Sensitivity Efficiency of the 64 Antenna Correlator Upgrade

Prepared by Omar Yeste Ojeda
January 2017

1 Summary of results

1. Either blanking or not blanking have similar worst case sensitivity loss (<0.4%), being not blanking preferable because:
    a. It is simpler to implement
    b. There is no sensitivity loss in FDM and it is generally smaller (than blanking) for TDM.
2. However, not blanking has a subtle windowing effect on the spectrum estimate in TDM (the spectral window, which is at least 1.7-MHz wide, is broadened 2 KHz) resulting in a resolution loss about 0.1%.
3. Special care must be taken in the ADC data conversion (pass-through option of TFBs).
4. Sensitivity efficiency for TDM is summarized in Table 1.
5. Sensitivity efficiency for FDM is gathered in Table 2.
6. Using 3-byte (24 bits) final results incurs a sensitivity loss:
    a. Less than 0.02% for a mode dependent truncation
    b. Less than 0.6% for a fixed, mode independent, truncation considering that:
        i. 4-bit scaling is such that 4 additional bits are truncated in 4-bit modes (i.e., 1×, 1/4×, and 1/16×).
        ii. Sum of different correlation planes is scaled by 1/16 (4 additional bits are truncated)

Table 1 Summary of sensitivity efficiency for TDM – To be multiplied to ADC's efficiency

ADC # bits	Correlator # bits	Pass-through	Packetizing		1-ms truncation	Total
			Blanking	η		
3	4	1.0000	Yes	0.9990 [a]	0.9997 [d]	0.9987
			No	1.0000 – 0.9993 [a,c]		0.9997 – 0.9990
	2	0.9088	Yes	0.9959 [b]	0.9997	0.9048
			No	1.0000 – 0.9973 [b,c]		0.9085 – 0.9061
4	4	1.0000	Yes	0.9990 [a]	0.9996	0.9986
			No	1.0000 – 0.9993 [a,c]		0.9996 – 0.9989
	2	0.8914 [e]	Yes	0.9959 [b]	0.9998 [e]	0.8876
			No	1.0000 – 0.9973 [b,c]		0.8912 – 0.8888

[a] *Computed at* $\tau_{max} = 512$.
[b] *Computed at* $\tau_{max} = 2048$.
[c] *Minimum efficiency computed for a spectral line between channels using a Hanning window.*
[d] *Assuming ADC data are zero-padded on the right to form a 4-bit word.*
[e] *Assuming ADC data are converted following Section 4.1.1.*

Table 2 Summary of sensitivity efficiency for FDM – To be multiplied to ADC's and TFB's efficiencies. Additional loss if blanking is implemented.

Correlator # bits	Re-quantization	Oversampling		1-ms truncation	Total
		Twice Nyquist	η		
4	0.9885	Yes	1.0060	0.9997	0.9941
		No	1.0000	0.9996	0.9881
2	0.8812	Yes	1.0611	0.9999	0.9349
		No	1.0000	0.9998	0.8810

2 Contents

3 Introduction

The purpose of this study is to quantify the sensitivity loss attributable to the ALMA correlator in order to support its upgrade design.

Unless otherwise stated, the signal at the ADC input is assumed stationary Gaussian noise, with zero mean and standard deviation such that the ADC quantization step is optimal in terms of sensitivity. Additionally, the cross correlation coefficient between different antennas is assumed small (low SNR).

Unless otherwise stated, sensitivity loss is quantified herein in terms of efficiency, defined as the ratio between the sensitivities of an ideal receiver operating at the output and the input of such device. Hence, the efficiencies of all stages multiplies previous to the correlator in order to get the overall efficiency. Both terms, *sensitivity loss* and *efficiency*, are used interchangeably in this document.

4 Time Division Mode

Figure 1 roughly sketches the functional block diagram (per baseline) of the correlator operating in TDM. Only the elements relevant in terms of sensitivity loss are shown, assuming all the neglected elements incur no sensitivity loss. The packetizer is implemented at the station card and splits every millisecond of input samples into 32 packets each made of 250,000 contiguous samples. Packetizing the data stream allows matching the signal data rate to the correlator clock frequency. However, it also destroys the correlation between samples on both sides of a packet edge, which may be solved through blanking at the cost of some sensitivity. Additionally, depending on the observational mode, the packetizer limits the input samples to 2-bit (4-level) quantization by direct truncation of the least significant bits (LSBs).

At the correlator portion of the system, every packet is correlated with its counterpart from the other antenna in the baseline pair. All baseline pairs (per baseband) are correlated in one correlator plane, computing up to 2048 leads (plus 2048 lags) per baseline pair. The accumulation of the correlation products is done in several stages, being the first one the only one at which the accumulation result is truncated. Specifically, the six LSBs are removed after accumulating 250,000 samples (1 packet), which produces additional sensitivity losses. For the rest of the correlator, full precision is kept, so no additional sensitivity loss is expected (the loss at the CDPs –single precision floating point– is negligible). Section 6 analyzes the additional sensitivity loss that would be incur by a hypothetical truncation immediately before data are transmitted to the CDPs.

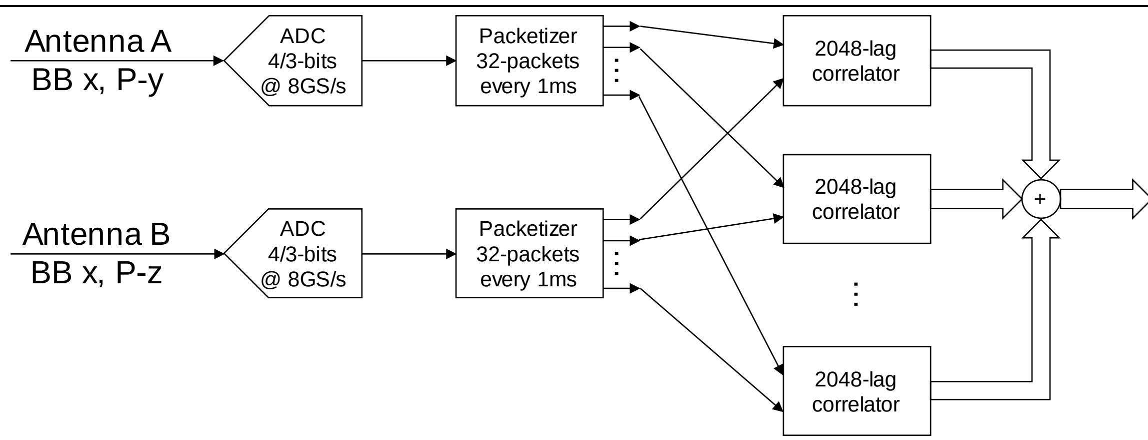


Figure 1- Block diagram of the ALMA correlator upgrade operating in FDM

In summary, the three different operations that incur some loss of sensitivity in TDM are:

1. Pass-through truncation
2. Packetizing (with or without blanking)
3. 1-ms truncation[1]

All of them are quantified in the next paragraphs.

4.1 Pass-through truncation

The term truncation, as opposed to re-quantization, is used herein to refer to a process whereby a given number of LSBs of the input are discarded. The sensitivity loss can be well approximated through quantization noise only when the standard deviation of the signal to be truncated is much greater than 1, which is not the case herein.

In this case, the sensitivity efficiency can be computed by considering the digitization and subsequent truncation as a whole. For example, a 3-bit ADC and truncation of one LSB can be thought as a 2-bit ADC with a quantization step twice the optimal for 3-bit quantization, i.e.

$$\Delta = 2 \cdot 0.586 = 1.172 \text{ rms} \tag{1}$$

where Δ is expressed relative to the input's standard deviation (or in "rms" units). As a reminder, the optimal step of a 2-bit ADC is $\Delta = 0.995$ rms, with an efficiency $\eta = 0.8812$ [1]. However, the quantization efficiency of a 2-bit ADC with the quantization step in (1) is [1]

$$\eta = 0.8748 \tag{2}$$

The sensitivity loss attributable to truncation is the ratio between the quantization efficiencies of the actual ADC (3-bit with optimal step) and that hypothetic ADC (2-bit with step in (1)), which results in

$$\eta = \frac{0.8748}{0.9626} = 0.9088 \tag{3}$$

For a 4-bit ADC and truncation of two LSBs, the quantization step of the equivalent 2-bit ADC becomes four times the optimal for 4-bit quantization, that is

$$\Delta = 4 \cdot 0.335 = 1.34 \text{ rms} \tag{4}$$

The total quantization efficiency is hence [1]

$$\eta = 0.8587 \tag{5}$$

and the sensitivity loss due to truncation

$$\eta = \frac{0.8587}{0.9885} = 0.8688 \tag{6}$$

[1] Despite the accumulation of 250,000 samples corresponds to 1/32th of a millisecond, the name 1-ms truncation is used for similarity to FDM.

4.1.1 Recommendation for a 4-bit ADC

In order to minimize the sensitivity loss due to a non-optimal quantization step, the best strategy is to adjust the signal level at the ADC input according to the effective number of bits at the input of the correlator, which itself depends on the observation mode. This might require some modifications in the total power digitizer, monitor and control board of the analog backend.

Otherwise, instead of direct truncation of the two LSBs, it would be better to keep the sign as the most significant bit (MSB), while setting the LSB of the truncated output to zero if the absolute value of the analog input is less than three quantization steps. Assuming two's complement representation of the ADC output,

$$\mathrm{LSB} = (\mathrm{X} > 2) \tag{7}$$

where X is the (3-bit) result of an exclusive OR of the MSB with the rest of bits. As a result, the quantization step of the equivalent 2-bit ADC is

$$\Delta = 3 \cdot 0.335 = 1.005 \text{ rms} \tag{8}$$

which is closer to the 2-bit optimal, yielding a total quantization efficiency

$$\eta = 0.8811 \tag{9}$$

And a sensitivity loss due to truncation

$$\eta = \frac{0.8811}{0.9885} = 0.8914 \tag{10}$$

that is, an improvement of 2.6% with respect to direct truncation at a small cost.

4.2 Packetizing

As mentioned, packetizing has the deleterious effect of decorrelating nearby samples at the packet edges. Therefore, when a new packet starts at the prompt signal in the correlator, the delayed signal contains samples from the previous packet, accumulating only noise in the correlation estimate. Two options are considered next:

1. Blanking (the noisy products)
2. Not blanking (do nothing)

4.2.1 Blanking

At the start of a new packet, the prompt signal starts feeding the new packet, but the values in the shift registers of the correlator for the delayed signal correspond to the old packet (except for lag zero). The product of samples from the new and the old packets are referred to as *noisy products* in this document. To prevent corrupting the correlation estimate, the implemented solution consists of resetting the internal accumulator to zero when all the samples of the old packet have been pushed out of the registers (in other words, the transient due to a new packet is blanked). Thus, the number of blanked samples is equal to the maximum lag of interest. For simplicity, this reset is done synchronously for all lags, thereby blanking is fully equivalent to decreasing the number of accumulated samples. As the sensitivity gain of the accumulator is proportional to the square root of samples used, the sensitivity loss due to blanking results

$$\eta(\tau_{max}) = \sqrt{1 - \frac{\tau_{max}}{N_{packet}}} \tag{11}$$

where N_{packet} is the number of samples within a packet, that is 250,000 samples. To quantify this result, the sensitivity loss for several values of τ_{max} are gathered here:

$$\begin{gathered} \eta(\tau_{max} = 512) = 0.9990 \\ \eta(\tau_{max} = 1024) = 0.9979 \\ \eta(\tau_{max} = 2048) = 0.9959 \end{gathered} \tag{12}$$

4.2.2 Not blanking

In case of not blanking, the number of noisy products accumulated depends on the specific lag being computed. For example, packetizing has no effect at all when computing a zero-lag, 1 noisy product is accumulated for a lag equal to 1, and so on. Therefore, the number of noisy products is equal to the lag being computed.

The average number of accumulated noisy products is hence

$$\xi_n(\tau) = \frac{|\tau|}{N_{packet}} \tag{13}$$

where τ is the lag being computed.

4.2.2.1 Biased estimate

If the noisy products are not blanked, the resulting correlation estimates become biased. However, this bias can be neglected as long as $\xi_n(\tau_{max}) \ll 1$ (or $|\tau_{max}| \ll 250{,}000$ samples). Since the noisy products exhibit no correlation, the expected value of the estimated correlation is proportional to $(1 - \xi_n)$. Therefore, the computed spectrum in the true spectrum convolved with a triangular window with half a million samples, resulting in a resolution bandwidth of 20 KHz. In TDM, the maximum lag is $\tau_{max} = 512$ samples, so the effective window is truncated to the innermost 1024 samples. This effective window is the result of convolving the 20-KHz triangular window with a 7-MHz rectangular window, therefore the result is accurately approximated by the 7-MHz rectangular window. It can be checked that the bandwidth of the effective window is increased 2 KHz[2] with respect to the rectangular window. Using any window other than rectangular for computing the spectrum will increase the bandwidth of the window, and thereby the accuracy of the approximation. Increasing the maximum lag to 2048 samples will reduce the rectangular window bandwidth to 1.7 MHz, still three orders of magnitude greater than the bandwidth of the triangular window.

4.2.2.2 Sensitivity loss

Since noisy products do not alter the noise component of the correlation estimates, the noise component of the spectrum is the same as for an ideal non-packetized correlator. That is to say, only the

[2] 2 Khz, not 20 KHz.

signal component is windowed by the triangular window. From the foregoing discussion, such effect can be neglected.

The specific sensitivity loss depends on the signal spectrum and is a function of frequency. It can be computed as the ratio between the signal spectrum with and without applying the truncated triangular window:

$$\eta(f, \tau_{max}) = \frac{\sum_{\tau=-\tau_{max}}^{\tau_{max}-1} [1 - \xi_n(\tau)]\, w(\tau)\, r(\tau)\, e^{-j2\pi f\tau}}{\sum_{\tau=-\tau_{max}}^{\tau_{max}-1} w(\tau)\, r(\tau)\, e^{-j2\pi f\tau}} \tag{14}$$

where $w(\tau)$ is the spectral window and $r(\tau)$ is the signal correlation function. To quantify it, two cases are considered next: a flat signal spectrum and a spectral line.

4.2.2.2.1 Flat signal spectrum

In case of a flat signal spectrum, the cross correlation function is a sinc function with pseudo-frequency equal to the symbol rate. For a rectangular window, the maximum sensitivity loss is produced when this sinc function is centered at the maximum lag[3] (degenerating into a Kronecker delta at that lag). The sensitivity loss is the ratio between the sum of the cross correlation function with and without the triangular window, that is:

$$\eta_{min}(f) = \eta_{min}(f = 0) = 1 - \xi_n(\tau_{max}) = 1 - \frac{\tau_{max}}{N_{packet}} \tag{15}$$

Quantifying this result for several values of τ_{max}:

$$\begin{gathered} \eta_{min}(\tau_{max} = 512) = 0.9980 \\ \eta_{min}(\tau_{max} = 1024) = 0.9959 \\ \eta_{min}(\tau_{max} = 2048) = 0.9918 \end{gathered} \tag{16}$$

In practice, the geometric delay is compensated to values close to zero, therefore not blanking produces negligible sensitivity loss for flat-spectrum signals.

4.2.2.2.2 Spectral line

In case of a spectral line, the spectral cross correlation function becomes a sinusoidal function. Therefore, the sensitivity loss results[4]:

$$\eta_w(f_L) = \left| \frac{\sum_{\tau=-\tau_{max}}^{\tau_{max}-1} [1 - \xi_n(\tau)]\, w(\tau)\, \exp(2\,\pi\,\tau[f_L - f_{CH}])}{\sum_{\tau=-\tau_{max}}^{\tau_{max}-1} w(\tau)\, \exp(2\,\pi\,\tau[f_L - f_{CH}])} \right| \tag{17}$$

f_L is the frequency of the spectral line, and f_{CH} is the center frequency of the closest frequency channel. For a spectral line centered on the channel and rectangular window:

[3] This is a very pessimistic possibility, so the minimum efficiency shall be extracted from the spectral line case.
[4] The effect of image frequencies is neglected, assuming a low-sidelobe window and/or high enough frequency.

$$\eta_{Rect}(f_L - f_{CH} = 0) = \frac{1}{2\,\tau_{max}} \sum_{\tau=-\tau_{max}}^{\tau_{max}-1} [1 - \xi_n(\tau)] = 1 - \frac{\tau_{max}}{2\,N_{packet}} \quad (18)$$

which yields approximately the same sensitivity loss as blanking ($\sqrt{1-x} \approx 1 - x/2$, for $x \ll 1$).

If the spectral line locates at the mid-point between two channels (i.e., $f_L - f_{CH} = 0.5$ bins), the sensitivity loss (at each of those channels) results:

$$\eta_{Rect}\left(f_L - f_{CH} = \frac{1}{4\,\tau_{max}}\right) = \left| \frac{\sum_{\tau=-\tau_{max}}^{\tau_{max}-1} [1 - \xi_n(\tau)]\ \exp\left(\frac{\pi\,\tau}{2\,\tau_{max}}\right)}{\sum_{\tau=-\tau_{max}}^{\tau_{max}-1} \exp\left(\frac{\pi\,\tau}{2\,\tau_{max}}\right)} \right| \quad (19)$$

Leading to:

$$\begin{aligned} \eta_{Rect}(\tau_{max} = 512) &= 0.9988 \\ \eta_{Rect}(\tau_{max} = 1024) &= 0.9976 \\ \eta_{Rect}(\tau_{max} = 2048) &= 0.9952 \end{aligned} \quad (20)$$

Those values are slightly worse than what obtained if blanking. However, this is not the case if other spectral windows are used. For instance, using a Hanning window results into:

$$\begin{aligned} \eta_{Hann}(\tau_{max} = 512) &= 0.9994 \\ \eta_{Hann}(\tau_{max} = 1024) &= 0.9989 \\ \eta_{Hann}(\tau_{max} = 2048) &= 0.9976 \end{aligned} \quad (21)$$

for a spectral line centered in a channel, and

$$\begin{aligned} \eta_{Hann}(\tau_{max} = 512) &= 0.9993 \\ \eta_{Hann}(\tau_{max} = 1024) &= 0.9986 \\ \eta_{Hann}(\tau_{max} = 2048) &= 0.9973 \end{aligned} \quad (22)$$

if the spectral line locates at a mid-point.

4.2.3 To blank or not to blank

Taking into account the foregoing results, and the fact that the maximum of the signal cross correlation function appears at short lags in practice, blanking increases the complexity with no clear benefit. Thus, not blanking appears as the most suitable option.

4.3 1-ms truncation

The last operation degrading the sensitivity in TDM is the truncation of the 1-ms accumulator result. Specifically, the six LSBs of the accumulation are not dumped to the subsequent stages. This introduces additional quantization noise which can be quantified through well-known results, as the standard deviation of the accumulation result is much greater than one [1]. Truncation translates to different sensitivity losses depending on the number of bits of the correlator and the input ADC bits.

4.3.1 4-bit correlator

4.3.1.1 4-bit correlator and a 4-bit ADC

The PDF at the output of the ADC is given by:

$$\mathrm{P}_{ADC}(\mathrm{x}) = \begin{cases} 0.1312 & x = \pm 1 \\ 0.1174 & x = \pm 3 \\ 0.0940 & x = \pm 5 \\ 0.0673 & x = \pm 7 \\ 0.0432 & x = \pm 9 \\ 0.0248 & x = \pm 11 \\ 0.0127 & x = \pm 13 \\ 0.0095 & x = \pm 15 \end{cases} \tag{23}$$

Thus, according to the conversion table in (Missing Ref.), the conversion to 2-bit pairs leads to

$$\mathrm{P}_{MSP}(x) = \begin{cases} 0.4099 & x = \pm 1 \\ 0.0901 & x = \pm 3 \end{cases} \tag{24}$$

For the most significant pair (MSP) of bits, while

$$\mathrm{P}_{LSP}(x) = \begin{cases} 0.2488 & x = \pm 1 \\ 0.2512 & x = \pm 3 \end{cases} \tag{25}$$

For the least significant pair (LSP).

Then, the PDFs of the cross correlation products become (see Missing Ref. for details on correlation product computation):

$$\mathrm{P}_{MM}(x) = \begin{cases} 0.0162 & x = 0 \\ 0.1478 & x = 3 \\ 0.3360 & x = 4 \\ 0.3360 & x = 5 \\ 0.1478 & x = 6 \\ 0.0162 & x = 9 \end{cases} \tag{26}$$

for the products of MSPs,

$$\mathrm{P}_{ML}(x) = \mathrm{P}_{LM}(x) = \begin{cases} 0.0453 & x = 0 \\ 0.2508 & x = 3 \\ 0.2040 & x = 4 \\ 0.2040 & x = 5 \\ 0.2508 & x = 6 \\ 0.0453 & x = 9 \end{cases} \tag{27}$$

for the cross-products between MSPs and LSPs, and

$$\mathrm{P}_{LL}(x) = \begin{cases} 0.1262 & x = 0 \\ 0.2500 & x = 3 \\ 0.1238 & x = 4 \\ 0.1238 & x = 5 \\ 0.2500 & x = 6 \\ 0.1262 & x = 9 \end{cases} \tag{28}$$

for the products involving LSPs only. These distributions exhibit the following means and variances:

$$\begin{gathered} \mu_{MM} = \mu_{ML} = \mu_{LM} = \mu_{LL} = 4.500 \\ \sigma^2_{MM} = 1.491\,, \qquad \sigma^2_{ML} = \sigma^2_{LM} = 3.064\,, \qquad \sigma^2_{LL} = 6.297 \end{gathered} \tag{29}$$

Assuming that the 250,000 correlation products accumulated each millisecond are uncorrelated[5] the mean and standard deviation of the accumulator results will be:

$$\begin{gathered} \mu_{A,MM} = \mu_{A,ML} = \mu_{A,LM} = \mu_{A,LL} = 1{,}125{,}000 \\ \sigma_{A,MM} = 610\,, \qquad \sigma_{A,ML} = \sigma_{A,LM} = 875\,, \qquad \sigma_{A,LL} = 1{,}254 \end{gathered} \tag{30}$$

which is much greater than the quantization step (unity) of the accumulator. Moreover, the PDF in all cases is approximately Gaussian due to the central limit theorem. If the results are not truncated, the final standard deviation becomes proportional to:

$$\sigma_{NT} \propto \sqrt{256\,\sigma^2_{A,MM} + 32\,\sigma^2_{A,ML} + \sigma^2_{A,LL}} = 11{,}023 \tag{31}$$

On the other hand, truncation increases the noise power by adding independent, uniformly distributed, quantization noise with standard deviation

$$\sigma_Q = \frac{\Delta}{\sqrt{12}} \tag{32}$$

where Δ is the quantization step after truncation. Truncation also biases the correlation estimate, but the bias should not affect sensitivity as it biases equally the probability distribution for uncorrelated inputs.

The final standard deviation of the truncated 1-ms accumulation becomes:

$$\sigma_T \propto \sqrt{256\,\sigma^2_{A,MM} + 32\,\sigma^2_{A,ML} + \sigma^2_{A,LL} + 289\frac{\Delta^2}{12}} = \sqrt{\sigma^2_{NT} + 289\frac{\Delta^2}{12}} \tag{33}$$

Thus, the sensitivity loss results into:

$$\eta(\Delta) = \sqrt{\frac{\sigma^2_{NT}}{\sigma^2_{NT} + 289\frac{\Delta^2}{12}}} = \frac{1}{\sqrt{1 + 289\frac{\Delta^2}{12\,\sigma^2_{NT}}}} \approx \frac{1}{\sqrt{1 + \left(\frac{\Delta}{2246}\right)^2}} \tag{34}$$

[5] Their correlation coefficient is approximately proportional to the squared SNR.

which for truncation of the six LSBs of the accumulation result becomes:

$$\eta(\Delta = 64) = 0.9996 \tag{35}$$

4.3.1.2 4-bit correlator and a 3-bit ADC

The analysis of the sensitivity loss when a 3-bit ADC is used can be carried out in a fashion completely analogous to the previous section. However, there is the question of how to fit these three bits into the four available bits.

The most immediate answer is to use the LSBs. However, this solution makes the variance of the MSPs too small, which also reduces the dominant terms of σ_{NT}. Thus, the correlator becomes more sensitive to the quantization noise introduced by truncation. Numerically, it can be obtained that

$$\sigma_{A,MM} = 250\,, \qquad \sigma_{A,ML} = \sigma_{A,LM} = 568\,, \qquad \sigma_{A,LL} = 1{,}292 \tag{36}$$

which leads to a sensitivity loss

$$\eta(\Delta = 64) = 0.9982 \tag{37}$$

worse than when a 4-bit ADC is used.

Alternatively, the 3-bit ADC word can be completed on the right with a zero (or a one). However, this will shift the bias introduced by the accumulator and needs to be taken into account by the software. Specifically:

$$\mu_{A,LL} = 1{,}225{,}000 \tag{38}$$

That should not be problematic, and in exchange the sensitivity improves, obtaining

$$\sigma_{A,MM} = 732\,, \qquad \sigma_{A,ML} = \sigma_{A,LM} = 957\,, \qquad \sigma_{A,LL} = 1{,}225 \tag{39}$$

and

$$\eta(\Delta = 64) = 0.9997 \tag{40}$$

4.3.2 2-bit correlator

4.3.2.1 2-bit correlator and a 4-bit ADC

Let’s first assume that the ADC data truncation is done following the recommendation in Section 4.1.1. Then, the probability of the truncated ADC data becomes:

$$\mathrm{P}_{ADC}(\mathrm{x}) = \begin{cases} 0.3426 & x = \pm 1 \\ 0.1574 & x = \pm 3 \end{cases} \tag{41}$$

which leads to a mean and standard deviation of the accumulator

$$\mu_A = 1{,}125{,}000\,, \qquad \sigma_A = 880 \tag{42}$$

Since only one accumulator is used to estimate each lag of the correlation function, $\sigma_{NT} = \sigma_A$, and the sensitivity efficiency of the truncation for the 2-bit correlator results:

$$\eta(\Delta) = \frac{\sigma_{NT}}{\sigma_T} = \sqrt{\frac{\sigma_{NT}^2}{\sigma_{NT}^2 + \frac{\Delta^2}{12}}} = \frac{1}{\sqrt{1 + \frac{\Delta^2}{12\,\sigma_{NT}^2}}} \approx \frac{1}{\sqrt{1 + \left(\frac{\Delta}{3048}\right)^2}} \tag{43}$$

which for 6-bit truncation yields:

$$\eta(\Delta = 64) = 0.9998 \tag{44}$$

On the other hand, if the ADC data are directly truncated, then we obtain the same results as for the accumulator of the most significant pair in the previous Section 4.3.1.1 (4-bit ADC and correlator), i.e.:

$$\mu_A = 1{,}125{,}000\,, \qquad \sigma_A = 610 \tag{45}$$

resulting into:

$$\eta(\Delta) \approx \frac{1}{\sqrt{1 + \left(\frac{\Delta}{2115}\right)^2}} \tag{46}$$

which for 6-bit truncation yields:

$$\eta(\Delta = 64) = 0.9995 \tag{47}$$

4.3.2.2 2-bit correlator and a 3-bit ADC

Analogously to the foregoing discussion, direct truncation of the 3-bit ADC data leads to the same results as for the most significant pair accumulator of a 4-bit correlator when the 3-bit ADC data are zero-padded on the right. Thus,

$$\mu_A = 1{,}125{,}000\,, \qquad \sigma_A = 732\,, \qquad \eta(\Delta) \approx \frac{1}{\sqrt{1 + \left(\frac{\Delta}{2537}\right)^2}} \tag{48}$$

and

$$\eta(\Delta = 64) = 0.9997 \tag{49}$$

5 Frequency Division Mode

In FDM, the Tunable Filter Banks (TFBs) split the signal from the ADC into up to 32 sub-bands, sampled at 250 MS/s. Each sub-band may contain a 125-MHz wide signal sampled at Nyquist rate, or a 62.5-MHz signal sampled at twice Nyquist. Although each sub-band can be tuned independently, in practice they are configured to cover a continuous band of frequencies, with a relative spectral overlap between sub-bands equal to 1/16 of the bandwidth [2]. This is transparent to the correlator, as each sub-band is processed independently. However, at most 93.75% of the digital bandwidth can be processed ($\eta_{\mathrm{TFB},max} = \sqrt{0.9375} = 0.9682$), which should be taken into account, along with the actual bandwidth of the analog signal, in a fair TDM-FDM sensitivity comparison.

Figure 2 outlines the processing chain (per baseline, per baseband) of the correlator operating in FDM. Compared to TDM, the packetizer is substituted by the TFBs. Throughout the TFB, the bit width increases, so a re-quantization is required for the correlator to be able to process the TFB output. Next, each correlator plane processes all baselines for a given sub-band of a baseband pair (same sky frequencies, two polarizations). The specific sub-band sent to each plane can be the same or different. As in TDM, each 2048-lag core accumulates cross-products for 1 ms, and truncates some LSBs before dumping the results. Full precision is maintained in all subsequent stages. If the accumulator is reset due to TDM blanking implementation, an additional loss is incurred. Moreover, when accumulating cross-correlation products from an oversampled sub-band, the correlation between nearby products increases the noise variance of the result.

Therefore, four operations incur some sensitivity loss:

1. Re-quantization
2. Oversampling
3. 1-ms truncation
4. Blanking

which are studied next. In addition, when TFBs output at twice the Nyquist rate, there is a sensitivity gain which is detailed as well.

5.1 Re-quantization

Assuming a standard deviation of each sub-band much greater than 1 (that is the quantization step previous to the re-quantization), the re-quantization sensitivity loss can be well approximated as if the input were analog. Thus, the optimal quantization level can be used for the re-quantization, and the sensitivity loss follows the well-known results [3]:

$$\eta = 0.9885 \tag{50}$$

for 4-bit quantization, and

$$\eta = 0.8812 \tag{51}$$

for 2-bit quantization.

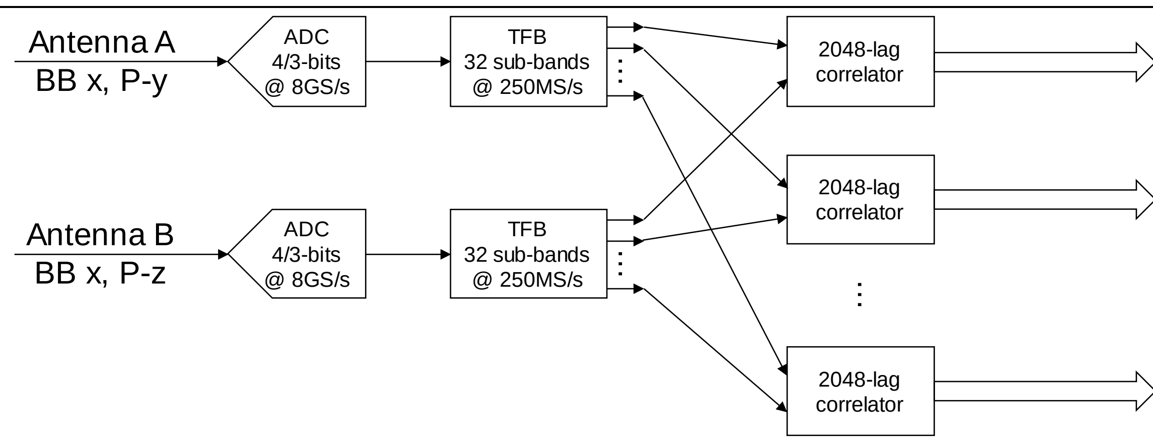


Figure 2- Block diagram of the ALMA correlator upgrade operating in FDM

5.2 Oversampling

The TFBs can be configured to halve their bandwidth while keeping the sampling rate at 250 MS/s. Although decreasing the bandwidth also decreases the sensitivity, it is common practice to refer the sensitivity of the correlator to what a correlator processing the same bandwidth at the Nyquist rate would obtain. In that case, there is a sensitivity gain because twice the samples are accumulated. However, this sensitivity gain is smaller than $\eta = \sqrt{2}$ because of the correlation between accumulated products. In fact, the sensitivity gain can be shown to be approximately equal to [4]:

$$\eta \approx \sqrt{\frac{2}{1 + 2\sum_{\tau=1}^{\infty} R^2(\tau)}} \tag{52}$$

where $R(\tau)$ is the normalized ($R(0) = 1$) autocorrelation function of the oversampled signal. For a non-quantized signal, the correlation is such that no sensitivity gain is obtained (i.e., $\eta = 1$). However, the correlation decreases with quantization and some gain is obtained from oversampling. The sensitivity gain when sampling at twice the Nyquist rate after a 4-bit re-quantization results[6]:

$$\eta \approx 1.0060 \tag{53}$$

while for 2-bit re-quantization is [4]:

$$\eta \approx 1.0611 \tag{54}$$

Note that the foregoing computations are based on a perfectly rectangular sub-band filter. In practice, the sensitivity loss due to the filter roll-off must be attributed to the TFBs, while computer simulations reveal that the oversampling gain is almost insensitive to the filter shape.

5.3 1-ms truncation

5.3.1 4-bit correlator, sub-band sampled at Nyquist rate

See Section 4.3.1.1

5.3.2 4-bit correlator, sub-band sampled at twice the Nyquist rate

As stated in Section 5.2, correlation increases the variance of the accumulated cross-products, thereby diluting the impact of quantization noise due to truncation. The variance increase is given by the denominator in (52) and, from Section 4.3.1.1, the bulk of the variance is given by the accumulator of the most significant pair of bits. As a result, the standard deviation of the 1-ms accumulation of the MSP becomes[7]:

$$\sigma_{A,MM} = 610\sqrt{1 + 2\sum_{\tau=1}^{\infty} R^2(\tau)} \approx 610 \times 1.3181 = 804 \tag{55}$$

[6] Obtained from (52) through computer simulation.

[7] Note that any relationship with the value in (54) vanishes as the quantization step is now four times the optimal for 4-bit quantization, but not the optimal for 2-bit quantization as in (54).

which yields:

$$\eta(\Delta = 64) = 0.9997 \tag{56}$$

5.3.3 2-bit correlator, sub-band sampled at the Nyquist rate

The sensitivity loss can be computed in the same fashion as in Section 4.3.2. In this case, the PDF of the 2-bit re-quantized signal is

$$\mathrm{P}_{reQ}(\mathrm{x}) = \begin{cases} 0.3401 & x = \pm 1 \\ 0.1599 & x = \pm 3 \end{cases} \tag{57}$$

Then, the mean and standard deviation of the 1-ms accumulation become

$$\mu_A = 1{,}125{,}000\,, \qquad \sigma_A = 889 \tag{58}$$

And the truncation efficiency:

$$\eta(\Delta) \approx \frac{1}{\sqrt{1 + \left(\frac{\Delta}{3081}\right)^2}} \tag{59}$$

which for 6-bit truncation results into:

$$\eta(\Delta = 64) = 0.9998 \tag{60}$$

5.3.4 2-bit correlator, sub-band sampled at twice the Nyquist rate

Analogously to Section 5.3.2,[8]

$$\sigma_A = 889\sqrt{1 + 2\sum_{\tau=1}^{\infty} R^2(\tau)} \approx 889 \times 1.3328 = 1{,}185 \tag{61}$$

where the standard deviation when sampling at the Nyquist rate was computed in (60). As a result, the incurred sensitivity loss is:

$$\eta(\Delta = 64) = 0.9999 \tag{62}$$

5.4 Blanking

Only if TDM blanking is implemented. See Section 4.2.1.

[8] With the difference that now the quantization step is the optimal for 2-bit quantization.

6 Truncation in the Final Adders

A second point of interest is to study how many bits could be truncated prior to the transmission of results to the CDPs, so that the bandwidth requirements are minimized. Assuming no correlation between the 1-ms results[9], the variance of the FA result is:

$$\sigma_{FA}^2 = M\,\sigma_{ms}^2 \tag{63}$$

where M represents how many 1-ms results have been integrated to obtained the FA result, and σ_{ms}^2 stands for the variance after integrating for 1-ms and truncating the 6 LSBs of the result. Then, the efficiency of the truncation becomes:

$$\eta(b) = \frac{1}{\sqrt{1 + \frac{4^b}{12\,\sigma_{FA}^2}}} \tag{64}$$

where b is the number of LSBs discarded from the final result. From the foregoing equations, it stems that the minimum sensitivity loss is achieved when the truncation is performed at the end [1].

To get an approximate number for the bits that could be truncated, Figure 3 plots the sensitivity loss incurred by truncation as a function of the integration time. The longer the integration time, the bigger the noise variance. Therefore, quantization loss becomes less important and more bits can be truncated for a given sensitive loss. However, the number of bits required to avoid overflow also increases. In fact, it increases faster than the number of bits that can be truncated, making the number of bits of the result increase with the integration time. For example, for an integration time of 1 second, 6 LSBs can be truncated

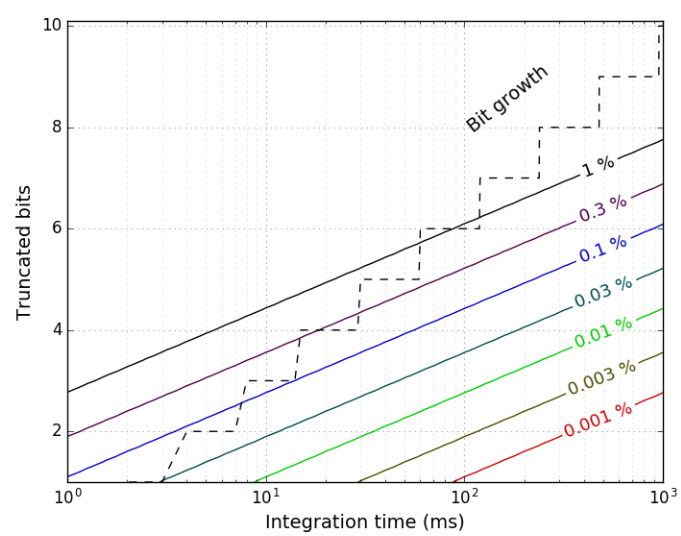


Figure 3- Sensitivity loss $(1-\eta)$ as a function of the integration time and the number of truncated bits. Dashed line indicates the bit growth as a function of the integration time. FDM, Nyquist sampling and 2-bit correlator.

[9] Even at twice the Nyquist rate, the correlation between samples vanishes in a time frame much shorter than 1 ms.

incurring a sensitivity loss less than 0.1%. However, the bit growth is 10 bits, so the final result needs 4 additional bits (for a total of 20 bits).

Due to implementation requirements, the number of bits transmitted to the CDPs per final result should be an integer number of bytes, and the same for all observation modes. Therefore, another conclusion from Figure 3 is that targeting for a final result word size of 2-bytes (16-bits), which requires truncating as many bits as the bit growth, will incur excessive sensitivity loss for the longest integration times (more than 10% at 1000 ms). Therefore, a study focusing on computing the sensitivity loss for 3 bytes has been carried out, the results thereof can be found in Appendix 1.

6.1 Main conclusions from the study

1. FDM:
    a. A maximum sensitivity loss less than 0.001% is achieved if the number of truncated bits depends on the total integration time, and whether a 2 or 4-bit mode.
    b. Truncating a constant number of bits (2 bits) would yield a maximum sensitivity loss less than 0.6% (including sub 16-ms modes) if:
        i. 4-bit scaling is such that 4 additional bits are truncated in 4-bit modes (i.e., 1×, 1/4×, and 1/16×).
2. TDM:
    a. A maximum sensitivity loss less than 0.02% is achieved if the number of truncated bits depends on the total integration time, and whether a 2 or 4-bit mode.
    b. Truncating a constant number of bits (2 bits) would yield a maximum sensitivity loss less than 0.6% (including sub 16-ms modes) if:
        i. 4-bit scaling is such that 4 additional bits are truncated in 4-bit modes (i.e., 1×, 1/4×, and 1/16×).
        ii. Sum of different correlation planes is scaled by 1/16 (4 additional bits are truncated)

7 References

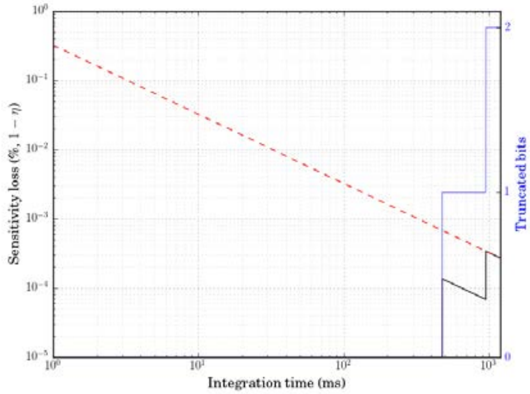
[1] O. A. Yeste Ojeda, "PMD-365-036-A-REP. Analysis of the Sensitivity Degradation in a Digital FXF Correlator," 2016.

[2] ALMA Partnership, 2016, S. Asayama, A. Biggs, I. de Gregorio, B. Dent, J. Di Francesco, E. Fomalont, A. Hales, E. Humphries, S. Kameno, E. Müller, B. Vila Vilaro, E. Villard, F. Stoehr, ALMA Cycle 4 Technical Handbook, 2016.

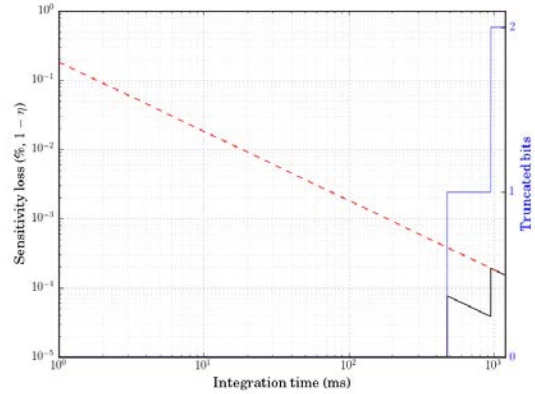
[3] A. R. Thompson, D. T. Emerson and F. R. Schwab, "Convenient formulas for quantization efficiency," *Radio Science,* vol. 42, 2007.

[4] A. R. Thompson, J. M. Moran and G. W. J. Swenson, Interferometry and Synthesis in Radio Astronomy, 2nd ed., New York: John Wiley and Sons, 2001.

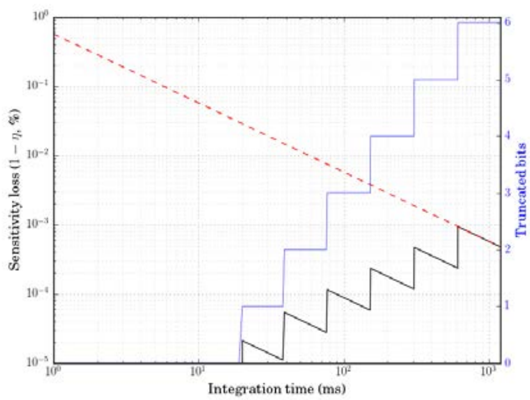
Appendix 1 Sensitivity loss results for 3-byte final truncation



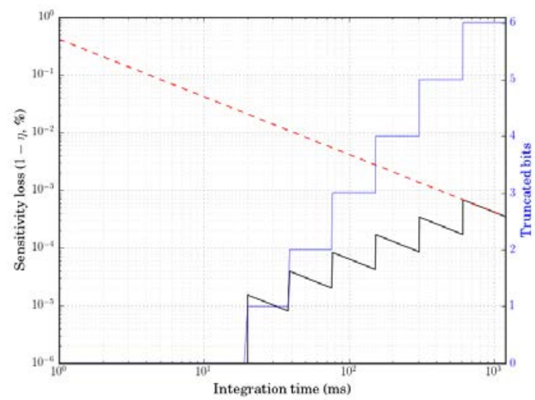
FDM, 2 bits, Nyquist



FDM, 2 bits, twice Nyquist

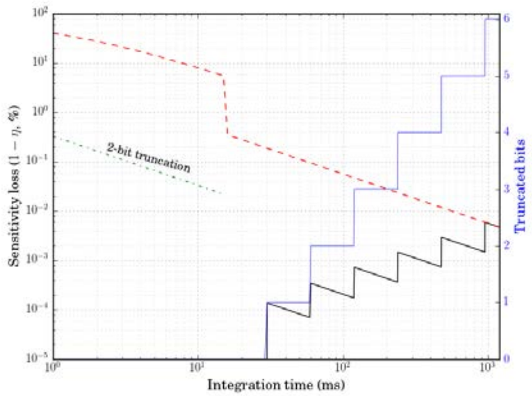


FDM, 4 bits, Nyquist

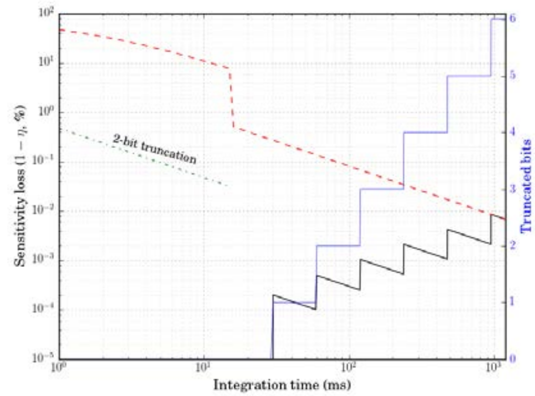


FDM, 4 bits, twice Nyquist

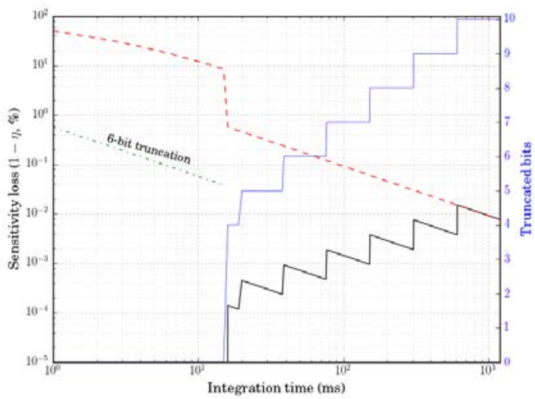
Figure 4- (Black thick line) Sensitivity loss $(1-\eta)$ as a function of the integration time for different FDM modes when only the LSBs in excess of 24 bits are truncated.. (Red dashed line) Sensitivity loss if the number of bits truncated is set regardless of the integration time. (Blue thin line) Number of bits in excess of 3 bytes.



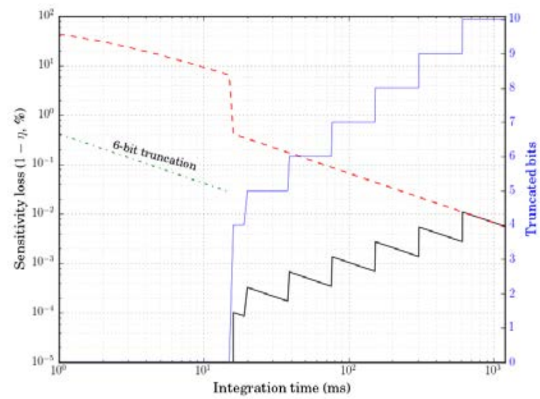
TDM, 2 bits, 4-bit ADC (as Section 4.1.1.)



TDM, 2 bits, 3-bit ADC

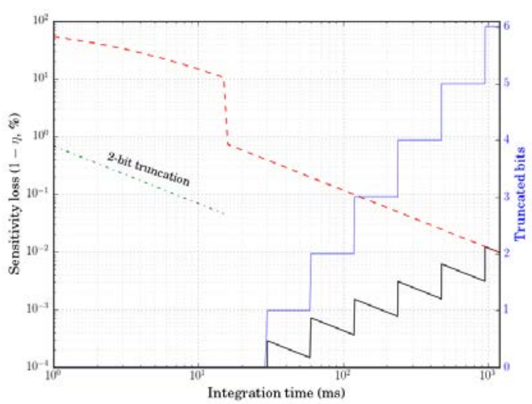


TDM, 4 bits, 4-bit ADC

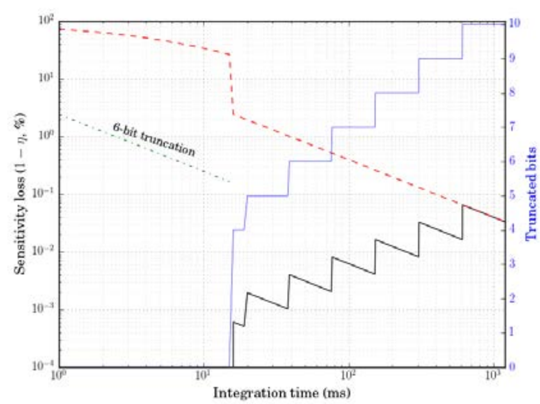


TDM, 4 bits, 3-bit ADC (zero-padded)

Figure 5- (Black thick line) Sensitivity loss ($1-\eta$) as a function of the integration time for different TDM modes when only the LSBs in excess of 24 bits are truncated.. (Red dashed line) Sensitivity loss if the number of bits truncated is set regardless of the integration time. Note the loss increase for sub 16-ms modes because different correlator planes do not accumulate to each other. (Blue thin line) Number of bits in excess of 3 bytes. Other TFB's pass-through implementations shown in Figure 6.



TDM, 2 bits, 4-bit ADC, Direct truncation



TDM, 4 bits, 3-bit ADC, no zero-padding

Figure 6- As Figure 5 but using other TFB's pass-through implementations.