VLBA Sensitivity Upgrade Memo 34

System Temperature Determination with DiFX

Walter Brisken

January 18, 2011

Abstract In anticipation of the need for an interim source of system temperature measurements for data acquired using the ROACH Digital Back End (RDBE), functionality to determine switched/total power from sampled baseband data has been added to the DiFX software correlator. This approach is shown to work well but results in a computed system temperature that is typically 6.5% *lower* than those derived within the analog VLBA baseband converters. Some possible origins of this discrepancy is discussed, however the observed effect is several times larger than can be easily explained. It is argued that the DiFX derived values are closer to the truth and are likely to be closer to the values to be computed in the RDBE firmware in the future. Some strategies for dealing with the two sources of system temperature are discussed.

1 Introduction

Often system temperatures are measured using electronics at a radio telescope. This has always been the case for the VLBA. The VLBA back-end electronics are amidst a significant upgrade with much of the electronics to be replaced with an FPGA-based digital back-end. The baseband converters, the current source of record channel system temperatures, are included in the equipment that is being replaced. This functionality is to be one of the last features implemented in the new firmware. This further motivated a long standing wish to have the capability for switched power detection implemented in the DiFX correlator. This functionality in DiFX is in no way intended as a replacement for hardware system temperature determination which has enormous value as a real-time performance indicator during observing.

2 Theory of operation

The goal of a switched power measuring device is to calculate two quantities, P_{on} and P_{off}. These quantities are proportional (with arbitrary and possibly time-variable proportionality) to signal power averaged respectively over periods when the injected calibration noise is on or off. The methods of computing these values are necessarily very different in the analog and digital cases. In the subsections that follow the strategy used on the existing VLBA system and that a implemented in DiFX are described. Note that the details of power measurement algorithm to be implemented in the RDBE may vary somewhat from that implemented in DiFX as the RDBE will have much less coarsely quantized data to work with and thus can probably use a more natural algorithm.

2.1 Analog power measurements

In the case of the baseband converter power meters, an unquantized analog signal is available for power measurement using a conventional square-law detector based circuit. In the case of the VLBA, the analog signal is fed into a tunnel diode which produces an output voltage proportional to the input voltage squared. This signal is amplified and digitized using a voltage-to-frequency (V-F) converter. Depending on the state of the switched power, the V-F cycles are counted into an `on` or `off` register. Every second a measurement of P_{on} and P_{off} is made available to the monitor and control system for longer timescale averaging.

2.2 Digital power measurements

The power detection method recently implemented in DiFX makes use of the fact that 2-bit (4-level) quantized baseband data is being processed. 1-bit quantized data contains no amplitude scale information (except possibly in pathological cases where the sampler threshold is offset from zero by a fixed amount!) so is not amenable to the digital power measurements described here. Quantization using 3 or more bits per sample could be supported with this or another algorithm but is not currently supported and won't be discussed further in this document. The algorithm here assumes that the samplers are balanced in the sense that there are equal numbers of positive and negative voltage measurements. See section 5.2 for some investigation into the outcome if this condition is not met. The algorithm further assumes that the threshold between the $-high$ and $-lo$ states and that between the $+lo$ and $+high$ states are symmetric about zero which should imply that there are equal numbers of samples in the $-high$ and $+high$ states and likewise for the $-lo$ and $+lo$ states. If the automatic digital gain control (see VLBA Sensitivity Memo 28 [1]) is working properly there will be a roughly 2:1 ratio in the population of lo and high states and adjustment of digital gain should only be made on 1-second ticks to ensure that any calculation of $(P_{\mathrm{on}}, P_{\mathrm{off}})$ made over a 1 second period is not affected by instrumental change.

Given the conditions assumed above, there exists a simple mapping from fractional population of high states to a proportional power measurement. Naïvely, in analogy to the unquantized case, this would take the form:

$$P \propto (1-f)v_{\mathrm{lo}}^2 + f v_{\mathrm{high}}^2, \tag{1}$$

where f is the fraction of samples in the high state and v_{lo} and v_{high} are respectively the reconstructed voltage values for lo and high states. This obviously breaks down at both extremes of the range of P and is thus not the correct prescription. Instead one can compute the power statistically by predicting the fraction of high state samples as a function of the power. Assuming a Gaussian distribution of voltages[1], the fraction of high states is given by

$$f = 1 - \mathrm{erf}\frac{\sigma}{\sqrt{2P}}, \tag{2}$$

where erf is the error function and σ is the magnitude of the lo-high sample threshold. Inverting this yields the desired power as a function of high-state fraction:

$$P \propto \frac{1}{2\left(\mathrm{erf}^{-1}(1-f)\right)^2}. \tag{3}$$

As long as the high-state fraction can be computed separately for the periods when the injected noise is on and off it is possible to compute the pair $(P_{\mathrm{on}}, P_{\mathrm{off}})$ from $(f_{\mathrm{on}}, f_{\mathrm{off}})$ with trivial application of equation 3 twice. The factor of 2 in the denominator of equation 3 is unncecessary, but makes the output a conveniently close to 1 for optimal state counts.

2.3 Calculation of system temperature

The system temperature, T_{sys}, is computed based on the measured $(P_{\mathrm{on}}, P_{\mathrm{off}})$ values and the magnitude, in temperature units, of the injected noise source, T_{cal}, which is carefully determined in the laboratory separately for each receiver as a function of frequency. If true system temperature is proportional to the measured time-averaged power $\bar{P} = \frac{1}{2}(P_{\mathrm{on}} + P_{\mathrm{off}})$,

$$T_{\mathrm{sys}} = \alpha \bar{P}, \tag{4}$$

where α is a proportionality factor, and the same proportionality relates the magnitude of the true noise power to the measured switched power:

$$T_{\mathrm{cal}} = \alpha \left(P_{\mathrm{on}} - P_{\mathrm{off}}\right), \tag{5}$$

then the system temperature can be measured as:

$$T_{\mathrm{sys}} = \frac{\bar{P}}{P_{\mathrm{on}} - P_{\mathrm{off}}} T_{\mathrm{cal}}. \tag{6}$$

It is useful to note that that the absolute scaling of the various P values is completely divided out in the final result and are thus completely unimportant as long as P_{on} and P_{off} are treated equally.

[1] This Gaussian-distributed voltage assumption is also made in the digital correlator (e.g., van Vleck) corrections.

3 Implementation

The implementation of system temperature calculation in DiFX is spread across several of the source code modules. Below are brief discussions of changes to each. These changes have all been made to the current *trunk* of the source code respository and will become part of the DiFX version 2.0.1 which should be announced in February 2011.

3.1 mark5access

The library mark5access is used within DiFX to decode the packed 1-bit and 2-bit sampled data from disk files containing baseband data in one of the following formats: VLBA, MarkIV, Mark5B and simple (single thread) variants of VDIF. Before decoding is done within mark5access a structure is initialized with enough information to select one of many (over 200!) different decoding routines. Each case of different fanout, number of channels, number of bits per sample is most optimally decoded with a hand-optimized routine. The decoding algorithm references this data structure and populates a series of arrays (one per channel) with the requested number of time samples.

A parallel infrastructure was set up for efficiently counting (without first decoding) the number of high states per baseband channel within the requested number of samples to be checked. Like the decoders a separate counting algorithm is chosen per format upon initialization. The new function to call to decode this information is `mark5_stream_count_high_states(...)`. This function returns the number of successfully investigated samples, which must be the same for each channel. Thus even in the case of a partial counting (due to data replacement headers, replaced data or other causes of blanking) a correct fraction of high state samples can be determined. Note that using the same `mark5_stream` for both decoding and high state counting will not work without an unpleasant workaround as each will increment the current read pointer with each operation, possibly in a non-reversible manner. An addition useful function `high_state_fraction_to_power(...)` implements equation 3. Full function prototypes for these functions can be found in file `mark5_stream.h` .

3.2 mpifxcorr

Within the heart of the DiFX correlator (a program called `mpifxcorr`) most of the changes to support system temperature extraction were placed in a new source file, `switchedpower.cpp`, and its header file, `switchedpower.h` . These files define a c++ class that maintains the necessary state information and exposes a very simple interface to the rest of `mpifxcorr`. A constructor `SwitchedPower(...)` is used to initialize switched power detection. Function `feed(...)` is passed a `mark5_stream` structure, which itself points to the in-core baseband data, to cause all of the samples to be appropriately counted. This feed function steps through the baseband data in chunks of time $\frac{1}{2\nu}$ long where ν is the switching frequency of the injected noise source, alternately assigning the statistics to P_{on} and P_{off}[2]. Function `flush()`, which is typically called once every 1 second, causes the accumulated statistics to be written do disk file. Finally a destructor, `~SwitchedPower()`, is automatically called for cleanup. These functions are called within the Mk5DataStream and NativeMk5DataStream classes. Formats not supported by mark5access currently cannot generate these power measurements.

The switched power detection is enabled on a per-datastream basis based on the optional `.input` file parameter `TCAL FREQUENCY` (which, if used, must be found immediately after the `FILTERBANK USED` parameter). This parameter expects an integer to assign to ν. If not specified, or if set to 0, switched power detection is disabled for the datastream in question.

An output file is made for each datastream with enabled switched power detection. These files are placed inside the `.difx/` output directory created to contain the output visibility data for the job in question. The names of these files are `SWITCHEDPOWER_`x, where x is the datastream id, which can range from 0 to $n_{\text{datastream}} - 1$. The ASCII text files have a simple format consisting of an indeterminate number of lines of text each with the following format. The first and second numbers are the start and stop epochs of

[2] It is silently assumed that the switching cycle has a 50% duty cycle with a transition to the on state on each 1 second tick, which further implicitly assumes that ν is an integer number of Hz.

accumulation in modified Julian day. For each baseband channel of the datastream another four numbers are listed. These numbers are $P_{\rm on}$, $\delta P_{\rm on}$, $P_{\rm off}$, and $\delta P_{\rm on}$, where the δ's are the statistical uncertainty of the measurements of their associated value. If insufficient statistics have been accumulated for robust calculation of either $P_{\rm on}$ or $P_{\rm off}$ these four numbers will be all zeros.

3.3 vex2difx

The program `vex2difx` is responsible for taking an VLBI experiment description file (`.vex`) and creating one or more sets of files used to drive the DiFX correlator. Currently the vex format [2] doesn't have a slot to describe switched power injection. An integer parameter, `tcalFreq`, has been added to the `ANTENNA` block. By default its value is zero. If set to a number greater than zero, switched power detection will be enabled with a switching frequency in Hz given by this parameter.

3.4 difx2fits

Program `difx2fits` is responsible for taking the output of DiFX and converting to `FITS-IDI` format. It is within this program that conversion from $(P_{\rm on}, P_{\rm off})$ to $T_{\rm sys}$ is done. One new command line parameter was added to `difx2fits` for this purpose: `--difx-tsys-interval` (or shortcut `-i`) followed by a number in seconds will allow an averaging time to be set; its default is 60 seconds. Unless disabled by setting the averaging time to zero (or less), any DiFX-derived swiched power measurements will override any system temperature measurements, presumably determined from on-line measurements, stored in the `tsys` file. The steps in the process are as follows:

1. Subdivide scans into reasonable sized segments. This is performed separately for each antenna in the array. The effective scan time range is determined by the time when the first swiched power measurement that is not explicitly flagged by `vex2difx`[3] is made until the end of the scan as determined originally by the `.vex` file. This time range is divided by the $T_{\rm sys}$ averaging interval and rounded to the nearest non-zero integer to determine the number of equal length segments.

2. $P_{\rm on}$ and $P_{\rm off}$ are each independently averaged, in a weighted sense using the uncertainty contained within the `SWITCHEDPOWER_`x file, to determine segment averages.

3. A nominal $T_{\rm sys}$ is formed using equation 6. See below for a discussion of the origin of $T_{\rm cal}$.

4. An emperically determined fudge factor is applied for VLBA antennas. See section 5 for details on this.

3.4.1 $T_{\rm cal}$

A text file pointed to by environment variable `TCAL_FILE` contains a table of $T_{\rm cal}$ values. The columns of this file are:

1. Antenna name

2. Receiver name

3. Frequency of measurement (MHz)

4. $T_{\rm cal}$ for RCP[4] (K)

5. $T_{\rm cal}$ for LCP (K)

The center frequency of each recorded channel is used as the reference in a linear interpolation of the tabulated values.

[3] Currently, on-line generated flags are ignored in this process.

[4] Note that the VLBA calibration files stored in `/home/jansky3/vlbaops/TCAL` have the LCP column first.

4 Performance

The switched power detection is implemented in the datastream process. In the VLBA's deployment of DiFX the datastream processes run on Mark5 units which are not specifically designed to have significant computing capabilities. In cases where the correlation bottleneck is the actual cross correlation calculations the extra burden on the datastream processes is not noticed. However, for jobs involving a small number of antennas or those with a very simple correlation, the switched power calculations can dominate the overall processing time.

The Mark5 units installed at the VLBA correlator have motherboards that span a range of processing speeds ranging from 2 GHz single-core to 2.5 GHz quad-core configurations. All of the Mark5 units can saturate their 1 Gbps network when no system temperature extraction is enabled. When their network interfaces are saturated correlation proceeds at about 880 Mbps per Mark5 unit[5]. Enabling switched power measurement resulted in fairly substantial data rate drops to between 470 Mbps and 570 Mbps depending on the Mark5 unit.

This rather serious performance loss may have algorithmic solutions that could involve making use of multiple concurrent threads. A simpler solution, one especially valid for the wide-band observations this capability is targetting, is to compute the system temperature on a subset of the samples. Using only half the samples resulted in throughputs of 835 Mbps for Mark5 units with intermediate processing capabilities. Reducing to one quarter the number of samples brings the rate back up to the level where the network is again the bottleneck. Sufficient accuracy would be attained even if only 1/16 of the samples were inspected.

The above tests were all performed on VLBA formatted data; Mark5B data is slightly simpler to decode and single-thread VDIF data may ultimately allow much faster detemination of the high-state statistics as 4 samples of the same channel would be contained in each byte of data.

5 Comparison with on-line measurements

Overall the DiFX-derived system temperatures agree very well with the determinations made in the VLBA baseband converters, but only after applying a 6.5% correction factor. It appears that there is some antenna-to-antenna, and even channel-to-channel within an antenna, variation in the needed correction factor, but to first order a single factor seems to work fairly well. Figure 1 shows one set of comparisons, chosen because of the fairly wide range of system temperatures experienced over a short amount of time. This comparison is representative of other comparisons even at different frequencies and on different antennas. Some possible origins of the 6.5% differences are discussed below.

5.1 Effect of leaky power measurements

If some fraction, ϵ, of the measurement of $P_{\rm on}$ were erroneously assigned to $P_{\rm off}$, and vice versa,

$$P'_{\rm on} = (1-\epsilon)P_{\rm on} + \epsilon P_{\rm off} \quad (7)$$

$$P'_{\rm off} = (1-\epsilon)P_{\rm off} + \epsilon P_{\rm on} \quad (8)$$

then the measured system temperature, $T'_{\rm sys}$, will be in error:

$$T'_{\rm sys} = \frac{T_{\rm cal}}{2}\frac{P'_{\rm on}+P'_{\rm off}}{P'_{\rm on}-P'_{\rm off}} = \frac{T_{\rm cal}}{2}\frac{P_{\rm on}+P_{\rm off}}{(1-2\epsilon)(P_{\rm on}-P_{\rm off})} = \frac{1}{1-2\epsilon}T_{\rm sys}. \quad (9)$$

The net effect will be an increase in reported system temperature by approximately 2ϵ. Some possible causes of the observed 6.5% discrepancy are described in the sections below.

[5] An upgrade to use the already installed 40 Gbps Infiniband system will occur shortly, which will eliminate this as a bandwidth bottleneck for the foreseeable future.

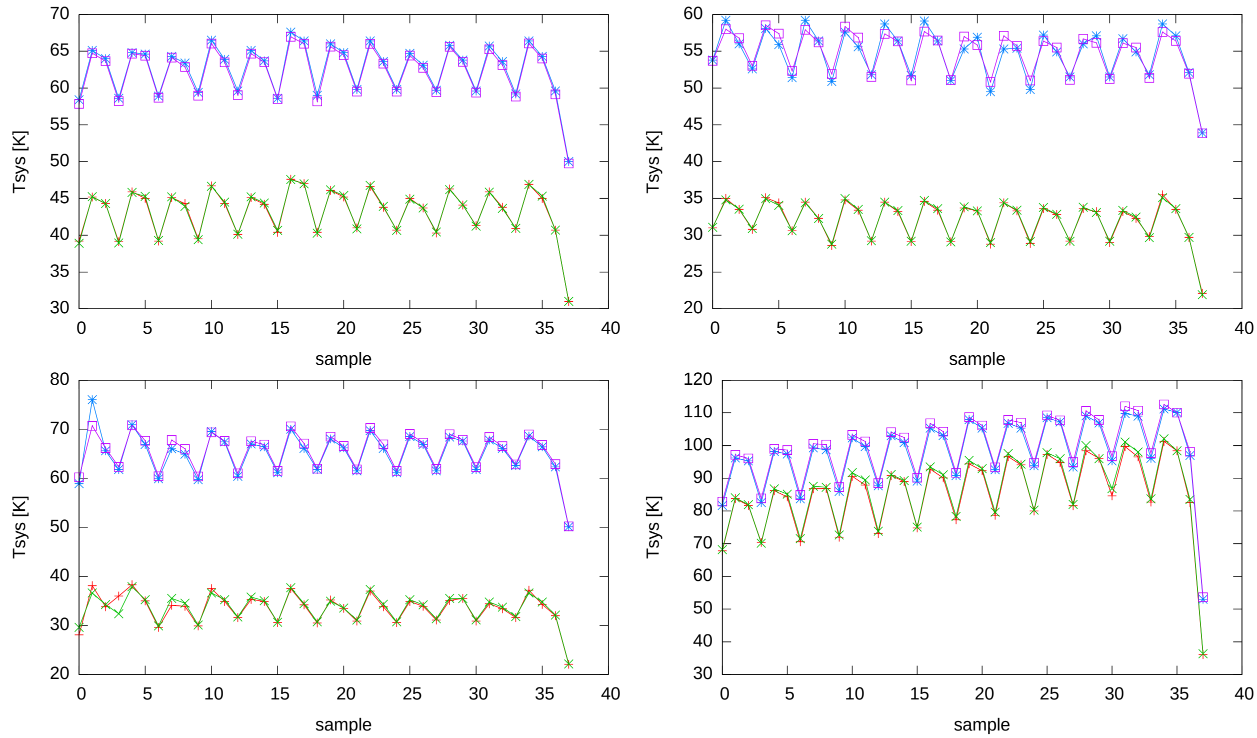

Figure 1: *Comparison of $T_{\rm sys}$ determined with VLBA hardware and DiFX.* This figure compares the system temperatures on 4 antennas. The top left panel shows Fort Davis; the top right shows Kitt Peak; the bottom left shows Los Alamos and the bottom right shows North Liberty. The red and blue data are derived from the hardware and the green and cyan come from DiFX. Right circular polarization is shown in red (hardware-derived) and green (DiFX-derived). Left circular polarization is shown in blue (hardware-derived) and cyan (DiFX-derived) and has 20K added to separate the plots. In all cases the Difx-derived values are divided by 0.935 so that a comparison after correcting for the empirical average 6.5% fudge factor can be made. Since all scans are less than 2 minutes long both sets of system measurements are effectively the duration of the scan; the entire time spanned is about 1 hour. This observation cycled between three low-elevation sources at ~ 1.5 GHz.

5.1.1 Timing offset

One possible cause of symmetric leakage in the power measurements would be a timing error of magnitude τ. For such an error, the leakage would be $\epsilon = 2\tau\nu$ for switching frequency ν. For special extreme values $\tau = \frac{1}{4\nu}$ and $\tau = \frac{1}{2\nu}$ one would achieve the expected measurements $T'_{\rm sys} = 0$ and $T'_{\rm sys} = -T_{\rm sys}$ respectively. The latter of these could occur if the sense of switching was flipped. To achieve the typical 6.5% observed over-estimate in $T_{\rm sys}$ the timing offset would need to be 0.4 ms, a value large enough to be difficult to accidentally introduce in the analog electronics.

5.1.2 Low pass filtering

Another possibility is that a low pass filter between the square law detector and the V-F converter would smear the effective switching time causing an effective timing shift. A simple R-C low-pass filter couples the tunned diode (CR1) to the amplifier in the VLBA total power detector circuit (see figure 2). The time constant, $\tau_{\rm RC}$, is given by the product of component values for R03 (470Ω) and C02 (0.1μF), or 47 μs. This value is a factor of over 8 too small to account fully for the discrepancy.

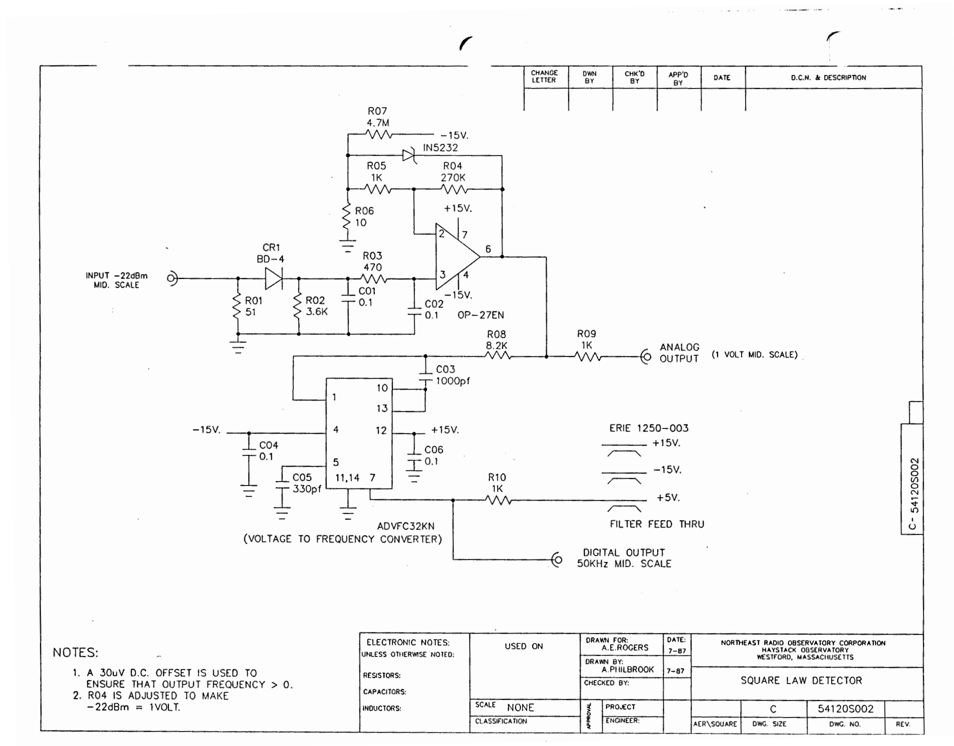

Figure 2: *The VLBA baseband converter power meter.*

5.2 Effect of a power meter DC offset

The addition of a constant unmodeled offset in a power measuring device, ΔP, would give rise to a measured system temperature of

$$T'_{\rm sys} = \frac{T_{\rm cal}}{2} \frac{(P_{\rm on} + \Delta P) + (P_{\rm off} + \Delta P)}{(P_{\rm on} + \Delta P) - (P_{\rm off} + \Delta P)} = T_{\rm cal} \frac{\bar{P} + \Delta P}{P_{\rm on} - P_{\rm off}} = \left(1 + \frac{\Delta P}{\bar{P}}\right) T_{\rm sys}. \quad (10)$$

A DC offset in the VLBA BBC square law detector circuit has been documented [3] as being around 1% for operation at the nominal −22 dBm input power level, resulting in a device-dependent, roughly 1%, error in the estimate of $T_{\rm sys}$.

5.3 Switched power waveform quality

One may wonder if the injected power differs greatly enough from the ideal 50% duty cycle square wave pattern to disturb measurements. Total power folded with a period of 80 Hz, the VLBA's switched power frequency, demonstrates that the injected noise indeed has a very clean square-wave form very well aligned to the 1 second tick. Figure 3 shows the typical square wave pattern of the injected switched power as determined from about 1 minute of averaged VLBA data. The rise of the signal from the off state to the on state is nearly instantaneous and very well aligned with the 1 second tick. The fall from the on to the off state is nearly perfect, perhaps happening with a centroid that is about 8μs (0.07% of a cycle) late and with a very slight visible exponential decay with timescale of 6μs. Note that careful investigation of the signal rise places it with a delay relative to the 1 second tick of less than 2.5μs with an exponential rise time of less than 1μs, both values small compared to the decay times. The net effect of this very slightly imperfect square-wave would affect power measurements at the $\sim 0.1\%$ level *and* should equally affect both the analog and digital system temperature measurements.

5.4 Other causes?

While discussions with engineers seem to rule out the following as probable causes, I mention here a few other mechanisms that might yield the T_{sys} discrepancy.

- Non-Gaussian distributed voltage statistics
- Unmodeled additive noise present only in the analog power meter path
- Non-square-law behavior of the tunnel diode
- Dead-time during measurements of P_{on} and P_{off} is non-zero and not equal

5.5 An unsolved mystery

None of the plausible causes, even all combined, seems to account for the fairly consistent (across different VLBA antennas, experiments and different receivers) $\sim 6.5\%$ discrepancy in the on-line and DiFX measurements of T_{sys}. A third, independent, measurement of this quantity will come once the FPGA core of the RDBE can make this calculation, and will hopefully vindicate at least one of the two existing sources of measurement. In the mean time the obvious course of action is to scale the DiFX-derived measurements to be, on average, consistent with those measured in the baseband converters. In the end proper calibration has been achieved regardless of the absolute correctness of the on-line power measurements as the antenna gains (as a function of elevation) would be computed in a way to absorb any incorrect, frequency-independent, scale factor.

The above discussions illustrate several possible pitfalls in analog measurements of power and system temperature and the mathematical accuracy that can be obtained through digital T_{sys} measurements. Thus I believe in the end the DiFX-based measurements will be closer to the truth and I expect the RDBE values to match. If this is the case, a careful transition to a newly calibrated set of antenna gains will be needed so that the fudge factor can be eliminated.

6 Future directions

If T_{sys} determination within DiFX is found useful the following improvements might be deemed worthwhile:

- Allow for more general switched power waveforms, such as non-integer-hertz frequencies, arbitrary phase with respect to the 1 second tick, or non-50% duty cycle.
- Include the T_{cal} values used in the resultant `.FITS` file.
- Optimize the calculations through use of multiple threads.
- Add swiched power waveform and magnitude information to the vex format.
- A generic strategy for maintaining current values for T_{cal} should be established.
- Respect relevant on-line flags.

Acknowledgements

Thank to Craig Walker, George Peck, Terry Cotter and Vivek Dhawan for interesting discussions, especially those related to understanding the discrepancy between the on-line (BBC) and DiFX-based system temperature measurements.

References

[1] *Some Thoughts on Gain Control in the DBE Era*, 2010, Walter Brisken, VLBA Sensitivity Upgrade Memo 28, http://www.vlba.nrao.edu/memos/sensi/sensimemo28.pdf

[2] *Vex definition version 1.5*, http://vlbi.org/vex/

[3] *Dynamic range of IFD square law detector bias subtraction*, 1989, Alan E. E. Rogers, VLBA Acquisition Memo 164

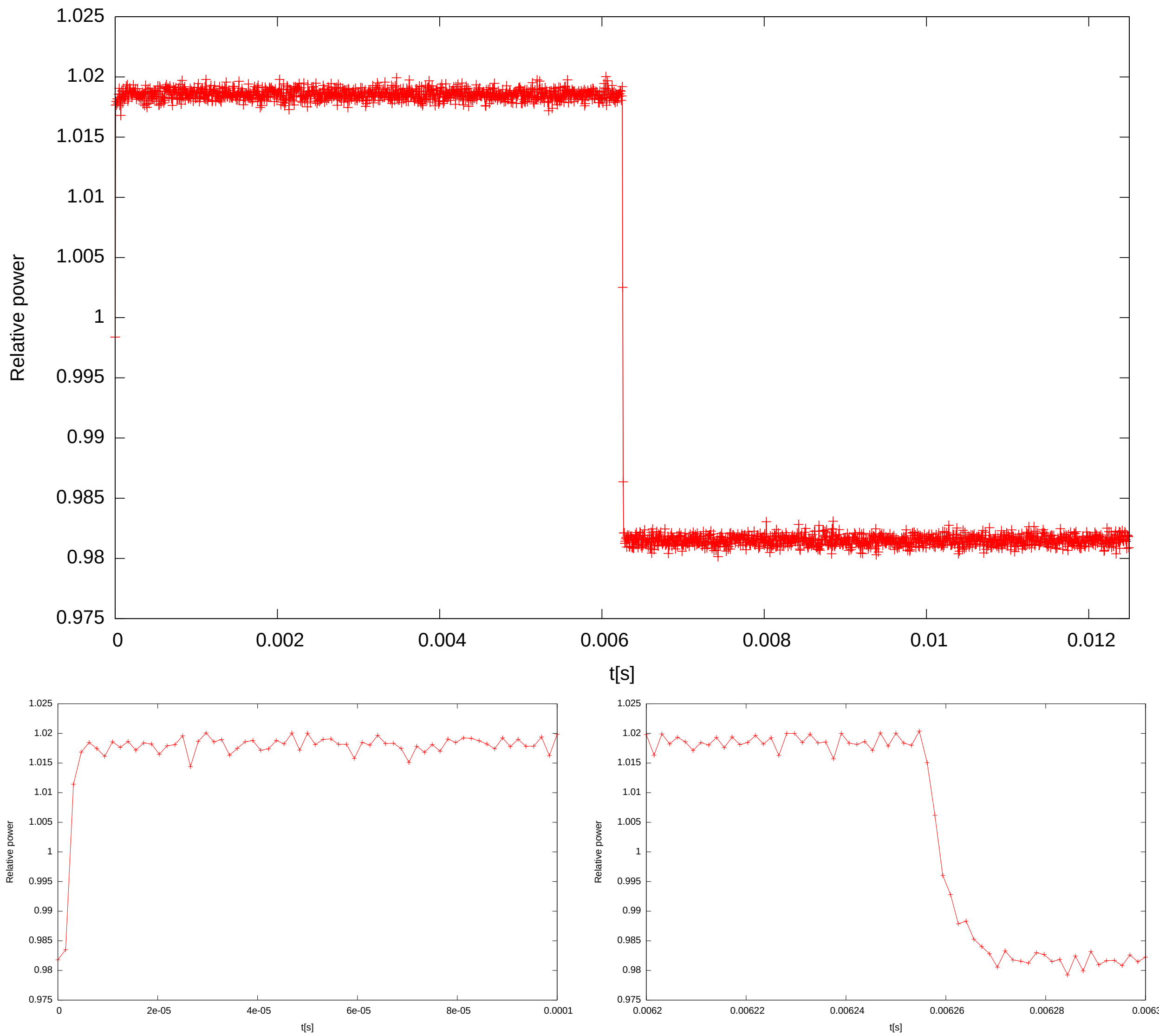

Figure 3: *The VLBA switched power waveform.* This figure shows about 1 minute of averaged cycles of the VLBA switched power as determined by test program `m5fold`. The top panel shows one full cycle (12.5 ms). The bottom left panel zooms in on the upward transition; the actual transition ideally would occur at the left edge of the plot at $t = 0$. The bottom right panel zooms in on the downward transition; the actual transition ideally would occur at the center of the plot at $t = 0.00625$ s. The bottom two plots are plotted on the same scale. Note that the bottom plots have 4 times the time resolution (and thus have twice the noise per point) as the top plot.