A Review of Filter Bank Techniques - RF and Digital

John Lillington
Director, Libra Design Associates Limited
Quarr Hill House, Quarr Hill, Ryde, Isle of Wight, PO33 4EH, UK
Tel: +44 (0)1983 611494
e-mail: john.lillington@ldal.eu

ABSTRACT

Filter banks and tuneable filters form key elements of system and receiver design. Many will be familiar with the various types of analogue filter but, with the advent of ever higher ADC conversion rates, digital filters are displacing analogue in an increasing number of applications. Although digital filters can offer performance parameters not available to analogue designs, there is often a price to pay in terms of power and cost and, in some cases, this approach may not even be the best option.

This paper aims to look at some of the trade-offs between the two types as well as typical examples of techniques used, particularly the less familiar types of digital filter bank.

1.   INTRODUCTION

Filters are obviously a very key element of any design and getting a proper definition which can actually be realised in practice is often far from simple. This is especially true when, as is commonly the case with modern systems, we have the choice between analogue and digital filters. For example, the decision on where to place the boundary between the two is a complex one, based on available sample rates, number of bits, cost and power consumption of the ADC or DAC.

There is a tendency to think (especially if you are a digital designer) that placing the ADC or DAC at the antenna port would be an ideal solution. Although there is no doubt that advances in technology are pushing things in this direction, there will always be other considerations. Placing an ADC at the antenna Rx port without any analogue filtering will invite all of the "out-of-band" signals and noise to be folded into the ADC's output, never to be removed. Conversely, placing an unfiltered DAC output at the antenna Tx port will fill the world with unwanted spurs which will <u>not</u> be welcomed by other users of the spectrum.

This is obviously a huge subject in which no one paper can hope to address more than a fraction of the picture. This paper, therefore, is limited to the author's personal experience of such matters (as in all authorship, it is best to write about what you know!). Apart from a few examples of analogue filter bank design, the main thrust of the paper is to delve into the world of digital filter and filter bank design.

2.   AN EXAMPLE RECEIVER SYSTEM DESIGN

The choice is somewhat arbitrary but it is based on a design actually carried out a few years ago and serves to illustrate some of the trade-offs between analogue and digital filter-bank design. Figure 1 below is a simplified schematic of a 100 kHz to 3 GHz scanning receiver.
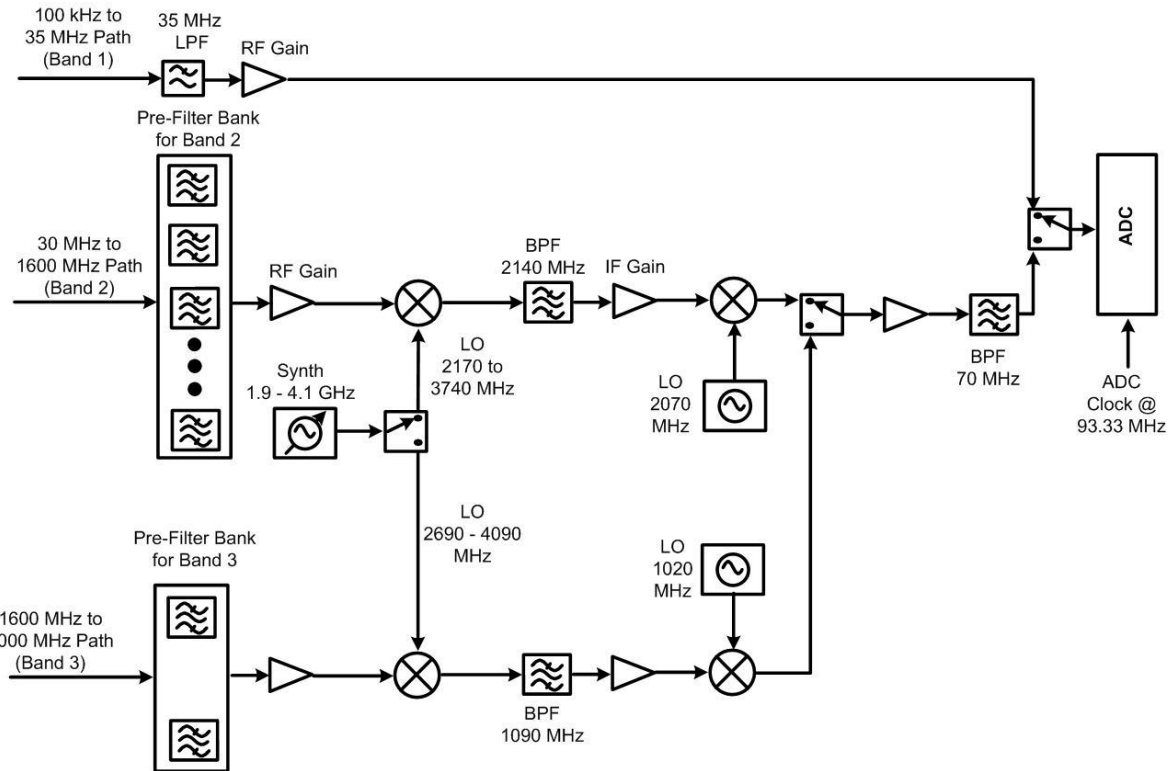
**Figure 1. Simplified Architecture of Existing 3 GHz Scanning Rx**

The frequency plan is chosen, amongst other factors, to ensure that neither the RF or LO breakthrough from the 1st Mixer(s) will fall in the IF band. The design of the analogue filter banks needs to take into account a number of criteria:-

1. All RF Image frequencies must be sufficiently attenuated (typically -60 to -80 dBc).

2. The individual filters should overlap sufficiently to accommodate largest bandwidth signals to pass at <u>any</u> centre frequency (smallest overlap at lowest frequencies).

3. Ideally, upper passband frequency < 2*lower passband frequency to avoid in-band harmonics.

The measured results for the filter banks used in this project are summarised in Figure 2 below and the actual filter bank together with the path switching is shown in Figure 3.
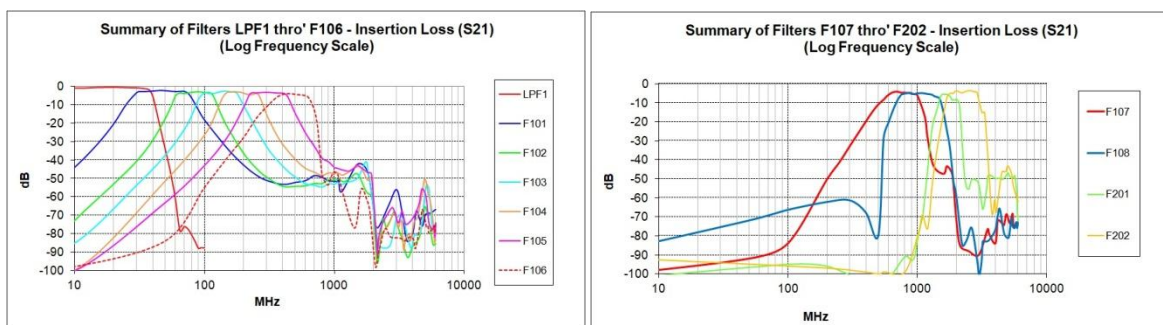

**Figure 2. Measured Filter Bank Responses, DC - 3 GHz**
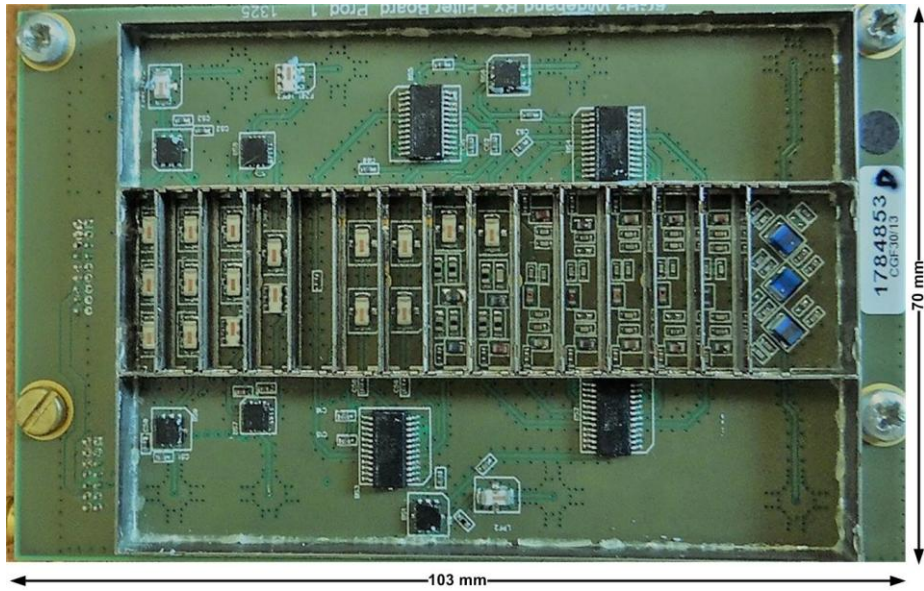
Figure 3. View of Actual Filter Bank, Screening Lid Removed

The lowest band, 100 kHz to 35 MHz, is designed to use the **1st Nyquist** zone whereas, all other bands use the **2nd Nyquist.** The reason for latter is simply that, in the 1st Nyquist zone, harmonics of lower frequencies will fall <u>in-band</u> and therefore in the ADC output as illustrated in Figure 4 below.

(a)   **(b)**
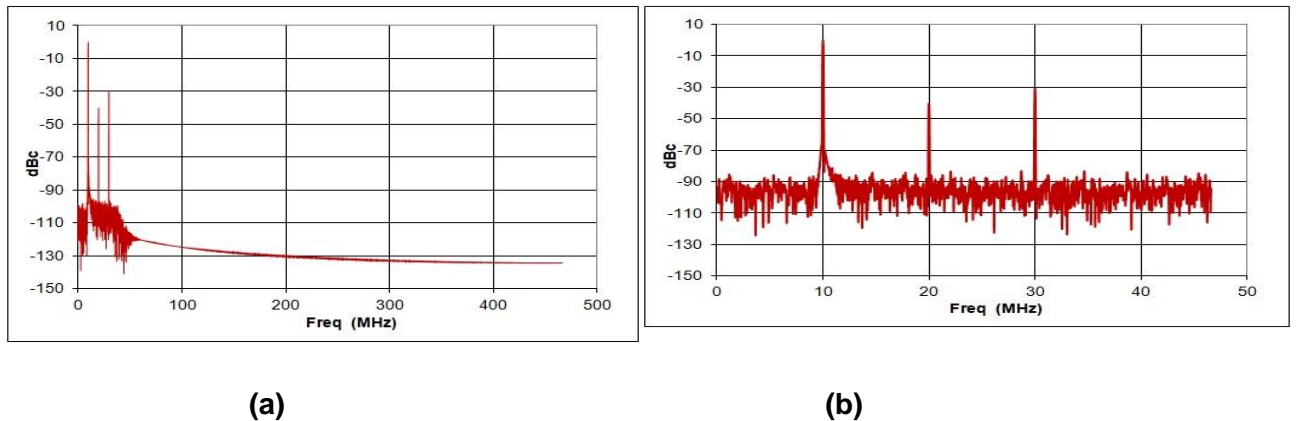
Figure 4. Signal with In-Band 2nd and 3rd Harmonics, 1st Nyquist, (a) Before and (b) After ADC

For signals in the 2nd Nyquist zone, this problem does not occur as illustrated in Figure 5 below. The harmonics lie outside the anti-alias filter and are thus attenuated before reaching the ADC. Note that the signal is <u>inverted</u> in frequency at the ADC output when using 2nd Nyquist.

3
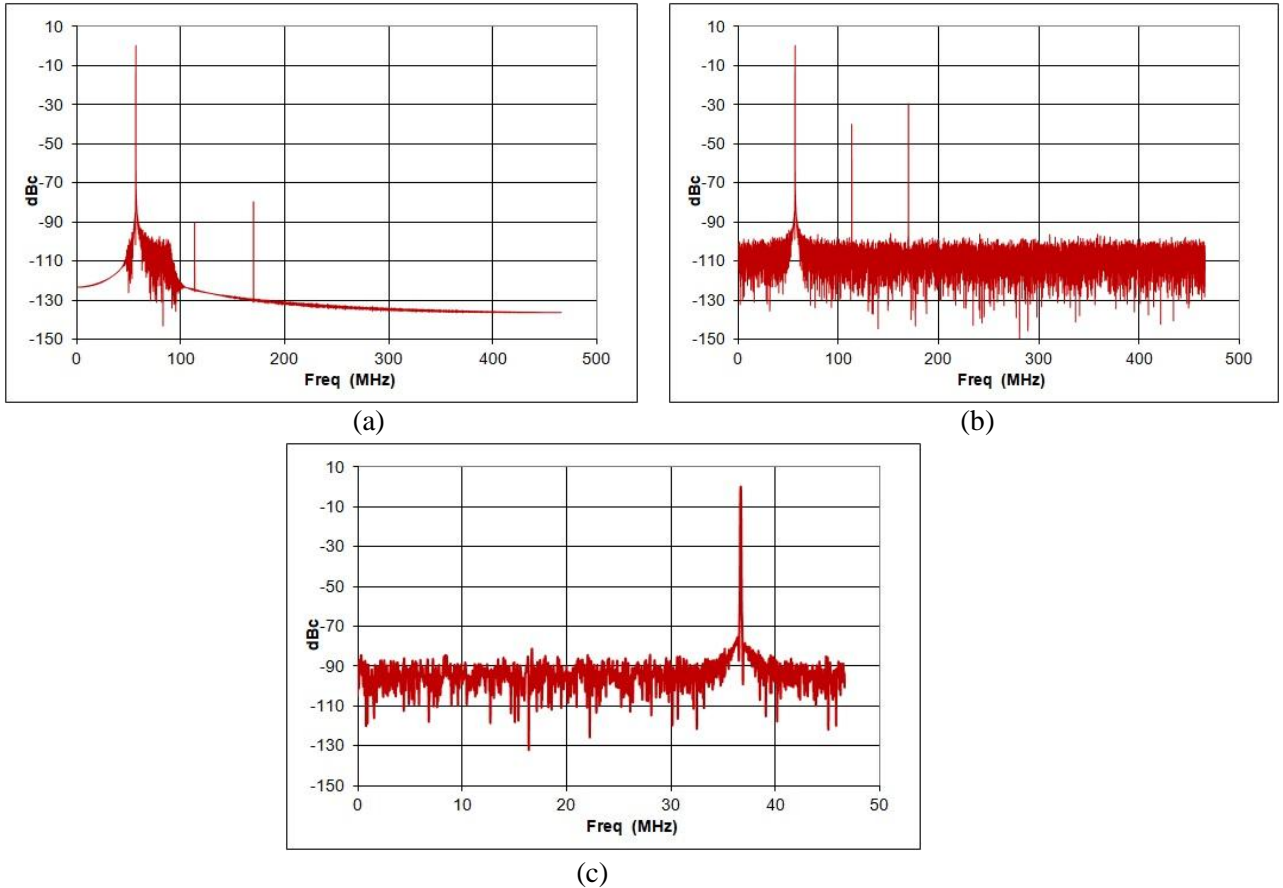
Figure 5. Signal with 2nd & 3rd Harmonics in 2nd Nyquist
(a) Before Filter (b) After Filter (c) After ADC

3. *AN EQUIVALENT DIGITAL RX DESIGN?*

With the advances being made in ADC / DAC design, we may soon be seeing sample rates in excess of 6 Gsps with maybe up to 12 bit resolution although it is hard to see the Effective Number of Bits (ENOB) exceeding about 7 - 8 bits. On this basis, it is worth seeing what a Digital Receiver covering the band 100 kHz to 3 GHz might look like. We will assume a sample rate of 6.66 Gsps and an EONB of 7 bits.

Figure 6 below shows a possible configuration where the whole of the 100 kHz to 3 GHz band is Lowpass filtered, amplified and sampled. The ouput of the ADC is then processed using a Digital Filter Bank equivalent to the analogue version shown in Figure 2 & Figure 3 above.

4
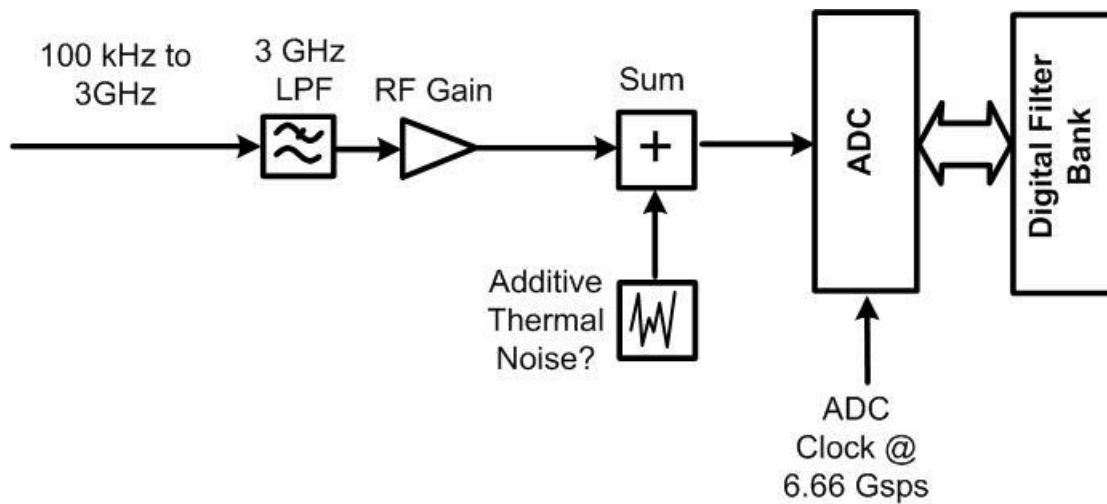
Figure 6. A Possible Digital Rx Equivalent to Figure 1

We can immediately see an issue in that, by using the 1st Nyquist zone and a wide lowpass filter, the same problem of harmonics of the lower frequency signals (as discussed above) will appear in band and cannot be subsequently filtered out by the Digital Filter Bank. This is clearly illustrated in Figure 7 below.

(a)   (b)
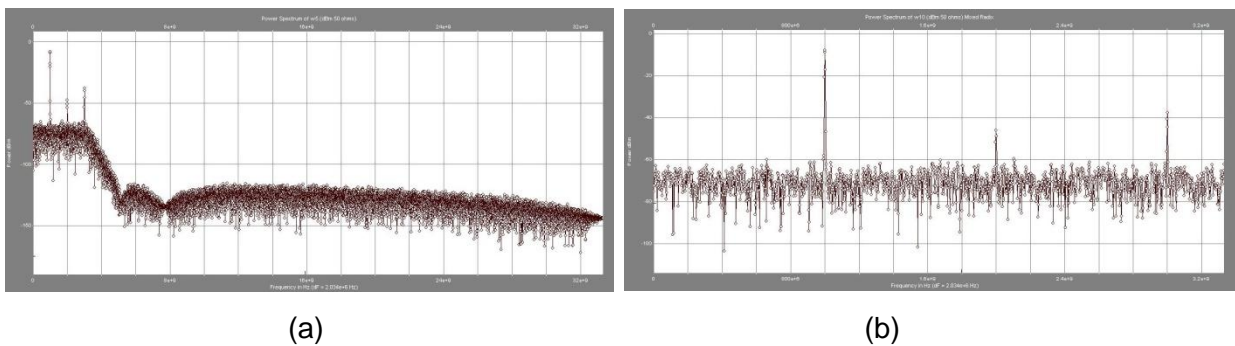
Figure 7. Digital Rx Signal with 2nd & 3rd Harmonics After LPF (a) and at ADC Output (b)

There is also the need to ensure that there is enough noise in the signal at the ADC input to "toggle" the LSB's, otherwise there be the risk of quantisation spurs appearing at a mix of signal and ADC clock frequencies. This problem is illustrated in Figure 8 below.
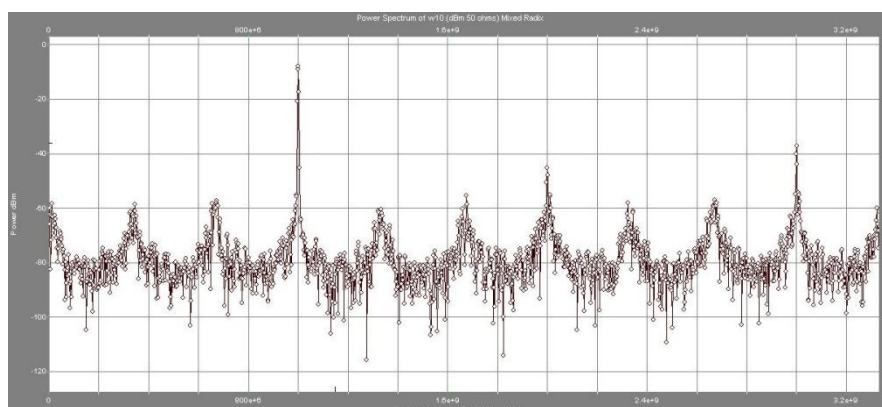
Figure 8. Effect of Insufficient Noise giving Quantisation Spurs

In many cases the Rx noise floor will be sufficient but with low noise figure and a lower number of bits, it may be necessary to add thermal noise, preferably in a frequency band

outside of the main Nyquist zones. This is known as "Dither Noise" and ensures the LSB's of the ADC are randomly toggling rather than causing narrow spurs. Figure 9 below shows an example of added dither noise (below 100 kHz) which removes the quantisation spurs without affecting the thermal noise floor in the wanted signal region.
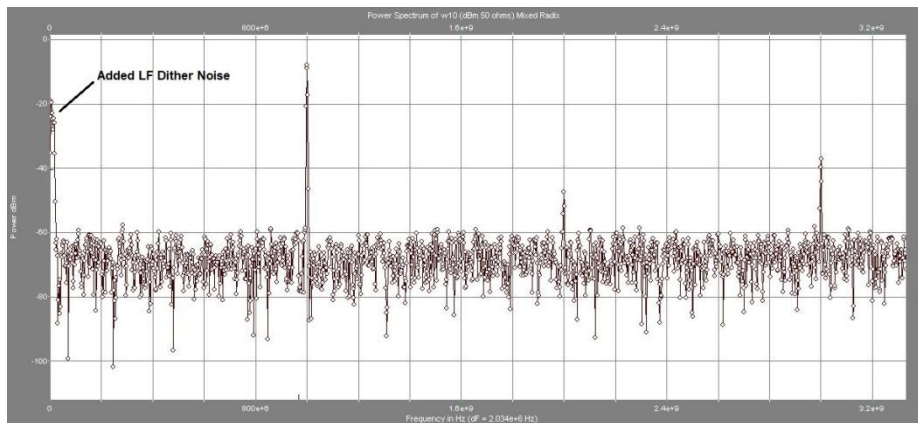
Figure 9. Using Low Frequency "Dither" Noise to Remove Quantisation Spurs

An alternative approach might be to use the 2nd or higher Nyquist zones to avoid harmonic problems. To do this with a 100 kHz to 3 GHz span would require a two stage up / downconverter to avoid RF and LO breakthrough and Mixer Intermod problems. This would rather defeat the object of using the ADC close to the Rx input and would be as complex in terms of analogue hardware as the analogue receiver described above.

The conclusion we reach from this is that it is not necessarily beneficial to use a high-speed ADC close to the RF input. It may well reduce circuit complexity and component count but the performance, in terms of harmonics and spurs may well be inferior to a standard superhet design with analogue filter banks and lower speed, higher dynamic range ADC.

Nevertheless, digital filtering is increasingly important so we will continue the process of comparison by opening up the "Pandora's Box" of Digital Filter and Digital Filter Bank design.

4.   *DIGITAL FILTER and FILTER BANK DESIGN*

    4.1 Main Filter Types
Again, this is a huge subject so that, in the space of this short paper, only those areas within the personal experience of the author will be covered. We probably need to start with a description of the basic building blocks of digital filters..

Very broadly, digital filters can be defined as Finite Impulse Response (FIR) or Infinite Impulse Response (IIR). As the name suggests, FIR filters have a response to an impulse which only lasts for a finite time and will depend on the number of Taps in the design. In general, the sharper the required frequency domain response, the greater the number of taps and the longer the time-domain response.

IIR filters, on the other hand, rely on feed back to minimise the number of taps so that, in theory, the time domain response can "ring" on, gradually decaying, forever, just like most analogue filters. In practice, the response eventually decays below the output bit width (just as in analogue filters, it decays below thermal noise). The advantage of IIR filters is that they need significantly fewer digital processing resources for a given frequency domain response but do have issues with both stability and bit-width so are rather harder to design.

There are many types of FIR filter (Lowpass, Halfband Lowpass, Bandpass etc.) and within this discussion we will limit ourselves to the Lowpass and Halfband Lowpass variety. A simplified diagram of a FIR filter is shown in Figure 10 below. The digitised signal is clocked through a delay line and multiplied by a coefficient at each tap. These are summed to provide the filtered output.
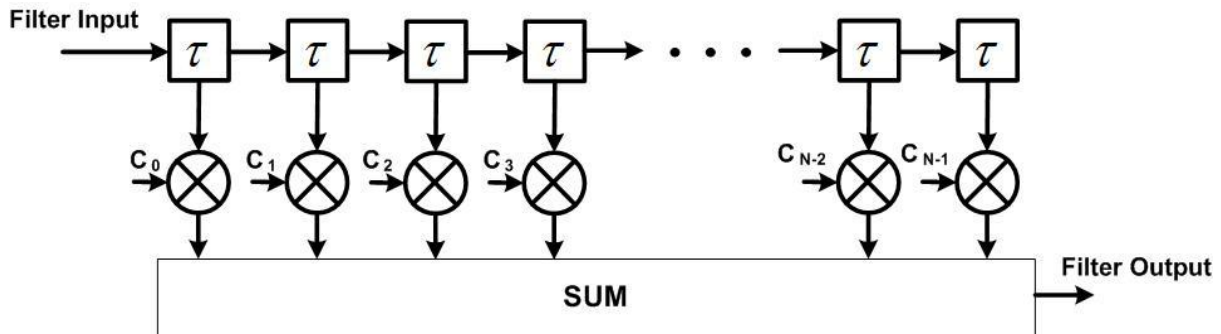
Figure 10. Basic FIR Filter Structure

A typical Lowpass FIR is shown in Figure 11 below, showing the normalised gain (Passband Loss 0 dB) and normalised frequency response (wrt Input Sample rate). The effects of bit-width restrictions on the Data and Coefficients are not shown but, provided these are adequate, the response and stop-band levels are completely predictable and repeatable.
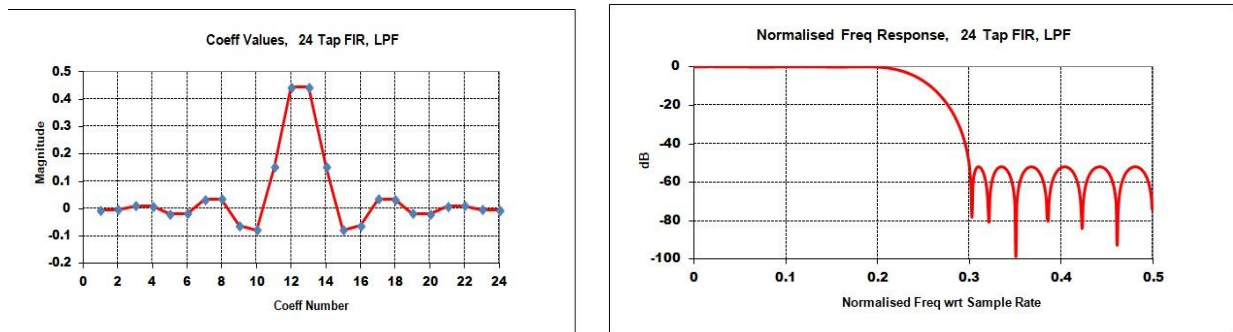

Figure 11 Typical Lowpass FIR Filter Coefficients and Response

4.2 More Complex Filter Types

Moving on to a slightly more complex filters, we will examine briefly the Half-Band Filter (HBF) and its useful companion, the Cascaded Integrator-Comb (CIC). Both of these are highly efficient forms of digital filter. The HBF is often core to a structure which takes a real stream of data from the ADC and turns it into complex (I & Q) data with lowpass filtering and decimation (reduction of sample rate). The CIC is frequently found near the front end of digital receivers as an efficient structure for signal bandwidth and sample rate reduction.

Looking first at the HBF, a typical 67 Tap design is shown below in Figure 12 below. At first sight, this looks very little different from the above FIR filter but there is a subtle, but important, difference in that, apart from the centre three taps, every alternate tap is zero. This leads to an elegant and efficient structure which will now be described.
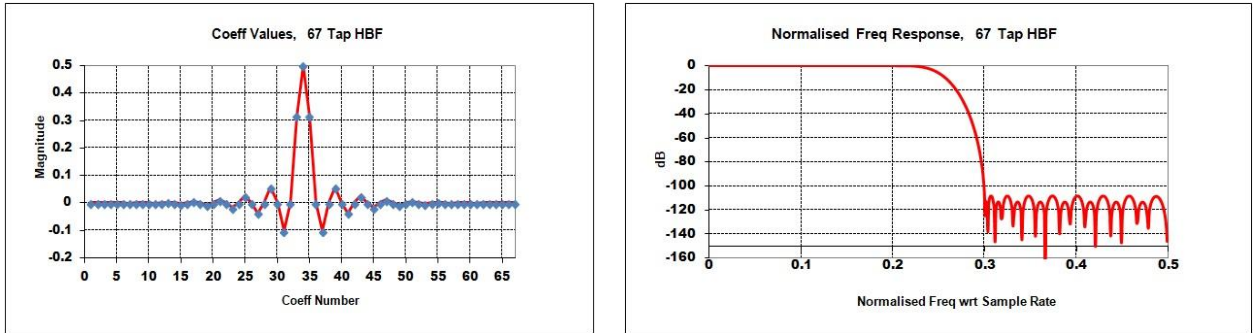
Figure 12. Typical Half-Band Filter Coefficients and Response

A common requirement is to take a stream of Real data (from an ADC for example) and convert it into a complex baseband signal (I & Q) at half the input sample rate. Figure 13 below shows how this may be done in a very elegant fashion.

The left-hand diagram shows the basic process of an I & Q downconverter which is much the same as its analogue equivalent. We take advantage of the fact that alternate samples for the Sine and Cosine LO,s are zero and that alternate taps of the HBF (except for the three centre taps) are also zero. The process simply reduces to taking alternate samples from the input into the I & Q path, passing the odd samples through a 33 tap FIR filter and the even samples through a delay and one tap multiply (i.e. gain). High speed ADC's often use interleaved samples anyway so are likely to provide an alternate sample output so that even the switch may not be necessary.
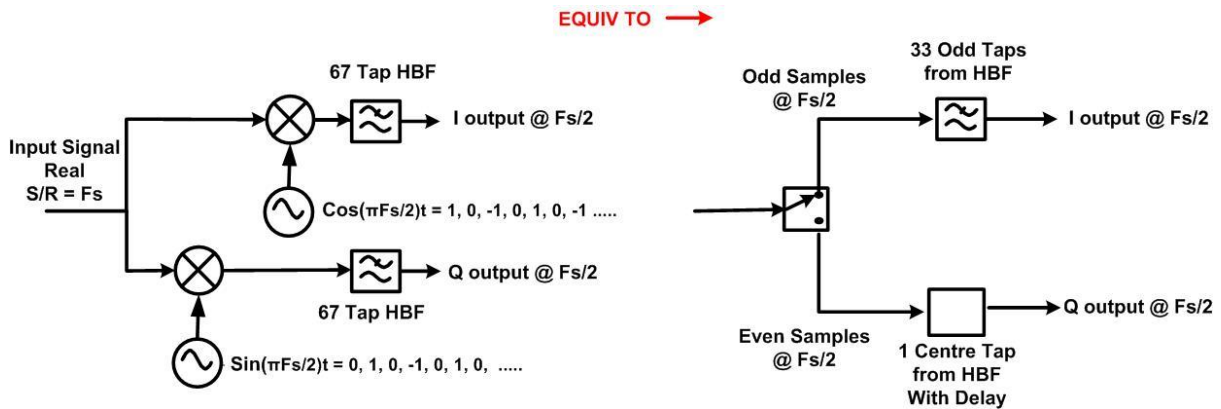

Figure 13. Half-Band Filter in Downconvert & Decimate Mode

From this point in the system, it would be normal to use further filtering, down-conversion, decimation (reduction of sample rate) and bit-width adjustment. One of the most useful filter types, owing to its high efficiency, is the Cascaded Integrator-Comb (CIC) filter (see Ref.[1]). A simplified schematic of this is shown in Figure 14 below. Basically it consists of two sections, the first being a number of cascaded integrators (a simple form of IIR filter) followed by a decimator and an equal number of cascaded comb sections. Both comb and integrator consist simply of delays and adders with no need for multipliers which makes for a very simple structure which is light on resources.

The question is, how does this help to build an efficient downconvert and decimating filter. This is best explained by using an example, in this case, a 5th order CIC (i.e. 5 integrator and 5 comb sections) with Decimate by 10, followed by (in this case), a 385 Tap FIR Filter.
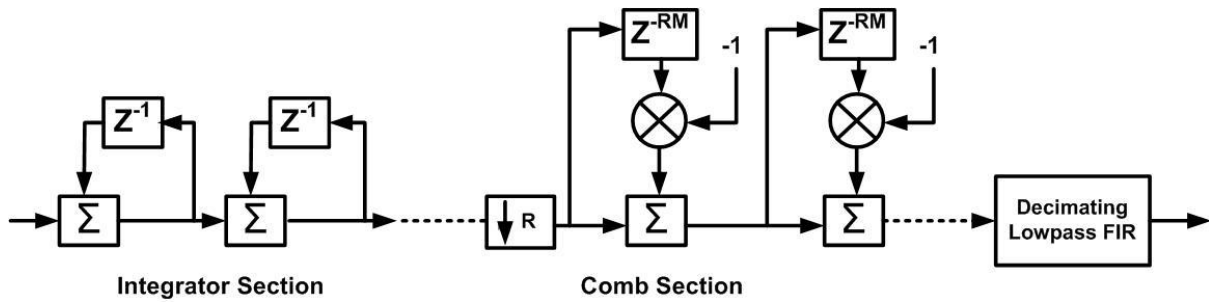
Figure 14. The Cascaded Integrator-Comb Filter

The frequency response (normalised) of the CIC is shown by the Blue curve in Figure 1Figure 15 below. By itself, it would not provide a very selective filter which is why, normally, this is followed by a FIR filter. The Red curve is that of the FIR alone and the Green curve is the combined response. Unfortunately, the CIC response does have significant "droop" in the pass-band which, in turn gives droop in the combined response, as may be seen most clearly in the right-hand diagram (green plot).
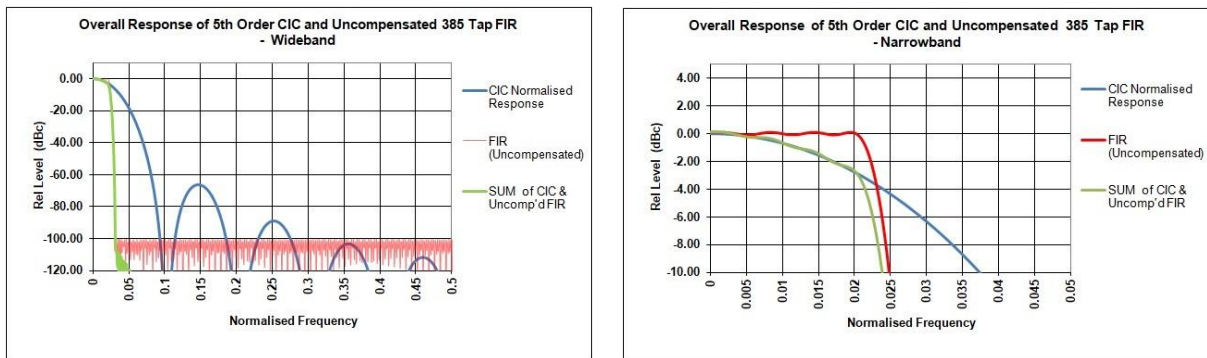
**Figure 15. Response of 5th Order CIC with Dec/10 and Uncompensated 385 Tap FIR**

Fortunately, there is a simple remedy which is to design a FIR which compensates for this droop. This is illustrated in Figure 16 below where the FIR (Red curve) is pre-compensated to make the overall combined response very flat in the passband. For narrowband filters, this can be done with very little degradation of the stop-band performance.
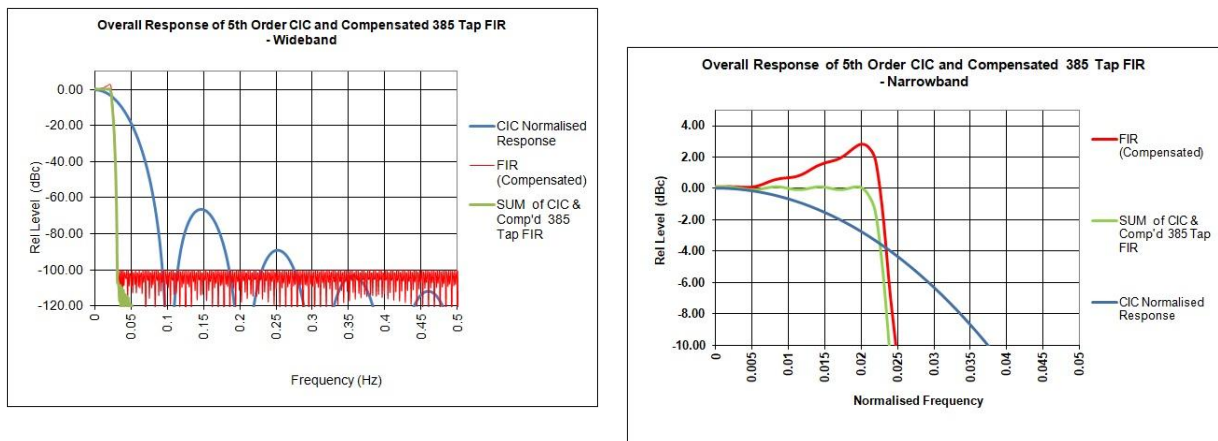
Figure 16. Response of 5th Order CIC with Dec/10 and Compensated 385 Tap FIR

To see just how economical this is, consider the alternative of a FIR filter on its own followed decimation by 10. For a similar passband and cut-off we would need something like 10 times

the number of taps, operating at 10 times the sample rate of the above example. In Figure 17 below, a typical design, using 4121 Taps is compared with the CIC / 385 Tap FIR above. If anything, the CIC /FIR (red plot) combination has a slightly better stop-band performance whilst being more-or-less identical in the passband. The saving in resources is very significant which is why, of course, CIC / FIR filter architectures are very widely used as decimating filters.
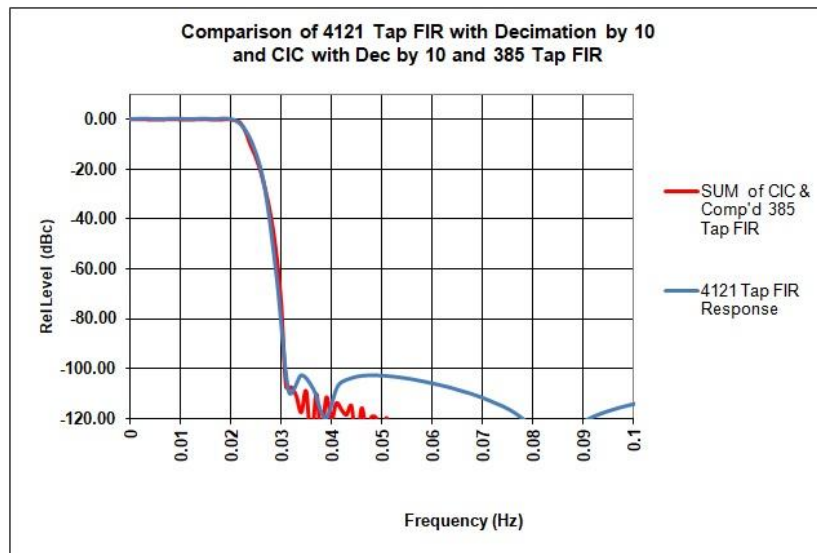

Figure 17. Comparison of 4121-Tap Decimating FIR and Combined CIC / 385-Tap FIR

4.3 Digital Filter-Bank Architectures

The above section dealt with individual filter structures. We will now move on to the subject of Filter Banks and how to create them efficiently. Filter banks can consist of a number of identical, contiguous filters, evenly spaced over the frequency span, or it can be a bank containing filters of different bandwidth, output sample rate, random centre frequency and filter response.

The commonest type of "filter bank" is, of course, the well known Discrete Fourier Transform (DFT) and its efficient implementation, the Fast Fourier Transform (FFT). The FFT may not normally be thought of as a filter bank but that is precisely what it is in the frequency domain.

| (a) | (b) |
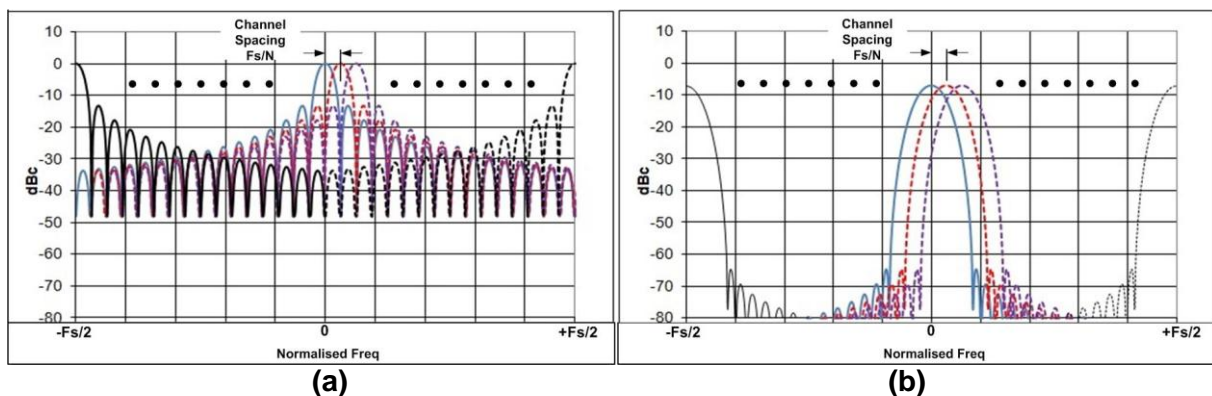
Figure 18. (a) Full Filter Bank for Unweighted FFT (b) For Kaiser Weighted FFT

Figure 18(a) above shows the effective filter bank formed by an N-Point unweighted FFT (i.e. no weighting function is applied to the input data samples). The individual filters are

spaced at Fs/N (where Fs is the sample rate) and the frequency response follows the familiar Sin(x)/x or "SINC" function with the first sidelobe at -13.3 dBc, so not a very selective filter!

The sidelobe performance may be improved by weighting the input samples, an example of Kaiser weighting being shown in Figure 18(b). Although the sidelobe performance is greatly improved with a first sidelobe at -52.7 dBc, the main lobe is very much wider (approx. doubled) with a relative peak gain loss of about -7.2 dB. Although this type of "filter bank" has its uses, in spectral analysis for example, a digital Rx will normally require much more selective channel filters, along the lines illustrated in Figure 19 below.
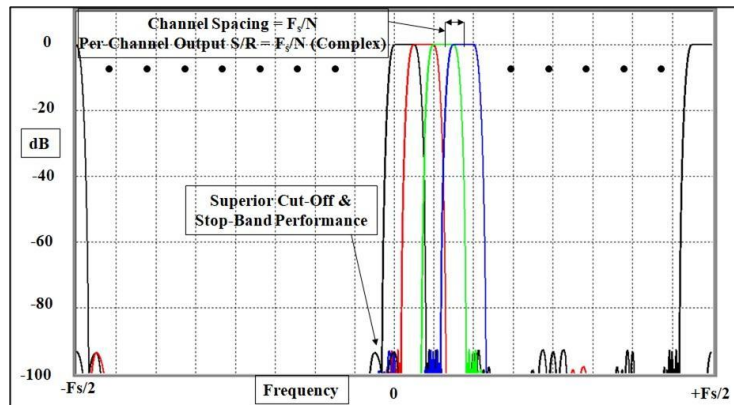

Figure 19. Typical Response of Filter Bank with More Selective Filters

We will now look at the various ways in which such filter banks may be realised in practice.

4.4 Polyphase or WOLA Filter Bank

The filters shown in Figure 19 above can be realised but would required a very large number of taps and a very long FFT. Fortunately, there is a much more elegant solution, as shown in Figure 20 below, known as the Polyphase or WOLA (Weight, Overlap and Add) FFT.
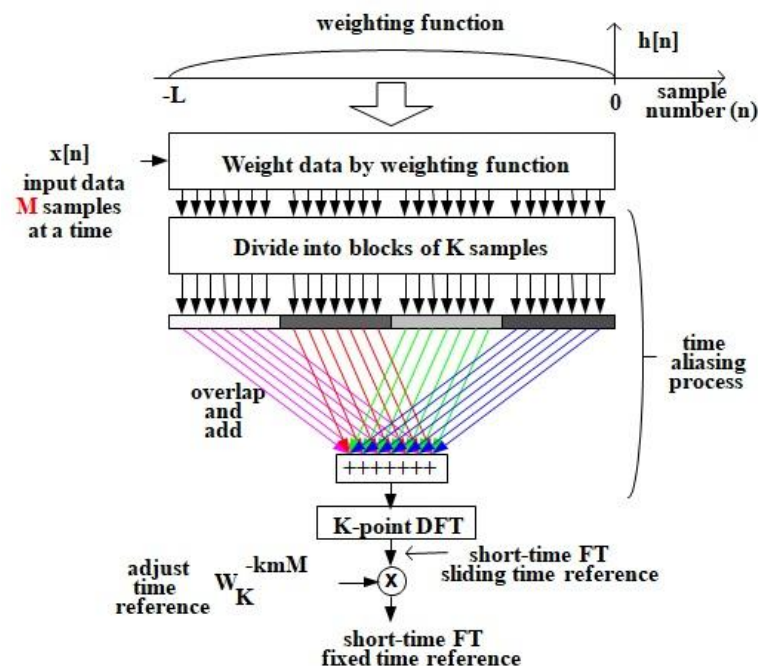

Figure 20. Polyphase or WOLA FFT

11

It may be shown quite readily (but not in this paper - for further reading see Ref. [2] below) that the long data steam of weighted samples can be split into shorter blocks, overlapped and then added before applying a shorter FFT. This very much reduces the required processing resource and is a very economical solution where a large number of identical contiguous filters are required.

4.5 The Pipeline Frequency Transform (PFT).

It is here that I have to admit to a personal interest since the PFT was my own patented invention which was used to launch my company, RF Engines Limited (now known as RFEL Limited, see Ref [3]). The basic concept is very simple and is shown in Figure 21(a) below, relying on successive band splitting into two at each stage using complex down-conversion (CDC) and up-conversion(CUC) to zero IF.
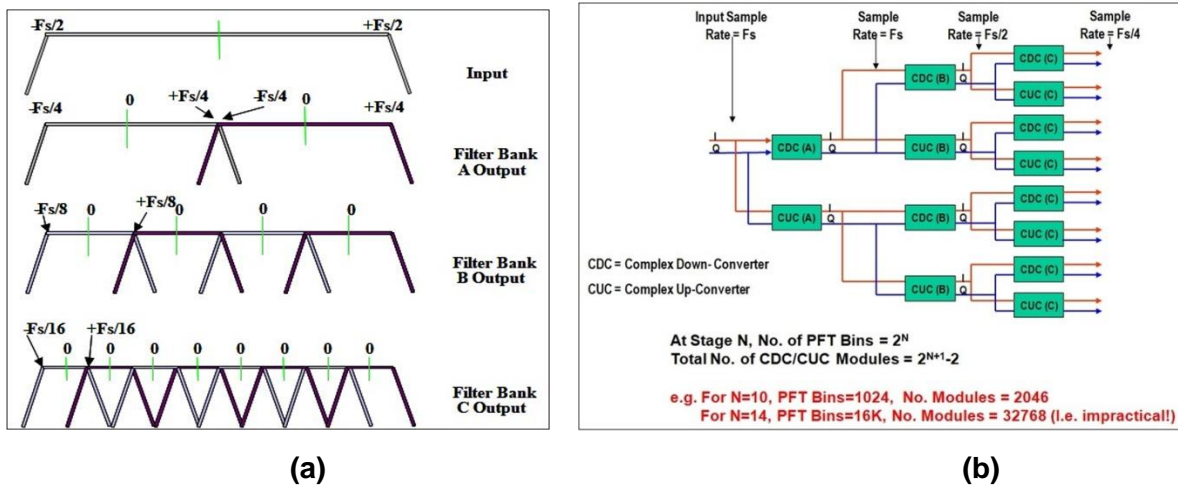

| (a) | (b) |

Figure 21 The Basic Pipeline Frequency Transform (PFT) Structure

Since the input to the filter bank is a complex (I & Q) data stream, it is necessary to use complex up and down-converters which are a modified version of the real to complex converters shown previously in Figure 13 above. Obviously, if a tree system were based on the simplistic approach shown in Figure 21(b), it would become quite impractical for larger numbers of "bins". A 14-stage process with 16K output frequency bins, for example, would require 32768 CDC / CUC modules!
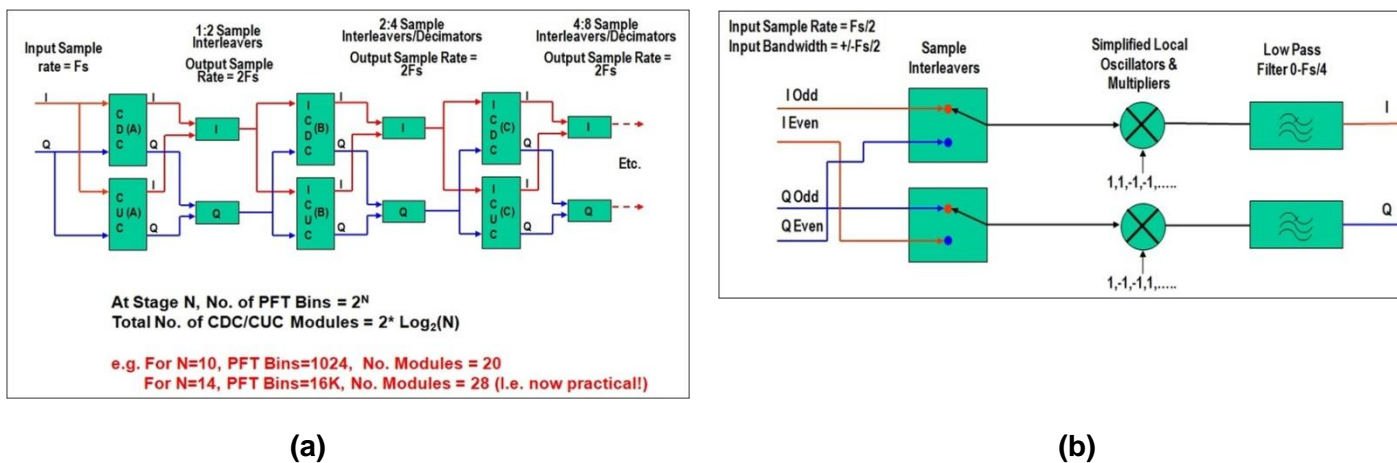

| (a) | (b) |

Figure 22. Simplified PFT by (a) Interleaving Samples and (b) Using Alternate Samples

Fortunately, there is a fairly simple level of simplification which makes use of the fact that the individual sample rates at each stage have been reduced by one half. Since the actual CUC / CDC process is identical for each branch of the tree, we can interleave samples, as shown in Figure 22(a), so that the aggregate sample rate at each stage remains the same at 2Fs. Furthermore, since the individual LO's have alternate values of zero (similar to that described in Figure 13 above) we can, once again, select alternate samples from the I & Q data stream, as shown in Figure 22(b) which halves the number of samples to be processed. There is a further level of simplification which involves the property of zero taps in the Half-Band Filter, similar to that discussed in Sec. 4.2 above (and which the reader can explore further in Ref.[4]). Overall, the structure is a very efficient means of realising a wideband filter bank where we need equal channel spacing and bandwidth and identical filtering in each channel.

Both the PFT and the WOLA filter-bank can be used in this way and have the same overall frequency response (as shown in the example of Figure 19 above). There is a cost in terms of latency (approx 2x) in the PFT because of the pipeline but the PFT does have the additional advantage that all stages in the pipeline are available so that different stages can be selected to match different channel bandwidth requirements.

4.5 Filter Banks with Mixed Bandwidths and Sample Rates

In many cases, of course, there will be the need to handle more complex spectra with a mixture of carrier types such as that illustrated in Figure 23 below. Apart from the obvious solution of building individual filter banks, tailored to match a particular carrier type and channel plan, a number of efforts have been made to build an efficient filter bank which can be adapted "on-the-fly" to provide a flexible channel plan.
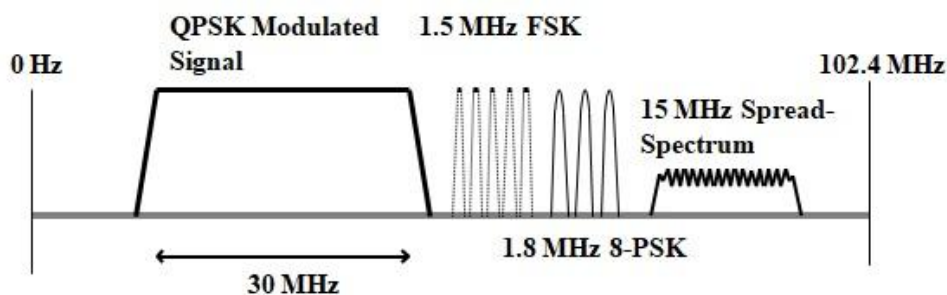

Figure 23. Spectrum with Mixed Carrier Types

Efforts have been made adapt the PFT in this way by selecting outputs from the various stages (known as the "Tuneable PFT") but, in general, it has proved difficult to implement in practice mainly due to the complex routing algorithms required.

A much more practical and efficient implementation is known as "ChannelCore Flex[TM]". This is a product from RFEL Limited (See Ref.[3]) and is shown in Figure 24 below. Although much of the detail is commercially confidential, in broad terms it consists of a number of "coarse" stages (up to 16) each of which can accept an input channel of multi-GHz bandwidth. These are followed by a multiplexer with gain and fine-tuning stages to give the required channel filter type, bandwidth and sample rate. The coarse stages can be based on different architectures including the PFT, WOLA or Frequency Domain Convolution, depending on the particular application. The fine stages use fairly standard DDC techniques together with advanced re-sampling to ensure the sample rate matches the required output symbol rates. All of this is programmable at run-time to adapt to changing requirements.
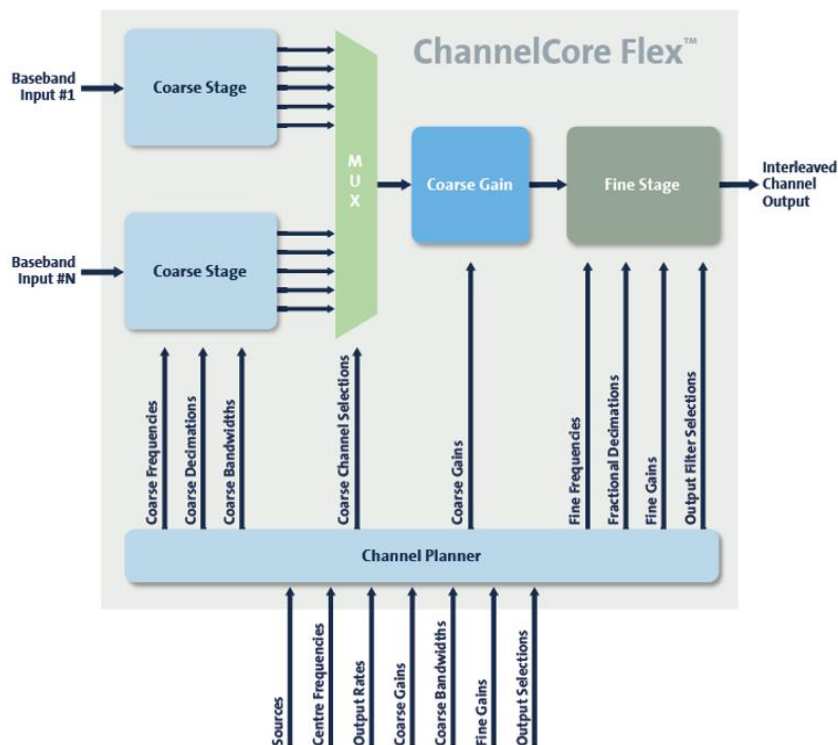
Figure 24. Functional Diagram of ChannelCore Flex Processing

5.   *Conclusions*

It is hoped that, within the constraints of this paper, the reader has gained some insight into the world of digital filters and advances in digital filter bank implementation. The author believes that analogue filters still have a very clear role in modern system design for reasons discussed above and, although digital filters will continue to displace many analogue functions as ADC and DAC sample rates and bit-width increase, the all-digital "antenna-to-bits" system is still some way off. At present, the dividing line seems primarily where digital systems can carry out such functions as channelisation, signal detection, modulation and demodulation, leaving the higher frequency filtering and up / downconversion functions to the analogue designer.

6.   *REFERENCES*

[1] E. B. Hogenauer. "An economical class of digital filters for decimation and interpolation", IEEE Transactions on Acoustic, Speech and Signal Processing, ASSP-29(2):155-162, 1981., http://read.pudn.com/downloads163/ebook/744947/123.pdf

[2] Robert F. Ackermann. "Understanding and Implementing the Weight(ed) Overlap Add Fast Fourier Transform (WOLA FFT)", (SETI Institute) 28-11-08, http://ackrman.net/seti/wola-fft/2008-11-28-wola-fft.pdf

[3] http://www.rfel.com

[4] J. Lillington. "Comparison of Wideband Channelisation Architectures", International Signal Processing Conference, Dallas, March 2003