LOCO EDGES REPORT #121

ARIZONA STATE UNIVERSITY
TEMPE, ARIZONA 85287

FASTSPEC spectrometer code to replace PXSPEC

Judd Bowman
June 27, 2018

REVISED June 30, 2018

Introduction

The PXSPEC code has been used to interface to the Signatec PX14400 digitizer board and provide spectrometer functionality for EDGES for several years. The code reads samples from the digitizer board and transfers them to RAM. Using the CPU, it then multiples the sample stream by a blackman-harris window function, performs an FFT, squares the output of the FFT to yield a spectrum, accumulates spectra, and writes spectra to disk. These steps are performed mostly sequentially. Modest parallel processing occurs by a) having data transfers alternate between to two buffers asynchronously and by b) using four FFT threads in parallel. Nevertheless, the code is limited to a low overall duty cycle by the largely sequential data flow, the limitation on the number of cores used for FFTs, and the use of only a single phase (tap) of the window function. PXSPEC running on a typical EDGES computer (high-end Core i7 or mid-range Xeon chip circa 2012) over the last ~5 years reaches a computational efficiency of about 50% for a 32768 channel spectrum and an overall integration efficiency of only about 20% (0.5 computation efficiency * 0.36 window function efficiency).

Modern processors can contain many cores and the opportunity exists to increase the overall integration efficiency to a nearly ideal 100% by utilizing more cores, more window function phase taps, and a fully parallel data flow architecture to keep up with the raw 400 MS/s data rate of the digitizer board. In 2017, we purchased and installed replacements for the EDGES computers at the MRO with specifications sufficient to support 100% efficiency (Xeon E5-2660V4 14-Core 2.0GHz 35MB L3 Cache 9.6gt/S).

Here, we introduce the FASTSPEC spectrometer code that re-implements the PXSPEC functionality in a more flexible architecture designed to scale to many-core machines. FASTSPEC utilizes the same dual-buffer asynchronous data transfer from the digitizer board as PXSPEC with a dedicated thread to coordinate processing of data transfers. As data arrives from the board, it is immediately copied and pushed into a FIFO queue, allowing the transfer thread to return to wait for the next transfer. A separate pool of threads is used to pop blocks of data from the FIFO queue and processes the samples by applying the blackman-harris window function at up to 4 four phase taps, FFTing, and squaring. Each thread adds the resulting spectrum to an accumulator shared between the threads and then returns to pop another block of data from the FIFO queue. When a specified amount of data has been accumulated, the accumulated spectrum is available to be written to disk. The process is repeated for each of the EDGES

internal switch states and, when all three switch states have completed, the spectra are written to disk together at the same time.

Running on the new data acquisition computers, FASTSPEC reaches 96% efficiency (see section on Blackman-Harris efficiency below) using about 8 of the 14 cores (when performing FFTs in single precision floating point, double precision needs additional cores). Noise in each of the tap outputs is uncorrelated with the noise in other taps at the expected levels. Comparison of FASTSPEC spectra to the output of PXSPEC are indistinguishable.

PXSPEC and FASTSPEC

The following default configuration values are used in PXSPEC:

Samples per FFT: 65536 = NSIZ
Samples per transfer: 2097152 = NBR
Transfers per spectrum: 1280 = nblock
Samples per accumulation = NBR * nblock = 2684354560

For FASTSPEC, the same mode can be set with the following configuration values:

num_channels = 32768
samples_per_transfer: 2097152
samples_per_accumulation: 2684354560
num_taps: 1

(additionally, a higher num_taps could be used along with write_taps_to_separate_files: true)

Both PXSPEC and FASTSPEC normalize the incoming voltage samples from the digitizer card to a range of +/-1.0 before continuing with subsequent calculations.

There minor differences between PXSPEC and FASTSPEC.

- FASTSPEC uses the following definition of the blackman-harris window function:

$$w(n) = a_0 - a_1 \cos\left(\frac{2\pi n}{N-1}\right) + a_2 \cos\left(\frac{4\pi n}{N-1}\right) - a_3 \cos\left(\frac{6\pi n}{N-1}\right)$$

$$a_0 = 0.35875; \quad a_1 = 0.48829; \quad a_2 = 0.14128; \quad a_3 = 0.01168,$$

  whereas PXSPEC applies a multiplicative factor of 0.5 to this expression for its blackman-harris window.
- PXSPEC calculates adcmin and adcmax statistics for each accumulated spectrum AFTER applying the window function to the stream of incoming normalized voltage samples. This limits the fraction of samples effectively considered in the adcmin and adcmax statistics. FASTSPEC corrects

this behavior and calculates adcmin and adcmax of incoming samples BEFORE applying any window function.

- Because PXSPEC calculates adcmin and adcmax after applying the window function, the adcmin and adcmax values are limited to +/- 0.5 due to the factor 0.5 in its blackman-harris expression. To be backward compatible with the PXSPEC adcmin and adcmax range, FASTSPEC applies a factor of 0.5 to its adcmin and adcmax statistics before recording them. FASTSPEC does not, however, apply the 0.5 factor to the window function or data.
- The FFT of a real input is even (same function on negative and positive sides of the origin). After performing an FFT, when squaring the resulting complex output, PXSPEC sums the spectrum value and its mirror (opposite side of origin) value. This effectively cancels out the factor of 0.5 applied in the blackman-harris window function. FASTSPEC does not sum both sides of the even function and instead only retains half of the output array. Because FASTSPEC does not apply the 0.5 factor in its blackman-harris window function, the net effective of these differences cancels out and both PXSPEC and FASTSPEC yield spectra with the same amplitudes.

During operation, FASTSPEC is able to process the same number of incoming samples as PXSPEC in about half the wall-clock time. Hence, the EDGES receiver internal switch cycle is accelerated by a corresponding amount. This results in a shorter time in each switch position, as well as more total switch cycles recorded per day.

FASTSPEC and PXSPEC codes are publically available in the EDGES repository at:

https://github.com/jdbowman/edges

https://github.com/jdbowman/edges/tree/master/src/fastspec

https://github.com/jdbowman/edges/tree/master/src/pxspec

Configuration

FASTSPEC adopts the .ini configuration file scheme that is used by the EDGES python monitor and control scripts. The .ini file enables changes to the primary configuration parameters (e.g. samples per FFT, number of taps to use in the window function, etc.) to made easily without recompiling the code. The configuration file for an EDGES installation is typical located at /home/loco/edges.ini and FASTSPEC looks there for the file by default.

Parameters

FASTSPEC looks for the following parameter:value pairs in the .ini configuration file:

```
[Installation]
datadir:  /home/loco/testdata
site: mro
instrument: low2

[Spectrometer]
switch_io_port: 0xE010
```

switch_delay: 0.5
input_channel: 1
voltage_range: 0
samples_per_transfer: 2097152
acquisition_rate: 400
num_channels: 32768
samples_per_accumulation: 2684354560
num_fft_threads: 12
num_fft_buffers: 60
num_taps: 3
write_taps_to_separate_files: false

Output path and files

FASTSPEC writes output files by default to a standardized path based on information in the .ini configuration file. The path is [datadir]/[site]/[instrument]/YYYY

If write_taps_to_separate_files is true, then output files are named: YYYY_DOY_tap[#]_[instrument].acq. One output file is written per window function phase tap. The spectra recorded in each of the tap files have the exact same timestamps as in any other tap files generated from the same observation.

If write_taps_to_separate_files is false, then output files are named YY_DOY_[instrument].acq and any taps are averaged together before being written to file.

Example output filenames for PXSPEC and FASTSPEC:

PXSPEC: 2018_178_00_low2.acq

FASTSPEC equivalent to the above PXSPEC file:

a) .ini file setting: write_taps_to_separate_files: true
   2018_178_tap0_low2.acq
   2018_178_tap1_low2.acq
   2018_178_tap2_low2.acq

b) .ini file setting: write_taps_to_separate_files: false
   2018_178_low2.acq

A new file is started (per tap) at every DOY change.

The output file format is backwards compatible with PXSPEC .acq files and the files written by FASTSPEC are drop-in replacements to PXSPEC files if the same spectrometer configuration settings have been used. (Although the FASTSPEC files will contain more spectra per unit time due to the improved computation efficiency).

Command line configuration overrides

A limited number of configuration settings can be overridden by command line arguments. In particular, the .ini configuration file can be specified manually by using "-i" and the spectrometer can be set to write output to a single filename base (per tap) using "–f", omitting the default directory structure and naming convention.

Blackman-Harris window function efficiency

The blackman-harris window function has an efficiency of 36% for a single tap. Using two taps, i.e. applying the window function to the data twice (and hence performing an FFT twice) with the second application of the window function 180 degrees out of phase of the first gives the combined two-tap filter 72% efficiency. Using three taps with the window functions shifted by 0, 120, and 240 degrees gives the combined three-tap filter about 96% total efficiency. Using four taps with the functions shifted by 0, 90, 180, and 270 degrees gives a nearly perfect 100% total efficiency.

Number of taps	Total filter efficiency	Computational cost (arbitrary units)	Total efficiency per cost
1	0.36	1	0.36
2	0.72	2	0.36
3	0.96	3	0.32
4	1.0	4	0.25

Using two taps maintains the same best-case efficiency/cost as using a single tap. Increasing to three taps slightly reduces the overall efficiency/cost. Using four taps significantly reduces the overall efficiency/cost because there is only ~4% improvement in total efficiency compared to three taps, but the computational cost increases 33%. Hence, we generally choose to use three taps for the default EDGES configuration.

Validation of performance

Data were acquired with both PXSEPC and FASTSPEC (separated by ~20 minutes) on 2018_178 using the low-2 instrument while its antenna was disconnected and a stable noise source was connected to the receiver input.

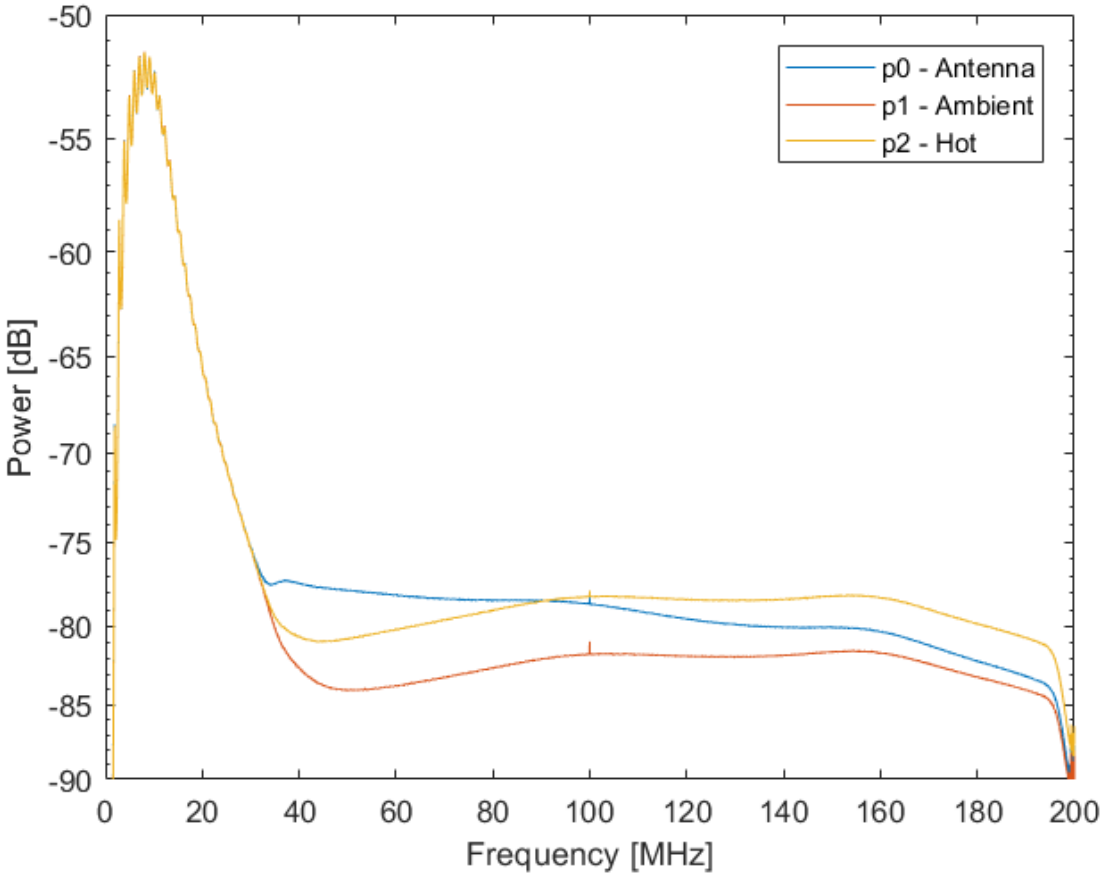

Figure 1. Example 3-position spectra from FASTSPEC (tap 0) connected to Alan's Noise Source on low-2.

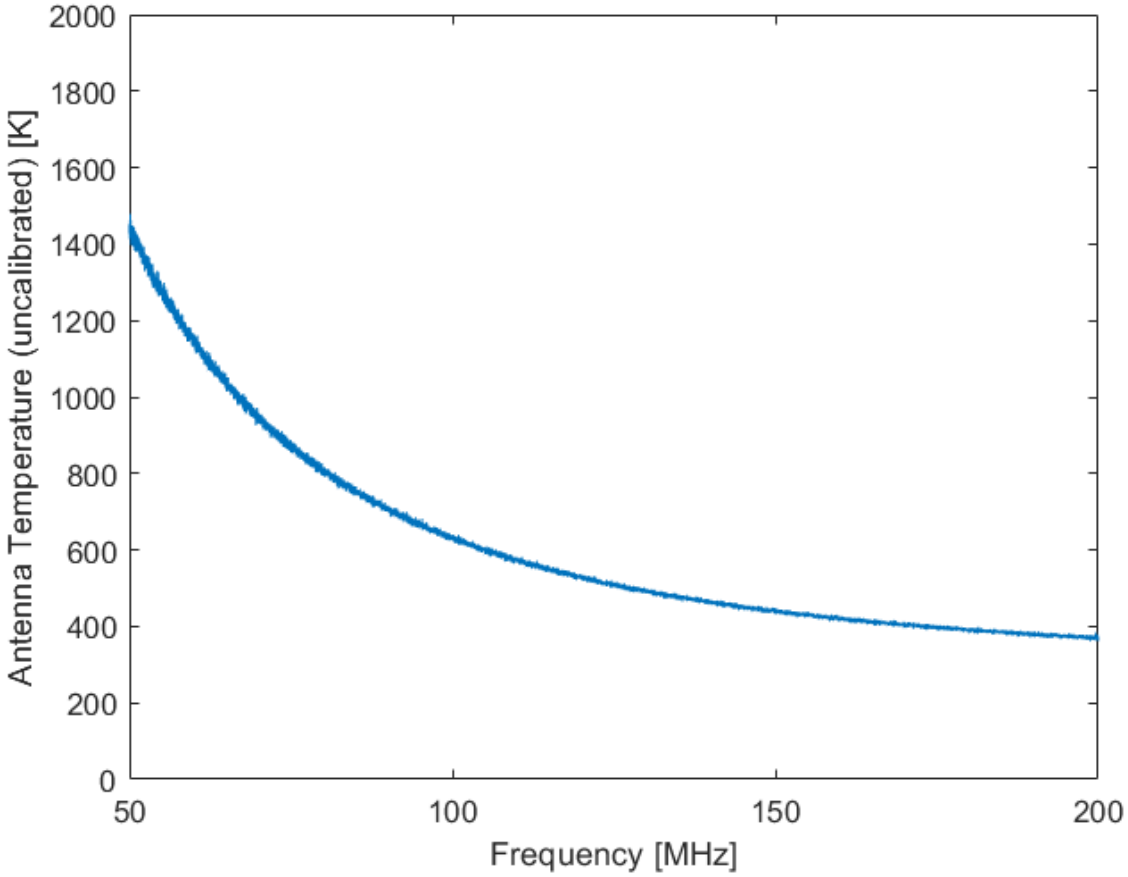

Figure 2. Antenna temperature spectrum (3-position switch corrected) from FASTSPEC calculated from the curves in Figure 1.

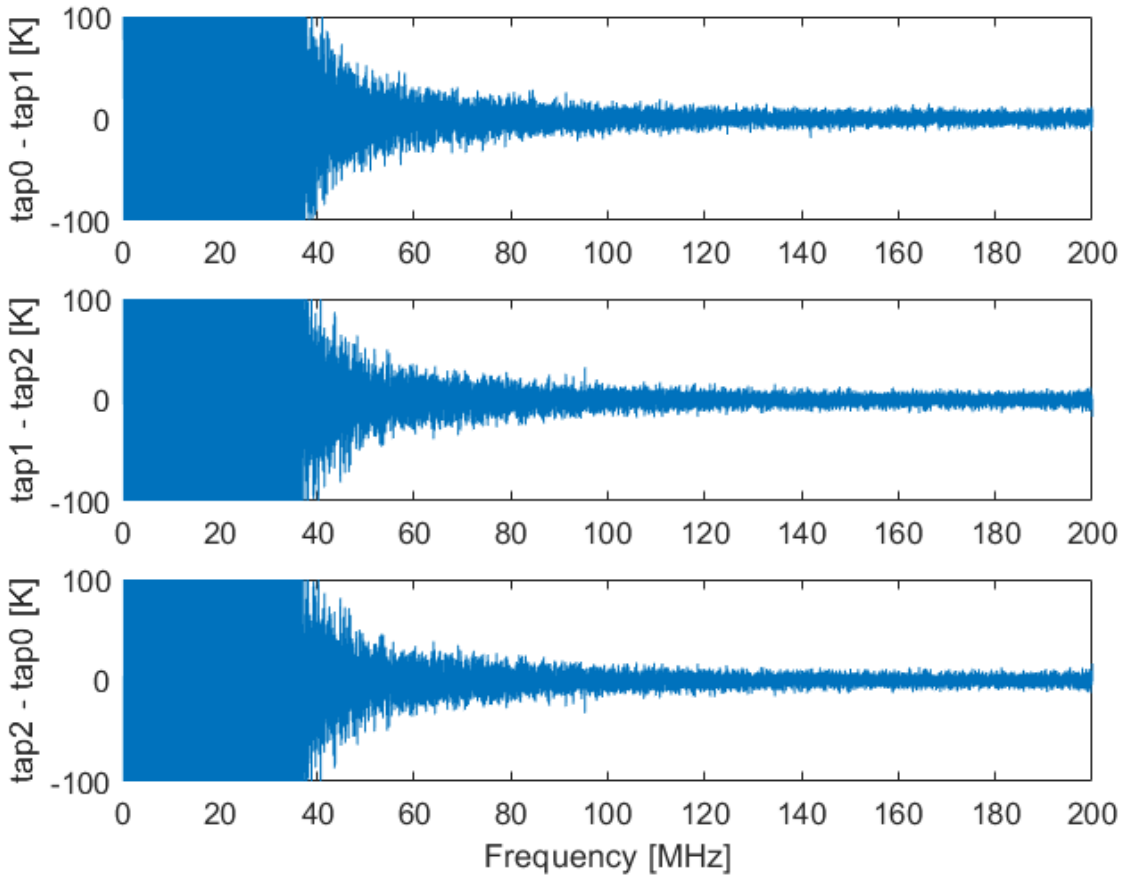

Figure 3. Antenna temperature differences between each of the three tap outputs from FASTSPEC. The differences are noise-like showing that the noise in the separate taps are uncorrelated.

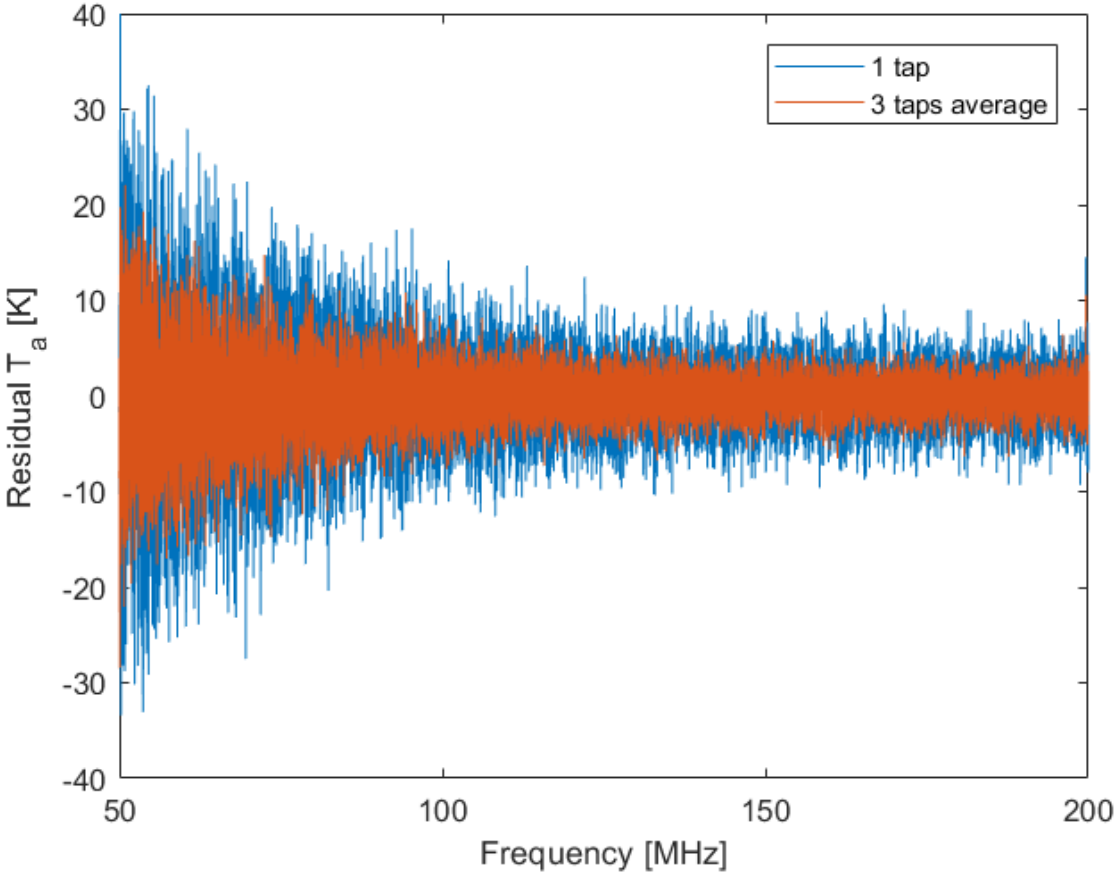

Figure 4. Residuals to a 10-term (9^{th}-order) polynomial fit to antenna temperature. The blue curve is for a single tap 0. The red curve is after averaging the antenna temperatures calculated independently from taps 0, 1, and 2. The RMS over a portion of the band for the blue curve is 7.9809 K. The RMS for the red curve is 4.9707 K. We expect a factor of sqrt(0.96/0.36) = 1.633 improvement from the blue to the red curve. We find 1.606, in reasonable agreement.

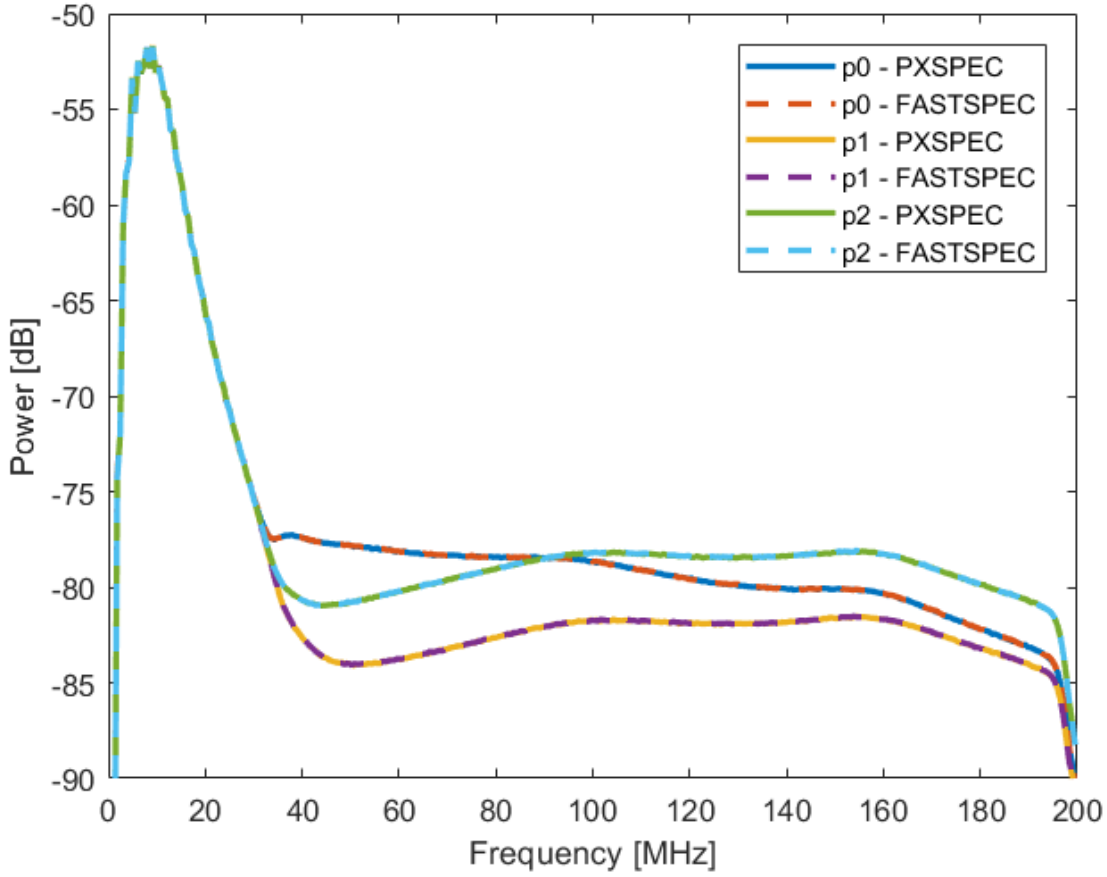

Figure 5. Comparison of 3-position switch spectra for FASTSPEC and PXSPEC. The two codes were used to acquire spectra of the same noise source on low-2. The spectra were acquired about 20 minutes apart.

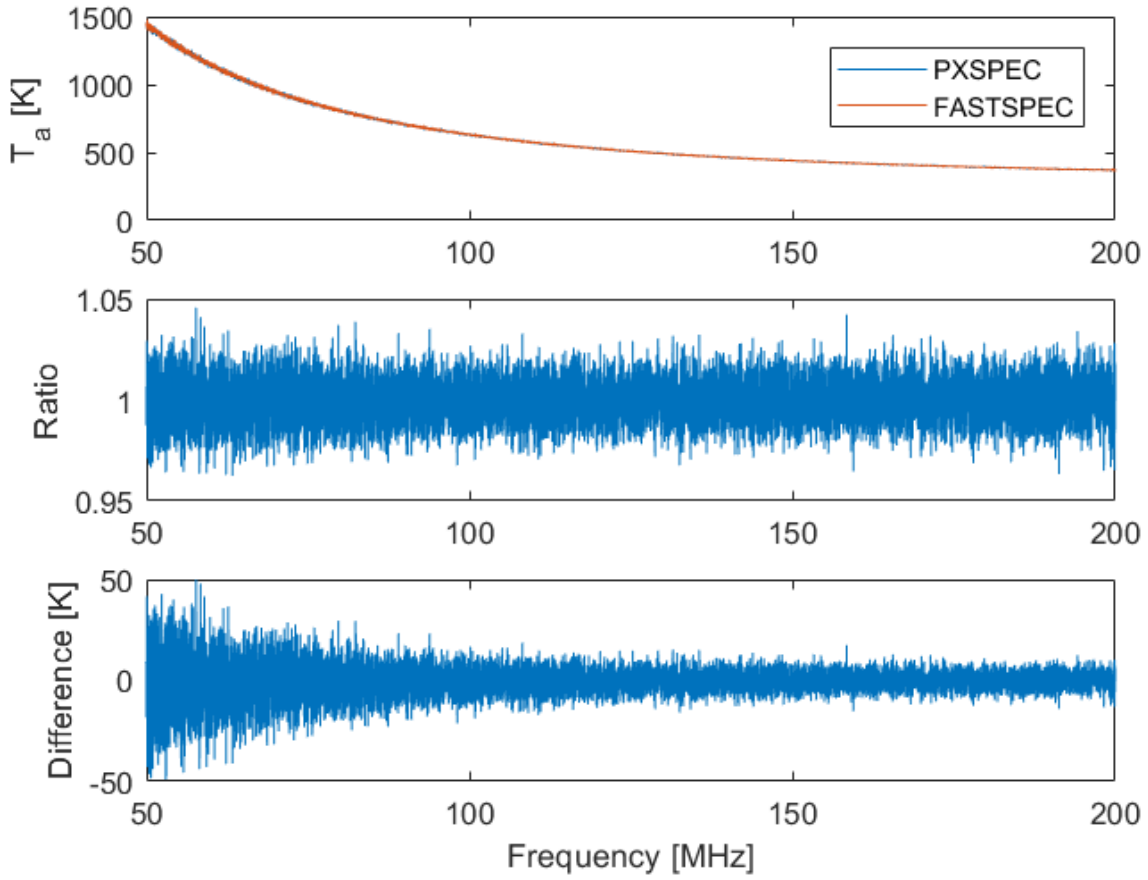

Figure 6. Comparison of 3-position switch spectra for FASTSPEC and PXSPEC. The two codes were used to acquire spectra of the same noise source on low-2. The spectra were acquired about 20 minutes apart.

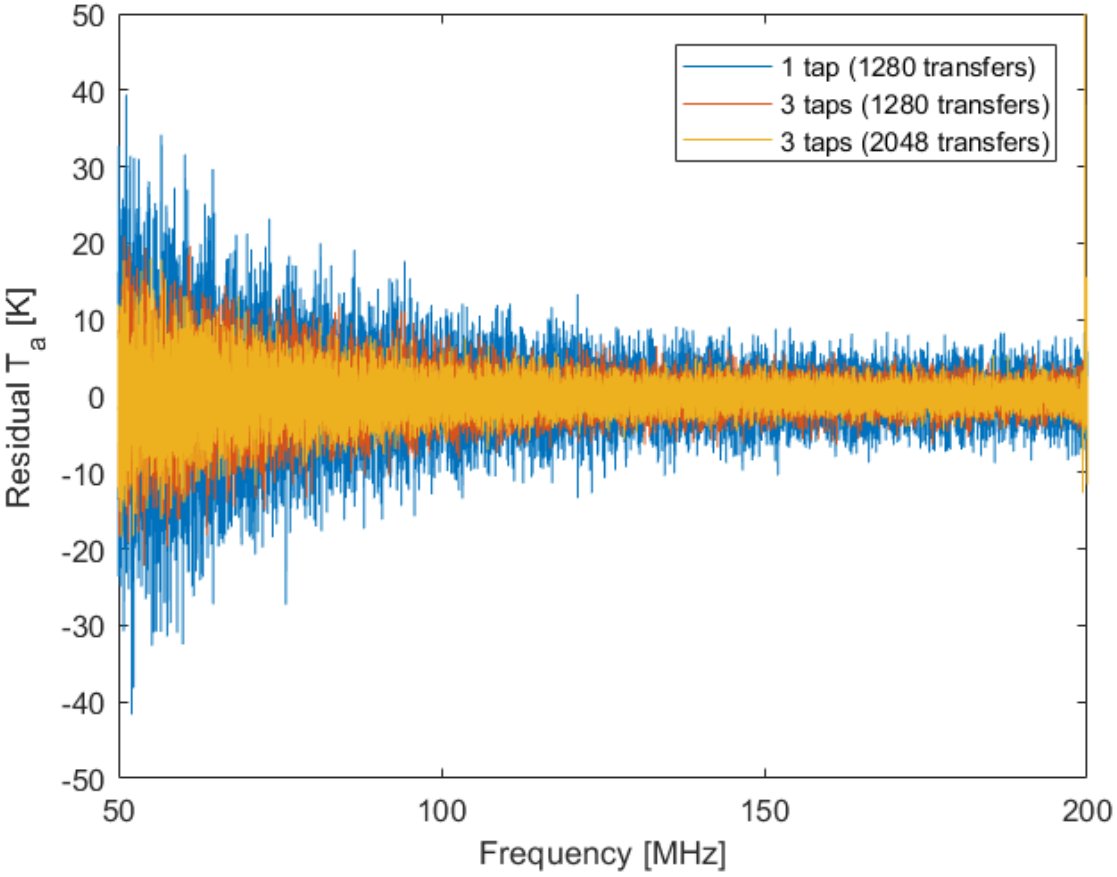

Figure 7. Comparison of internal accumulation options. The residual after subtraction of a best-fit to each curve follows the expected behavior as more data is included in the internal accumulation. The residual RMS over a portion of the band for the three curves are: a) 1280 transfers of data processed with a single tap is 7.3440, b) 1280 transfers of data processed with three taps averaged together is 4.3227, and c) 2048 transfers of data processed with three taps averaged together is 3.4980. We expect an improvement of sqrt(0.96/0.36) = 1.633 from case (a) to case (b) and an additional factor of sqrt(2048/1280) = 1.265 from case (b) to case (c). We find 1.699 and 1.236, in good agreement.