B-factor of FX correlator for different pairs of digitizers (two- and four- levels quantization)

L. Kogan[1]

(1) - National Radio Astronomy Observatory, Socorro, New Mexico, USA

October 13, 1993

Abstract

The B-factor of the FX correlator has been analyzed for different pairs of digitizers (two- and four- levels) and for different tapering functions at the FFT stage. Problems of reconstructing the spectrum of the original analog signals from measurements of an FX correlator in the case of strong correlation(auto correlation in particular) are formulated. It is shown that zero padding can not be just an option of FX correlator but *must* be applied in the high correlation coefficient case (auto correlation in particular).

1 The approach to the problem

Let's follow the path of the signals through all stages of the FX correlator. The signal at the FFT output can be described by equation (1).

$$\eta(f) = \sum_{n=1}^{N} H(n)\xi(t)\exp(-j2\pi ft)\exp(-j2\pi f_{fr}t)exp(-j2\pi f\delta\tau(t)) \qquad (1)$$

where $\xi(t)$ is the digitized sampled input signal; $t = n\Delta t + K(L-1)$ is a time;
$K \leq N$ is a shifting of the beginning of the FFT sequence relative to the previous one;
There is no overlapping if $K = N$; L is a sequence number of FFT;
$\Delta t = \frac{1}{2\Delta f}$ is a sampling period; Δf is the bandwidth of the original analog signal;
N is the number of input points at FFT; f_{fr} is an estimated fringe rate relative to a reference point (Earth center for example);
$\delta\tau(L)$ is a fractional sample correction relatively to the reference point; $H(n)$ is a tapering function.

The outputs of the FFT's from two stations are multiplied and accumulated separately for each frequency. The second output is taken in conjugate form. The expectation of this accumulated product can be represented by the triple sum:

$$\zeta(f) = A\sum_{L}\sum_{i=-N}^{i=N}\sum_{n=1}^{N} H(n)H(n+i) < \xi_1(n)\xi_2(n+i)\exp(-j\Phi_{fr}(L,n,i) > \cdot\exp(-j2\pi f(i\Delta t + \delta\tau_{12}(L))) \qquad (2)$$

where A is a coefficient which is applied to the data on this stage;
$\Phi_{fr}(L,n,i)$ is the difference phase of the fringe lobes;
$\delta\tau_{12}(L)$ is the difference in fractional sample correction between two antennas;

Having changed the sequence of calculation of the sums we can rewrite equation 2 :

$$\zeta(f) = A \sum_{i=-N}^{i=N} \sum_{n=1}^{N} H(n)H(n+i) \sum_{L} < \xi_1(n)\xi_2(n+i)\exp(-j\Phi_{fr}(L,n,i) >$$
$$\cdot \exp(-j2\pi f(i\Delta t + \delta\tau_{12}(L))) \qquad (3)$$

The expectation $< \xi_1(n)\xi_2(n+i) >$ is connected with the correlation coefficient of the input analog signals ρ through the conversion function $F(\rho)$ associated with the digitization level used (Kogan, 1993). The correlation coefficient of the analog input signals is represented by a sine wave with phase $\Phi_{fr}(L,n,i) + \arg(b(i)) + \phi$ and amplitude $r \mid b(i) \mid$, where r is the maximum value of correlation coefficient occuring when delay $i = 0$; $b(i)$ is the Fourier transform of the spectrum of input analog signals:

$$b(i\Delta t) = \frac{1}{\Delta f}\int_0^{\Delta f} F(f)\exp(j2\pi\frac{f}{2\Delta f}i)df \;=\; \int_0^1 F(x)\exp(j\pi xi)dx; \;\; x = \frac{f}{\Delta f} \qquad (4)$$

The conversion function $F(\rho)$ distorts the sine wave. This distortion is symmetrical because the function $F(\rho)$ is odd. So the first harmonic (only this one is relevant for the accumulated expectation of (3)) has the same phase but less amplitude. This suppression of the amplitude is described by function $F_1(\rho)$ which has been analyzed for different combinations of digitizers in [1]. So equation (3) can be rewritten:

$$\zeta(f) = \frac{1}{2}A \sum_{i=-N}^{i=N} \sum_{n=1}^{N} H(n)H(n+i) \sum_{L} \frac{F_1(r \mid b(i\Delta t + \delta\tau_{12}(L)) \mid)}{r \mid b(i\Delta t + \delta\tau_{12}(L)) \mid}$$
$$\cdot rb(i\Delta t + \delta\tau_{12}(L))\exp(j\phi)\exp(-j2\pi f(i\Delta t + \delta\tau_{12}(L))) \qquad (5)$$

The fractional bit correction $\exp(-j2\pi f\delta\tau_{12}(L))$ to be applied to the Fourier transform of the function $b(i\Delta t + \delta\tau_{12}(L))$ shifts it and converts it to $b(i\Delta t)$.

2 A linear approach

The function $F_1(\rho)$ is almost linear for the whole range of ρ from 0 to 1 for 2 of the 3 digitizer combinations accepted by VLBA correlator (TWO-FOUR and FOUR-FOUR) (Kogan, 1993). Let us define the slope of the function $F_1(\rho)$ in the linear limit as α. Then taking into account our reasoning above about the fractional bit correction we can rewrite equation (5):

$$\zeta(l) = c \cdot r \sum_{i=-N}^{i=N} h(i)b(i)\exp\left(-j2\pi\frac{il}{N}\right)$$

$$where\; h(i) = \frac{1}{N}\sum_{n=1}^{n=N} H(n)H(n+i); \;\; l = 1,2,\ldots,\frac{N}{2}; \;\; f = l\frac{\Delta f}{N/2} \qquad (6)$$
$$c = \frac{1}{2}A \cdot r_m \cdot \alpha \cdot N \cdot \exp(j\phi)$$

The constant r_m in (6) is the maximum value of correlation of digitized signals. The values of this constant are given for different pairs of digitizers in [1]. Having substituted the expression $b(i)$ from (4) to (6) we receive the following expression for the output of the FX correlator:

$$\zeta(l) = c \cdot r \int_0^1 F(x) f\left(x - \frac{1}{N/2}\right) dx$$

$$where\ f(z) = \sum_{i=-N}^{N} h(i) \exp(j\pi i z) \tag{7}$$

The function $f(z)$ describes the pattern of an equivalent filter corresponding to the weighting function $h(i)$.

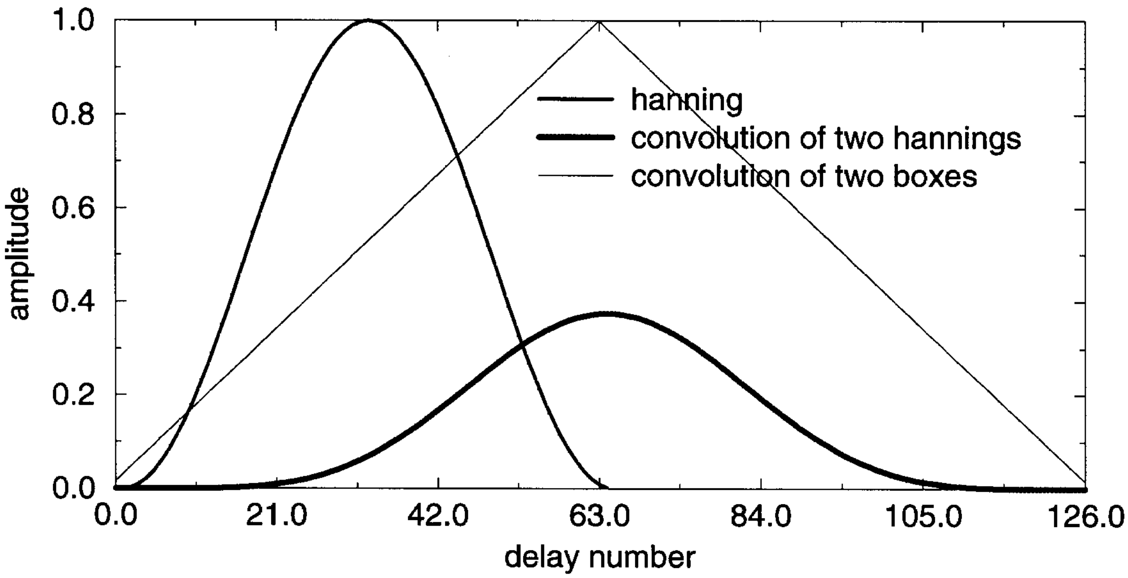

Figure 1: Equivalent weighting function at delay axis for a box and hanning tapering function at the original FFT

It is necessary to note that the pattern of an equivalent filter is not determined by the original tapering function $H(n)$ but its auto convolution (see equation (6)). If the original tapering function is represented by a box then the weighting function h(i) has a triangular form:

$$h(i) = \begin{cases} \frac{1}{N}(N- \mid i \mid) & \text{if } \mid i \mid < N \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

For Hanning original tapering function $H(n) = 0.5 - 0.5\cos\left(2\pi\frac{n-1}{N}\right)$ the weighting function $h(i)$ is determined by the expression:

$$h(i) = \begin{cases} \frac{1}{4}\left[\left(1 - \frac{i}{N}\right)\left(1 + \frac{1}{2}\cos\left(2\pi\frac{i}{N}\right)\right) + \frac{3}{4\pi}\sin\left(2\pi\frac{i}{N}\right)\right] & \text{if } \mid i \mid < N \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

The original tapering functions $H(n)$ and corresponding weighting function $h(i)$ are shown in Fig.1 for the box and Hanning case.

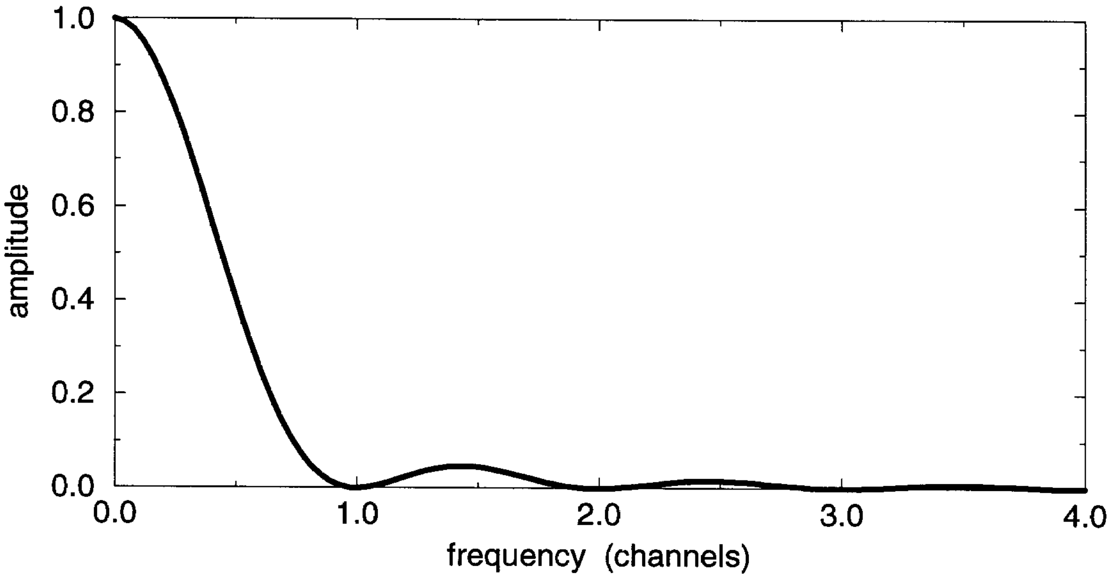

Figure 2: Filter corresponding to a Box weighting function at FFT of a FX correlator

The equivalent filter patterns calculated using (7) are shown at Fig.2-3. Now calculate the sum of all measured harmonics at output of FX correlator to estimate the whole correlation coefficient.

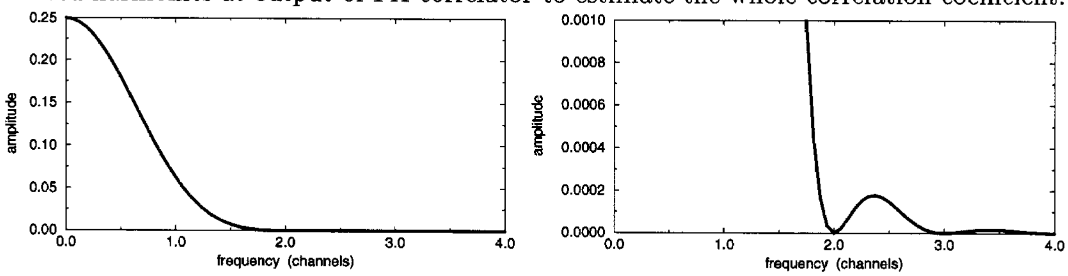

Figure 3: Filter corresponding to a Hanning weighting function at FFT of a FX correlator

The equivalent filters shown in Fig.2-3 have a very narrow bandwidth (< 1 channel) so we can be sure that the output of the FX correlator would be zero at negative frequencies as well as original spectrum. That means we can add $\zeta(l)$ for l from 1 till N. The first $\frac{N}{2}$ coefficients are measured. The second $\frac{N}{2}$ have to be equal to zero. Using (7) we receive the expression for this sum:

$$\sum_{l=1}^{N/2}\zeta(l) \simeq \sum_{l=1}^{N}\zeta(l) = c\cdot r\int_0^1 F(x)\left[N\cdot h(0)+2\sum_{i=1}^{N}h(i)\left(\sum_{l=1}^{N}\cos\left(\pi i\left(x-\frac{l}{N/2}\right)\right)\right)\right]dx \qquad (10)$$

The internal sum at (10) is 0. The $\int_0^1 F(x)dx = 1$ because $b(0)=1$. So expression (10) can be simplified:

$$\sum_{l=1}^{N/2}\zeta(l) \simeq \sum_{l=1}^{N}\zeta(l) = c\cdot r\cdot N\cdot h(0) \qquad (11)$$

Having substituted the expression for c from (6) we receive a final expression for the estimation of a whole correlation coefficient:

$$r = \frac{\left|\sum_{l=1}^{N/2}\zeta(l)\right|}{\frac{1}{2}\cdot A\cdot r_m\cdot\alpha\cdot N^2\cdot h(0)} \qquad (12)$$

The estimation of the correlation in the individual frequency channel is given by the next equation:

$$r\frac{F(l)}{\sum\limits_{l=1}^{N/2} F(l)} = \frac{|\zeta(l)|}{\frac{1}{2}\cdot A\cdot r_m\cdot\alpha\cdot N^2\cdot h(0)} \tag{13}$$

3 The problem of extremely high correlation.

Now let us consider a case when the function relating the input correlation and correlation of digitized signals can not be considered linear. It occurs when the input correlation is close to unity especially for two level digitizers, and definitely for the auto correlation case. Then equation (6) can be rewritten:

$$\zeta'(l) = c\cdot r\sum_{i=-N}^{i=N} h(i)b'(i)\exp\left(-j2\pi\frac{il}{N}\right) \tag{14}$$

where $b'(i)$ is the normalized cross correlation function of the digitized signals. The measured spectrum $\zeta'(l)$ does not coincide with the cross correlation spectrum of original analog signals $\zeta(l)$ because $b'(l)$ and $b(l)$ are not identical. How can we restore the original spectrum $\zeta(l)$ from a measured one $\zeta'(l)$ using the knowledge of the conversion function relating $b'(l)$ and $b(l)$? The standard way is to transform back to the delay domain by taking the inverse Fourier transform of the measured spectrum $\zeta'(l)$. So we calculate the function $R'(i)$:

$$R'(i) = \frac{1}{N}\sum_{i=-N}^{i=N}\zeta'(l)\exp\left(j2\pi\frac{il}{N}\right) \tag{15}$$

The NRAO VLBA correlator measures the spectrum only for positive frequencies ($N/2$ points). The negative components (i.e. the components corresponding to $N/2 < l < N$) are supposed to be zero in accordance with the spectrum of original analog signals or they are symmetrical in the auto correlation case. But the spectrum of the digitized signals is not zero for the negative frequencies. So generally we don't have any idea about the spectrum at negative frequencies since we only have information about the spectrum at positive frequencies.

That means we can not remove the nonlinear effect due to the digitizing of signals by transforming back to the delay domain because absence of information about negative components in the spectrum.

Now suppose we have this information, that is we know the spectrum for the whole range $1 \leq l \leq N$. This case is realized in particular for an auto correlation spectrum because the auto correlation spectrum of a digitized signal is a real and symmetrical function, as well as the auto correlation spectrum of the relevant analog signal. Then we meet another problem. It is clear that $R'(i)$ is related to the nuclei of the direct Fourier transform $R(i) = h(i)b'(i)$ at (14) by the expression:

$$R'(i) = R(i) + \sum_k R(i+kN) \tag{16}$$

where k is any integer for which $i + kN$ belongs to the range of i at the sum of expression (14). In our case k can be equal to 1 or -1. How can we provide the identity between $R'(i)$ and $R(i)$? We can select the special tapering function $H(n)$ at the FFT which is zero for half of the points ($N/2$) (zero padding). It provides nonzero values of $R(i)$ only for N points and so provides the identity between $R'(i)$ and $R(i)$. With this supposition we can restore $R(i) = h(i)b'(i)$ without ambiguity for $-N/2 < i < N/2$. We have to divide calculated values of $R(i)$ by $h(i)$ to find the relevant values of normalized correlation function of digitized signals $b'(i)$. Now we can apply a known conversion function to $b'(i)$ to restore the correlation function of the original analog signal $b(i)$. But we can do it only for a reduced range of i - approximately $-N/2 < i < N/2$ instead of original one $-N < i < N$. So the restored spectrum would have worse resolution by factor ≥ 2 (An additional problem can appear at the division $R(i)$ by $h(i)$ because $h(i) \simeq 0$ near the edges of its range). So zero padding can not be just an option of FX correlator but *must* be applied in the high correlation coefficient case (auto correlation in particular). This is a repetition of John Granlund's conclusion [2]. *An auto correlation spectrum of original analog signal can be restored from the measurements of FX correlator by transforming back to the delay domain at the cost of loosing the resolution by more than a factor of two.*

References

1. L. Kogan, VLBA scientific memo 5, 1993
2. J. Granlund, VLBA correlator Memo 66, 1986