## Synthetic VLBI Experiment

Fringe stopped data from a DiFX VLBI run was used for the analysis below. A 1.6 GHz VLBI setup contained the short 330km EF-WB baseline and correlated synthetic 2-bit Mark5B files from both stations. A GPS satellite carrier tone with amplitude 1.0 was placed at 6 MHz of the 16 MHz wide bands. It was optionally BPSK modulated by a Coarse Acquisition (CA) pseudo-random chip sequence, to see the effect of modulation on RFI filter performance. No Doppler shift or geometric delay was applied; the GPS satellite is at the baseline zenith and co-rotates with the Earth and baseline. Uncorrelated Gaussian noise with sigma=1.0 was added to each station. The final signals were 2-bit quantized and written to Mark5B files. DiFX correlated the synthetic VLBI experiment using 128 channels and an empty phase center at RA=0h DEC=+45d with ~80d elevation at the stations. Raw DiFX floating point data after fringe stopping, still before time averaging, was written to very large files.

The new filter infrastructure in DiFX was verified by loading 2.094 seconds (262144 samples) of raw fringe-stopped data of one channelized 16 MHz band into Matlab. The sampling rate was 125 kHz in each of the 128 DiFX channels.

Two filters were used for the CA-modulated data. The first is labeled “Time-integrated” or “Mean-filtered” in the figures and is a cumulative mean filter function or normalized integrator commonly used in VLBI correlators. The second filter is a 12th order Chebyshev Type I low-pass IIR filter with either 10 Hz or 16 Hz cut-off. It is labeled “Low-pass filtered” in the figures below. When this filter is cascaded with a second stage, in this case a cumulative mean filter, the resulting plots carry a “Low-pass filtered, integrated” label.

The Figure 5 ripple pattern in absence of low-pass mean filtering allows the following theories. DiFX FITS-IDI files contain visibilities sampled at certain integration times, say 2.0 seconds. Visibilities picked at these intervals from the output of the usual mean filter will sample the sinc(x) pattern. These visibilities would thus have high variance. On the other hand, visibilities picked from a low-pass filtered time series have a much lower variance. A side by side comparison of DiFX visibilities with and without filtering is complicated by the ripple.

In the interferometric UV plane (or its sparse spatial 2D covariance matrix) the mean filtered data samples would exhibit a similar ripple pattern. In the final UV image this causes characteristic ringing. The same image ringing should be absent, and RFI mitigated, in all low-pass mean filtered UV plane data. This still remains to be verified. For spectral line data sets the impact of visibilities with reduced RFI variance is still unclear.

# Figures for GPS Signal with CA Modulation

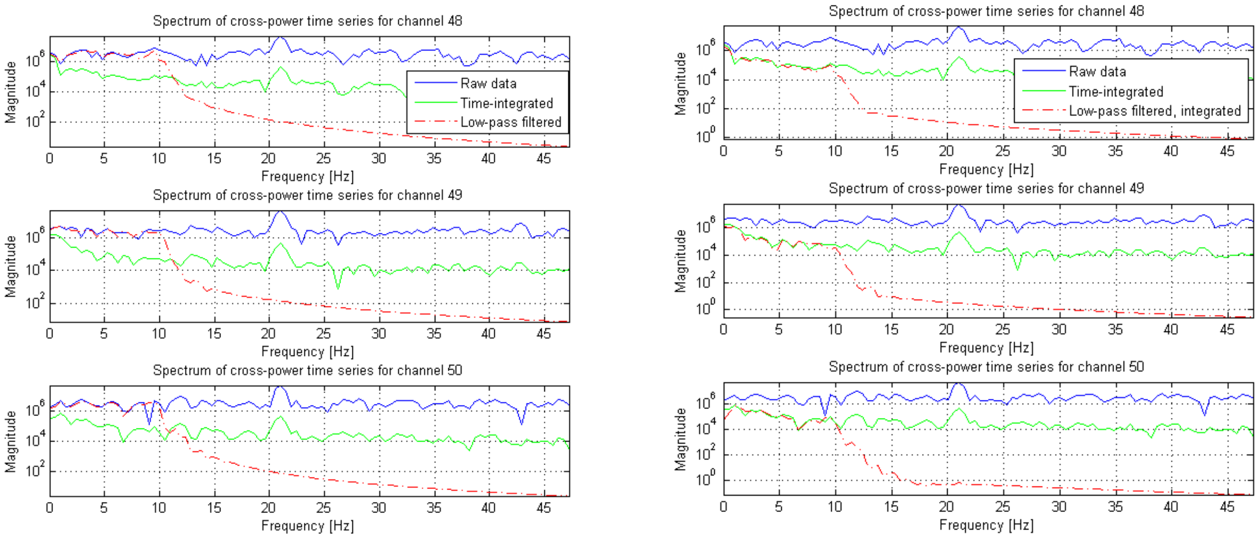

**Figure 1 – Fringe frequency spectra zoomed to 0-45 Hz for channels 48 to 50. Spectra are Hanning windowed Fourier transforms of the cross-correlation time series of a channel. 1 MHz BPSK modulation by the GPS CA chip sequence was enabled. Modulation spreads the GPS carrier in channel 49 across neighboring channels. Spectra of raw cross-correlation data (solid blue) show GPS at a 22 Hz fringe frequency. Mean filtering reduces GPS levels in the time series by ~20 dB (solid green). A 10 Hz low-pass filter alone suppresses GPS by ~50 dB (left panel, dashed red), but if followed by a mean filter, GPS is suppressed ~70 dB over the original (right panel, dashed red).**

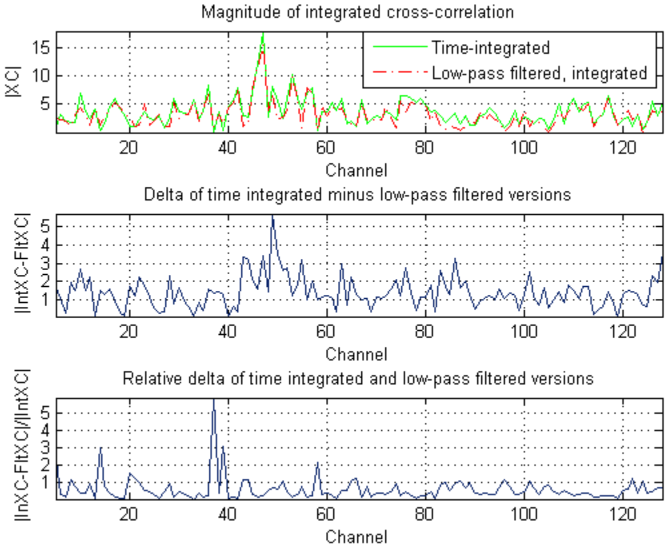

**Figure 2 – Cross-correlations in all 128 channels after T=2.094 seconds, the final baseline visibility. Although Figure 1 demonstrates low-pass mean filtering with 50 dB improved GPS suppression compared to mean filtering alone, the final visibility still contains significant amounts of GPS (dashed red, top). There is almost no reduction in level compared to mean filtering (solid green, top). The absolute difference |int(xc) – flt(xc)| is sometimes larger because low-pass mean and mean results have different phase (middle). The final low-pass filtered visibility fails to demonstrate the expected 50 dB decrease in GPS level. The reason is not yet entirely understood.**

# Figures for Unmodulated GPS Carrier Signal

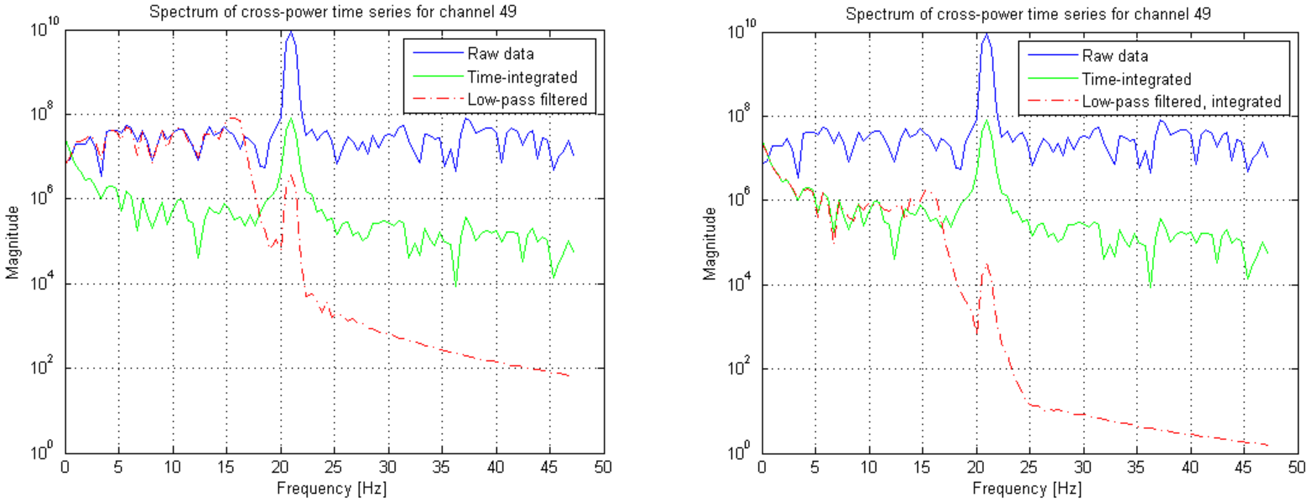

**Figure 3 – Fringe frequency spectra zoomed to 0-45 Hz for channel 49 data with a clean GPS carrier and no modulation. The spectrum of the raw data time series (solid blue) is overlaid with spectra from mean filtered (solid green) and 16 Hz low-pass filtered data (left panel, dashed red). Using 16 Hz low-pass mean filtering increases the GPS suppression from ~30 dB to ~60 dB (dashed red, right).**

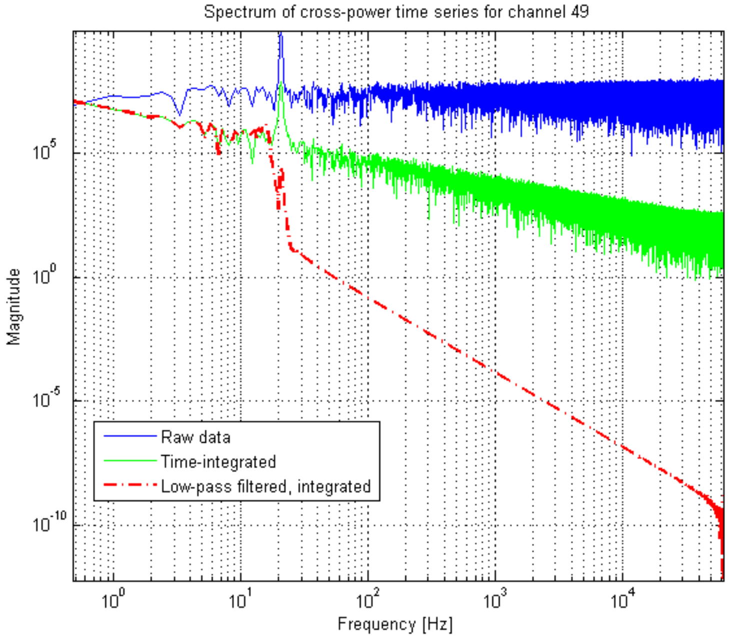

**Figure 4 – Full fringe frequency spectrum 0-62.5 kHz of channel 49 with a clean GPS carrier and no modulation. Spectra are shown for raw unprocessed data (solid blue), mean filtered data (solid green), and 16 Hz low-pass mean filtered data (dashed red).**

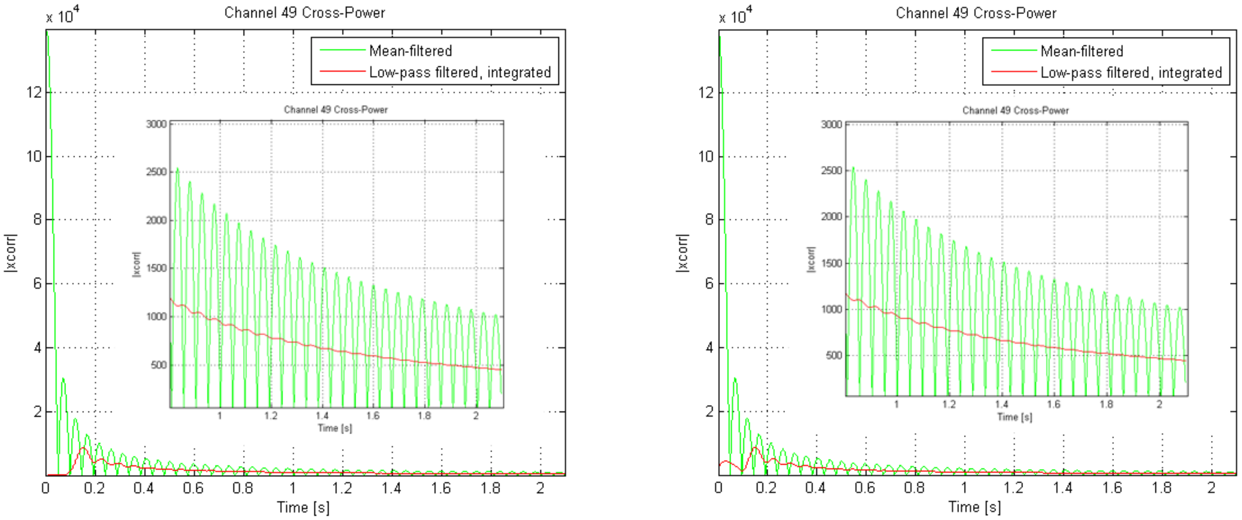

Figure 5 – Absolute value of filtered time series when using uninitialized filters (left panel) and initialized filters (right panel). Channel 49 contains a clean GPS carrier and no modulation. The DiFX mean filtered output exhibits a large sinc(x) ripple pattern due to the poorly attenuated GPS tone (green curve). You may compare this ripple to the 22 Hz fringe frequency in Figure 3. The filter output from a two stage filter that consists of a 16 Hz low-pass followed by a mean filter reduces the ripple by ~60 dB (red). For an uninitialized low-pass the output starts from zero (left panel, red). Before filtering actual data the low-pass section may also be initialized first by Gaussian noise equal to the expected later level. In this case the output starts from the estimated noise mean (right panel, red). In both cases the filter output after 100ms is essentially identical and converges equally fast.

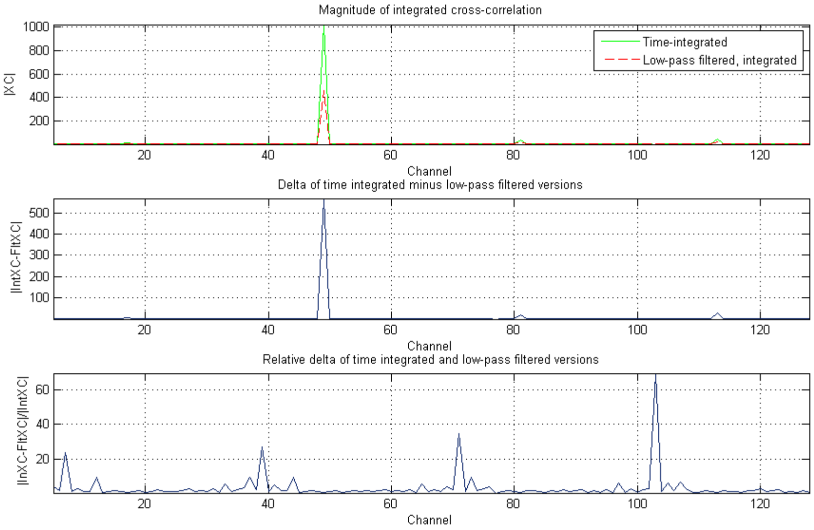

Figure 6 - Cross-correlations in all 128 channels after T=2.094 seconds, the final baseline visibility. No GPS carrier modulation. As in Figure 1 with CA chip modulation, even now the final visibility has more GPS signal than anticipated. For the low-pass mean filtered data the final visibility (dashed red, top) happens to be reduced by factor 2 (3 dB << 40 dB) compared to the mean filtered (solid green, top). This is depends however on the choice of time T. Further, the integrated and filtered complex visibilities have different phase, resulting in a possibly larger absolute difference |int(xc) – flt(xc)| (middle). The normalized absolute difference (bottom) shows certain channels such as channel 104 with noise seemingly enhanced by the filter.

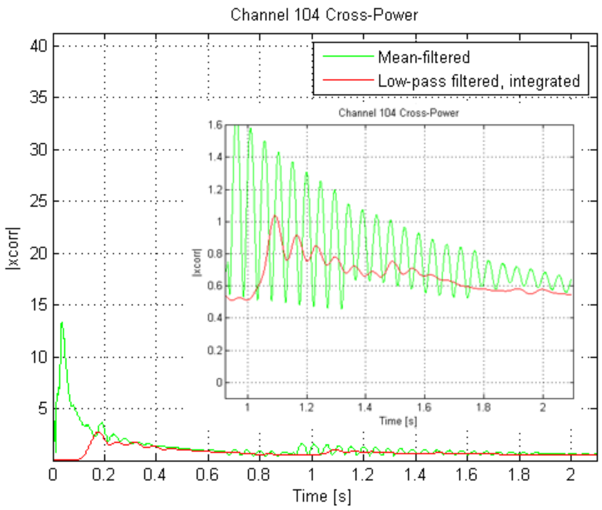

**Figure 7 – Time series of filter output. Channel 104 contains no GPS signal, no added noise, and only the 2-bit quantization noise (Mark5B input file) of the GPS carrier signal. The 16 Hz low-pass mean filtered data is largely noise free (red). The mean filtered data (green) has residual noise from the 16 Hz to 125 kHz signal range.**

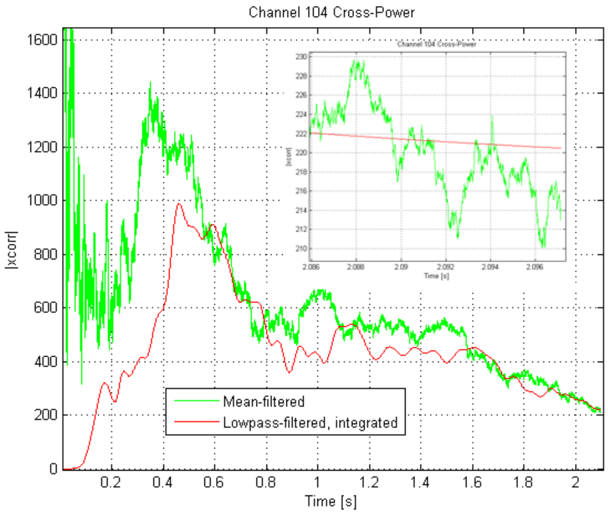

**Figure 8 – Identical to Figure 7 except complex Gaussian noise was added in Matlab. Noise real and imaginary parts had mean 0 and sigma=$10^5$, which after 262144 samples at T=2.096s reduces to sigma=195 for standard deviation of the mean. The low-pass mean filter tracks the plain mean filter with ~100ms delay, a filter characteristic. It does not account for the lower than expected GPS suppression as seen in Figures 2 and 6.**