per clock cycle. However, it is crucial to know the energy spent on the processing of each sample of data, and the apparent benefit of using *par* structure is lost when the energy per operation or energy per sample is considered. Fig. 4(b), shows the energy dissipation per sample for different structures. Reason being that unfolded circuits perform twice, four and eight times as much operations per clock cycle, therefore the over all energy per sample for these circuits is reduced when compared to a single sample implementation. Fig. 4(b), shows that the most efficient architecture is *uf-2* as it dissipates 35.8 fJ per sample which is 45 % less than the

energy dissipated by the *par* structure. Here, we may observe that the *uf-8* architecture is loosing to *par* even in energy dissipation per sample at lower voltages and is almost equal at near threshold voltages. The reason for this behaviour is that the area of *uf-8* architecture is much larger, therefore, has much higher leakage. The maximum frequency attainable with respect to V_{DD} is shown in Fig 4(c), the maximum frequency for both *par* and *uf-2*, is always higher than their counterpart due to a shorter critical path, and the *uf-8* has the slowest maximum clock because of higher critical path, see Table I. Fig 4(d), shows the energy dissipation of all the structures with respect to throughput. Table II, presents the characteristics of all the presented architectures at EMV, showing the maximum frequencies attainable, energy dissipated per clock cycle, as well as per sample. These simulations, show that we benefit from unfolding technique as we process more samples in the same amount of time and after processing the data we stay in idol or shut-down mode to save energy.

In the project discussed in section I, we need a chain of four HBD filters, that helps reduce the high frequency data with the rate of 4 Msamp/s from the ADC to the actual massage with data rate of 250 Ksamp/s. The first HBD filter has a requirement of processing the input data stream with the rate of 2 Msamp/s. This throughput requirement is only fulfilled by using *uf-8* structure of HBD near 390 mV as shown in Table III and Fig 4(d). The throughput requirement of data with the rate of 1 Msamp/s for the second HBD is fulfilled by using three of the unfolded structure, however, *uf-4* gives the minimum energy dissipation. The throughput requirement of data with the rate of 500 Ksamp/s for third HBD is fulfilled by all four structures as shown in Table III and Fig 4(d), however, *uf-4* gives the minimum energy dissipation. The throughput requirement of data with the rate of 250 Ksamp/s for last HBD is again fulfilled by all structures, however, *uf-4* gives the minimum energy dissipation as shown in table III and Fig 4(d). In Fig 4(b), the *uf-2* structure appears to be the least energy dissipating circuit. However, when stringent throughput requirements are in-placed the *uf-4* structure proves to be the best option as shown in Fig 4(d). This analysis shows that its crucial to identify the best suitable architectures for the given throughput and energy requirements. In [12] it is argued that Low leakage Low Threshold cell are more beneficial in term of higher throughput rates in sub-V_T domain, which needs to be investigated for these filter implementation.

In [5] it was shown in that the supply voltage of sub-V_T circuits may be reduced down to 50 mV. However, in practical terms at such voltage values functional failures occur due to the process variations. It was found in [13] that the supply voltage value which realizes operation with less than 0.001 failure rate for a 65 nm LL-HVT process is 250 mV and this value is taken as the minimum reliable operating voltage (ROV), indicated in the Fig 4(b) by a line at 250 mV. The simulations show that the EMV points of these circuits are already close to ROV points as shown in 4(b), and for the required throughput we are operating above ROV, Table III.
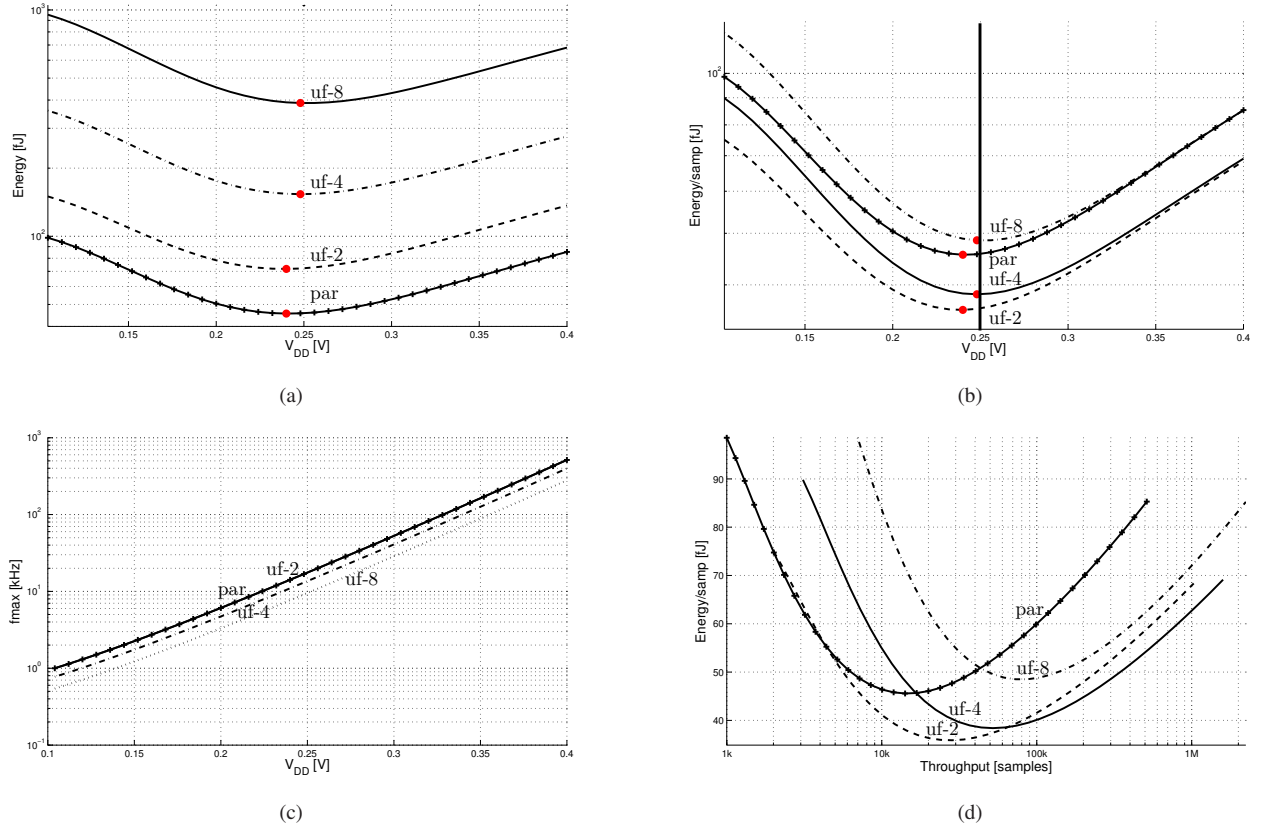
Fig. 4. Simulation Plots of HBD filter architectures, (a) Energy vs V_{DD} per clock cycle, (b) Energy vs V_{DD} per sample. (c) Frequency vs V_{DD}, (d) Energy vs Throughput

TABLE III
PERFORMANCES OF CIRCUITRY AT REQUIRED THROUGHPUTS

Throughput	Circuits	Op V [mV]	E/Cyc [fJ]	E/smp [fJ]
2 Msamp/s	uf-8	390	656	**82.2**
1 Msamp/s	uf-8	368	586	73.3
	uf-4	376	246	**61.5**
	uf-2	400	136	68.3
500 Ksamp/s	uf-8	344	525	65.2
	uf-4	352	226	**54.7**
	uf-2	368	116	58.4
	par	400	85.2	85.2
250 Ksamp/s	uf-8	300	434	55.0
	uf-4	320	188	**47.0**
	uf-2	344	126	51.8
	par	368	72.9	72.9

V. CONCLUSION

This paper presents four HBD filter architectures that are evaluated for minimum energy dissipation in the sub-V_T domain for a throughput constrained system. All architectures i.e., the unfolded by 2,4,8 and the basic single HBD filter, are implemented and simulated using 65 nm LL-HVT standard cells. It is shown that it is beneficial to use unfolded implementation to achieve low energy dissipation per sample at EMV, when compared to the energy dissipated by a basic single sample implementation of the same design.

REFERENCES

[1] J.-J. Kim and K. Roy, "Double gate-mosfet subthreshold circuit for ultra low power applications," *Electron Devices, IEEE Transactions*, 2004.
[2] P. van der Meer, *Low-Power Deep Sub-Micron CMOS Logic*. Kluwer Academic Publishers, 2006.
[3] J. M. Rabaey and *et all*, *Digital Integrated Cicuits*. Prentice Hall, 2003.
[4] H. Soeleman and *et all*, "Robust subthreshold logic for ultra-low power operation," *VLSI Systems, IEEE Transactions on*, 2001.
[5] E. Vittoz, *Low-Power Electronics Design*, ch. 16.
[6] O. C. Akgun and Y. Leblebici, "Energy efficiency comparison of asynchronous and synchronous circuits operating in the sub-threshold regime," *Low Power Electronics, Journal of*, vol. 4, OCT 2008.
[7] P. Nilsson and M. Torkelson, "Method to save silicon area by increasing the filter order," in *Electronic letters*. ACM, NY, USA, 1995.
[8] H. Ohlsson and *et al.*, "Arithmetic transformations for increased maximal sample rate of bit-parallel bireciprocal lattice wave digital filters," in *The IEEE International Symposium on Circuits and Systems*.
[9] K. K. Parhi, *VLSI Digital Signal Processing Systems*, ch. 5.
[10] A. Pontus and *et all*, "Power reduction in custom cmos digital filter structures," *AICSP Journal*.
[11] J. Rodrigues and *et all*, "A <1 nj sub-vt cardiac event detector in 65 nm ll-hvt cmos," *VLSI-SOC*, 2010.
[12] D. Markovic, J.M.Rabaey, and *et al.*, "Ultralow-power design in near-threshold region," *Proceedings of the IEEE*, 2010.
[13] J. Rodrigues and *et all*, "Digital implementation of a wavelet-based event detector for cardiac pacemakers," *Circuits and Systems I: Regular Papers, IEEE Transactions on*.