LOW COST APPROXIMATE ADDER SUBTRACTOR FOR FIR FILTER

Dr. W.A.Mahmoud Dhafer R. Zaghar Dr. Mohannad K. Sabir
Department of Electrical Engineering / College of Engineering / University of Baghdad /Iraq

ABSTRACT
The cost of the FIR filter depends on the cost of its components that are classified into three types namely buffer, multiplier and adder/subtractor.
The adder/subtractor cover the most FIR filter components and any reduction in its cost means high reduction in the total cost of FIR filter. In addition, FIR filters with pipeline multiplier the speed of FIR depends on the speed of adder/subtractor and any improvement in speed will result to increasing in the total speed of such filter.
This paper shows a proposed method to increase the speed and reduce the cost of the adder/subtractor by neglecting the rotated carry output. This method reduces the accuracy of FIR filter, in other words it adds small noise to the output of the filter, but the level of this noise is inversely proportional with the number of bits in data bus of FIR filter. Therefore this method will be used with midum noise output filters or with wide data bus (12-bits or more) FIR filter.

الخلاصة
إن الكلفة في بناء مرشح الترددات ذو الاستجابه النبضية المحدودة (FIR) تعتمد على كلف مكوناته الثلاثة (وحدات الخزن، الضارب و الجامع/الطارح). يعد الجامع/الطارح من اكثر المكونات استخداما لذا فان أي تقليل لكلفته سيقلل كثيرا من الكلفة الكلية لمرشح الترددات ذو الاستجابه النبضية المحدودة (FIR). إضافة إلى ذلك فان السرعة القصوى لهذا المرشح تعتمد على السرعة القصوى للجامع/الطارح.
إن هذا البحث يقترح طريقه لزيادة سرعة الجامع/الطارح وتقليل كلفته وذلك عن طريق إهمال الزيادة الراجعة (rotated carry). إن هذا الإهمال سوف يزيد من الضوضاء الإضافية في الإشارة الخارجة، وهذه الضوضاء الإضافية تتناسب عكسيا مع عدد خطوط الدخل، لذا فان هذه الطريقة تستخدم مع المرشحات ذات الضوضاء المتوسطة أو تلك التي تملك عدد خطوط دخل عالية (تزيد على 12 خط).

KEY WORDS
FIR, adder/subtractor, DSP, synchronous, asynchronous, Short Processing Time, 1'complement, 2'complement, pipeline.

INTRODUCTION
Digital FIR filters that process data in speeds less than 100 MSPS can be implemented by using three methods.

First by the use of a spatial programmable devices such as adders and multipliers or one chip programmable FIR. This approach is easy and fast in design but it has three disadvantages, which are:

1- High cost.

2- It is restructured by the characteristics of the available of the standard ICs and components.

3- Need to interfacing with other components because the different sources of ICs mean different speeds and technologies (TTL, CMOS, ECL ….).

The second approach is to build FIR using processors such μP, μC or TMS with suitable software. This technique is suitable for real time applications in speeds less than 20 MSPS by using the modern technologies of TMS, and is used generally for simulations [Balasubramanian, 1997].

The third approach uses general DSP components such as general programmable devices like FPGA. This approach is complex in the design and needs to build all components in a form suitable with architecture of base (general programmable) devices. It is commercial and very flexible in comparison with the first approach, also it is suitable for real time applications for speeds less than 200 MSPS which is an advantage over the second approach.

Third approach is more suitable for large digital system for speeds less than 70 MSPS, gives low cost and universal chips system which is another advantage [Pratt, 1991] [Mintzer, 1987] [Frerking, 1994].

SHORT PROCESSING TIME

The digital systems are be classified depending on the timing as synchronous and asynchronous systems.

Asynchronous systems execute their jobs or transfer function between input and output without any internal storage and take a long time to reach a stable case. These systems are used in low speed and small controls circuits [Brown2000] [Melham 1993].

Synchronous systems built as a multi-layers process/storage systems as shown in **Fig.(1)**. The storage layers have a common clock control on the time storage and input/output the data in process layers are consumed a short constant time compared with the processing layer time.

The maximum speed of the clock system depends on the minimum speed in the processing layers in the system, because the time of the clock must be larger than the maximum processing time layer plus the storage time.

Thus the full adder circuit as a processing layer in large synchronous system must have a short delay to give a high-speed adder hence a high-speed digital system [Brown 2000].

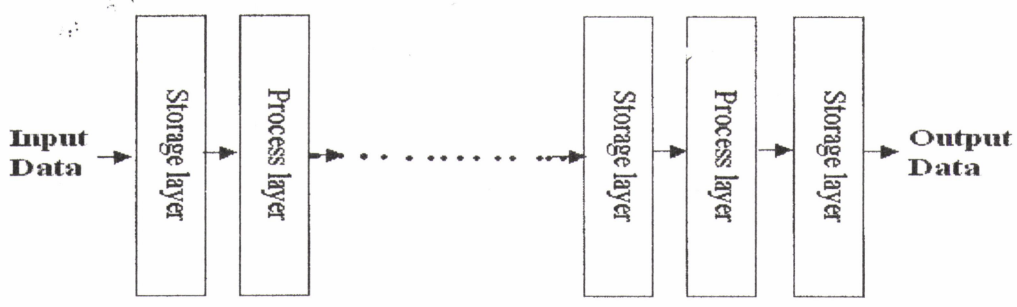

Fig.(1) Synchronous Systems Architecture

FIR FILTER COMPONENTS

Fig.(2) shows an example of FIR filter (16-tap) built using three types of components: buffers, multipliers, and adder/subtractors. However, an n-tap FIR filter needs (n-1) buffers, (n/2) multipliers and (n-1) adder/subtractors.

Generally, the buffers are low cost and high speed with fixed design (D flip-flop), when the multipliers and adder/subtractors have complex digital circuits that have different design approaches. These approaches give a different cost and speed circuits that result in different costs and speeds of the FIR filters.

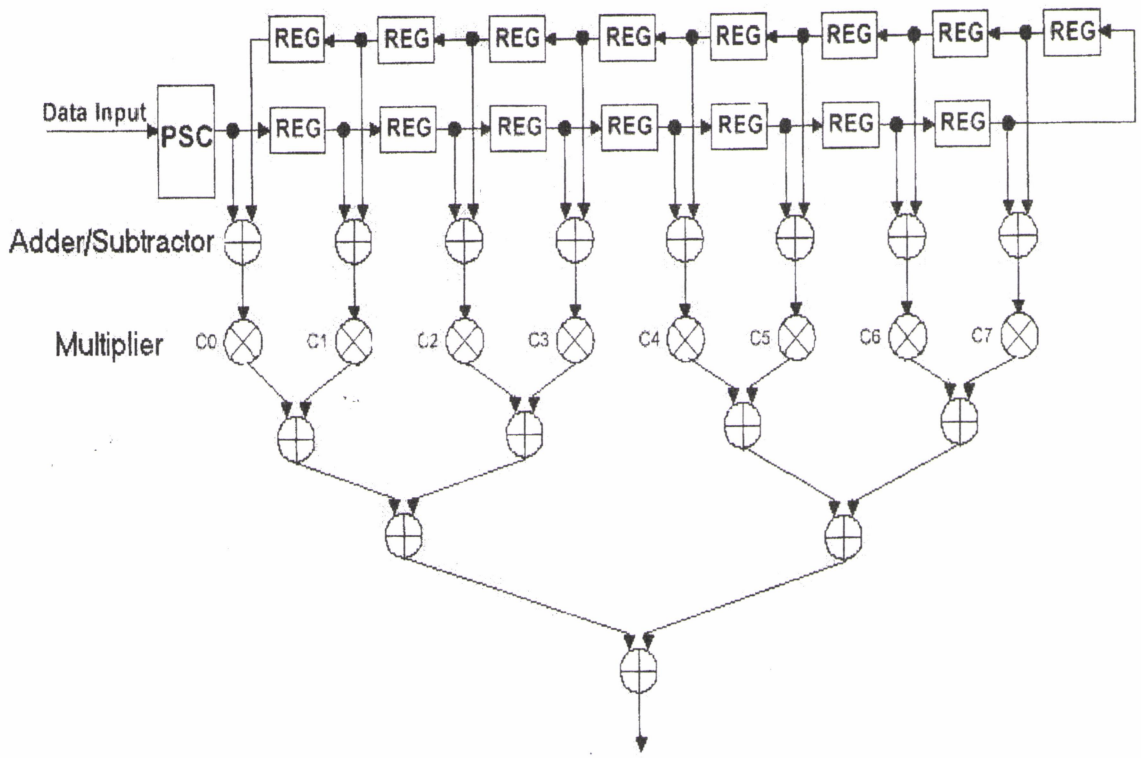

Fig.(2) 16-taps FIR filter.

The adder/subtractor is the most used component in FIR filters design and any reduction in its cost represents a high reduction in the total cost of FIR filter. The multiplier is a large and complex circuit and it is the slowest component in FIR filter.

There are a number of methods to design high-speed multiplier such as parallel shift and add or pipeline shaft and add ... etc. The pipeline parallels shift and add multiplier has a high speed but need a multi-bit latency. [Balasubramanian 1997] [Pratt 1991]

The FIR filters use a pipeline parallels shift and add multiplier (PPSAM-FIR) has a high speed multiplication therefore the speed of FIR depend on speed of adder/subtractor. In PPSAM-FIR any improve to its speed will be increasing the total speed of the filter.

The adder/subtractor use the pipeline approach can be increase the speed of PPSAM-FIR, but this approach increase the total bit latency and in result increase the total delay of FIR filter to a very large value.

ADDER/SUBTRACTOR COMPONENTS

There are two approaches to built adder/subtractor firstly by use the 2'complement as shown in **Fig.(3)**, secondly by use the 1'complement as shown in **Fig.(4)**. [Brown 2000][Wakerly 2001] [Pirsch 1998]

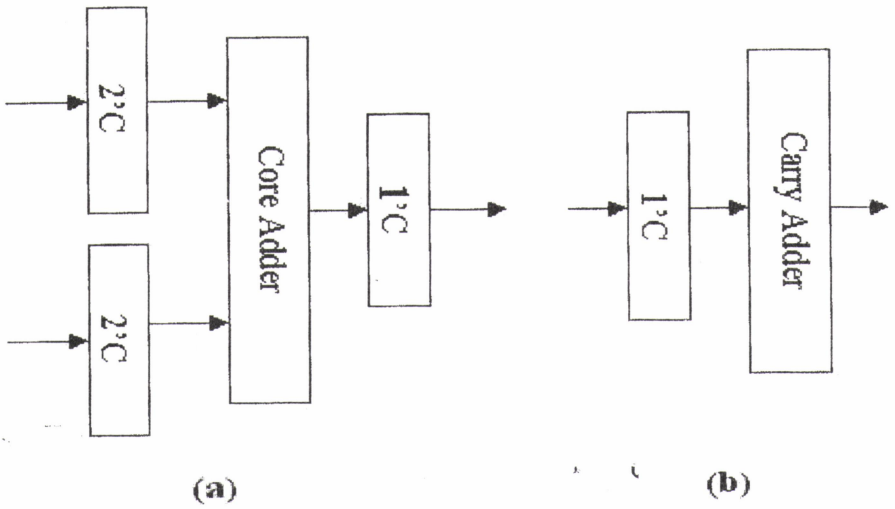

Fig.(3) Block diagram of 2'C adder/subtractor
a) Total circuit b) 2'C circuit.

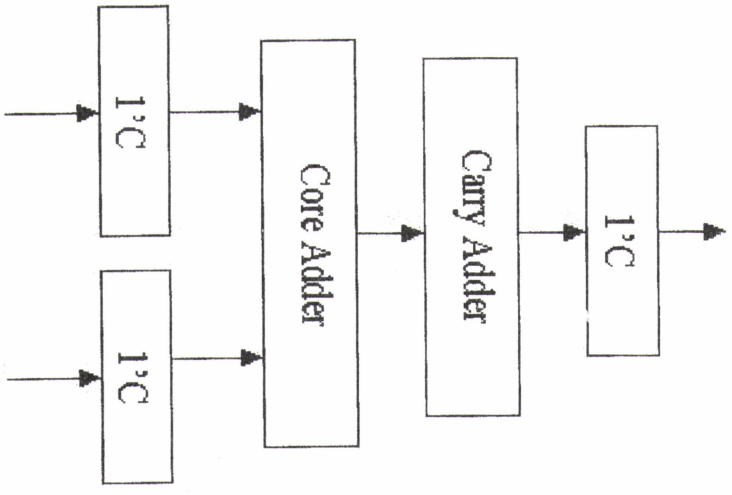

Fig.(4) Block diagram of 1'C adder/subtractor.

Table (1) show the 1'complement method has same speed but it is lower cost from 2'compement because it need to one carry adder when 2'complement need to two carry adder outputs.
The processing of the rotated carry in 1'C adder/subtractor in **Fig.(4)** represent about 30% from the total delay and about 30% from the total rational cost of 1'complemnt adder/subtractor circuit.

Table (1) The rational delay and cost of 1'C and 2'C Add/Sub

Component	1'C Circuit	Carry Adder	2'C Circuit	Core Adder	1'C Add/Sub	2'C Add/Sub
Rational Delay	1	5	6	10	17	17
Rational Cost	1	10	11	20	33	44

MATHEMATICS OF ADD/SUBTRACT [Brown 2000] [Wakerly 2001] [Pirsch 1998]
The add/subtract process has four cases as shown in **Fig.(5)**, that represents an example to these cases for 2'C in (a) and 1'C in (b). The approximate adder/ subtractor cover all mathematics process in **Fig.(5-b)** except it will neglect the rotated carry in case of opposite sign as shown in **Fig.(6-b)**.

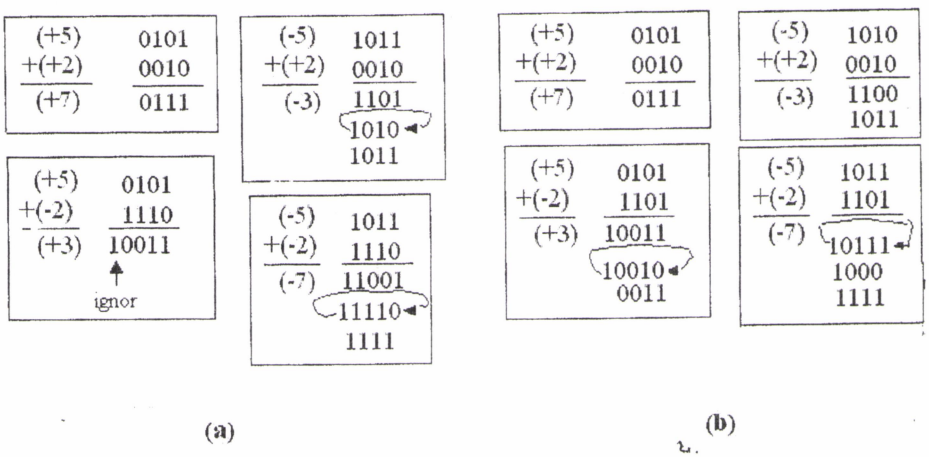

Fig.(5) An example to add/subtract for 2'C in (a) and 1'C in (b).

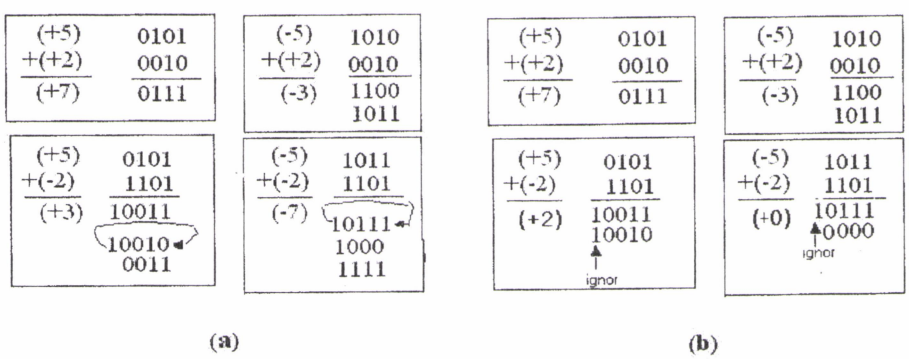

Fig.(6) An example to add/subtract for
a) Correct values by 1'C b) the approximate adder/ subtractor.

PROPOSED METHOD OF ADD/SUBTRACT
The proposed method depend on neglecting the rotated carry of 1'C as in **Fig.(6-b)** that show the example for the four cases of 1'C approach with neglecting the rotated carry that will called approximation adder/subtractor.
The approximation add/subtract approach in **Fig.(6-b)** give two errors, first is the approximation in case 3 that give the value 2 but the real value is 3 (case 3 in **Fig.(6-a)**), this error represent an additional noise to the output signal of FIR filter. The second error shows in case 4 give the value 0 when the correct value is –7, this error is very large and found in the case of two negative inputs.
The proposed method will be modified of approximation add/subtract approach to convert case 4 (two negative signs) to become an ordinary addition with negative sign in output, the algorithm of proposed approach has the following steps:
a) Check the sign of the two inputs.
b) If the two signs are symmetrical the process is add with output sign like sign of first input.
c) If the two signs are different the process is normal subtraction.

Fig.(7) Show how the proposed adder/subtractor correct the result of case 4 without any change on the results for the three other cases.

The proposed approximate adder/subtractor generate an error equal to one in decimal or LSB in binary with probability of error equal to 0.25 statistical average.

The effects of proposed method on the speed and cost shows in **Fig.(8)** and **Fig.(9)** show the cost and delay of 1'C, 2'C and proposed methods with number of input bits when all the circuits built by same type of logic gates.

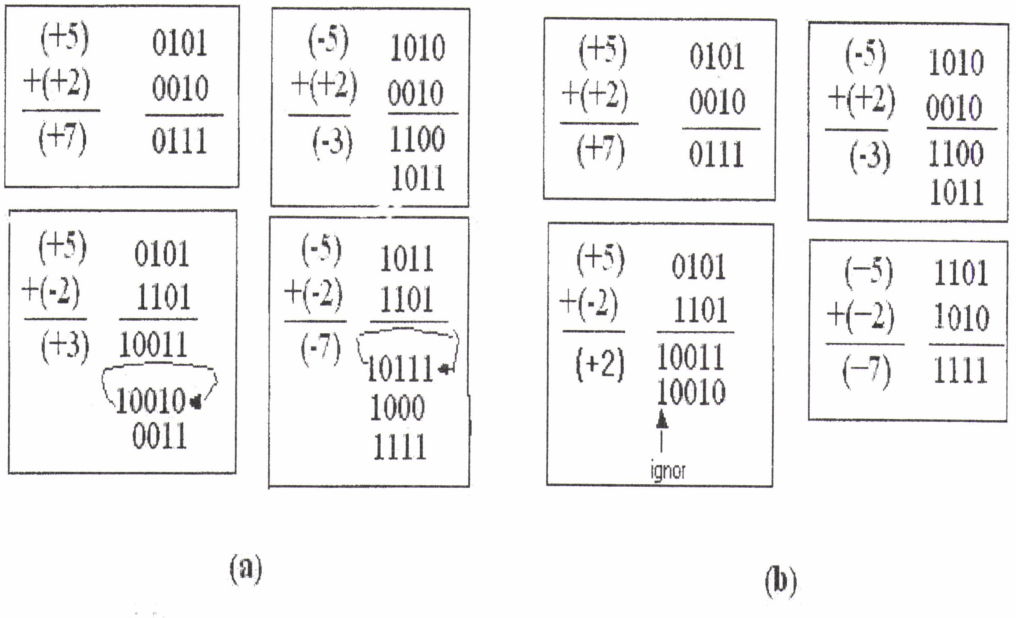

Fig.(7) The result of adder/subtractor
a) Correct values b) Values of proposed adder/subtractor.

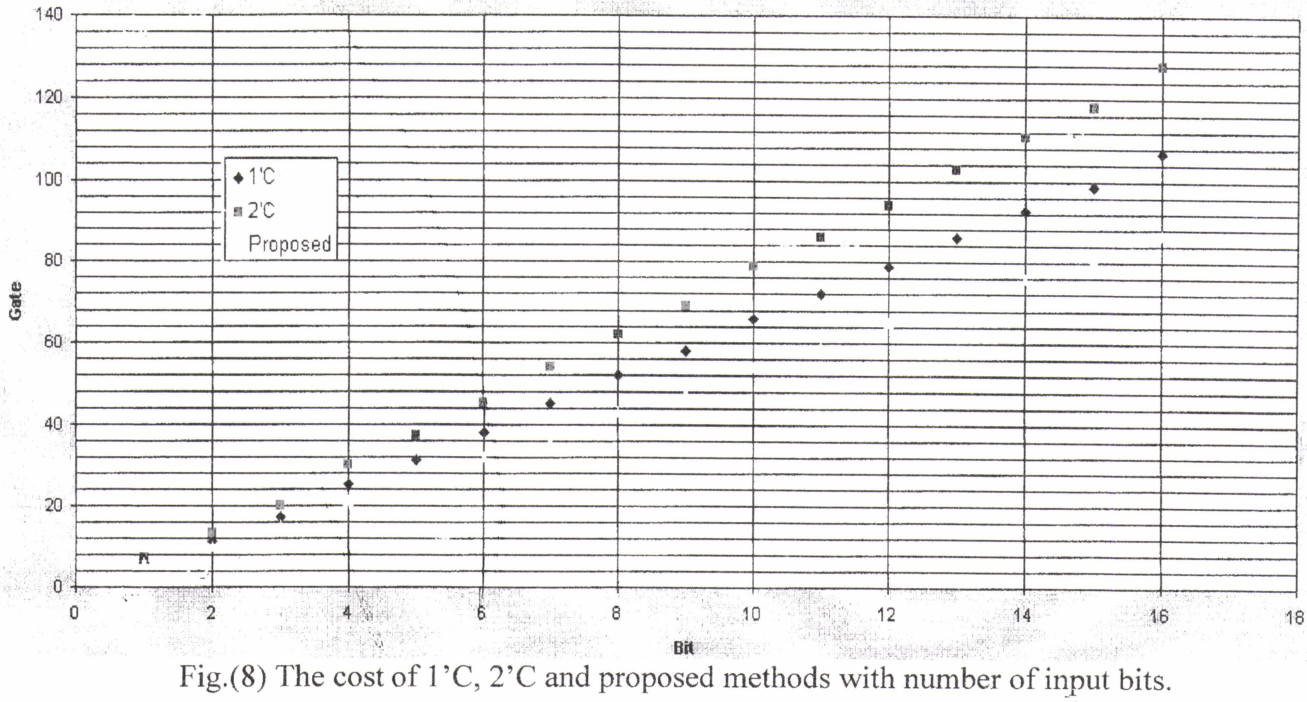

Fig.(8) The cost of 1'C, 2'C and proposed methods with number of input bits.

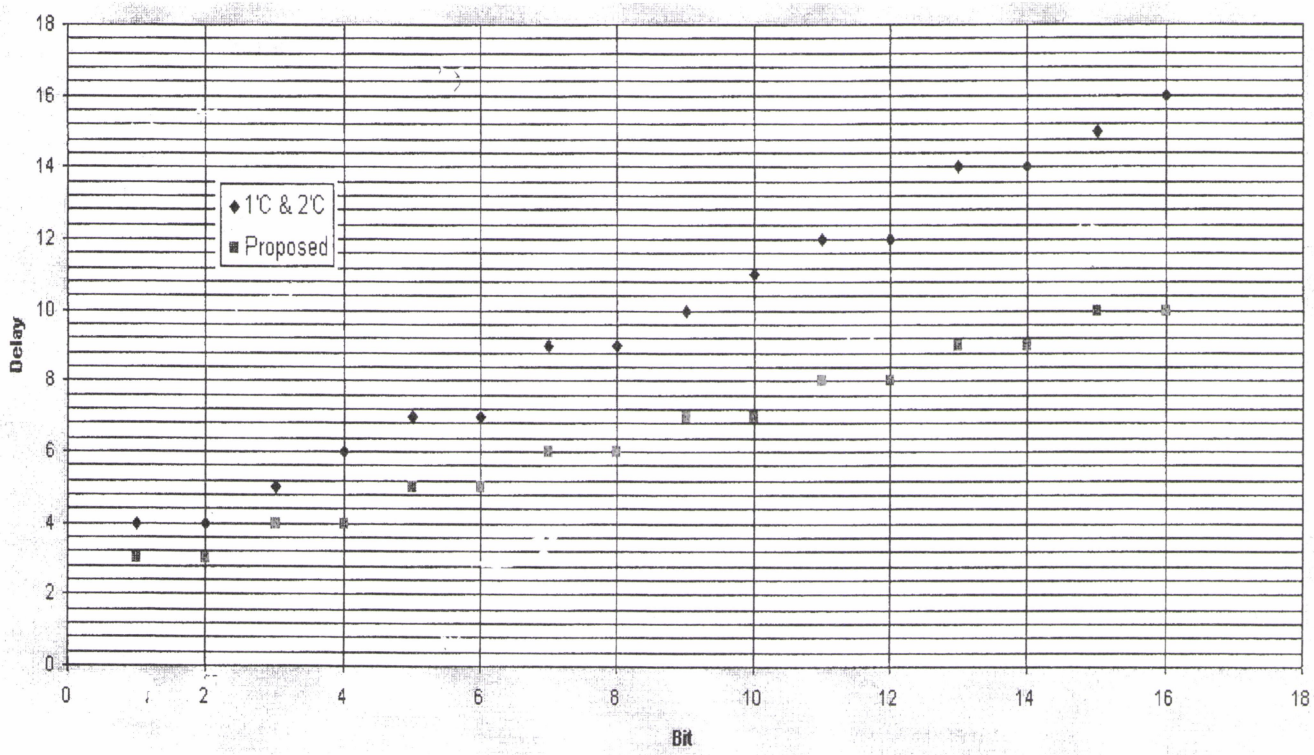

Fig.(9) The delay of 1'C & 2'C and proposed methods with number of input bits.

ADDITIONAL NOISE

The additional noise in the proposed approximate adder/subtractor can be calculate from the following equation:

$$Na = -20\log_{10}^{2^n} = -6n \quad \text{dB} \tag{1}$$

When Na is the additional noise and n is the number of bits in data bus of filter.
However, the average for the probability of error equal to 0.25 and eq.(1) simplified to the following equation:

$$Na = -6n - 12 \quad \text{dB} \tag{2}$$

The additional noise in FIR filter use proposed approximate adder/subtractor would be calculated from eq. (3) and eq.(4). [Yarmolik, V. 1988] [Dixon, R. 1984]

$$Nf = -6n - 12 + \log_{10}^{m} \tag{3}$$

$$\lceil \text{Number of taps} \rceil = 2^m \tag{4}$$

When Nf is total additional noise in FIR filter that used the proposed approximate adder/subtractor
For example a 128 taps FIR filter using the proposed approximate adder/subtractor has a data bus equal to 8-bit and 16-bit has an additional noise equal to –42 dB and –90 dB respectively.

CONCLUSIONS

The proposed method of approximation adder/subtractor reduces the cost of adder/subtractor to about 65% from the original cost as shown in **Fig.(8)**. So this method will be increase its speed to

about 150% or in other words reduces the delay as shown in **Fig. (9)**, i.e. the gain vector of approximation adder/subtractor method is about 2 for 16-bit adder/subtractor.
The gain vector increase with increasing the number of bits in data bus, when the noise decreases with increasing the number of bits in data bus as in equations 1-4.
However, this method is active for 16-bit or more data bus FIR filter, that will reduces the total cost of PPSAM-FIR filter to 70-85% and increase the speed to 150-200% from the original speed with very low additional noise in output. But the approximation case is very bad with ALU and other logic circuits, therefore the approximation add/subtract method is spatial for FIR filter and it cannot be used in other applications.

REFERENCES

Balasubramanian, (1997), N. B. "Optimal design of digital filters using CSD coefficients",UNIVERSITY OF CALIFORNIA.

Pratt, W. K. (1991),"Digital Signal Processing, Second Edition," Wiley-Interscience Publication,

Mintzer, (1987), L. "Mechanization of Digital Signal Processors," Handbook of Digital Signal Processing, pp. 941-973.

Frerking, (1994), M. E. "Digital Signal Processing in Communications Systems" Van Nostrand Reinhold.

Brown, S. and Vranesic.Z. , (2000), " Fundamental of Digital Logic with VHDL Design", McGraw-Hill.

Melham, (1993), T. F. " Higher Order Logic and Hardware Verification ", Cambridge Tracts in Theoretical Computer Science. Cambridge University Press.

Wakerly, (2001), J. F. "Digital Design Principles and Practices", Prentice Hall.

Pirsch, (1998), P. ,"Architectures for Digital Signal Processing", University of Hannover, Germany, John Wiley & Sons.

Yarmolik, (1988), V. N. and Demidenko, S. N. ,"Generation and Application of Pseudorandom Sequences For Random Testing", John Wiley and Sons.

Dixon, (1984), R.C., "Spread Spectrum Systems", 2nd Edition, Wiley, New York.

LIST OF SYMBOLS

FIR: finite impulse response.
DSP: digital signal processing.
MSPS: magi sample per second.
PPSAM-FIR: pipeline parallels shift and add multiplier.
1'C: ones complement.
2'C: twos complement.