ABSTRACT

CONSTRUCTION OF FASR SUBSYSTEM TESTBED AND APPLICATION FOR SOLAR BURST TRAJECTORIES AND RFI STUDY

by
Zhiwei Liu

The construction of the Frequency Agile Solar Radiotelescope (FASR) Subsystem Testbed (FST) and observational results are described. Three antennas of Owens Valley Solar Array (OVSA) have been upgraded with newly designed, state of art technology. The 1-9 GHz RF signal from the antenna feed is transmitted via broadband (45 MHz-9.5 GHz) optical fiber links to the control room. The RF is then downconverted to a 500 MHz, single-sideband signal that can be tuned across the 1-9 GHz RF band. The data are sampled with an 8-bit, 1 GHz sampling-rate digitizer, and further saved to a computer hard disk. The full-resolution time-domain data thus recorded are then correlated through offline software to provide phase and amplitude spectra. An important feature of this approach is that the data can be reanalyzed multiple times with different digital signal-processing techniques (e.g., different bit-sampling, windowing, and RFI excision methods) to test the effects of different designs. As a prototype of the FASR system, FST provides the opportunity to study the design, calibration and interference-avoidance requirements of FASR. In addition, FST provides, for the first time, the ability to perform broadband spectroscopy of the Sun with high spectral, temporal and moderate spatial resolution. With this three-element interferometer, one has the ability to determine the location of simple sources with spectrograph-like time and frequency resolution.

The large solar flare of 2006 December 6 was detected by the newly constructed FASR Subsystem Testbed, which is operating on three antennas of Owens Valley Solar

Array. This record-setting burst produced an especially fine set of fiber bursts–so-called intermediate-drift bursts that drift from high to low frequencies over 6-10 s. According to a leading theory (Kuijpers 1975), the fibers are generated by packets of whistler waves propagating along a magnetic loop, which coalesce with Langmuir waves to produce escaping electromagnetic radiation in the decimeter band. With this three element interferometer, for the first time fiber burst source locations can be determined relative to the background even though the absolute location is still unkown for the lack of phase calibration information. The radio information over a 500 MHz band (1.0-1.5 GHz) was used to determine the trajectories of the bursts.

Since the digital data are recorded with full resolution and processed offline, a key advantage of it is that one can process the data in different ways in order to simulate and test hardware implementations. FST data provides a unique testbed for studying methods of RFI excision. RFI is observed to be present in every one of the 500 MHz bands, and the high time and frequency resolution provided by FST allows one to characterize it in great detail. The use of time-domain kurtosis, and a variant of the kurtosis method in the frequency domain were explored to identify the presence of RFI and flag bad channels in simulated real time (i.e., we play back the raw, full-resolution recorded data and flag the bad channels during play-back just as a real-time system would do). The ability to select alternate RFI excision algorithms during play-back allows one to compare algorithms on an equal basis. From the same data set, the two kurtosis (time domain and frequency domain) RFI excision algorithms were compared . The results are compared quantitatively to show that the spectral kurtosis is more effective than time domain kurtosis algorithm for detecting the RFI contamination, as expected from theoretical considerations.

CONSTRUCTION OF FASR SUBSYSTEM TESTBED AND APPLICATION FOR SOLAR BURST TRAJECTORIES AND RFI STUDY

by
Zhiwei Liu

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology and
Rutgers, the State University of New Jersey - Newark
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Applied Physics

Federated Physics Department

May 2007

APPROVAL PAGE

CONSTRUCTION OF FASR SUBSYSTEM TESTBED AND APPLICATION FOR SOLAR BURST TRAJECTORIES AND RFI STUDY

Zhiwei Liu

Dr. Dale E. Gary, Dissertation Advisor — Date
Chairperson of Physics, Director of Owens Valley Solar Array,
Associate Director of Big Bear Solar Observatory, NJIT

Dr. Stephen M. White, Committee Member — Date
Professor of Astronomy, University of Maryland

Dr. John F. Federici, Committee Member — Date
Associate Chairperson of Physics, Director of Optical Science & Engineering, NJIT

Dr. Haimin Wang, Committee Member — Date
Distinguished Professor of Physics, Associate Director of the Center for
Solar-Terrestrial Research and Big Bear Solar Observatory, NJIT

Dr. Carsten J. Denker, Committee Member — Date
Assistant Professor of Physics, NJIT

Dr. Martin Schaden, Committee Member — Date
Assistant Professor of Physics, Rutgers University, Newark

BIOGRAPHICAL SKETCH

Author: Zhiwei Liu

Degree: Doctor of Philosophy

Date: May 2007

Undergraduate and Graduate Education:

- Doctor of Philosophy in Applied Physics,
  New Jersey Institute of Technology, Newark, New Jersey, 2007
- Master of Science in Physics of Semiconductor Device,
  Wuhan University, Wuhan, China, 1997
- Bachelor of Science in Semiconductor Physics,
  Wuhan University, Wuhan, China, 1994

Major: Applied Physics

Publications and presentations:

Liu, Zhiwei; Gary, Dale E.; Nita, Gelu M.; White, Stephen M.; Hurford, Gordon M., A Subsystem Testbed for the Frequency Agile Solar Radiotelescope, PASP 2007, 119, 303.

Nita, Gelu M.; Gary, Dale E.; Liu, Zhiwei; Hurford, Gordon M.; White, Stephen M., Radio Frequency Interference Excision Using Spectral Domain Statistics, PASP 2007, 119, accepted.

Wang, Xiaodong; Li, Baoqing; Liu, Zhiwei; Roman, Harry T.; Russo, Onofrio L.; Chin, Ken K.; Farmer, Kenneth R., Analysis of Partial Discharge Signal Using the Hilbert-Huang Transform, IEEE Transactions on Power Delivery, vol 21, No.3, July 2006.

Liu, Zhiwei; Gary, Dale E.; Nita, Gelu M.; White, Stephen M.; Hurford, Gordon J., FASR Subsystem Testbed, AAS SPD Meeting, 2006.

Gary, Dale E.; Hurford, Gordon J.; Liu, Zhiwei; Nita, Gelu M.; White, Stephen M.; First Solar Results with the FASR Subsystem Testbed, AAS SPD Meeting, 2006.

Liu, Zhiwei; Gary, Dale E.; Nita, Gelu M.; White, Stephen M.; Hurford, Gordon J., Imaging Type IIIdm Burst Trajectories, AGU Spring Meeting, 2005.

Guo, Kangzhu; Ou, Haijiang; Grycewicz, Thomas; Joshi, Abhay; Wang, Xinde; Thomas, Gordon; Wang, Haimin; Greene, Richard; Misra, Durga; Liu, Sheng; Liu, Zhiwei; Zhen, Jianjun; Wang, Xiaodong; Li, Baoqing; Xiao, Zhixiong; Ciampa, Nicholas A.; Opyrchal, Jan; Opyrchal, Halina; Chin, Ken K., Proceedings of the SPIE, Volume 5406, pp. 64-72, 2004.

Bao, Z.; Boer, B. de; Erbe, A.; Zhitenev, N.; Abusch-Magder, D.; Meng, H.; Perepichka, D.F.; Jiang, W.; Ermakov, A.; Garfunkel, E.; Frank, M.; Chabal, Y.J.; Liu, Z.; Chin, K., Synthesis, Characterization, and Electrical Measurements of Self-assembled Molecular Wires, Invited Paper, MRS Spring meeting, San Francisco, Apr. 2003.

Liu, Zhiwei; Chin, Ken K.; Bao, Zhenan; Zheng, Jie, Nano-Scale Organic Field Effect Transistor, APS March Meeting, 2003.

To
My Family

ACKNOWLEDGMENT

First and foremost I would like to gratefully and sincerely thank Dr. Dale Gary for being my advisor during my graduated study. I was impressed by his broad knowledge and passion for this field. He has always been extremely generous with his time, knowledge and ideas and giving me much valuable guidance. I would not overcome many crisis situations and finish this thesis without his patience and support.

My gratitude goes to Dr. Haimin Wang. He is very amiable and easy to access. His Magnetic Hydrodynamics class is very important for the students, like me, without solar study background. He was always so patient to explain all aspects of concepts and gave a lot of practical examples which make the materials more easy to be understood.

I would like to acknowledge Dr. Stephen White and Dr. Gordon Hurford for numerous conversations and valuable advices. It is always full of happiness working with them.

I am thankful to Dr. Gelu Nita for the insights he provided over the course of innumerable conversations.

Sincere thanks go to Kjell Nelin for help with the observations and his contribution to keep the instrument running well.

I thank Dr. John Federici, Dr. Carsten Denker, and Dr. Martin Schaden for serving as my thesis defense committee member.

I am grateful to the Center for Solar and Terrestrial Research for the learning-conducive atmosphere it provided. Christine Oertel deserves the special thanks for her support I received.

I am deeply thankful to my wife, Honglin Min, for her love, meticulous care through these years. Without her encouragement and understanding, it would have been impossible to finish this thesis.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

CHAPTER 1

INTRODUCTION

1.1 Radio Emission Associated with Solar Flares

Solar flares are the most intense energy release phenomena in the solar system. A solar flare is a violent explosion that takes place in the solar corona and chromosphere. Radiation is emitted across the entire electromagnetic spectrum, from radio waves to hard x-rays and gamma rays. Solar radio bursts generally are associated with solar flares and originate from all levels of the solar atmosphere between the lower chromosphere (millimeter and centimeter waves) and the outer corona to heights of several solar radii (meter and decameter waves). They also occur at lower frequencies coming from regions throughout the heliosphere. The bursts occur on all wavelengths, but they exhibit very different properties on different wavelengths. Generally, coherent plasma radiation, incoherent gyrosynchrotron, and free-free (Bremsstrahlung) emission are considered to be the main emission mechanisms. Gyrosynchrotron emission dominates the radio spectrum at frequencies larger than 3 GHz.

Meter-wave bursts are among the best studied solar bursts. They are well reviewed by McLean & Labrum (1985) in their book 'Solar Radiophysics'. They are classified as type I through V bursts. Decimetric emissions, in contrast, occur in a wide range of types that defy simple classification. While often studied spectroscopically, they are among the least imaged radio emission of solar flares. The radio emission of solar flares at decimeter wavelengths includes a variety of plasma emission processes. They are situated in the spectrum between the metric range and the centimetric range, although very intense coher-

ent emissions may be observed at frequencies up to 9 GHz (Benz 2004). They are believed to be caused by plasma instabilities driving various wave modes which in turn emit radio waves (McLean & Labrum 1985). The emission frequency of coherent emission is tightly related to the source plasma. The plasma frequency is given by

$$\omega_p = 2\pi f_p = \sqrt{\frac{4\pi e^2 n_e}{m_e}} = 2\pi \times 900 \sqrt{\frac{n_e}{10^{10}\ \mathrm{cm}^{-3}}}\,[\mathrm{MHz}] \tag{1.1}$$

where n_e is the electron density. The emission from 0.3 to 9 GHz corresponds to source densities of 10^9 to 10^{12} cm^{-3}. That is the density range where the primary energy release of flares is expected to take place. Some of the decimeter wave observations have been reviewed by Bastian et al. (1998) and Benz et al. (2002). The analysis for decimetric solar emission provides the potential opportunity for the study of flare physics, CME origin, particle acceleration and propagation, coronal heating (Benz 2004).

One of the more interesting types of bursts is the fiber bursts. Typically the drift rate of fiber bursts is roughly an order of magnitude faster than shock driven type II bursts at the same frequency and an order of magnitude slower than beam driven type III bursts (Benz & Mann 1998). For this reason fiber bursts are called intermediate drift bursts. Fiber bursts are fine structure within decimetric type IV bursts, and are caused by coupling of travelling whistler waves and Langmuir waves, which combine to form electromagnetic emission according to a leading theory (Kuijpers 1975). Treumann & Bernold (1981) developed an alternative theory that suggested that fiber bursts originate through the interaction between whistler solitons and Langmuir waves. Fiber bursts have been considered very important because they can be used to measure the coronal magnetic field in a particularly precise and useful manner as long as the fiber bursts model is known. Spatial observations are the

key to realizing the potential of fiber bursts, because they can be used to directly measure the velocity of fiber emission sources, and thus offer a way to distinguish between these two models.

The lack of imaging due to sparseness of imaging instruments in these wavelengths limits the progress in the field. Thus when it is put into operation, the solar dedicated array FASR (Frequency Agile Solar Radiotelescope) will provide exciting possibilities for the diagnostics of the acceleration of particles and the associated energy release process.

1.2 Instrumentation for Solar Observations

Observations of radio emission associated with solar flares provide diagnostics for flare physics, CME origin, particle acceleration and propagation, and coronal heating. Historically, two approaches to solar observations have been imaging observations at a few discrete frequencies and spectroscopy without spatial resolution. The types and frequency coverage of instruments that are currently used for solar observations are summarized in Table 1.1. Imaging (interferometric) arrays have been implemented at discrete frequencies, such as the VLA (USA; 0.075, 0.3, 1.4, 5, 8.5, 15 and 23 GHz; Napier et al. 1983), the Nobeyama Radioheliograph (Japan, 17 and 34 GHz; Nakajima et al. 1994) and the Nancay Radioheliograph (France; 150, 164, 237, 327, 410 MHz; Kerdraon & Delouis 1997). Interferometers require a large number of antennas for good UV plane coverage and high angular resolution (see Section 1.3). Broadband high-frequency-resolution spectroscopy, on the other hand, usually exploits swept-frequency or broadband digital spectrographs, such as those at Zurich (Switzerland; 0.1-4 GHz; Messmer et al. 1999), Nancay (France; 20-70 MHz; Lecacheux 2000), Culgoora (Australia; 18-1800 MHz; Prestage et al. 1994),

and Green Bank, WV (USA; 12-1000 MHz; Bastian et al. 2005). Two radio facilities with simultaneous spatial and spectral resolution are RATAN 600 (Russia; 1-20 GHz; Bogod et al. 1999) and Owens Valley Solar Array (USA; 1-18 GHz; Hurford et al. 1984), but neither have the image quality necessary to fully exploit the diagnostic potential of radio emission from the Sun. For this, both imaging and spectroscopy data must be obtained simultaneously over a large bandwidth with angular, time and spectral resolution corresponding to the properties intrinsic to solar radio emissions. The solar dedicated Frequency Agile Solar Radiotelescope (FASR) was proposed to meet the challenge (Bastian 2003).

FASR is a major advance over existing solar radio telescopes, and is expected to remain the world's premier solar radio instrument for two decades or more after completion. In 2002, the NAS/NRC Committee on Solar and Space Physics decadal survey ranked the FASR project first among small projects. FASR will utilize hundreds of antennas to perform broadband imaging spectroscopy over a frequency range of 0.05-20 GHz with spatial, spectral, and temporal resolutions designed to exploit radio diagnostics of the wide variety of physical processes that occur in the solar atmosphere (Gary & Keller 2004).

FASR will be designed and built around the time of the coming solar maximum. Table 1.2 lists the main specifications for FASR. These instrument requirements are mainly driven by the scientific demands. A dynamic range $\geq$ 1000:1 and angular resolution of $\simeq$ 1" at a frequency of 20 GHz are considered reasonable to image radio emission on the Sun. To resolve the time scale of interest, a high time resolution of $\leq$ 1 sec is needed for centimeter wavelengths. At decimeter wavelengths, $\sim$ 0.1 sec or even higher rates are required. FASR will include three imaging synthesis subarrays corresponding to the

Table 1.1 Capabilities of Solar Radio Instruments around the World

Observatory	Country	Angular Resolution	Frequencies	Type
VLA	USA			2D mapping
Nancay	France		150 - 450 MHz	2D mapping
RATAN-600	Russia	$240'' - 15''$	1 - 20 GHz	fan beam
OVSA	USA	$90'' - 5''$	1 - 18 GHz	2D mapping
Nobeyama	Japan	$15'', 8''$	17, 34 GHz	2D mapping
Green Bank	USA	-	12 - 1000 MHz	Spectrograph
Ondrejov	Czech Republic	-	3 - 20 MHz	Spectrograph
Zurich	Switzerland	-	0.1 - 8 GHz	Spectrograph
Nancay	France	-	10 - 40 MHz	Spectrograph
Culgoora	Australia	-	18 - 1800 MHz	Spectrograph
Huairou	China	-	1.0 - 2.0 GHz	Spectrograph
		-	2.6 - 3.8 GHz	
		-	5.2 - 7.6 GHz	
Nobeyama	Japan	-	1.0 - 86 GHz	7 fixed freqs
FST	USA	$220'' - 25''$	1.0 - 9.0 GHz	Imaging spectrograph
FASR	USA	$20''/f_{GHz}$	0.03 - 24 GHz	Imaging spectrograph

Table 1.2 FASR Instrument Specifications

Angular resolution	$20/f_{GHz}$ arcsec
Frequency range	50MHz $\sim$ 20GHz
Number channel pairs	2
Total instantaneous BW	1 GHz
Frequency resolution	1%
Time resolution	0.05-3 GHz: 100 ms
	3-20 GHz: 1 s
Polarization	IQ/UV
Number antennas	3-20 GHz: 100
	0.3-3 GHz: 80
	<0.3 GHz: 60
Size antennas	3-20 GHz: 2 m
	0.3-3 GHz: 6 m
	<0.3 GHz: LPDA
Maximum antenna spacing	4.1 km

nearly three decades of spectrum, each of which will employ different antenna size. The desire for full disk imaging implies the use of small size antenna. The range from 3-20 GHz will be covered by 100 antennas of 2 meter diameter, the range of 0.3-3 will be covered by 80 antennas of 6 meter diameter, and the meter wave will be will be covered by log-periodic dipole antennas (LPDA). FASR is designed to exploit a broad program of solar physics, such as understanding energy release, particle acceleration, understanding the evolution of coronal magnetic fields, understanding the nature of the chromosphere and corona in three dimensions, coronal heating and the evolution of filaments (Bastian 2003). After construction, FASR will provide the unique opportunity to study these fundamental problems, and make numerous unforeseen discoveries.

This thesis work centers on the design, construction, and use of a prototype instrument for FASR, the FASR Subsystem Testbed (FST), which was constructed to prototype and characterize the broadband RF transmission system and digital signal processing to be used for FASR. As a testbed for the FASR system, FST provides the opportunity to study the design, calibration and Radio Frequency Interference (RFI) mitigation requirements for FASR. FST is the first system having the ability to perform broadband spectroscopy with high spectral and temporal resolution, combined with spatial locating ability.

1.3 Radio Imaging Interferometry

The diffraction limit is defined by the wavelengh of interest and the telescope aperture (diameter of the antenna in a radio instrument).

$$\theta = \frac{1.22\lambda}{D} \tag{1.2}$$

where λ is the observing wavelength, and D is the angular diameter of telescope. A single element antenna has limited spatial resolution. At a frequency of 5 GHz, even the Arecibo dish (the largest single dish in the world, D=306 m) has an angular resolution of only about 50 arcseconds. Because of the limited spatial resolution of single element telescopes, imaging interferometers have been developed to combine multiple elements working together to form a single telescope, whose spatial resolution is determined by the maximum separation of the elements rather than the size of a single element. The diffraction limit of an interferometer is $\theta \sim 1/B_\lambda$, B_λ is the distance in wavelengths between two antennas, and is referred as baseline length. Figure 1.1 shows a single baseline, the basic unit of an interferometer. If the incident of plane wave arrives at the pair of antennas in phase or even number of wavelengths out of phase, the response will be maximum. If the plane wave arrive the antennas slightly out of phase, the response will be less. So the response of the antennas is a fringe pattern across the sky-the familiar interference pattern for a pair of slits or apertures. The fringe spacing $\theta = sin^{-1}\lambda/B$ ($= 1/B_\lambda$ for small θ), where B is the baseline length, is the same expressions as above. The baseline length is normally expressed with the spatial frequencies, u, v, in number of wavelengths. Figure 1.2 illustrates the fringe space and uv position for one baseline. For an array with multiple antennas, there are a total of $N(N-1)/2$ baselines (N is the number of antennas). Each of these baselines can be represented as a single uv point in the uv plane.

The maximum response measured is the cosine component (x) of incident signal. By inserting a $\pi/2$ phase delay on one antenna signal path, the sine component (y) can be obtained from that response. The complex response function of a point source can be formed from the two components, $\vec{a} = x + iy = a\exp(-i\phi) = a\exp(-i\theta_1/\theta)$, where a and

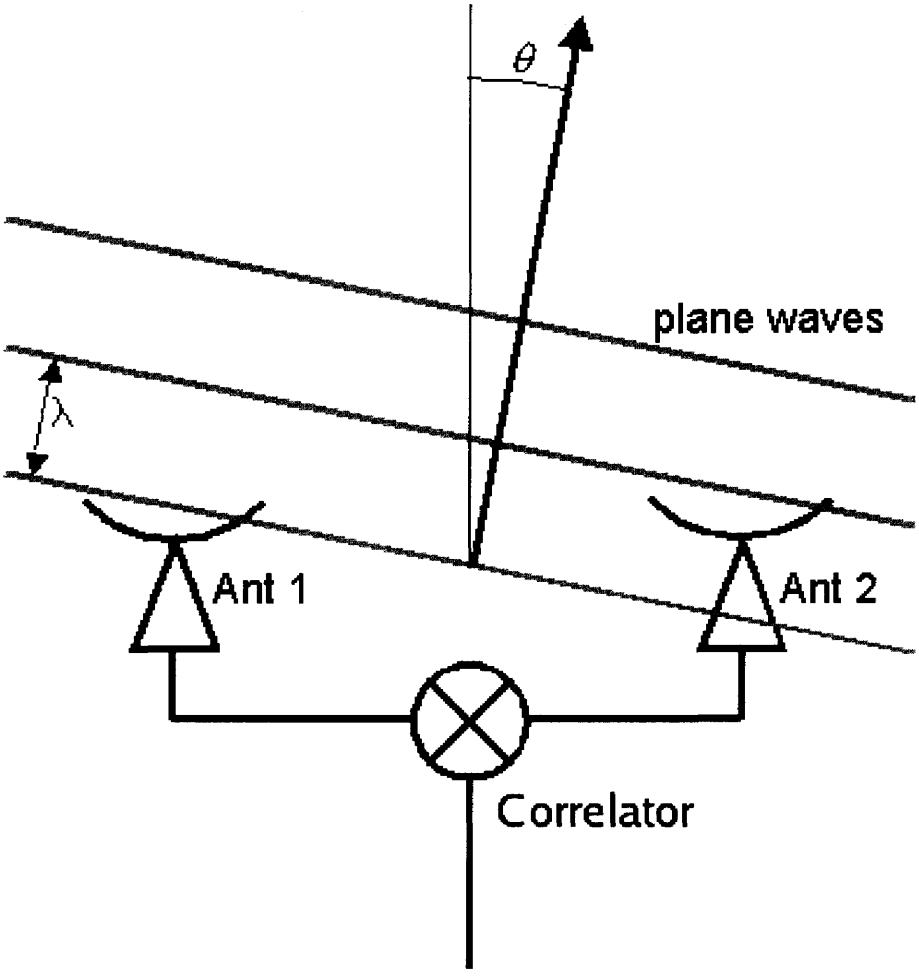


Figure 1.1 Geometry of an interferometer baseline, with plane waves incident from an angle θ. The waves arrive at the antennas exactly 1 wavelength out of phase, the response will be maximum.

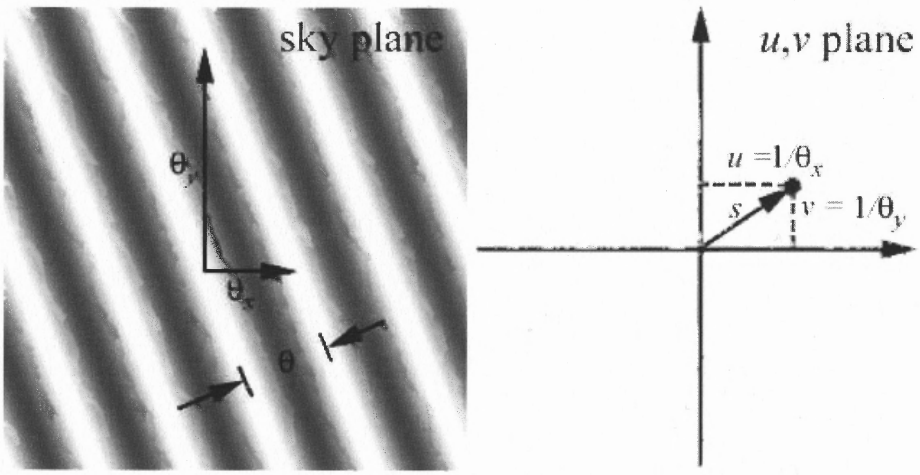


Figure 1.2 A point in the u, v plane a distance s from the origin (right panel) has components u and v, this corresponds to a single baseline. The Fourier transform of this sampling corresponds to the fringe (θ) in the sky plane (left panel). The two components θ_l and θ_m are the fringe separations in the l and m angular directions.

ϕ are the amplitude and phase of the wavefront, respectively, θ_1 is the point source angular distance from the phase center, and θ is the fringe spacing mentioned above. The brightness measured for multiple sources can be expressed as the vector sum of point sources,

$$a_r \exp(i\phi_r) = \sum_i a_i \exp(-i2\pi\theta_i/\theta) \quad . \tag{1.3}$$

Using the u, v coordinates for the fringe spacing, and l, m coordinates for the sky components, θ_i, writing the response of the interferometer, $a_r \exp(i\phi_r)$, as the complex visibility, $V(u,v)$, and sky brightness distribution as $I(l,m) = a_i$, equation 1.3 becomes

$$V(u,v) = \int I(l,m) \exp(-j2\pi(ul+vm))dldm \tag{1.4}$$

for a continuous source brightness distribution. So the sky brightness distribution (the map) can be obtained by taking an inverse Fourier transform of the measured complex

visibility (interferometer response),

$$I(l,m) = \int V(u,v) \exp(j2\pi(ul+vm)dudv. \tag{1.5}$$

An interferometer of more antennas (more baselines or more *uv* plane coverage) will have a better complex visibility measurement, which gives a better image.

FST has three antennas, which only includes three baselines, so true imaging is not possible, but source location is possible using three baselines. Figure 1.3 demonstrates how to determine the centroid from a three-antenna interferometer. The figure shows a row of images along the bottom, and two uv sampling functions on the upper row. The first figure on the bottom represents a small Gaussian source in the real space. The three element interferometer will measure three points in uv space. The point source, when sampled with the three-element array, will give the 3-baseline map shown in the middle of the bottom row. This is calculated by FFT^{-1}(FFT(source)*UV), where FFT(source) means do a forward FFT, and FFT^{-1}(source) means do an inverse FFT. Obviously, one gets "a wall-paper" pattern in which the real source cannot be distinguished. A black circle has been placed around the location of the source, and there is a peak there, but there are sidelobes of equal strength in several other places. However, when the array is measuring different frequencies, the uv points (baseline lengths in wavelength units) scale with frequency, so when the data are combined (frequency synthesis) then each baseline samples a line of points, as shown in the last panel in the upper row. The corresponding map is shown in the last panel of the bottom row. In the last map, one can see which source is the real one. After locating the source with frequency synthesis, one can go back to the single frequency map and apply the CLEAN algorithm to get the map of the source

at each frequency. For 1-5 GHz, the resolution is about 40"-200", which gives a centroid resolution of 4"-20" at 1/10th beam. Depending on the calibration and phase stability, a centroid location to within 1/10 or even 1/100 of the beam can be measured.

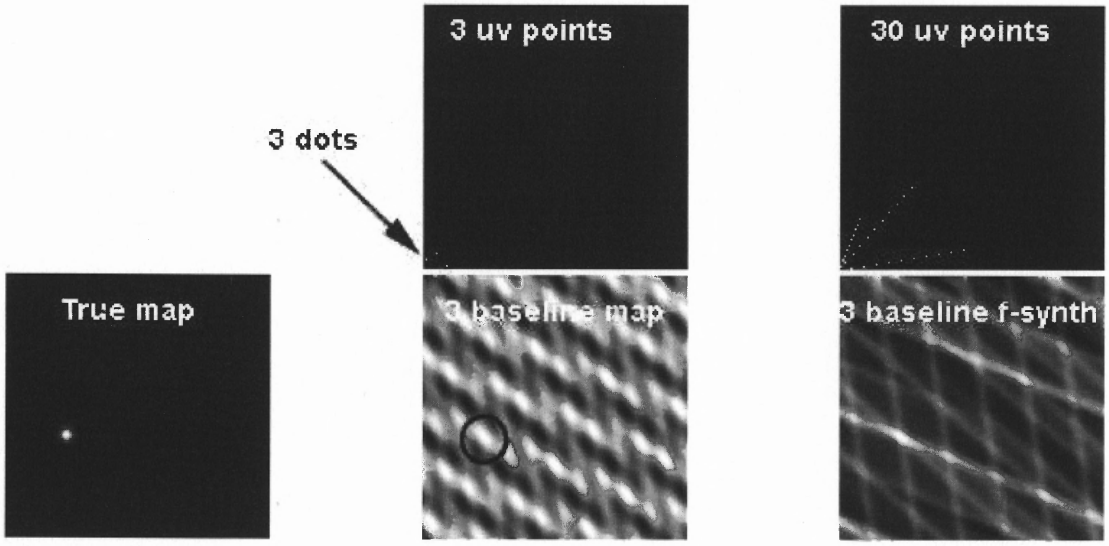


Figure 1.3 Demonstration to determine the centroid of point source from a three-antenna interferometer. The middle of the top row is the three baseline sample points in the *uv* plane, the right panel of upper row shows the baseline samples for ten different frequencies. The bottom row are the true map, three baseline map, and three baseline frequency synthesis map, respectively.

1.4 The Bursts Observed by FST During Six Months

The FST instrument was built at NJIT from 2004-2006, and installed at Owens Valley Radio Observatory in February 2006. Beginning with the first solar burst captured by FST on April 5 (Gary et al. 2006), more than 12 events have been observed during about six months of daily observation (Table 1.3). The solar observations of this 6 months took place in solar minimum, so a calculation to predict the expected number of bursts was done prior to observations. The number of events observed agreed with the prediction based on Nita et al. (2002) who investigated the peak flux distribution of 40 years of solar radio burst

data as a function of frequency and time over a wide range of frequencies. The power-law fits to the differential (density, dN/dS in events sfu^{-1}) and cumulative [N(>S) in events] distributions as a function of frequency, time and phase of the solar cycle was obtained:

$$N(S > S_1) = -\frac{1}{\lambda + 1}\frac{dN}{dS}|_{s=1} = N(S > 1)S_1^{\lambda+1} \tag{1.6}$$

where $N(S > S_1)$ is the number of events above certain flux limit S_1. $N(S > 1)$ and λ were chosen to be 4479 and -1.70. Based on his analysis and statistical data, 12 events with flux larger than 10 sfu in frequency range $1.0 \sim 1.7$ GHz in six months during solar minimum year should be observed. Although the table below indicates a time period longer than six months, in fact after the initial burst in April, regular observations did not start again until mid-June 2006. However, the exact match between the prediction and the number of events observed was coincidence since the observation band (1.0~1.5 GHz) is not exactly the same with the prediction condition (1.0~1.7 GHz).

Because of the configuration of the system, only a 500 MHz band of the signal can be observed at one time. Most solar events in the observing band ($1.0 \sim 9.0$ GHz) should occur in the lowest 500 MHz band ($1.0 \sim 1.5$ GHz), primarily type III bursts. The system was setup in mode 3 with observing band $1.0 \sim 1.5$ GHz during the six months observation. Many of the events captured are decimetric type III burst as expected when planning the observations, but several strong events with type IV continuum, including the intermediate drift fine structure appeared unexpectedly. Especially surprising was the record-setting burst of 2006 December 6 was also captured, which produced the largest solar radio flux ever measured, around one million solar flux units. This event had a severe impact on the reception of GPS signals, almost all of the GPS receivers on the sunlit hemisphere stopped

tracking for 10 minutes. This brought more attention to the impact of solar radio energy release on the satellite and wireless communication systems at Earth. In this thesis, the author concentrate on one feature of the 2006 December 6 event, a fine set of fiber bursts that were captured in unprecedented detail with FST.

1.5 Objectives of Thesis

The thesis covers three research topics. The first described in Chapter 2 is the construction of the unique FASR Subsystem Testbed, which involves a fairly large amount of hardware design and testing. The basic idea is to bring the broadband signal (1.0~9.0 GHz) from three of the OVSA antennas to the control room and down convert to 500 MHz (500 MHz-1.0 GHz) bandwidth signals. These signals are then digitized with three of the four channels of a Acqiris 1 GHz sampling-rate digitizer. I tested all of the individual components and subsystems from which the FST system was constructed (see Chapter 2 for details), which includes the optical fiber transmission system, the band switch and other components of the block downconverter, the tunable local oscillator, and the spectral line downconverter. I designed the block downconverter chassis which is the first down conversion stage (down converts the 5-9 GHz band signal to 1.0~5.0 GHz). I also adapted or wrote the Visual Basic and LabVIEW testing software to verify the functionality of the second stage downconversion, operation of the spectral-line downconverter, and digitizer, which formed the basis of the final operation software. Due to the huge amount of data taken every day, some daily data processing is needed prior to the detailed analysis. I wrote an IDL program that runs every night to search for the flare-potential data based on the daily NOAA event list and OVSA flare meter data. The FST system started taking data

in February, 2006 although further testing and software development was needed before daily solar observation began in mid-June 2006. The performance and capabilities of FST were demonstrated by my observations of satellites and solar radio bursts, and my analysis of the interferometry data. This work was written up and published in Liu et al. (2007).

Chapter 3 is the second part of this thesis and concerns scientific analysis of the fiber bursts observed in the large solar flare of 2006 December 6, which was detected by the FST system. As described earlier, this record-setting burst produced an especially fine set of fibers bursts–so-called intermediate-drift bursts that drift from high to low frequencies over 6-10 s. According to a leading theory (Kuijpers 1975), the fibers are generated by packets of whistler waves propagating along a magnetic loop, which coalesce with Langmuir waves to produce escaping electromagnetic radiation in the decimeter band. With this three element interferometer, for the first time fiber burst source locations can be determined. I use the radio information over a 500 MHz band (1.0-1.5 GHz) to determine the trajectories of the bursts relative to the background emission. Further phase calibration is needed to obtain the absolute location of these fiber bursts in order to compare with the data from other instrument, such as TRACE and Yohkoh, but I discuss the extent to which relative locations can be used to shed light on these interesting bursts, and describe the potential for FST observations when absolute calibration is available.

One of the main purposes for FST construction is to study Radio Frequency Interference (RFI) and possible approaches to RFI mitigation. Chapter 4 is the third part of this thesis. A radio frequency interference (RFI) excision algorithm based on spectral kurtosis, a spectral variant of time-domain kurtosis, is proposed and implemented in software by Nita et al. (2007). My contributions to this paper were, of course, construction of the

instrument, taking of the observations, and some analysis. The clever algorithm itself is due to Gelu Nita. The algorithm works by providing a robust estimator for gaussian noise that, when violated, indicates the presence of non-gaussian RFI. The expected efficiency and robustness of the algorithm were tested by Nita et al. (2007) using data from FST, and the ability of the algorithm to discriminate RFI against the temporally and spectrally complex radio emission produced during solar radio bursts was demonstrated. In Chapter 4, I only repeat those parts of Nita et al. (2007) that are necessary to understand the algorithm. As a quantitative comparison, I researched and implemented the similar RFI excision algorithm (Ruf et al. 2006) based on classical time domain kurtosis theory (Kenney & Keeping 1962), which channelize the time domain signal and determines the RFI presence for each sub-band. The comparison shows that the two algorithms are largely equivalent except that the spectral kurtosis estimator is more efficient than classical time domain kurtosis. Quantitatively, the time domain algorithm performs somewhat worse than expected. The thesis concludes in Chapter 5.

Table 1.3 FST Observed Events During Six Months Observation in 1.0–1.5 GHz Band

Date	Time	Type
2006 Apr 05	21:06	Reverse-drift type IIIs (about 20s duration)
2006 Aug 20	16:30	Type III bursts + continuum (Ondrejov/Zurich data)
2006 Aug 30		Very brief type III-like bursts
2006 Aug 31	17:14-19:04	Spikes, continuum, long lasting fibers, strong noise storm
2006 Sep 29	19:51-19:58	Just low-end of gyroemission
2006 Dec 05	20:53	Short burst with zebra pattern
2006 Dec 06	18:25-20:30	Long lasting X6.5 burst with excellent fibers, spikes (record flux>1 million sfu)
2006 Dec 09	19:15	Type III bursts
2006 Dec 10	19:58	Type III burst
2006 Dec 11	17:20:32	Type III-like, three dots reverse drift?
2006 Dec 14	22:05-23:15	Long lasting X1.5 burst (solely continuum?)
2006 Dec 16	20:20	Single type III-like burst

CHAPTER 2

CONSTRUCTION OF FASR SUBSYSTEM TESTBED

As a prototype for FASR, the FASR Subsystem Testbed (FST) is far more modest than FASR in terms of performance. Its goal is to provide a platform for the study of the design of some elements of the FASR RF and signal processing system, as well as approaches to calibration and RFI excision for FASR. The basic idea of FST is to bring back from the antennas a full 1-9 GHz RF, and do all further signal processing in the control room. It is anticipated that FASR will use this general approach, which provides a highly flexible design with minimal hardware (and associated, self-generated RFI) in the field. Once in the control room, the RF is downconverted to a single sideband IF of 500 MHz bandwidth, which is digitized and recorded as digital data at 8-bit amplitude resolution and full frequency resolution. Such data can then be digitally processed off-line with alternate algorithms, including downsampling, channelization via polyphase filters or other methods, RFI detection and excision, and correlation. As digital signal processing approaches for FASR are contemplated, one can implement each in software and quantitatively investigate the performance of each, using data recorded under actual solar and RFI observing conditions.

2.1 System Configuration

Figure 2.1 shows a block diagram overview of one channel of the FST system, and its relationship to the existing OVSA infrastructure. OVSA (Hurford et al. 1984) is a solar-dedicated interferometer array and takes daily observations roughly covering the time range

16:00-24:00 UT. After some upgrades (Gary & Hurford 1999), the array now consists of two 27 m antennas and five 2 m antennas that can observe the Sun at up to 86 frequencies in the range 1-18 GHz. For FST, antennas 5, 6, and 7 were chosen from the seven existing OVSA antennas, they are three identical 1.8 m antennas arranged in a nearly right triangle (shown in Figure 2.3). The longest baseline is 280 m. The output of the front-ends of antennas 5, 6, and 7 are split to simultaneously feed the OVSA receivers and FST.

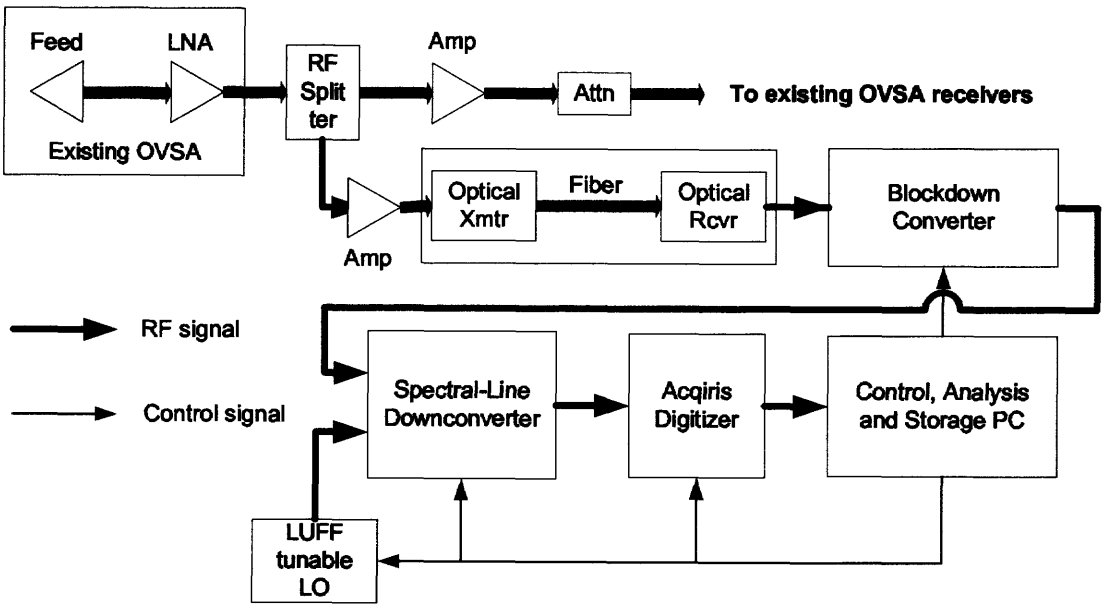


Figure 2.1 Block diagram of one channel of the FASR Subsystem Testbed. The signal is split to feed to the OVSA receiver and FST optical links simultaneously, so there is no impact on OVSA's operation.

Thus, there is no impact on OVSA's normal operation, and one can also utilize the existing antenna pointing and tracking system. The OVSA RF is amplified by 25 dB and then attenuated by 20 dB to isolate the OVSA signal path from the FST path. After the power divider, the FST signals are amplified by 30 dB and sent via broadband (11 GHz) RF optical links (link gain -30 dB; Figure 2.2; the same links are used for the Allen Telescope Array DeBoer et al. (2004)). The direct burial hybrid fiber cable (Corning Altos armored

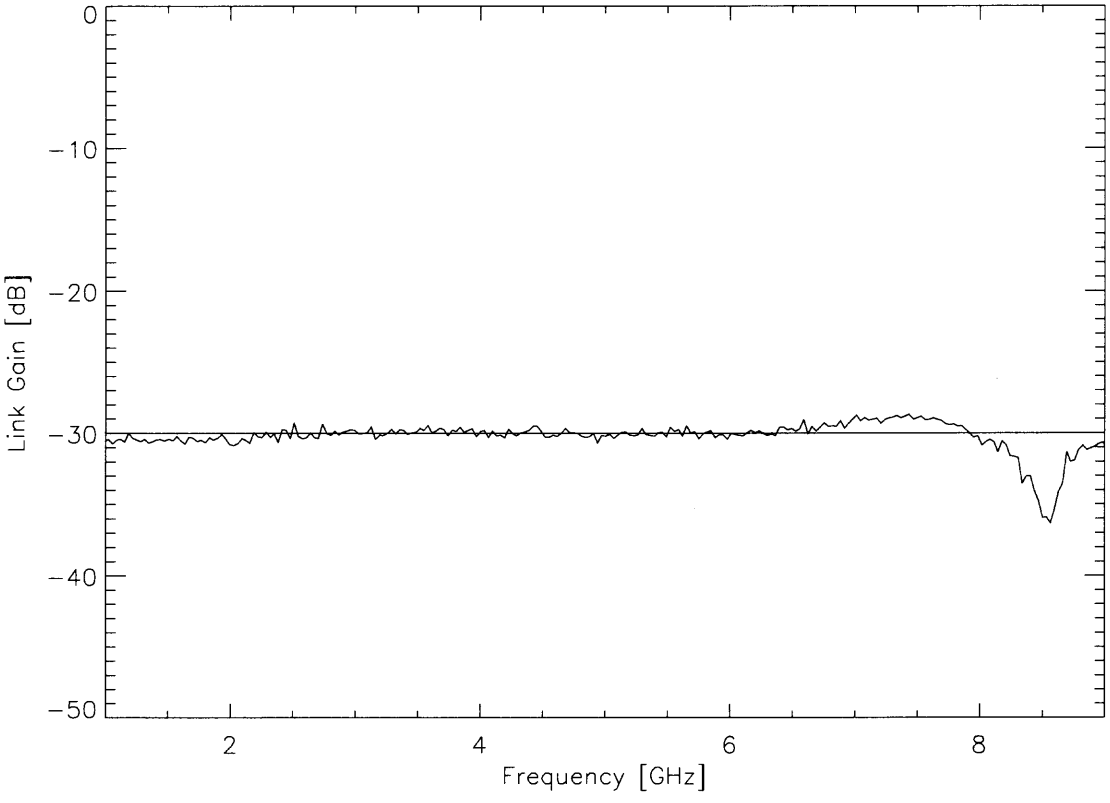


Figure 2.2 Link gain test of broadband RF optical fiber transmission system provided by PSI Inc. The link gain is about -30 dB. Notice that the 3 dB rolloff is seen at 8.5 GHz, but that can be calibrated out at the final stage.

cable) from each antenna comprises three single-mode fibers (one used for the RF, and two spares) and three multi-mode (for future use). The cable lengths from antennas 5 and 7 are precisely matched (232.9 m), while the length from antenna 6 is longer (402.0 m). This allows us to investigate how thermal changes in the fiber affect phase stability, since such effects should be minimal on the matched baseline 5-7, and have precisely opposite behavior on baselines 5-6 and 6-7.

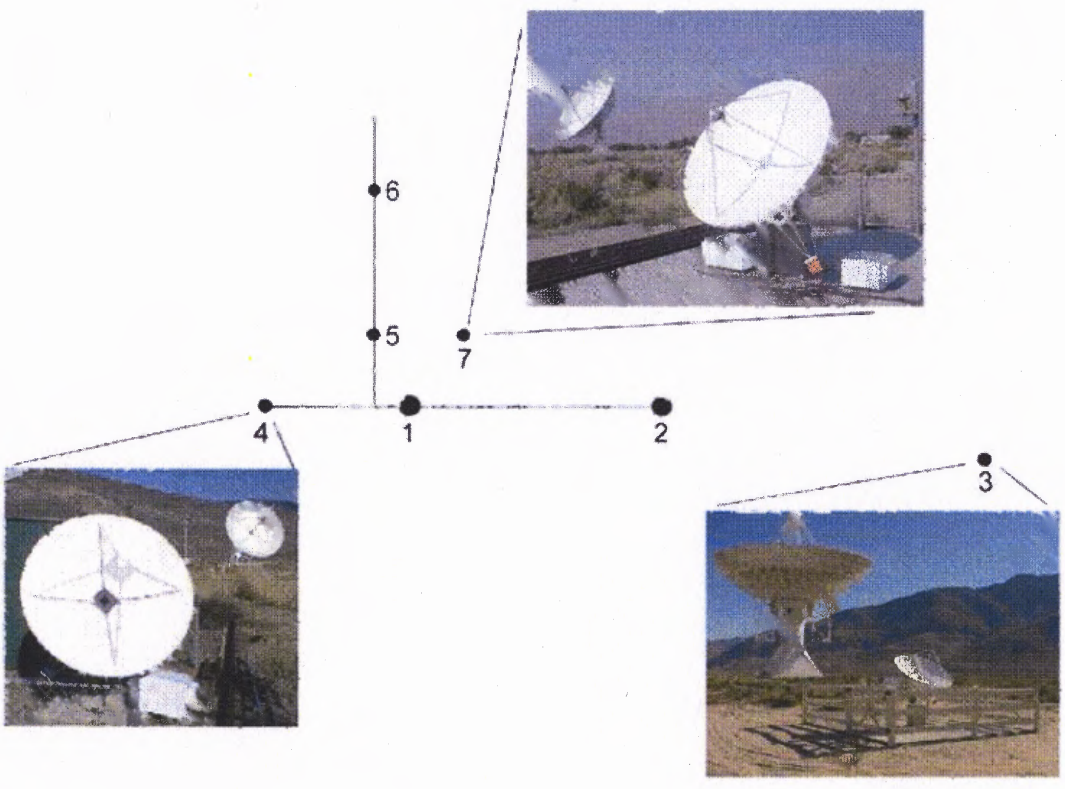


Figure 2.3 Antenna configuration of the FASR Subsystem Testbed (1.8 m antennas, number 5, 6 and 7 of the Owens Valley Solar Array. The longest baseline is 280 m in this nearly right-triangle configuration.

Because the RF is transmitted to the control room before any RF downconversion, it is not necessary to provide separate phase locking during downconversion—the LO from

a single source can be distributed to each of the mixers directly. The task of converting the broadband RF to a single sideband IF signal suitable for digitizing is accomplished in two stages, as shown in Figure 2.4. First the RF from each antenna is block downconverted (§ 2.2) to a 1-5 GHz RF that is switchable over the 1-9 GHz range. Then the 1-5 GHz signal is input into a spectral-line downconverter module (§ 2.3) designed for the CARMA array (Bock et al. 2006) to produce a single-sideband 500 MHz bandwidth IF that is tunable anywhere over the 1-5 GHz band. The single-sideband IF is digitized at 1 GS/s and transferred to a hard disk for storage (§ 2.4) and post processing (§ 2.6).

2.2 Downconverter

The 1-9 GHz broadband spectrum is split into two 4 GHz-wide channels, as shown in Figure 2.4. Channel 1 directly goes to the 2-way switch, while channel 2 goes to a 10 GHz mixer. The local oscillator (LO) signal is provided by a 10 GHz phase locked oscillator (Herley-CTI XPDRO) fed by a 10 MHz reference provided by a Rubidium frequency standard (Stanford Research Systems PRS10). The mixer output, after filtering, is the difference of the LO frequency and the incident RF spectrum, and hence inverts the spectrum over the 5 GHz frequency point. Thus after the 1-5 GHz band pass filter, channel 1 is the lower 1-5 GHz band and channel 2 is the upper 1-5 GHz band (containing the 9-5 GHz RF). The upper and lower bands are selected by means of an RF switch in response to a control signal from the control PC. The two 1-5 GHz channels are referred to as the "upper band" and "lower band" in order to differentiate from the similar "upper sideband" and "lower sideband" provided by the spectral-line downconverter (see below). The selected 1-5 GHz band from each channel is input to each of the three spectral-line downconverter

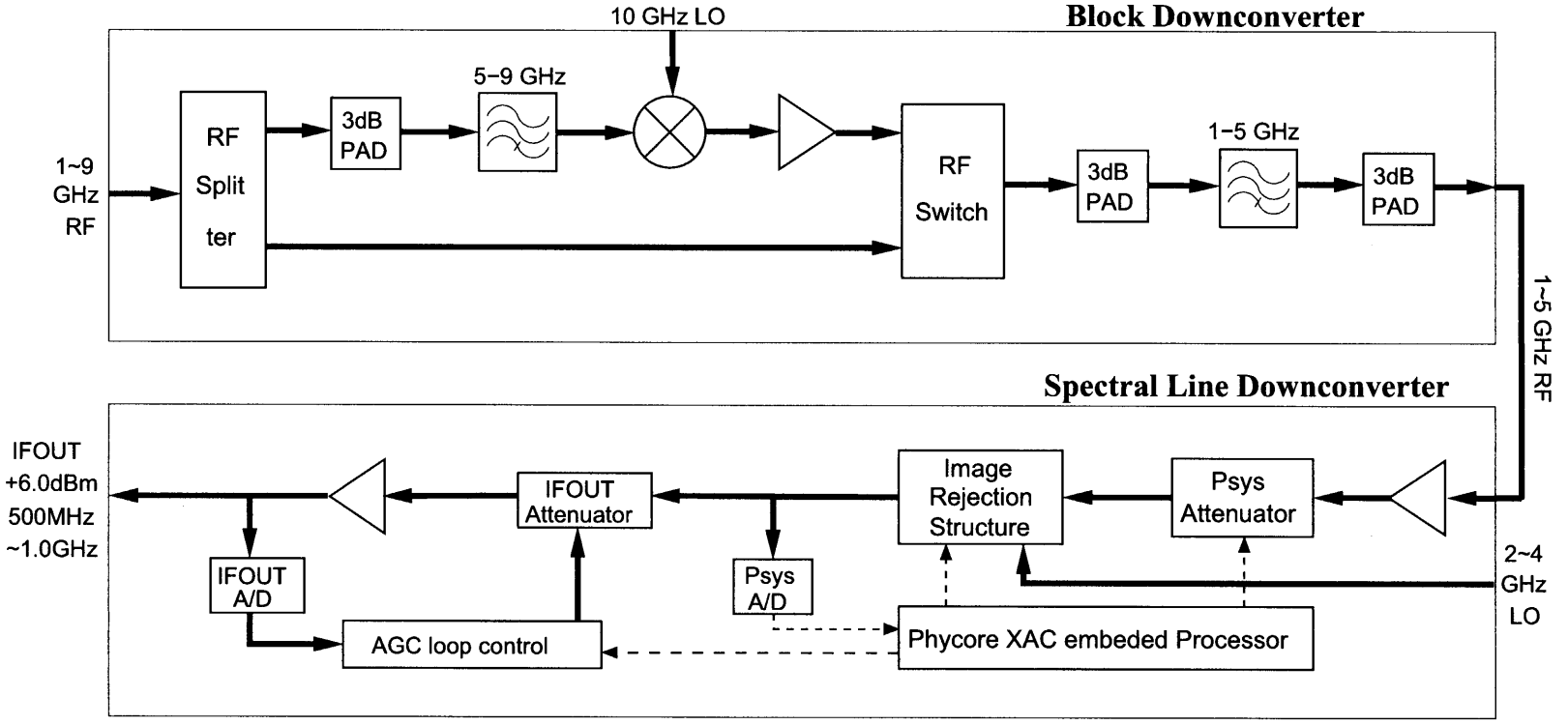


Figure 2.4 The two stages of downconversion: block downconverter (upper panel) and spectral-line downconverter (lower panel). The digitizer performs an implicit third stage of downconversion due to its sampling below the Nyquist frequency.

modules (see the lower panel of Figure 2.4) to produce a single-sideband IF of 500 MHz bandwidth (ranging from 500-1000 MHz), which is tunable anywhere within the 1-5 GHz band. The tunable LO is provided by a frequency synthesizer comprising a 10 MHz reference (provided by the same Rubidium frequency standard as above) and a tunable oscillator (Luff Research) which can be locked to a frequency between 2 and 4 GHz with 20 MHz step. The FST system uses only the 2, 2.5, 3, 3.5 and 4 GHz LO settings. The tunable oscillator's frequency switching speed is < 2 ms.

The frequency range of the selected 500 MHz bandwidth is determined by band B, sideband S and LO frequency ν_{LO} parameters according to

$$\nu_{base} = \sigma[10000(1-B)(1-2S) - \nu_{LO}(1-2S) - 1000] \tag{2.1}$$

where the unit is MHz, $B = 0, 1$ for 5-9 and 1-5 GHz, respectively, $S = 0, 1$ for upper and lower sideband of the LO, respectively, and σ is a parameter that determines if the frequency order is inverted: $\sigma = 1$ if $B = S$, $\sigma = -1$ otherwise. The frequency range observed then ranges from ν_{base} to $\nu_{base} + 500\sigma$.

2.3 Spectral Line Downconverter

The spectral-line downconverter module can be divided into the RF signal processing, automatic gain control (AGC) and card control functions (Figure 2.4). The 1-5 GHz signal that goes into the RF part passes through a Hartley image rejection architecture, which provides at least 25 dB rejection of the image signal to yield a single sideband 500 MHz bandwidth signal (500-1000 MHz). The variable voltage AGC maintains a fixed output

power level of 6 dBm, optimum for the following digitizer. The card control includes the power supply, the CANbus card and the embedded CPU PHYCore-XAC module needed to control the downconverter module, which communicates with the control PC through the CAN bus. In order to suppress the image frequency, the standard Hartley image rejection architecture (Hartley 1928) is used. The incoming RF signal is divided into I and Q channels and mixed with a quadrature phase signal, $\cos \omega_{LO} t$ and $\sin \omega_{LO} t$, derived from the LO. The resulting signals are low-pass filtered, and the Q-channel (I-channel) signal is shifted by 90° (−90°) before final summation of the two channels. The choice of which sideband will be suppressed is software programmable through a CANbus message.

This module has two levels of attenuation: P_{sys} and IF_{out} attenuation, each having an attenuation range from 0 to 31.5 dB. P_{sys} attenuation must be manually adjusted according to the incoming signal power level by resetting the board or by sending a P_{sys} level-set command. Once set, this level is held fixed. Its purpose is to keep the signal coming into the AGC loop in the middle of the IF_{out} attenuation range ($\sim$ 15 dB). In contrast, the IF_{out} attenuation is changed dynamically by the AGC to keep the output power level at 6 dBm. The digitized samples are scaled to total power by applying the measured P_{sys} power level, adjusted for the current value of P_{sys} attenuation. P_{sys} is the signal power level right before the AGC loop, and is measured every 100 ms and broadcast in CANbus "monitor packets" (5 measurements arrive every 0.5 s). Each 0.5 s set of monitor packets also contains one measurement of the current value of IF_{out}, with which one can verify that the signal is in range (i.e., that IF_{out} is 6 dBm).

2.4 Digitization and Storage

The 500-1000 MHz output of each of the three spectral-line downconverters is digitized using an 8-bit, 1 GS/s digitizer (comprising a 4-channel Acqiris DC271 digitizer with 2 MB per channel on-board memory, although only three channels are needed). By the Nyquist theorem, the sampled digital signal frequency ranges from DC to 500 MHz. The total 6 MB of data for three channels is transferred to the PC memory, and then transferred to a SATA RAID 0 hard disk at speed better than 40 MB/s. Using two independent computational threads, once the transfer to PC memory is complete, a new acquisition can be started and transferred while another thread writes the previously transferred data to the disk. The new acquisition typically finishes before the data writing is complete, and the two loops are synchronized by pausing the acquisition and transfer thread until the writing thread is ready. This sequence runs continuously with a duty cycle near one percent. The top panel of Figure 2.5 shows the timing of the digitizer and recording system when operating in this mode. In many cases it is desirable to avoid the long time gaps of order 200 ms between data acquisitions. As an alternative to filling up the 2 MB on board memory in one continuous acquisition, the digitizer can operate in a sequence acquisition mode (Mode 2, as shown in the middle panel of Figure 2.5). In sequence acquisition mode the acquisition memory is divided into a pre-selected number of segments. The waveforms are stored in successive memory segments as they arrive, until the 2 M-point/channel memory is filled (e.g., 20 segments of 100 *k*-points each). In this mode, each waveform requires its own individual trigger, and the trigger delay time is controllable. This makes it possible to smooth out the time gaps, and although the duty cycle does not improve, the time sampling becomes more regular. In this mode, the timing of the waveforms is precisely controlled

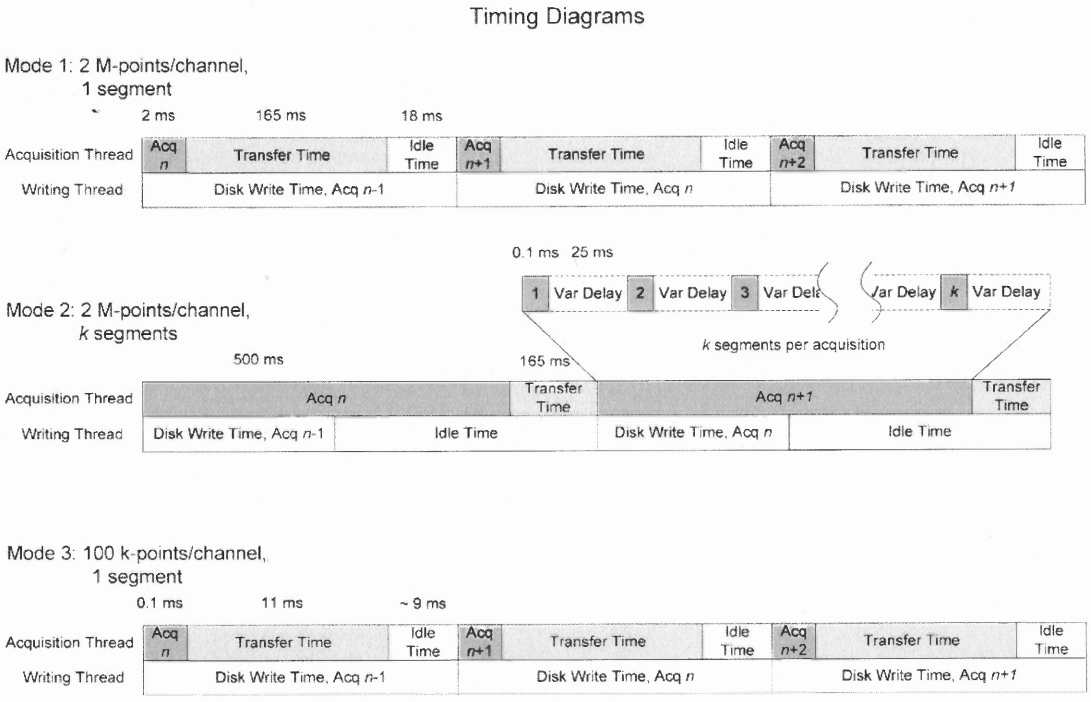


Figure 2.5 Timing diagram of acquisition (dark gray), data transfer (light gray) and data recording (white) for the three most commonly used modes of digitizer operation. The acquisition/data transfer and disk writing tasks are implemented as two independent execution threads. The typical times for each task, as actually used by the researchers, are shown above each acquisition thread. In mode 2, typically $k = 20$.

within an acquisition, with a variable time gap between acquisitions, as shown in the figure.

A third useful mode (a trivial modification to Mode 1), is shown in the lower panel of Figure 2.5. Here one simply reduces the number of points in an acquisition, thereby also reducing the transfer time. By taking 100 *k*-points, the system can transfer and write the data and be ready for a new acquisition in as little as 12.5 ms. To limit the data volume, one may set the cycle time to a longer time. The timing is then similar to Mode 2, but avoids the longer (165 ms) time gaps due to data transfer. The disadvantage with respect to Mode 2 is that the time between samples is not precisely controlled. Figure 2.6 shows the typical timing between acquisitions in Mode 3 for a nominal cycle time of 20 ms.

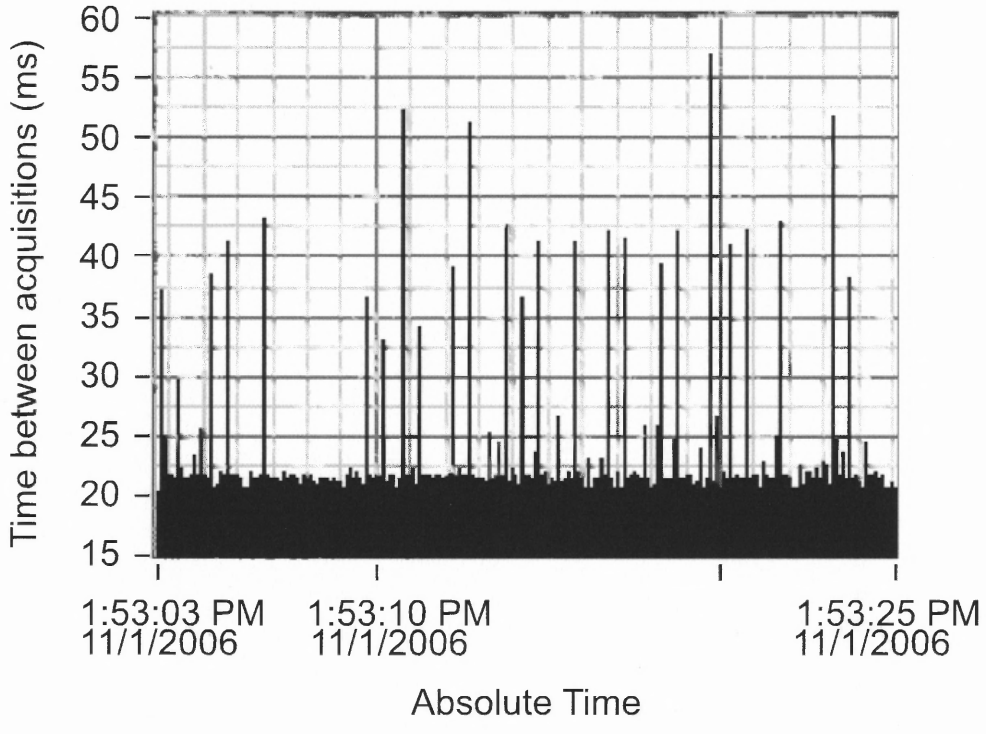


Figure 2.6 Typical performance of the acquisition and data recording for Mode 3 of Figure 2.5. The loop time is set to 20 ms, but due to overhead ranges typically between 20 and 22 ms, while occasional lags of up to 57 ms are seen. Mode 3 is the mode used for daily solar observing.

2.5 Control System and Data Structure

The cooperation of the different modules of the system requires the development of a control system. The antenna pointing is controlled by the existing OVSA control system, so pointing to calibration sources and communication satellites, or tracking the Sun, do not require new hardware and software development. The other parts that need to be controlled or monitored are the RF switches of the block downconverter, the tunable LO, the spectral-line downconverter, the digitizer and the data storage system. The embedded CPU in the spectral-line downconverter module provides the voltage AGC status (IF_{out} and IF_{out} attenuation), the measurement of IF total power (P_{sys}) level, and other monitor information through CANbus messages. The control and acquisition software is implemented in LabVIEW. There are two reasons that LabVIEW is selected as the control platform. As mentioned above, using LabVIEW's multi-threaded programming one can use separate acquisition and writing threads running in parallel to maximize the duty cycle. The second reason is that the graphic interface makes it easy to view intermediate observation results, such as the multi-channel auto- and cross-correlation spectrum, P_{sys}, IF_{out}, and other system settings. That allows us to monitor the health of the system and locate and respond to problems quickly. Tabbed views in the interface allow a dense and complete, yet well-organized display of information.

Synchronization between the FST system and the OVSA antenna control is done through intermediary "scheduling" software running on the OVSA data analysis PC. Through the LabVIEW data socket server and client, the scheduling software controls FST actions such as setting the observing parameters and frequency switching sequence, sending a scan description, and starting and stopping of data recording.

The data are written to disk in a self-describing LabVIEW binary data format. An IDL software was written to allow this self-describing format to be read transparently by IDL—the data language used for all subsequent analysis. Changes to the data format at the LabView level (such as additions to the file or record header structure) therefore require no changes to the IDL reading software. The file sizes may be set within the LabVIEW acquisition program, and are typically set at 100 MB each. In the typical solar observing mode (Mode 3 of Figure 2.5), each file contains 6-7 s of data, comprising a file header containing information that cannot change within the file, followed by a series of data records. Each data record represents one acquisition (which may contain multiple segments if taken in Mode 2 of Figure 2.5), and has a header containing information about the band, sideband, polarization state, noise diode state, and other settings pertaining to that record, followed by an array of 8-bit digitized data samples (1 signed byte/sample) for each recorded channel. The data samples are later to be scaled to true power units by application of the measured P_{sys} recorded in each record header.

2.6 Software Correlator

The recorded data are processed offline by FOCIS (FASR Offline Correlator Implemented in Software). A block diagram for the flow of the processing is shown in Figure 2.7. A key advantage of recording the full-resolution digital data and processing it offline is that one can process the data in different ways (e.g., with different levels of bit-sampling, methods of frequency channelization, RFI excision algorithms, frequency and time averaging, and correlation) in order to simulate and test hardware implementations for the future FASR digital system. This testbed function is the main goal of FST, but the approach also

provides rich datasets for side-by-side comparison of RFI excision algorithms, calibration approaches, and optimum trade-off of time and frequency resolution with signal-to-noise for solar bursts. The software correlator, written in the IDL language, allows the setting of key parameters at run-time, such as the number of time samples in each block (which sets the frequency resolution), the integration time (number of blocks to accumulate), and whether to apply amplitude and/or phase corrections. Other functions are controlled internally by selecting "plug-in" modules for frequency channelization (FFT or polyphase filter, for example) with different windowing functions or FIR filter coefficients, and different modules for RFI excision.

2.7 Calibration and Daily Operation

For FST, the small number (three) of small-diameter antennas (1.8 m) precludes the direct use of cosmic sources for calibration. However, the same antennas are used in the OVSA array, for which there are baselines involving the 27-m antennas. Therefore, it is possible to transfer the OVSA calibration as follows: First observe cosmic sources such as 3C84 with the OVSA array to determine the antenna-based amplitude and phase calibration for antennas 5, 6 and 7, then observe solar bursts and the stronger active regions with both OVSA and FST. The OVSA baseline-based amplitudes and phases are compared with those measured by FST on the same baselines, and the FST amplitudes and phases are corrected to agree with those from OVSA. To do this accurately requires using only those frequencies of the FST band that correspond with the smaller bandwidth of OVSA. This technique yields a band-averaged complex gain correction, but does not allow a bandpass calibration. A general procedure for transferring the OVSA calibration to FST in this way

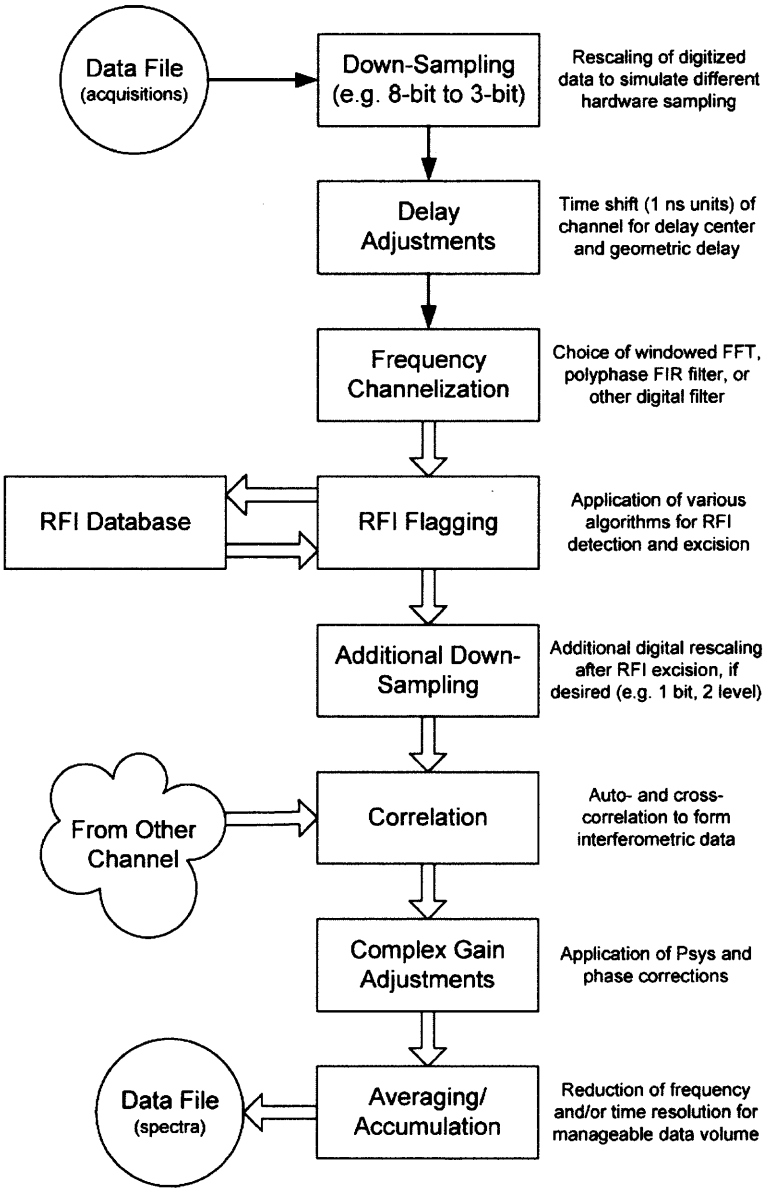


Figure 2.7 Block diagram for a single channel of the offline software correlator, FOCIS (FASR Offline Correlator Implemented in Software). The fact that the full time-resolution waveforms are recorded allows one to run FOCIS repeatedly with different parameters and function blocks to simulate any hardware implementation.

is still under development.

The absolute flux calibration is done by referencing the measured flux to tabulated solar flux values available from other radiometers. A bandpass amplitude calibration is first carried out through the intermediary of an injected cal source (noise diode) at the front end which is used as a secondary calibration source. The procedure is to observe the blank sky (in both R and L circular polarization), alternately turning the noise diode on and off. This yields the quantities T_{sky1} and $T_{sky1} + T_{nd1}$, the difference of which provides T_{nd1}. Then one can point the antennas at Sun center and repeat the measurement, yielding the quantities $T_{sky2} + T_{Sun}$ and $T_{sky2} + T_{Sun} + T_{nd2}$, the difference of which provides T_{nd2}. The expected noise diode temperature is 1200 K, so the contribution of the Sun to the system temperature is $T_{sys} = 1200[(T_{Sun} + T_{sky2})/T_{nd2} - T_{sky1}/T_{nd1}]$ K. This is related to the flux density of the Sun through the well-known relation $S = \frac{2kT_{sys}}{\eta A}$ where A is the antenna area and η is a frequency-dependent efficiency factor less than unity. To the extent that η and the noise diode temperature are known, this procedure yields the solar flux density directly. The calibrated data can be compared with measurements at specific frequencies in the FST band by the US Air Force Radio Solar Telescope Network and radiometers at Penticton (Canada) and Nobeyama (Japan), and with the known shape of the quiet-Sun temperature spectrum (Zirin et al. 1991). Figure 2.8 shows such a comparison for one particular day when the Sun was quiet. The noise diode calibration procedure does result in remarkably continuous total power spectra, balanced in both polarizations, with antenna temperature contributions close to the value expected in the middle of FST's range, but some small ripples remain in the spectrum and the calibrated data may be too bright at high frequencies. The antenna efficiency and the noise diode coupling into the system have

some uncertainty and this may contribute to the residual calibration differences.

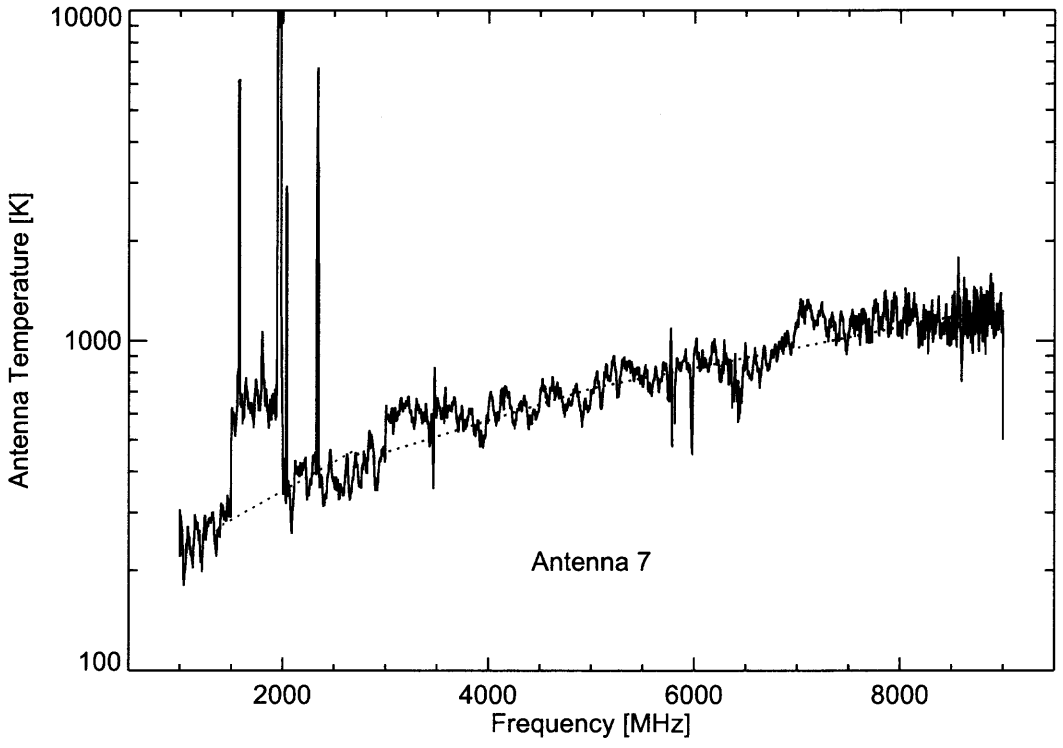


Figure 2.8 Total power spectrum on the Sun on 2006 August 31, measured as antenna temperature (on OVSA antenna 7) in RCP, and calibrated by injecting a noise diode of assumed temperature $T_cal = 1200K$ into the signal path. For comparison, the dotted line shows the antenna temperature contribution deducted from the quiet-Sun flux spectrum tabulated by NOAA Space Environment Center for the same day. Prominent narrow interference features are visible across the spectrum; the band between 1.5 and 2.0 GHz is dominated by a very broad and strong interference line at 1.96 GHz that causes nonlinearity and hence does not calibrate well.

Since about 2006 June, the FST has been operating daily in solar observing mode targeting solar bursts, while doing special observations such as RFI surveys and satellite observations as needed. Due to the huge data volume, the daily observation of solar bursts requires careful coordination as follows. A typical day's observations take up about 300 GB of storage, while the data-recording PC has 500 GB total storage. Therefore, at the end of each day's observing an automated procedure must determine the periods of likely

solar activity and delete all data outside of those periods. Because OVSA is observing in parallel, and has a solar burst monitor providing a record of periods of radio burst activity, it is possible to examine the OVSA flare file and keep all periods flagged as flare periods, and also examine the daily solar event list provided by NOAA's Space Environment Center and keep other periods as appropriate. It was found that the FST data themselves are not reliable indicators of solar bursts due to variable RFI and other variations that may swamp the weaker bursts. Because the year 2006 is in a solar minimum, solar bursts are relatively rare so on most days no data is kept at all, except for the integrated P_{sys} records. On some days, however, as much as 10-20% (30-60 GB) is saved and transferred to an external HD via USB. When the external HD is full, it is shipped from the OVRO site in California to NJIT, where the data are inspected visually and pared down to usually a much smaller volume for permanent storage and further analysis. The data to be saved are transferred off the external disk, and the disk is returned to OVRO for reuse.

2.8 Observation of Geosynchronous Satellites

First fringes with FST were obtained on a GPS satellite in early March 2006, which allowed a rough determination of the phase center of the system. A more precise determination was made by tracking the geosynchronous satellite Galaxy 10R at local coordinates $HA = 5.359°$, $Dec = -5.897°$ on 2006 March 14. Three 2-ms acquisitions were taken in each of two bands, $3500 - 4000$ MHz and $4000 - 4500$ MHz, encompassing the C-band transmission range of the satellite. Figure 2.9 shows the cross-correlation phase on each of the baselines (5×6, 6×7 and 7×5), and the closure phase, after vector averaging the three acquisitions. The analysis was done with 10000-sample blocks, yielding a frequency reso-

lution of 100 kHz. This broad range of transmitted frequencies (3700-4200 MHz) is useful

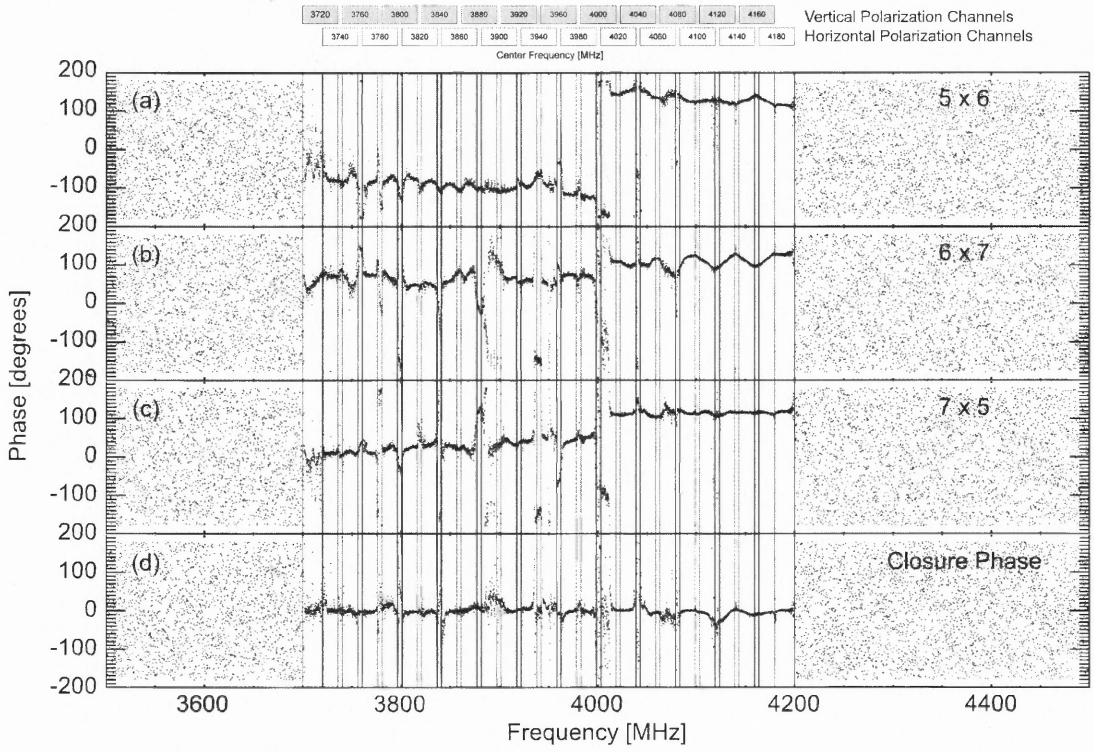


Figure 2.9 Correlated phase on all three baselines (panels *a-c*) and closure phase (panel *d*) on geosynchronous satellite Galaxy 10R, as measured in two 500 MHz bands of FST. Phases outside the signal band (3700-4200 MHz) are random. Phases within the signal band vary within the 24 overlapping channels of the satellite, which are shown schematically by rectangles above the plot. Channel center frequencies are giving in each rectangle, and the channels shown by gray rectangles are vertically polarized while the channels shown in white are horizontally polarized. Vertical lines show the channel edges to guide the eye.

for determining delay centers. The delay centers applied in Figure 2.9, using antenna 7 as the reference antenna, are [0, 959, 0] ns for antennas 5, 6 and 7, respectively. It is obvious that these delay center values result in a relatively flat phase variation versus frequency. A further change of 1 ns causes worse phase slopes, so the delay centers can be taken as [0, 959, 0] ns to within the 1 ns resolution of the 1 GS/s digitizer. Note that there is a phase jump from one FST band to another (on either side of 4000 MHz), which is expected

due to a change in system phase, uncorrected in Figure 2.9. Note also that the antenna 5 delay center is 0 ns (relative to antenna 7), which shows that the matched fiber lengths and other cable and electrical lengths in the system are precisely equal. The lengths were matched purposely (§ 2.1) in order to investigate the effects of thermal variations on the phase. Thermal variations in the fiber will affect baselines involving antenna 6, but should be absent on baseline 7×5.

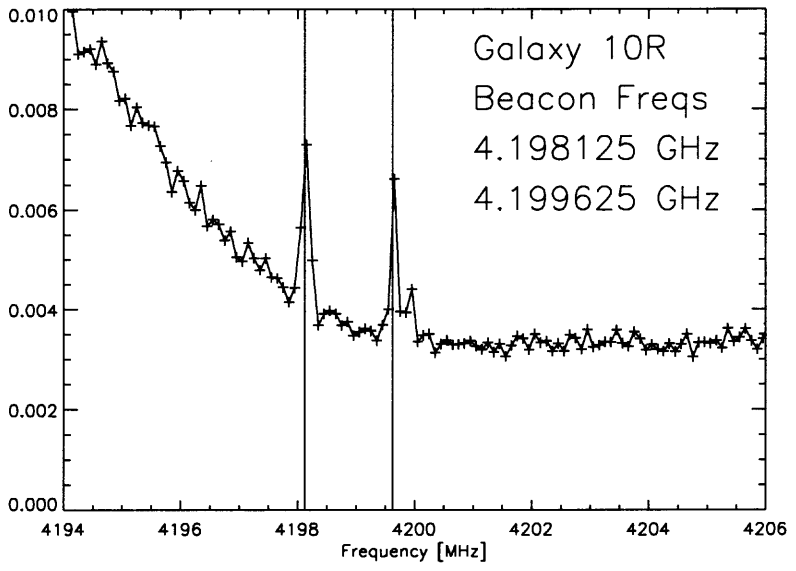
Some smooth phase deviations across the band in Figure 2.9 could be variations in complex gain over the band, but other variations and abrupt jumps are clearly associated with the broadcast channels of the satellite. The closure phase (lower panel of Figure2.9), should be zero for all geometrically defined phases, and while the overall trend is flat at zero, there are clearly significant deviations due to the character of the signals on the channels. The C-band channel distribution for Galaxy 10R is shown at the top of Figure 2.9. The channels are distributed with overlapping channels in V and H linear polarizations. It can be investigated whether the phase variations seen in Figure 2.9 are due to antenna-based phase differences between V and H polarization on the different antennas, by reconfiguring the front ends to bypass the 90° hybrid to allow the direct measurement of linear polarization. It was found that the phase variations and non-zero closure phase remain. Further investigation shows that there are two other satellites (Galaxy 14 and EchoStar 9/Intelsat Americas 13) with $\pm 2°$ hour angle difference and opposite channels polarization with Galaxy 10R within the beam, indicating that geosynchronous satellite signals are not suitable for phase calibration.

The geosynchronous satellites sometimes have "beacon" frequencies that are useful for phase and frequency calibration. Galaxy 10R is a PanAmSat satellite, whose beacon

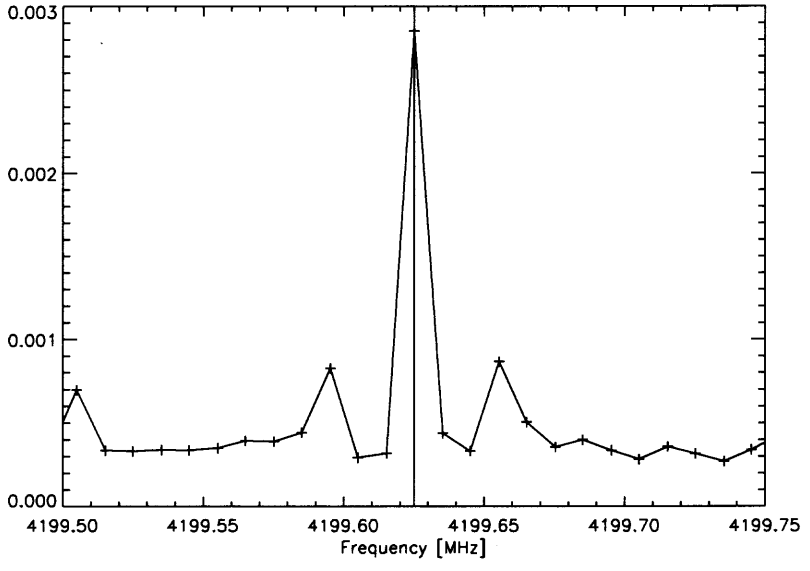
frequencies are at 4.198125 GHz and 4.199625 GHz. These beacon signals are easily seen in the data at 100 kHz frequency resolution, as shown in Figure 2.10*a*. The data were reanalyzed using 100000-pt blocks, for 10 kHz frequency resolution in order to check the FST frequency calibration. Figure 2.10*b* shows that the line at 4.199625 GHz appears at precisely the correct frequency as expected.

2.9 Observation of GPS Satellites

GPS satellite 22779 was tracked starting at 14:26 UT on 2006 April 21, using a position calculated from the standard "two-line elements" (TLE) data from NORAD. The GPS L1 signal is centered at 1.57542 GHz and spans 10 MHz bandwidth due to the satellite's P-code modulation. Figure 2.11 shows the correlated phase as a function of frequency using the delay center of [0, 959, 0] ns found previously. The phase is relatively flat over the 10 MHz band of the L1 signal, except at frequencies with non-GPS interference at 1576 MHz and 1577 MHz. Figure 2.13 gives the correlated phase as a function of time over about 28 minutes, at 1574.5 MHz. The analysis was done with 1000-pt blocks, yielding 1 MHz resolution. The left column shows the phase for the three baselines and closure phase, before geometric phase correction. Note the rapid phase wind in the three channels, while the closure phase remains zero. The rapid phase rotation is due to applying the geometrical delay to the IF (intermediate frequency, DC-500 MHz) digitized data instead of in the RF signal path at which the delay actually occurs (Thompson et al. 2001). The correction for this phase variation is called "fringe stopping," and depends on the various stages of down-conversion in the FST system. The FST uses two explicit levels of downconversion, plus one additional implicit downconversion by the digitizer. The downconversions introduce



(a)



(b)

Figure 2.10 (*a*) The beacons for geosynchronous satellite Galaxy 10R, overlaid with vertical markers showing the nominal frequencies of the lines. These beacons are V polarized. The vertical axis is total power in arbitrary units. (*b*) The upper-frequency beacon at 10 kHz resolution, overlaid with a vertical marker showing the nominal frequency of the line. It is remarkable that the line is symmetrically centered on the marker, indicating that FST frequency calibration is precisely correct to within a small fraction of 10 kHz. The vertical axis is total power in arbitrary units.

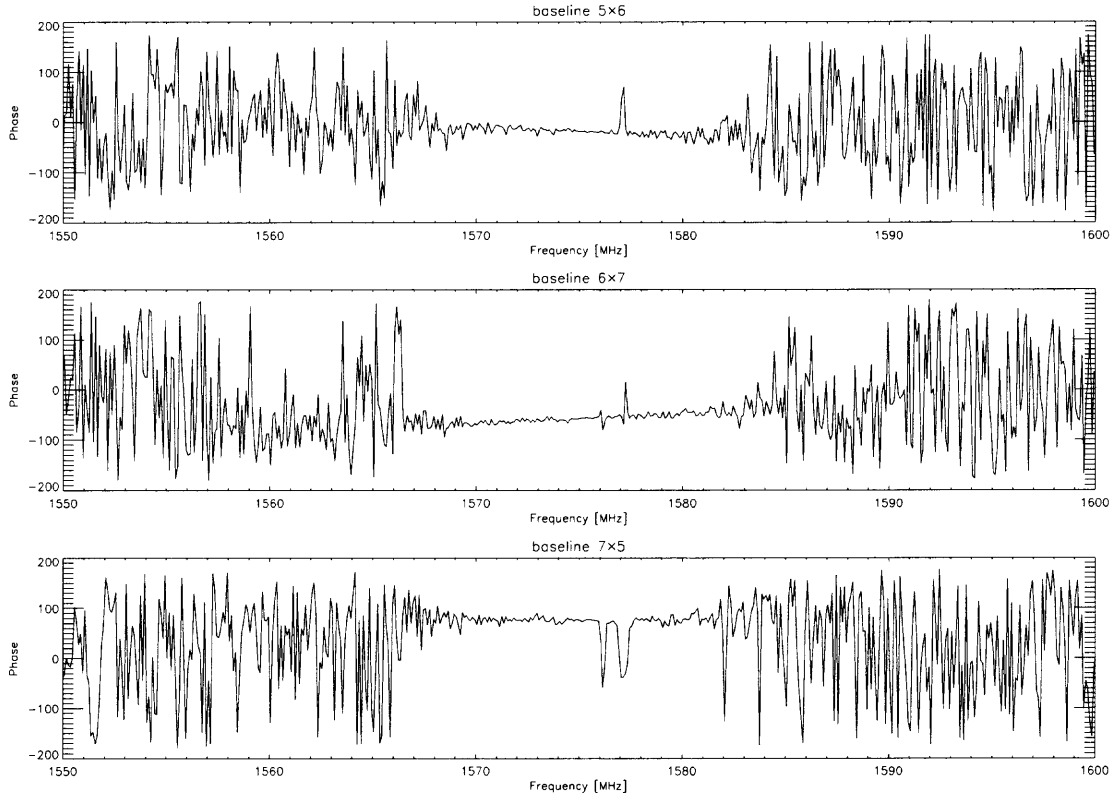


Figure 2.11 GPS satellite phases versus frequency for the three baselines. Note the phases across the 10 MHz band centered at the L1 frequency 1575.27 MHz are flat. Also note that there is interference at 1576 and 1577 MHz.

phase rotation (natural fringes) at the LO frequency. The general expression is developed here and given in equation 2.5. Figure 2.12 provides a schematic view of a single baseline, in which the incoming wavefront of frequency ω_{RF} arrives at one of the antennas after some geometric delay τ_g that depends on the baseline length and the direction to the source. To compensate for this delay, a delay $\tau = \tau_g + \tau_c$ is introduced in the other antenna just prior to correlation, where τ_c is a time- and frequency-independent delay (called the delay center) that takes into account unequal cabling and other system delays not related to the source-baseline geometry. At the antennas, the time-dependent voltages have the dependencies $\cos(\omega_{RF}t)$ and $\cos(\omega_{RF}t - \omega_{RF}\tau_g)$. Any additional antenna-based phase shifts can be ignored, which are corrected during calibration. After the first downconversion (at block-downconverter frequency ω_{BL}), the voltages are proportional to $\cos[(\omega_{BL} - \omega_{RF})t]$ and $\cos[(\omega_{BL} - \omega_{RF})t - \omega_{RF}\tau_g]$. Here,

$$\omega_{BL} = \begin{cases} 2\pi f_{BL}, & \text{upper band;} \\ 0, & \text{lower band.} \end{cases} \tag{2.2}$$

where $f_{BL} = 10$ GHz. Likewise, after a second stage of downconversion (at variable-LO frequency ω_{LO}) within the spectral line downconverter modules, the voltages are proportional to $\cos[\sigma(\omega_{BL} - \omega_{RF} - \omega_{LO})t]$ and $\cos[\sigma(\omega_{BL} - \omega_{RF} - \omega_{LO})t - \sigma\omega_{RF}\tau_g]$, where

$$\sigma = \begin{cases} +1, & \text{upper sideband;} \\ -1, & \text{lower sideband.} \end{cases} \tag{2.3}$$

represents the selection of the sideband. Finally, the selected 500-1000 MHz IF is digitized at 1 GS/s, which is below the Nyquist frequency, hence there is a further implicit downcon-

version at frequency $\omega_D = 2\pi f_D = 2\pi(1\text{ GHz})$, after which the voltages are proportional to $\cos\{[\omega_D - \sigma(\omega_{BL} - \omega_{RF} - \omega_{LO})]t\}$ and $\cos\{[\omega_D - \sigma(\omega_{BL} - \omega_{RF} - \omega_{LO})]t - \sigma\omega_{RF}\tau_g\}$. These voltage waveforms are the raw data that are recorded by the FST system. In the

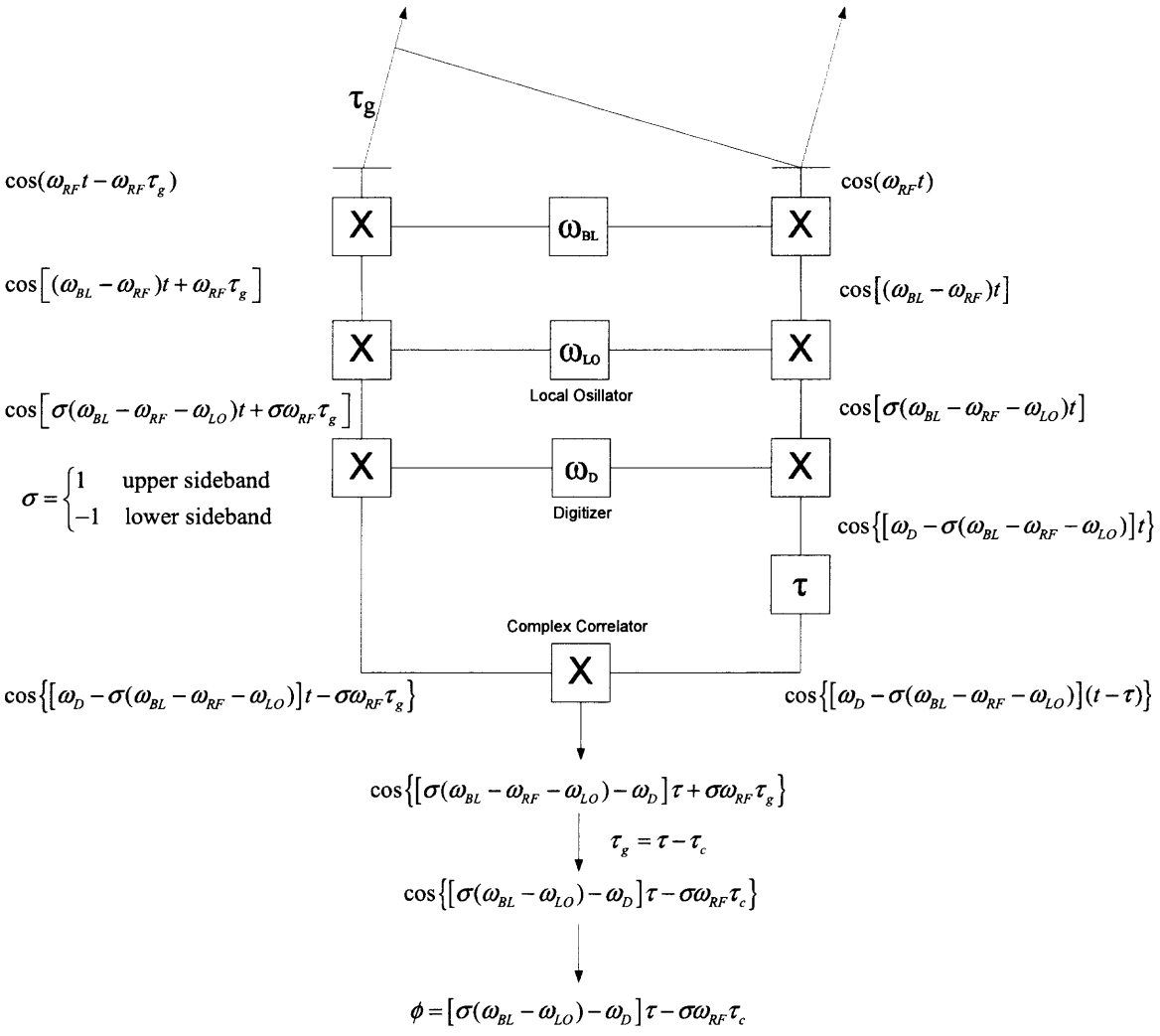


Figure 2.12 Origin of the phase correction needed due to two stages of explicit downconversion and one stage of implicit downconversion due to digitizer.

software correlator, the first channel is then further delayed by $\tau = \tau_g + \tau_c$, resulting in the final voltage waveform $\cos[\omega_D - \sigma(\omega_{BL} - \omega_{RF} - \omega_{LO})(t - \tau)]$ for that antenna. These signals are then correlated (multiplied and integrated), removing the dependence on t, but there remains the time- and frequency-dependent phase shift $\phi = [\sigma(\omega_{BL} - \omega_{RF} - \omega_{LO}) -$

$\omega_D]\tau + \sigma\omega_{RF}\tau_g$. Replacing τ_g with $\tau - \tau_c$, the terms involving $\sigma\omega_{RF}\tau$ cancel, yielding the expression

$$\phi = [\sigma(\omega_{BL} - \omega_{LO}) - \omega_D]\tau - \sigma\omega_{RF}\tau_c. \tag{2.4}$$

This expression becomes equation 2.5 upon replacing τ with $\tau_g + \tau_c$ and ignoring the constant term $[\sigma(\omega_{BL} - \omega_{LO}) - \omega_D]\tau_c$ (which becomes part of the complex gain correction, to be removed through calibration).

The above derivation gives the general equation for the phase correction, which is

$$\phi = (\sigma\omega_{BL} - \sigma\omega_{LO} - \omega_D)\tau_g - \sigma\omega_{RF}\tau_c \tag{2.5}$$

where τ_g is the geometric delay, ω_{BL} is the angular frequency of the block downconverter ($2\pi \times 10$ GHz), ω_{LO} is the angular frequency of the LO, ω_D is the angular frequency of the digitizer downconversion (2π GHz), and σ is -1 for lower sideband and $+1$ for upper sideband. However, an additional phase correction is needed due to the fact that the delays are implemented by shifting the digital data on each channel by an integral number of 1 ns samples, as discussed in § 2.8. The combined phase correction is obtained by replacing the true delay τ_g by the integer delay $\tau_1 = \text{round}(\tau_g)$ and applying a frequency-dependent term proportional to the non-integer part of the delay, $\delta = (\tau_1 - \tau_g)$:

$$\phi = (\sigma\omega_{BL} - \sigma\omega_{LO} - \omega_D)\tau_1 \pm \sigma\omega_{RF}\delta \tag{2.6}$$

Here the upper sign is for an IF in the 1-5 GHz band, and the lower sign is for the 5-9 GHz band, and ω_{RF} is the RF angular frequency ($2\pi f_{GHz}$). This calculated phase must be

subtracted from the raw phase.

The middle column of Figure 2.13 shows the phase of the three baselines and closure phase after the above phase correction. Note that the phases are nearly flat, but the baselines involving antenna 6 (upper two panels) are still drifting significantly with time over the 28-minute observation. This can be interpreted as due to a limitation on the precision of orbital positions provided by the TLE-derived coordinates (Chan & Navarro 2001, e.g.,). A best fit determination of the error between the phase center (TLE-derived coordinates) and the position of the satellite is 4.81 arcmin in hour angle, and 6.39 arcmin in declination. After correction for this pointing error, the phases are nearly constant over the observation, as shown in the right column of plots in Figure 2.13. Similar observations on other satellites indicate a typical error of order 10 arcmin. Unfortunately, such a large error precludes the use of GPS satellites as absolute phase calibrators when TLE data are used to determine the satellite position.

2.10 Observations of Solar Bursts

Prior to completion of FST, it was envisioned that the most effective science results would come from observations of solar type III bursts (due to beams of electrons travelling at speeds as high as $0.5c$ upward and downward in the solar corona). Type III burst emission is generated at the plasma frequency which is proportional to the square root of coronal electron density, based on Newkirk (1961) coronal density model, the height of the emission source can be estimated. Various authors have given the frequency drift rate of type III bursts at different frequencies (Staehli & Benz 1987; Allaart et al. 1990; Benz et al. 1992, 8.1 ± 4.3 GHz/s at 3.47 GHz, 11 ± 9 GHz/s at 6.0 GHz, 11.5 ± 8.5 GHz/s at 7.5 GHz).

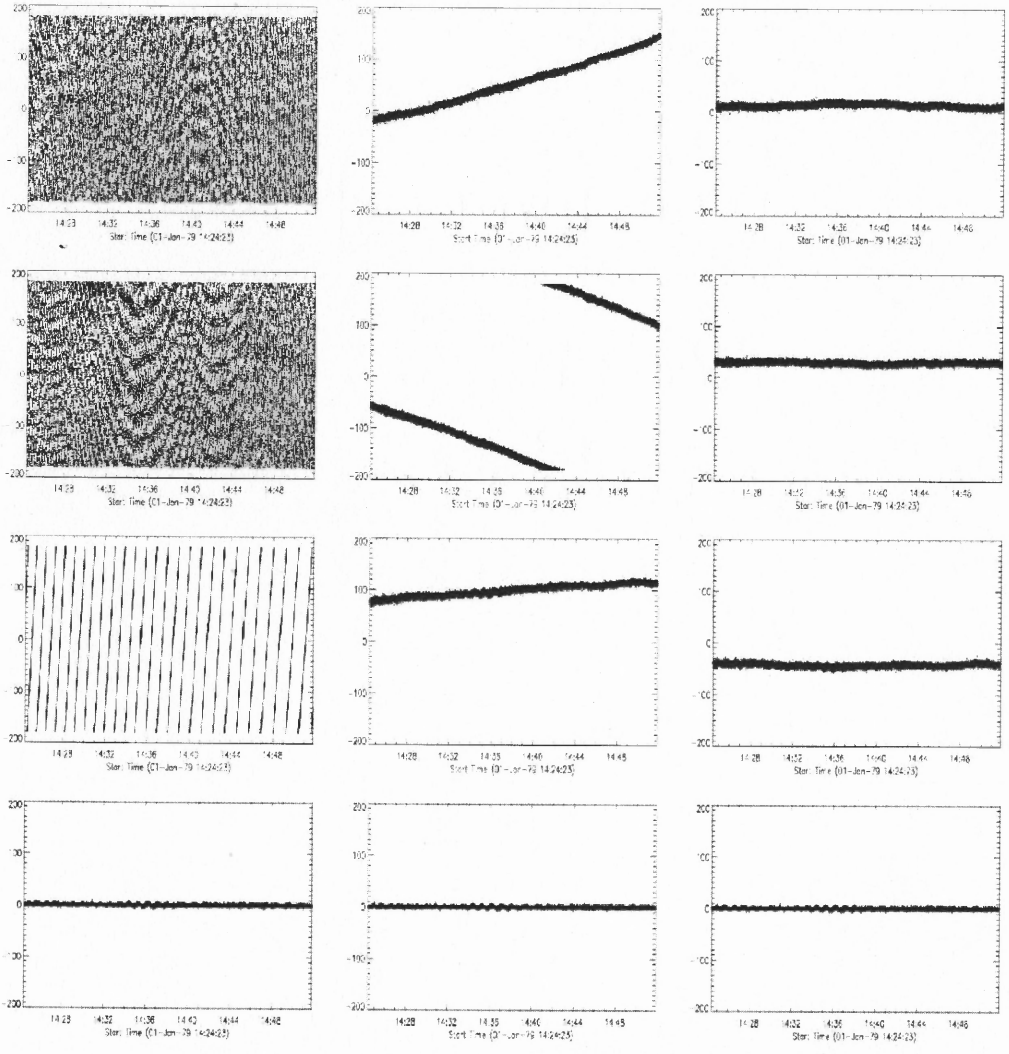


Figure 2.13 The four rows are correlated phases for three baselines and closure phases, respectively. The left column is the raw phase. The middle column is the phase corrected for fringe-stopping correction. The right column is the phases with fringe-stopping and refined estimate for satellite location.

Also the relation between duration, drift rate and observing frequency for the decimetric range (0.3-3.0 GHz) has been found (Meléndez et al. 1999) to be

$$t_{1/2} = 1.7 \times 10^4 f^{-0.60} \tag{2.7}$$

$$|df/dt| = (0.09 \pm 0.03) \times f^{1.35 \pm 0.10} \tag{2.8}$$

where $t_{1/2}$ is the half-power duration in ms, and f is in MHz. The half power duration ranges between 35 and 67 ms in the overlapping FST frequency range (1-3 GHz). This duration suggests that either Mode 2 or Mode 3 of Figure 2.5 should be used, both of which have 20-25 ms between acquisitions appropriate for resolving the bursts. After initial observations in Mode 2, Mode 3 was chosen as the optimum choice due to its elimination of the data transfer time gap, coupled with the fact that a precise time interval is not required between acquisitions. The ability of FST to measure phase (and hence positions) in two dimensions as a function of frequency should allow each type III's trajectory to be traced. Because FST can measure only one 500 MHz band at a time, and in a single polarization, it was decided to take data in the lowest frequency band (1-1.5 GHz), and to switch polarization alternately every few seconds. This setup has been used exclusively for the daily solar observations until recently (2006 November), to maximize the chance to observe type III bursts, which are most numerous at lower frequencies and may be highly polarized in either sense of circular polarization.

The first burst seen by the FST system, associated with a B5.1 X-ray class flare, was recorded in Mode 2 at 2030 UT on 2006 April 5. Figure 2.14 shows the correlated dynamic spectrum in both amplitude and phase. The gaps between the spectrum strips are

due to the dump time between acquisitions (see figure 2.5). The dynamic spectrum shows the presence of several type III bursts, some of which show a clear reverse drift behavior (drifting from low to high frequencies). That indicates the event is a type III burst generated by downward-going electron beams. The duration of individual bursts is $\sim$ 175 ms, and the drift rate is 700 MHz s^{-1}, which are consistent with the above empirical relations for type III burst duration and drift rate in the decimetric range (Meléndez et al. 1999). A full study of this event is underway, but one may immediately see from Figure 2.14*b* that all of the individual type III bursts have precisely the same phase behavior, whether seen in RCP or LCP. This implies that each burst follows the same trajectory (at least in the one dimension measured by the single baseline) within the roughly 10" spatial resolution of the observation. Four additional events have been recorded since the first event, none of which proved to be dominated by type III bursts. However, one of the events, a mere B7.0 X-ray class event seen on 2006 Aug 31, proved to be an extremely strong and rich event lasting nearly 2 hours and showing spike bursts (Bastian et al. 1998), weak type III bursts, fiber bursts (Kuijpers & Slottje 1976; Aurass et al. 2005), and strong continuum activity. Figure 2.15 shows a short 30 s period during the burst, showing fiber burst activity. This burst is also under study, and promises to provide the first spatially resolved observations of these enigmatic drifting bursts. The radio diagnostics available from the high time- and frequency-resolution FST data, coupled with the unique interferometric spatial information, are a powerful combination provided by this FASR-prototype instrument that shows a small part of the capability of the complete FASR.

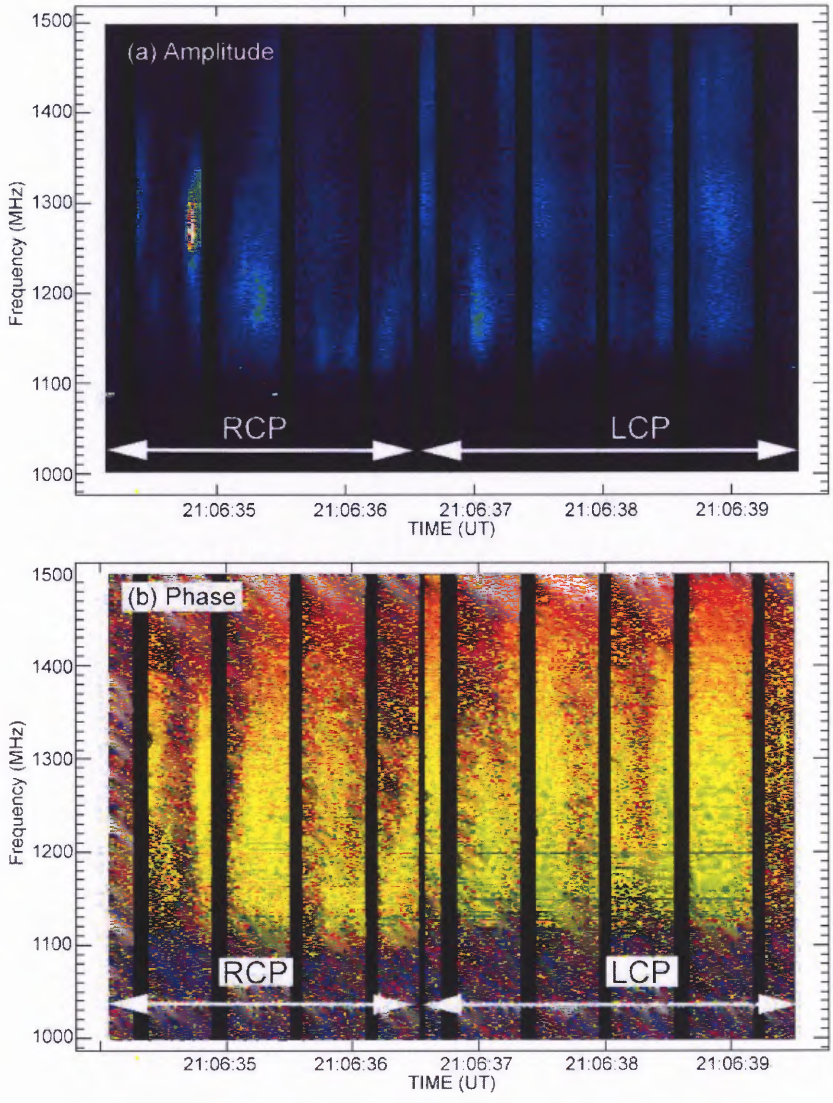


Figure 2.14 Correlated amplitude and phase of a type IIIdm burst with reverse drift indicative of downward-going electron beams. The phases are relative to the pre-flare (active region plus system phase). The polarization changes from RCP to LCP at the point shown, but the phases are unchanged, indicating that both polarizations arise from the same location to within about 10 arcsec. This suggests a low degree of polarization, as expected for second-harmonic plasma emission.

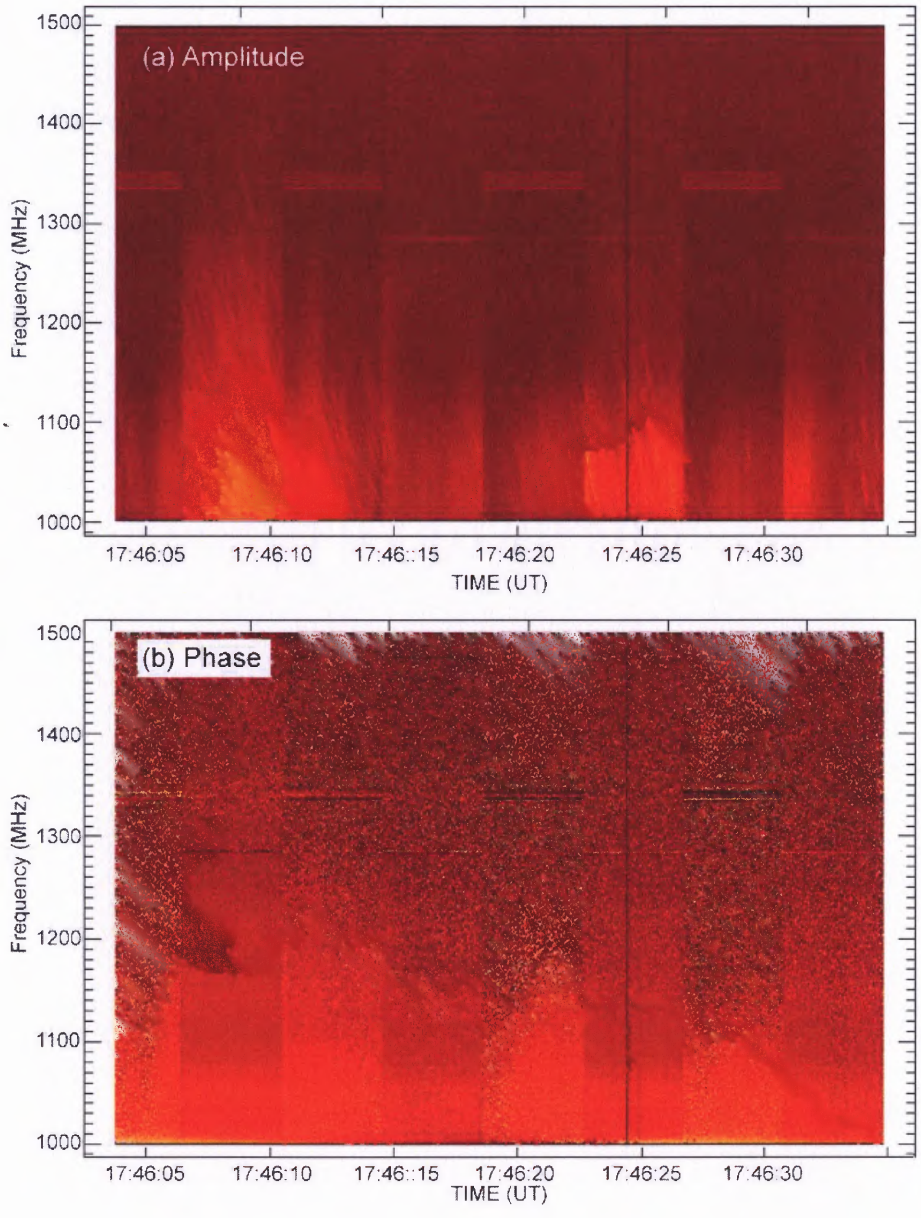


Figure 2.15 Correlated amplitude and phase for a short period (30 s) of a long-lasting (nearly 2 h) type IV continuum solar burst with fiber bursts superimposed. This is a period of emission in both RCP and LCP, although most of the burst was essentially 100% LCP. The fiber bursts are the prominent features, each drifting from as high as 1300 MHz to the low-frequency limit of 1000 MHz over about 3-4 s. There is little phase (i.e., position) difference between the type IV continuum source and the fiber burst emission, but more careful study is underway to determine both the continuum and fiber burst locations as a function of frequency.

2.11 Conclusions

The architecture of the FASR Subsystem Testbed has been described. For the first time it provides the capability to observe the Sun with three dimensional (temporal, spectral and spatial) information simultaneously for a broad bandwidth (500 MHz). As a prototype for FASR, FST also provides the opportunity to study the design, calibration and interference-mitigation requirements of FASR. The initial observational results demonstrate the extent to which satellite observations can be exploited to calibrate the system. Observations of geosynchronous satellites provide precise determination of the delay center and frequency calibration, but neither geosynchronous nor GPS satellites are suitable for phase calibration unless their positions are better known than those calculated from the NORAD Two Line Element sets. All observations and tests show that FST performs its expected function extremely well. Several solar bursts captured by the FST system nicely demonstrate the promise of FST, and hence that of the future FASR, for further solar observations. The instrument will continue daily solar observations, with the expectation that it will allow one to trace type III electron trajectories as well as locate the source positions of other decimetric bursts, applicable to the height range in which the primary energy release of solar flares is expected to take place. The huge amount of raw data makes the daily operation a rather challenging data-management problem, and the operations model will break down during solar maximum when flares are more numerous, unless more extreme data-selection can be done.

Since the correlated spectra are derived through offline software from full-resolution data, FST provides the flexibility to process the data multiple times with alternate algorithms and parameters, to simulate various hardware designs using real data. It is this

unique capability that makes FST so valuable as a testbed instrument. In addition, FST observations provide an excellent database for testing RFI mitigation algorithms under real-world conditions such as one might expect for FASR. The study shows that the estimator based on the frequency domain statistical analysis appears to be more stable and less re-source consuming than the one based on the time-domain Kurtosis (Kenney & Keeping 1962). Further details will be presented in Chapter 4. A number of studies are under-way using the first six months of data from the instrument, and there will be continued observations in the future.

CHAPTER 3

STUDY OF SOLAR BURST TRAJECTORIES

3.1 Introduction

Since the first solar burst captured with FST on 2006 April 5 (Gary et al. 2006), more than 12 events during six months operation of FST were observed. The initial expectation was that type III bursts would offer the best opportunity to find trajectories as described in Chapter 1, capitalizing on the FST capability of obtaining temporal, spectral, and spatial information simultaneously. And it is anticipated that most solar events in the observing band (1.0–9.0 GHz) should occur in the lowest 500 MHz band (1.0–1.5 GHz). The system was set up in mode 3 (see Chapter 2) during the six months observing period. Table 1.3 summarizes all of these events. Several decimetric type III bursts were captured as expected, but it was not expected that fiber bursts would be observed, which turn out to be even better for finding trajectories due to their slower drift rate.

Fiber bursts also called intermediate drift bursts, are sometimes observed as a fine structure during solar type IV radio burst, although it was surprising to find them in all three type IV events listed in Table 1.3. They show drift rates intermediate between shock driven type II and nonthermal electron beam driven type III bursts and frequently show an absorption edge at the low frequency side of the emission stripe.

The main characteristics of single fibers are as following (Benz & Mann 1998; Kuijpers 1975):

- The drift rate is roughly one order of magnitude large than shock driven type II bursts at the same frequency and one order of magnitude smaller than beam driven type III bursts.

- The drift rate is negative and the absolute value of the drift rate of a single fiber increases with increasing frequency.
- The instantaneous bandwidth is about two percent of the observing frequency.
- Typically there is an absorption ridge at the low frequency edge of the emission ridge, and they begin and end simultaneously.

Fiber bursts were first reported by Young et al. (1961), followed by reports by Slottje (1972), Elgaroy(1973), and Bernold(1980). Kuijpers (1975) first proposed the whistler wave model in which the fibers are generated by coalescence of packets of whistler waves with Langmuire waves into escaping electromagnetic waves. The frequency of whistler waves determined by the frequency difference between the emission ridge and absorption edge can be used to calculate the local magnetic field in the corona. Treumann & Bernold (1981) developed an alternative theory that suggested the fiber bursts originate by an interaction between whistler solitons and Langmuire waves. One difficulty for this model is that whistler solitons can only exist for $\omega_w > 0.25\omega_{ce}$ (ω_w is whistler frequency, ω_{ce} is electron cyclotron frequency) under coronal circumstances (Karpman & Washimi 1977; Spatschek et al. 1979; Mann 1985). The magnetic field strengths in corona derived from this model are lower than expected, and required the plasma-beta $\beta \geq 1$ (Dulk & McLean 1978). In this thesis, the whistler wave model by Kuijpers (1975) was adopted to interpret the fiber bursts.

3.2 Fiber Burst Whistler Wave Model

The whistler wave model of fibers states the coalescence of whistler wave packets and Langmuire waves should lead to an increase in emission (the emission ridge) at $\omega_{pe} + \omega_w$ and decrease in emission (the absorption ridge) at the local electron plasma frequency ω_{pe},

where ω_{pe} is the electron plasma frequency which is defined by

$$\omega_{pe}^2 = \frac{4\pi n_e e^2}{m_e} \tag{3.1}$$

(n_e–electron number density, m_e–electron mass, e–elementary charge). The absorption ridge is not always observed. Elgaroy (1982) found that the instantaneous bandwidth of the emission ridge has the same value as the frequency distance between the both ridges. Therefore one could use the instantaneous frequency width of the fiber as a good estimation of whistler wave frequency when the absorption ridge is not present.

With the whistler wave model, numbers of coronal physical parameters can be estimated with the observed quantities of fiber bursts, such as fiber burst frequency, frequency drift rate, and emission ridge width. One of impressive applications for this model is that it can be used to estimate the local coronal magnetic field strength B in a flaring loop.

Following Aurass et al. (2005), The whistler wave group velocity along the magnetic field lines is

$$v_g = 2c\frac{\omega_{ce}}{\omega_{pe}}\sqrt{x(1-x)^3} \tag{3.2}$$

where

$$\omega_{ce} = \frac{eB}{m_e} \quad x = \frac{\omega_w}{\omega_{ce}}$$

From the definition of the drift rate,

$$D_f = \frac{df}{dt} = \frac{1}{2\pi}\frac{d\omega_{pe}}{dt}$$

thus,

$$D_f = \frac{1}{4\pi\omega_{pe}}\frac{\omega_{pe}^2}{dt} = \frac{\omega_{pe}}{4\pi}\frac{1}{n_e}\frac{dn_e}{ds}v_g \tag{3.3}$$

where d/ds represents the derivative along the magnetic field direction, the expression

$$x(1-x)^3 = (\frac{2\pi m_e}{ce})^2(\frac{\lambda_s D_f}{B})^2 \tag{3.4}$$

is obtained immediately, where

$$\lambda_s = \frac{\lambda_r}{\cos\theta} = (\frac{\cos\theta}{N_e}\frac{dN_e}{dr})^{-1} \tag{3.5}$$

λ_r is the local scale height. θ is the inclination between magnetic field line and the electron density gradient. With the proper solar atmosphere electron density model, the coronal local magnetic field can be easily retrieved by combining Equations 3.2, 3.3, 3.4.

3.3 Observation

The large solar flare of 2006 December 6 was detected by the newly constructed FST system (Liu et al. 2007). The event was associated with a soft X-ray class X6.5 solar flare and lasted for more than one hour. Figure 3.1 shows the single frequency flux density at 1415 MHz. The strength of this solar radio burst is 340,000 solar flux units (sfu–1 sfu=10^{-22} $Wm^{-2}Hz^{-1}$). Actually at other frequency the strength is near 1 million sfu. For about 1 minute near 18:47 UT, This record setting decimetric continuum (type IV) event shows intermediate drift fine structure (also called fiber bursts) that drifts from high to low frequencies over 6–10 s. The FST system, operating in mode 3, saved the digitally

sampled raw data, which lets us analyze the data with any possible frequency resolution up to 2/T, where T is the duration of a sample (100 μs in mode 3). Figure 3.2 shows the correlated amplitude and phase dynamic spectrum with a resolution of 0.98 MHz in frequency and 20 ms in time. Notice that the alternating change of amplitude is due to the changing of polarization between RCP and LCP every 4 seconds. The fiber burst occurred at the starting phase (18:46 UT) of the continuum which is rare according to Benz & Mann (1998).

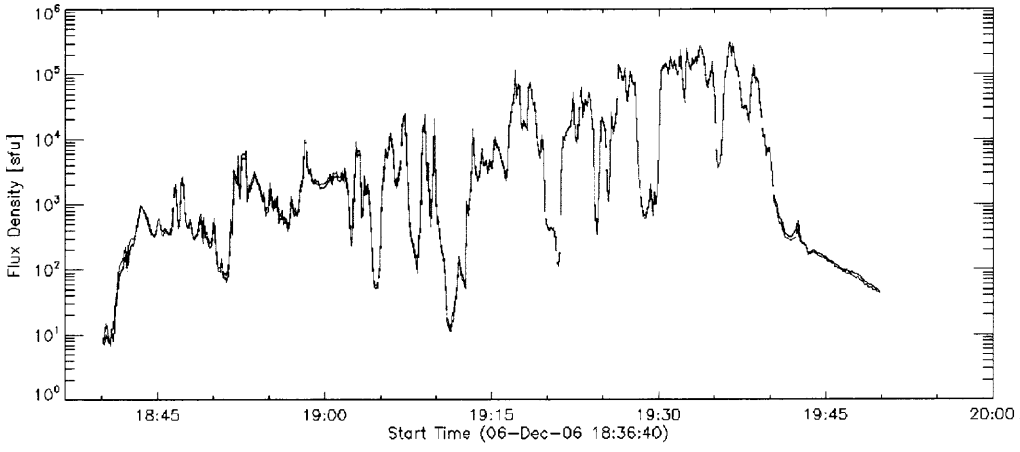


Figure 3.1 The calibrated lightcurve of 2006 December 6 event using FST, in right-hand circular polarization.

For the single fibers, the burst drift peak was fitted by second order polynomials. From the fitting function $f = F(t)$ the fiber burst frequency drift rate is given by the first order time derivative. Figure 3.3 shows the drift rate as a function of frequency. The frequency almost spans the whole observing band (1.0–1.5 GHz). The drifting rates of fiber bursts were found to be -70– -25 MHz/s, the absolute value of which are lower than the extrapolated mean values (Benz & Mann 1998) by a factor of 2, based on mostly on

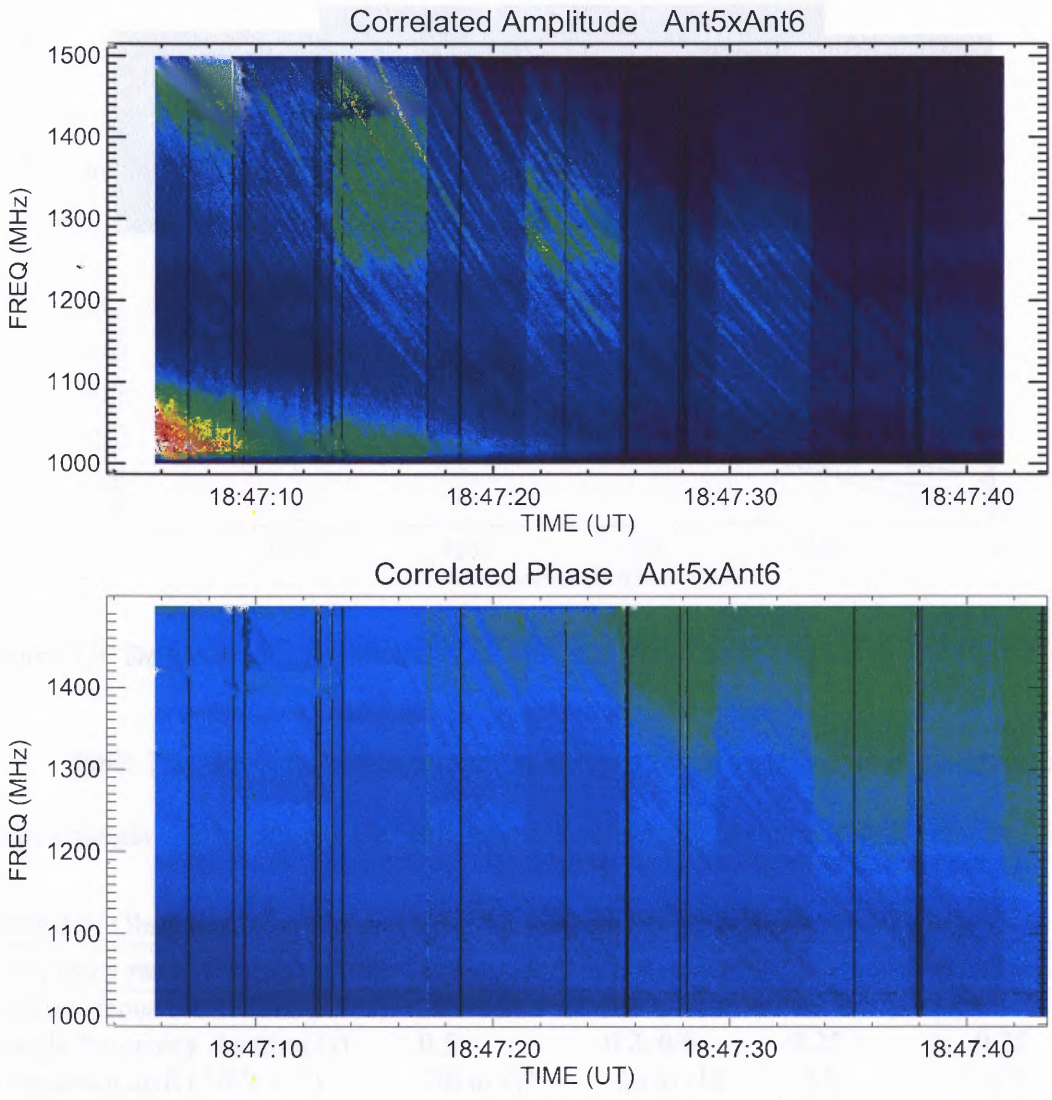


Figure 3.2 Correlated dynamic amplitude and phase spectrum of the fiber burst on 2006 December 6 with a resolution of 0.98 MHz in frequency and 20 ms in time.

observations at lower frequency.

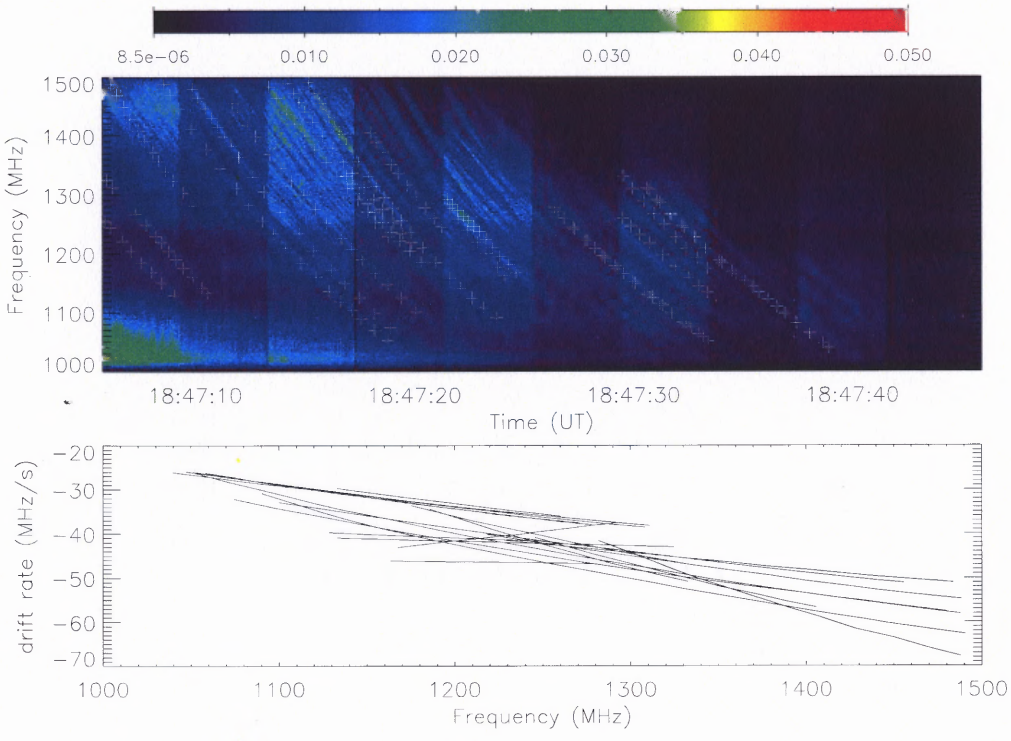


Figure 3.3 Drift rate of single fibers.

Table 3.1 shows the observational characteristics of fiber bursts at different frequency ranges. Interestingly the instantaneous bandwidth, frequency drift rate, and frequency extent all increase with increasing frequency, but the duration of single fiber frequency at a given frequency stays at around 0.2–0.6 s.

Table 3.1: Observed Fiber Bursts Characteristics at Different Frequency Range

Frequency range (*MHz*)	$1500-1000$	$950-500^{(a)}$	$340-310^{(b)}$	$175-145^{(b)}$
Instantaneous bandwidth (*MHz*)	20	10	1.75	0.5
Single frequency duration (*s*)	0.5	0.2–0.6	0.25	0.25
Frequency drift ($MHz\ s^{-1}$)	-70 to -25	-50 to -10	-9.5	-3.5
Frequency extent (*MHz*)	350–500	50–150		

[a] Young et al. (1961)
[b] Elgaroy (1973)

3.4 Analysis and Discussion

The purpose is to get the fiber burst location information from FST data of 2006 December 6. Figure 3.2 shows the correlated amplitude and phase spectrum of fiber bursts for the first baseline (5×6). Notice that the fibers are also visible in the correlated phase spectrum. One can obtain the location information from the phase spectrum on three baselines. Because of the lack of phase calibration information across the whole 500 MHz band, it is impossible to get the absolute location for the fiber burst sources. The next best thing to be done is to retrieve the drifting burst source locations relative to the background emission. The basic idea is to subtract the phase of background continuum emission from the correlated baseline phase in order to subtract the unknown system phases of the three antennas. Essentially active region emission is a good choice as background as long as one can treat active region emission as point source. However, for the active region before the flare, it is found that the phase closure of three baselines was not zero, the biggest derivation is almost 60 degrees, which is probably because the active region emission is not strong enough compared with the whole solar disk, so one cannot treat the active region as a point source. Apparently there is no automatic way to subtract the background phase. Instead, one has to manually select the background region for each peak emission along the fiber burst trajectories.

Figure 3.4 gives the waterfall plot of the fiber burst. The abrupt jump at time 18:47:13 is due to the alteration of circular polarization. Figure 3.5 gives the amplitude evolution along a single fiber. The peak emission of one single fiber is marked with an asterisk. One can also select the background manually, corresponding to each peak emission. The background selection should be unbiased which seems to be very hard to achieve. No-

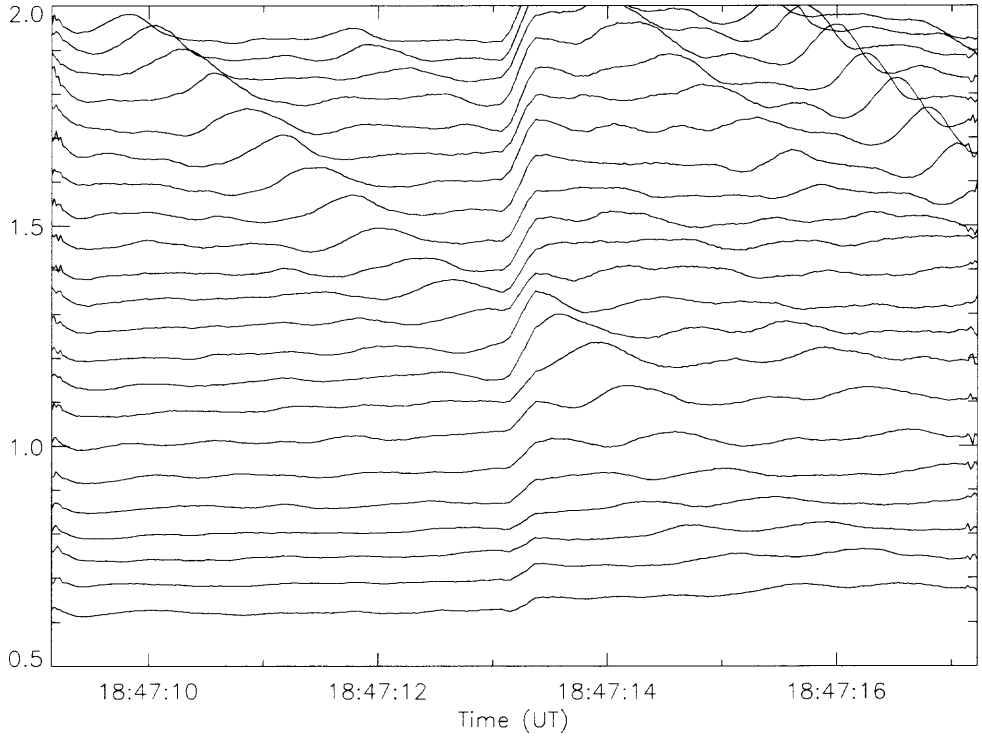


Figure 3.4 The waterfall plot of the fiber burst. The left half plot is in LCP, the right half plot in is RCP.

tice that there is a valley before every peak curve, which should be the absorption edge at the low frequency side of emission stripe. One should carefully select the background between the emission and absorption edge. Figure 3.6(a) shows the particular fiber analyzed, where the '+' signs mark the emission peak location. The data were reanalyzed with lower frequency resolution (15.6 MHz) and longer integration time (200 ms) for smoothing the noisy signal. Figure 3.6(b) illustrates the three baseline correlated phase evolution and phase closure with frequency after background subtraction.

One of the indications that the background selections are reasonable is that the phase should not have the abrupt jump and the phase closure should be around zero.

For each baseline, the phase variation indicates the amount of the emission source shift perpendicular to the baseline fringe direction. Based on three baselines phases, one can calculate the shift relative to the background. The three shifted fringes should join to one point if the closure phase is zero.

The trajectory of a single fiber using the above method is illustrated in Figure 3.7. The '+' signs are the locations relative to the background emission, which are determined using the techniques mentioned above. The color circles around the '+' signs are an indication of the phase errors. The drift is from blue to red (high to low frequency). It is obvious that the trajectories are not the absolute positions, in which this fiber burst occurred on the Sun. However, the relative positions with respect to the background emission are of major interest. The trajectories maybe stretched and shifted based on the background properties, but the burst drift topology remains the same. The reconstruction of source locations in three dimensions need two conditions: absolute position and the coronal density model. Assuming the fiber emission is at the plasma frequency ($\omega_w \ll \omega_{pe}$), the electron density is

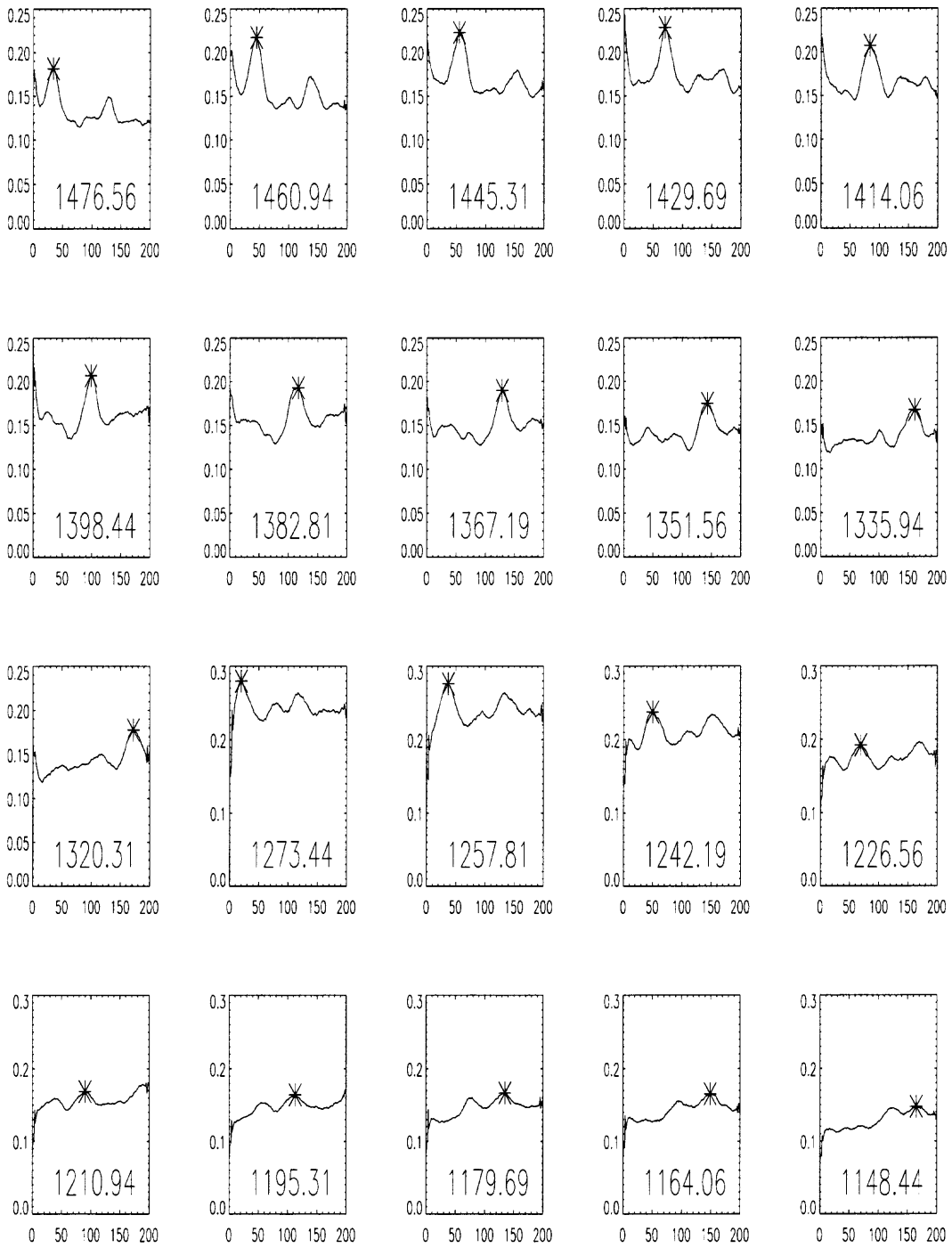
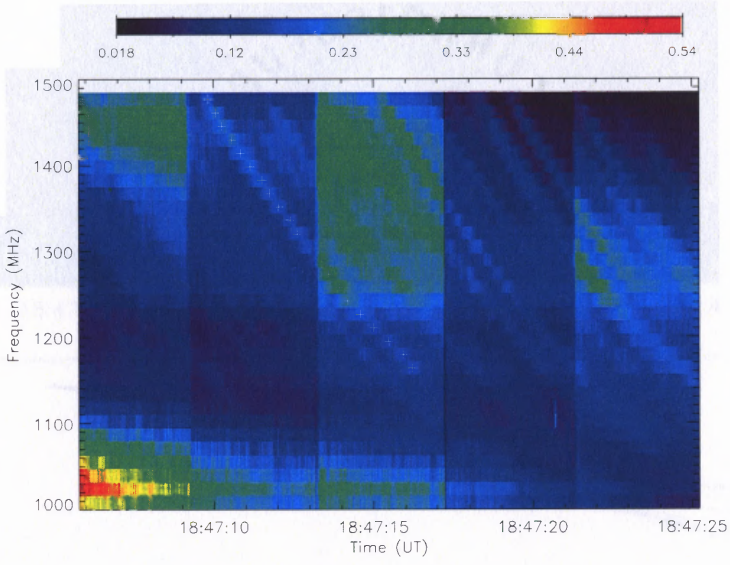
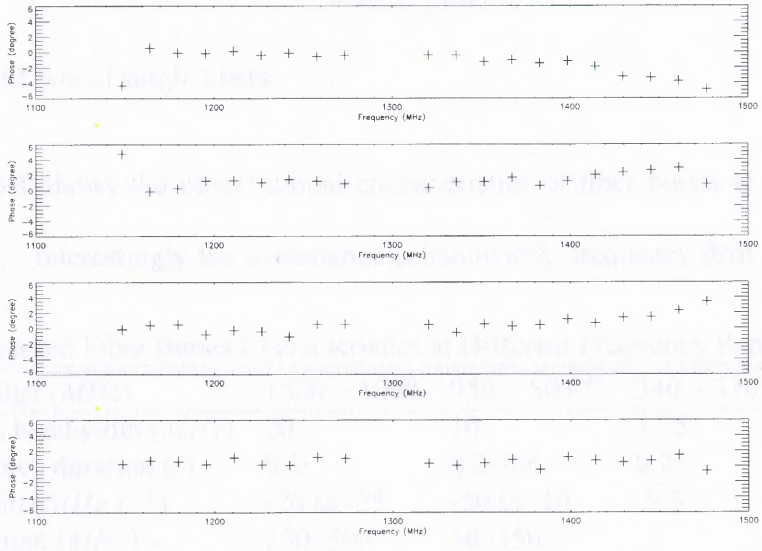


Figure 3.5 The amplitude evolution along the single fiber. The x axis is time sample in 20 ms units, y axis is amplitude in arb. units. The number labeled in each window is the frequency for that plot.



(a)



(b)

Figure 3.6 Fiber burst dynamic spectrum with 15.6 MHz frequency resolution (blockns=64 ns) and the location of single fiber burst in the spectrogram.

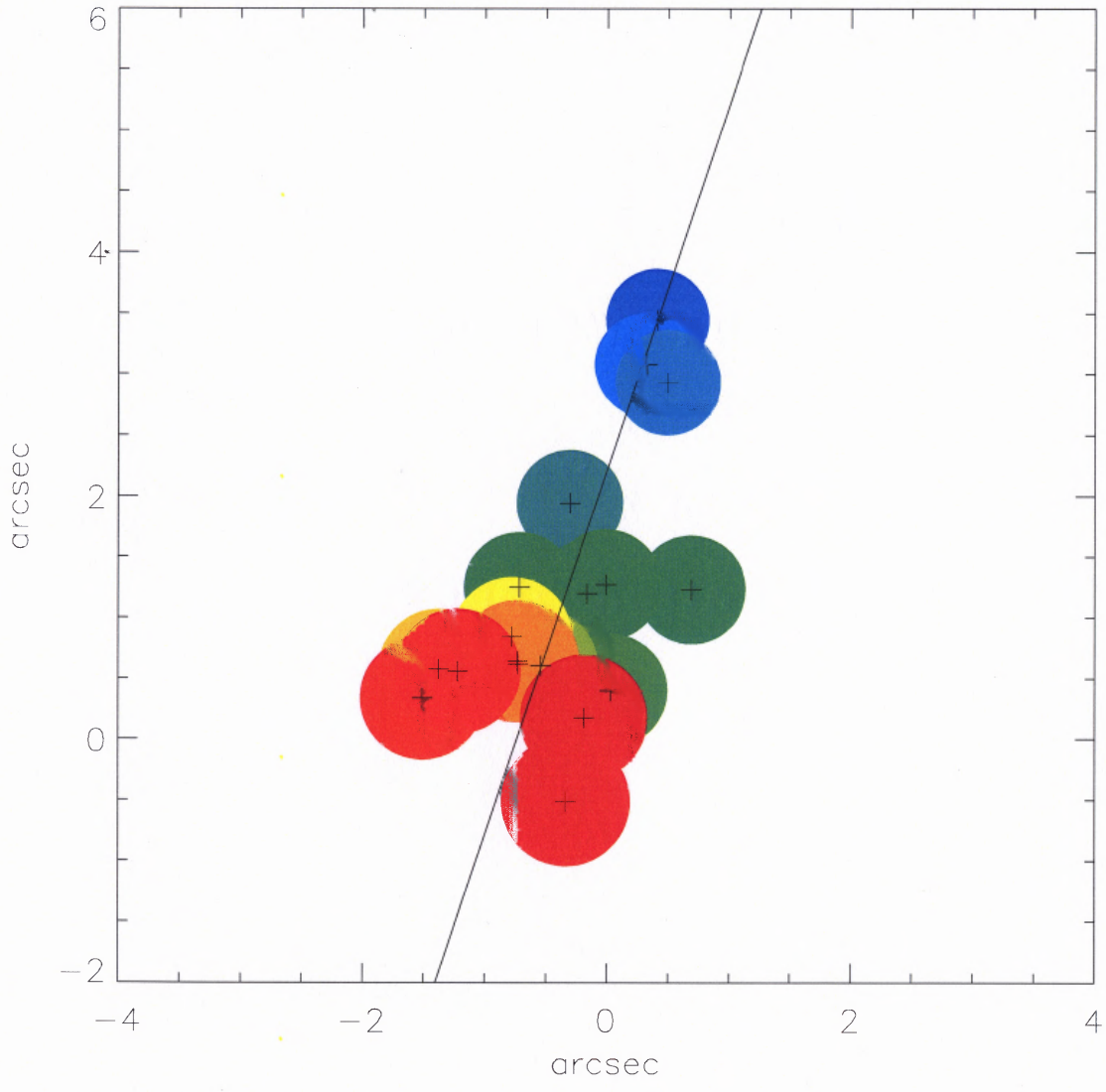


Figure 3.7 Single fiber burst drift trace relative to the background emission, the color circle around each point is the uncertainty based on three baselines error bar, the drift is from color blue to red, the black solid line is roughly the average drifting direction.

determined by Equation 3.1, and the height of fiber source can be estimated using a coronal density model. Three different solar atmosphere models have been compared, which are the exponential atmosphere model, Newkirk (1961) active region density model, and Lang (1980) density model with the atmosphere in hydrostatic equilibrium with thermal conductivity. Those density models are given as following,

$$n_e(h) = n_0 \exp\left(-\frac{h-h_0}{R_s}\right) \tag{3.6}$$

$$n_e(h) = \alpha N_0 10^{4.32R_s/(h+R_s)} \tag{3.7}$$

$$n_e(h) = n_0\left(\frac{h}{h_0}\right)^{-2/7} \times \exp\left\{-\frac{7h_0}{5H_n}\left[1-\left(\frac{h}{h_0}\right)^{-5/7}\right]\right\} \tag{3.8}$$

where h is the height above the solar photosphere, R_s is the solar radius, n_0 and h_0 are reference values of electron density and corresponding height, α is the geometry related quantity for Newkirk (1961) model, and H_n is the scale height of the exponential density model. The reference values were chosen to be $H_n = 7.5 \times 10^9$ cm^{-3}, $n_0 = 3.36 \times 10^8$ cm^{-3} and $h_0 = 2.21 \times 10^{10}$ cm^{-3} (Trottet et al. 1982; Suzuki & Dulk 1985; Paesold et al. 2001). The resulting electron densities from these three models are illustrated in Figure 3.8 The only reasonable result is given by Lang (1980) density model since none of the density models except Lang (1980) model can reach to the density 10^{10} cm^{-3}.

The frequency range of this fiber is between 1.16-1.46 GHz, which is the density range: 1.68×10^{10} cm^{-3} to 2.65×10^{10} cm^{-3} according to the Equation 3.1. According to the Lang (1980) model, the sources at 1.16 and 1.46 GHz emission correspond to heights of 9.09×10^9 cm and 8.42×10^9 cm, respectively. The travelling time is 6.15 s. The radial component of the average travelling velocity along a field line is about 1.09×10^3 kilometer

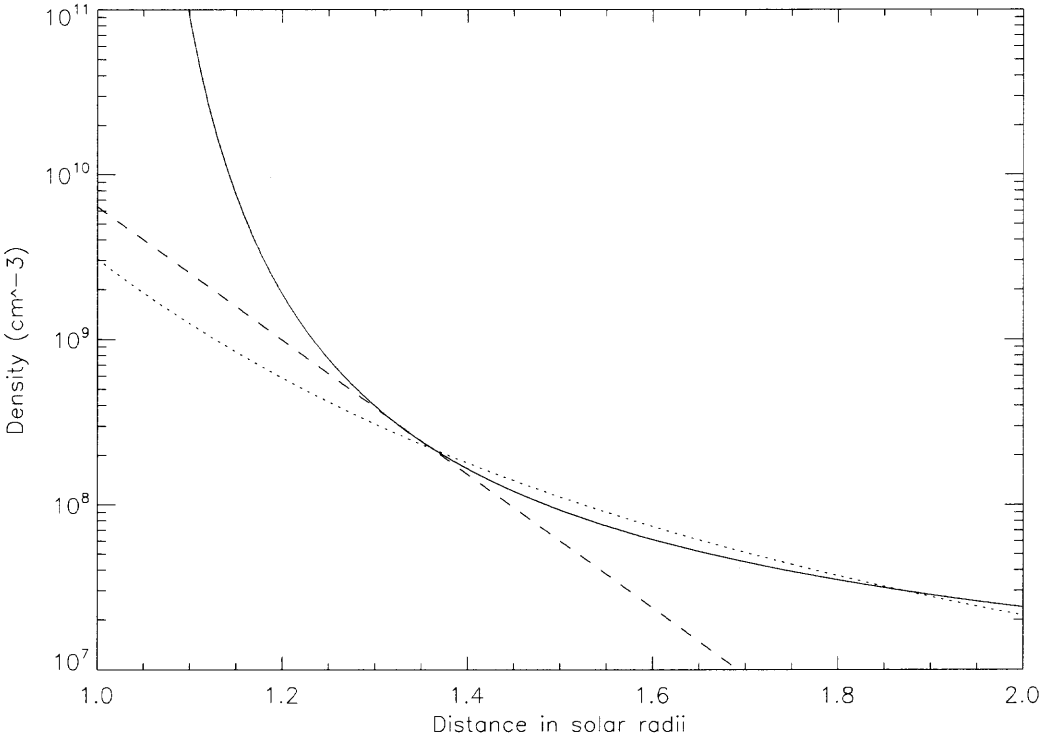


Figure 3.8 Coronal electron density vs. distance from the Sun center. The scale height and reference values of density and height are selected according to Paesold et al. (2001). α has been chosen to be 3.5 (Aurass et al. 2005). The solid line is the Lang (1980) model, the dotted line is the Newkirk (1961) model, and the dash line is exponential density model.

per second. Moreover, the radial velocity as a function of height can be obtained based on the FST observation and electron density model. The radial velocity at different height is shown in Figure 3.9. The predicted velocity slows down slightly with increasing height, which may be partially due to the deviation from radial direction with increasing height.

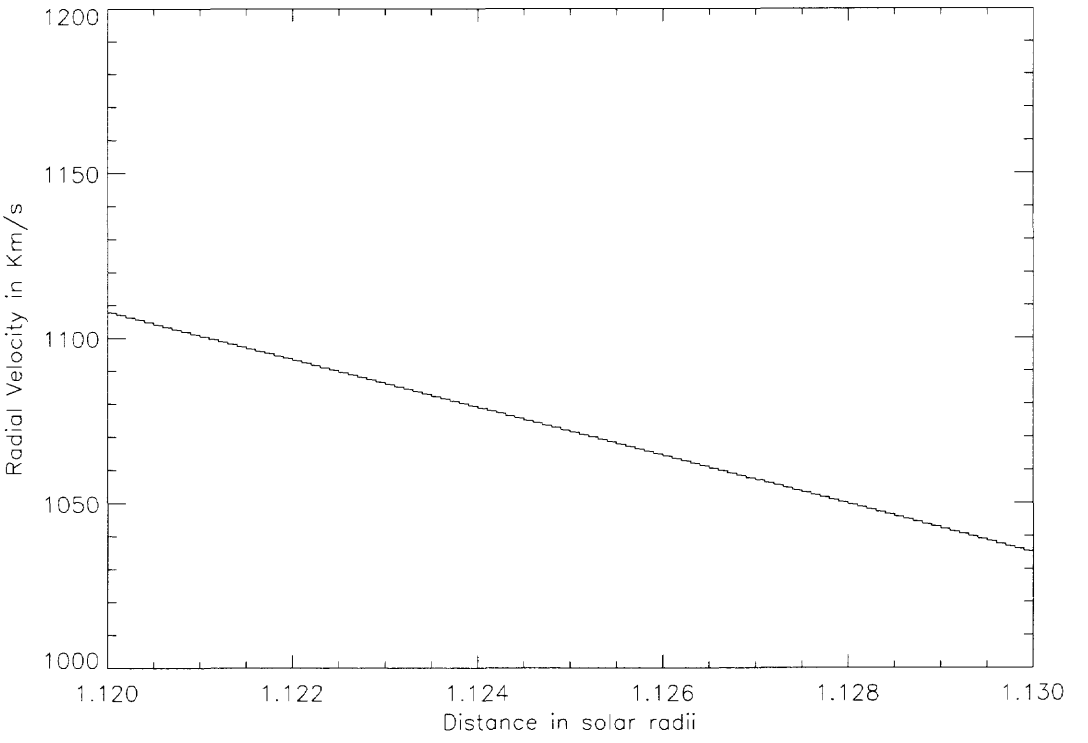


Figure 3.9 Fiber source radial velocity vs. distance in solar radii.

These velocities can be combined with fiber source location information to determine the group velocity in Equation 3.2. It is an advantage to measure the velocity directly contrasting with other methods. Since the velocity are measured in the whole drifting frequencies range, it is not necessarily to make such assumption, "If a fiber burst passes two neighboring NRH frequencies we demand that the corresponding source sites are uniquely connected by the same magnetic field line" (Aurass et al. 2005). Once these group velocity

are measured, the local magnetic field strength can be easily determined by Equation 3.2.

$$v_g = 2c\frac{\omega_{ce}}{\omega_{pe}}\left[\frac{\omega_w}{\omega_{ce}}\left(1-\frac{\omega_w}{\omega_{ce}}\right)^3\right]^{1/2} \tag{3.9}$$

using the measured values as, whistler frequency is about 20 MHz, $\omega_{ce} = 2\pi \times 2.8B$ in MHz, and assuming the group velocity deviate from electron density gradient by 45°, the magnetic field strength values are found to be 9–10 G for this single fiber.

Meanwhile, one can also examine another popular fiber burst model, whistler soliton model (Treumann et al. 1990), which considers the fiber bursts as plasma emission modulated by whistler solitons. For this model, the fiber source velocity is $\sim v_A$, v_A is Alfvén speed defined by

$$v_A = \frac{B}{(\mu\rho)^{1/2}} \tag{3.10}$$

where μ is the magnetic permeability, B and ρ are the local magnetic field strength and density, respectively. It maybe written numerically as $v_A = 2.8 \times 10^{12} B \cdot n^{-1/2}$ km s^{-1}, with B in Gauss and n in cm^{-3}. Assuming the angle between source velocity and its radial component is 45°, thus, the magnetic field is about 71 to 90 G.

Unfortunately, the absolute location can not be retrieved due to the lack of phase calibration information. This is a problem that remains to be solved for FST, because of the lack of point calibration sources bright enough to be measured with only three 2-m antennas. One possible solution is to use GPS satellite for phase calibration as long as the more accurate satellite position are known (ref. Chapter 2). Calibration will not be a problem for FASR with of order 100 of these antennas, since the accurate position can be easily retrieved by hundreds of baselines. Despite the uncertainty in absolute position

of the fibers, the measurement and analysis above demonstrate the technique to locate the burst position once the phase calibration information is provided. This work also shows the promise of FST and even more so for FASR, to provide detailed physical information on the density and magnetic field structure of flaring loops through the tracking of fiber bursts and other drifting bursts.

3.5 Conclusion

The large solar flare of 2006 December 6 was detected by the newly constructed FST system. This record-setting burst produced an especially fine set of fibers bursts that drift from high to low frequencies over 6-10 s across the whole observing frequency band (1.0–1.5 GHz) with frequency drift rate -70– -25 MHz/s. With this three element interferometer, a technique to determine the fiber burst source locations relative to the background in a large bandwidth was developed. In order to obtain the absolution locations information, phase calibration across the 500 MHz is needed. Combining with Lang (1980) density model, the burst radial travelling velocity is derived. Local magnetic field B is calculated to be 9–10 G assuming whistler wave model (Kuijpers 1975) and 71–90 G assuming whistler soliton model (Treumann et al. 1990). The accuracy of these quantities can be largely improved when the phase calibration problem is solved. The problem of lack of phase calibration information remains to be solved for FST in order to determine the absolute burst location and further provide the local density, velocity and magnetic field information. The technique to determine the trajectories of fiber burst was discussed earlier. This method can be used for other drifting bursts, such as type III burst and further examine the extensively used theoretical models in solar coronal. It shows the promise of FST and even more so for FASR.

CHAPTER 4

EXCISION OF RADIO FREQUENCY INTERFERENCE

4.1 Introduction

The OVRO site is relatively radio quiet, but there are local sources of interference in the L band, and near 6 GHz. In addition, satellite signals and signals from surveillance radars for air traffic control are always present. The quality of radio astronomy scientific data may be greatly affected by radio frequency interference (RFI) contaminating the natural radio emission produced by astrophysical objects. With the increasing demand for wireless communication systems operating in frequency bands of scientific interest, the RFI environment becomes ever more hostile. At the same time, astronomers seek to observe over an ever broader part of the radio spectrum. These trends make it necessary to find efficient and robust methods able to discriminate and excise the RFI contamination, while preserving as much of the underlying useful information as possible.

The techniques in RFI mitigation may be broadly categorized according to whether the excision is based on time-domain or frequency-domain analysis, and on whether a reference signal is available to provide adaptive noise cancellation (Fridman & Baan, 2001). The adaptive cancellation approach uses a reference channel to receive only the interference and then subtracts interference from the received data. If the interference is present for only a small fraction of the time, one can use time blanking to excise time intervals that contain interference. Or, if the interference is relatively narrowband in nature, one can use frequency blanking, which excises the frequency band containing interference. There is no universal solution for the problem of RFI mitigation, and one must find or develop

methods that are most suitable for a particular instrument and the RFI environment in which it operates (Fridman & Baan 2001). Among these methods, RFI mitigation algorithms based on higher order statistics in both the time domain (Ruf et al. 2006) and the frequency domain (Fridman 2001; Fridman & Baan 2001) have become increasingly popular in recent years due to the feasibility of their implementation using field programable gate arrays (FPGA), with which many modern digital instruments are equipped.

Based on the fact that system and radio source noise is thermal noise which has a gaussian distribution and a zero mean, whereas signal sources other than thermal noise have non-gaussian distribution, high order of statistics can be efficient in discriminating gaussian and non-gaussian signals. In the time domain, the fourth moments of signal $x(t)$ divided by the square of the variance of the probability distribution of x is defined as the kurtosis of x.

$$K = \frac{\langle x^4(t) \rangle}{\langle x^2(t) \rangle^2} \tag{4.1}$$

For the gaussian distribution, one can get the kurtosis using the generating function for higher order moments of a gaussian distribution. The generating function is given by

$$\langle x^n(t) \rangle = 1 \cdot 3 \cdot \ldots \cdot (n-1)\sigma^n \tag{4.2}$$

where σ is the standard deviation of signal x. The kurtosis of a gaussian distribution should equal three. In practice the kurtosis estimator has a variance based on the finite number of signal samples. In the following, K represents kurtosis based on the parent distribution, and $\widehat{K}$ represents a kurtosis estimator based on a finite subset of samples taken from that

distribution. The variance of the kurtosis estimator is (Kenney & Keeping 1962)

$$var\{\widehat{K}\} = \frac{24}{N} \tag{4.3}$$

where N is the number of signal samples, which can be replaced by the product of the signal bandwidth B and the available integration time (Ruf et al. 2006). Practically, the signal first needs to be channelized with a polyphase filter or filter bank, then the kurtosis of the time domain signal for each subband can be measured. Ruf et al. (2006) successfully developed a microwave radiometer detector with the capability of identifying RFI using the time domain kurtosis algorithm. But it can only handle eight subbands of total 24 MHz bandwidth. That is considerably deficient for broadband solar radio interferometry, especially for the coming FASR radio telescope, which will employ 4096 subbands over 500 MHz.

Although relatively new in the context of RFI mitigation for astrophysical radio instruments, the idea of using higher order statistics in the spectral domain was originally proposed more than two decades ago by Dwyer (1983), in the context of underwater acoustic signals. Pointing out that the real and imaginary parts of the Digital Fourier Transform (DFT) for a gaussian signal separately obey gaussian distributions, he proposed using the Kurtosis of the real part of the DFT as a tool for non-gaussian signal detection. He named this estimator "frequency-domain kurtosis" to distinguish it from the traditional time-domain kurtosis estimator. However, as Dwyer (1983) himself pointed out, a similar estimator can be defined for the imaginary part of the DFT, which should be considered for obtaining a complete description of the signal. Since the work of Dwyer (1983), others have investigated the statistical properties of power spectral density (PSD) estimates by treating

both the real and imaginary parts of the DFT, and even a theory of circular random complex variables has emerged from such efforts (see Amblard et al. 1996a,b, and references therein). In a more recent work, Vrabie et al. (2003) provided a comprehensive study, from both the theoretical and practical points of view, of the newly emerged concept of Spectral Kurtosis (SK), and provided a cumulant-based SK estimator (Kenney & Keeping 1962) and its expected variance, as well as the theoretical expectations for the SK parameter corresponding to some particular types of stationary time domain signals. In a subsequent study, Vrabie et al. (2004) revisited the subject and presented a practical application of the SK estimator in detecting bearing faults in induction motors. More recently, Antoni (2006) refined the theoretical definition of the SK parameter, provided a direct link of this parameter with the classical time domain Kurtosis parameter, and pointed out its value as an estimator of non-stationary time domain signals. Subsequently, Antoni (2007) proposed an algorithm for SK estimation based on a multi-rate filter-bank structure, and presented some experimental results related to the detection of transient faults in mechanical systems.

Despite its general applicability, the extent to which the frequency domain statistical analysis has been implemented as a RFI detection tool in astrophysical radio instruments is rather modest, possibly due to the lack of a clear and unifying picture of the subject. The study by Nita et al. (2007) provided a general framework that should be useful across disciplines, and applied it to the specific problem of RFI mitigation.

Another reason that there has not been more rapid progress in radio astrophysics may be the lack of existing instruments technically able to perform broadband spectral analysis with the temporal and frequency resolution needed for efficient implementation and testing of such RFI excision algorithms. Such instruments are useful in setting design

requirements for the new generation of radio instruments, such as FASR (Bastian et al. 1998). From this perspective, the newly designed FASR Subsystem Testbed (FST, Liu et al. 2007) operating in parallel with three of the radio antennas of the Owens Valley Solar Array (OVSA, Gary & Hurford 1999), offers a unique testbed capability. Its special design allows the direct comparison of various RFI excision algorithms by playing back in a simulated real-time mode the data recorded with full 1-ns temporal resolution, which provides a 500 MHz instantaneous bandwidth that may be tuned anywhere in the 1–9 GHz frequency band.

This chapter introduces the new, powerful RFI excision algorithm based on spectral variability (Nita et al. 2007), which is proved largely equivalent to spectral kurtosis defined by other authors (Dwyer 1983; Vrabie et al. 2004) and compares it with time domain kurtosis. The terms spectral variability and spectral kurtosis will be used interchangeably. The performance and effectiveness was demonstrated by Nita et al. (2007) by applying the algorithm to the real solar data for several cases, strong RFI contaminated quiet Sun, weak solar radio burst contaminated by moderate RFI, strong solar radio burst with fine time-frequency structures. These results are included also in this chapter, in order to provide a comparison with the effectiveness of the classical time domain kurtosis algorithm, which is implemented in software. A side-by-side comparison for the strong RFI contaminated quiet Sun case is given, along with a quantitative analysis.

4.2 Spectral Kurtosis Estimator

The high order statistics of a time domain signal can be used for detection of non-gaussian distribution signal has been understood for quite a long time. The feasibility of using high order statistics based on the power spectral density (PSD) estimates has been studied by several authors from different fields.

To obtain a PSD estimate of a time domain signal, one has to compute first the complex DFT coefficients given by

$$X_k = \sum_{n=0}^{N-1} w_n x_n e^{-2\pi ikn/N}, \quad k = 0, \ldots, N-1 \quad , \tag{4.4}$$

where $\{w_n\}$ is a set of N real coefficients of an arbitrary time window such as usually employed in order to reduce the spectral leakage among the PSD frequency bins (Press et al. 1989). Dwyer (1983) pointed out that the real and imaginary parts of the Digital Fourier Transform (DFT) for a gaussian signal separately obey gaussian distributions, he proposed using the Kurtosis of the real part of the DFT as a tool for non-gaussian signal detection. He named it frequency domain kurtosis as following:

$$K(f_k) = \frac{\langle A_k^4 \rangle}{\langle A_k^2 \rangle^2} \quad , \tag{4.5}$$

(A_k is the real part of the DFT coefficients) which is expected to be exactly 3 for a gaussian process just like time domain kurtosis. Vrabie et al. (2003) investigated the statistical properties of power spectral density (PSD) estimates by treating both the real and imaginary parts of the DFT, and provided a cumulant based definition of the newly emerged concept

of "spectral kurtosis" (SK) as

$$SK(f_k) = \frac{\langle |X_k|^4 \rangle - 2\langle |X_k|^2 \rangle^2}{\langle |X_k|^2 \rangle^2} \tag{4.6}$$

which is expected to be exactly zero for a gaussian time domain signal.

More recently, Nita et al. (2007) defined the dimensionless magnitude

$$V_k^2 = \frac{\sigma_k^2}{\mu_k^2}, \tag{4.7}$$

as spectral variability to quantify the relative statistical fluctuation of the PSD estimate $\widehat{P}_k$ with its mean $\mu_k = \langle \widehat{P}_k \rangle$, and variance $\sigma_k^2 = \langle (\widehat{P}_k^2) \rangle - \langle \widehat{P}_k \rangle^2$. Taking into account that $|X_k|^2 \sim \widehat{P}_k$, it immediately follows that

$$V_k^2 = SK(f_k) + 1 \tag{4.8}$$

which proves, up to a constant, the spectral kurtosis and spectral variability are equivalent. Consider the effect of window function ω_n applied in order to reduce the spectral leakage among the PSD frequency bins, The spectral variability is larger than one (Nita et al. 2007). Nita et al. (2007) derived the spectral variability in the general case of a windowed time domain signal,

$$V_k^2 = 1 + |W_{2k}|^2, \tag{4.9}$$

where

$$W_{2k} = \frac{1}{\sum w_n^2} \sum_{n=0}^{N-1} w_n^2 e^{-4\pi ikn/N} \tag{4.10}$$

is the normalized DFT of the squared time domain window evaluated at the even-indexed discreet frequencies $f_{2k} = 2f_k$. Since by definition, the normalized DFT of the squared time domain window, $|W_{2k}|^2$, is bounded by the $[0,1]$ interval, the spectral variability of a windowed gaussian signal can range, at least in principle, anywhere between 1 and 2, depending on the particular shape of the time domain window. However, because windows (e.g., Parzen, Hanning, Hamming, etc.) are chosen to reduce spectral leakage, they are invariably well-behaved, symmetric time-domain windows that gradually approach zero at both ends of the time-domain interval. This has the effect of drastically reducing $|W_{2k}|^2$ in most frequency bins, so that $V_k^2 = 1$ in all but a few of them—those near the DC and Nyquist frequency bins.

The goal is to decide with some level of confidence whether or not the observed PSD estimate $\widehat{P_k}$ corresponding to a physical process may be assumed to be driven by an underlying gaussian distribution. The statistical parameter defined by equation (4.7) is to be estimated from the data, and subsequently compared with the expected value, equation (4.9) for an ideal gaussian process. The ideal estimator would be an unbiased one, in other words an estimator having an expectation value identical to the expected value of the magnitude being estimated. Due to its linearity, the PSD estimate $\widehat{P_k}$ is an unbiased estimator of the true spectral power density of the signal. However, the statistical parameter defined by equation (4.7) as the ratio of the variance and the squared mean of the PSD estimate is no longer a linear combination of its components. Therefore, the best option is to define the

spectral kurtosis estimator as:

$$\widehat{V_k^2} = \frac{\widehat{\sigma_k^2}}{\widehat{\mu_k}^2}, \tag{4.11}$$

where $\widehat{\sigma_k^2}$ and $\widehat{\mu_k}$ are unbiased estimators of their associated parameters, i.e. $\langle\widehat{\sigma_k^2}\rangle = \sigma_k^2$ and $\langle\widehat{\mu_k}\rangle = \mu_k$. In practice, these two unbiased estimators will be computed from a number of M adjacent blocks of N time domain samples used to obtain M spectral estimates $\widehat{P_{ki}}$ $(k = 0,\ldots,N/2; i = 1,\ldots,M)$.

Introducing short hand notations for the sums,

$$S_1 = \sum_{i=1}^{M} \widehat{P_{ki}}; \qquad S_2 = \sum_{i=1}^{M} \left(\widehat{P_{ki}}\right)^2, \tag{4.12}$$

the unbiased estimators of the true σ_k^2 and μ_k are given by the k-statistics unbiased estimators of the first two cumulants of the $\widehat{P_k}$ distribution as (Kendall & Stuart 1958, p. 280):

$$\widehat{\mu_k} = \frac{1}{M} S_1; \quad \widehat{\sigma_k^2} = \frac{MS_2 - S_1^2}{M(M-1)}, \tag{4.13}$$

where S_1 and S_2 should be understood as having an implicit dependence on f_k. Using these expressions, the spectral kurtosis estimator becomes

$$\widehat{V_k^2} = \frac{M}{M-1}\left(M\frac{S_2}{S_1^2} - 1\right), \tag{4.14}$$

which is similar in form, but only asymptotically equivalent, to the SK estimator obtained by Vrabie et al. (2003) in terms of the second- and fourth-order cumulants of the complex X_k distribution.

Nita et al. (2007) present a detailed derivation of the variance of the spectral kurtosis

estimator defined in equation (4.14). The result, dependent on the particular shape of the time domain windowing function, is

$$var\left(\widehat{V_k^2}\right) = \frac{4M}{(M-1)^2}\left(1+5|W_{2k}|^2-|W_{2k}|^4+|W_{2k}|^6\right). \tag{4.15}$$

Since at the DC and Nyquist frequency bins $|W_{2k}| = 1$, and at most of the other bins $|W_{2k}| \simeq 0$, hence the $1/M$-order approximation of equation (4.15) simplifies to

$$var\left(\widehat{V_k^2}\right) \simeq \begin{cases} 24/M, & k = 0, N/2 \\ 4/M, & k = 1, \ldots, (N/2-1) \end{cases}, \tag{4.16}$$

which may be used as a statistical detection threshold for practical applications.

One may recognize the variance for the DC and Nyquist frequency bins are that for the time domain kurtosis estimator (Ruf et al. 2006), and note that it is about six times larger than the variance at other frequency bins. This predicts that the variance for the SK estimator, $\widehat{V_k^2}$, which is derived from an DFT-based spectrometer is smaller than that of the classical kurtosis estimator, $\widehat{K}$, derived using a filter-bank spectrometer.

To correctly quantify the difference between the variances of $\widehat{V_k^2}$ and $\widehat{K}$ estimators, it has to be taken into account that, in the case of a Finite Impulse Response (FIR) filter-bank spectrometer, the amplitudes of the filtered signal at $N/2+1$ frequencies are obtained directly as real linear combinations of $N/2+1$ time-domain samples. Therefore, from the same $N \times M$ data samples as used by the DFT-based PSD estimator, one would obtain, in each filter-bank frequency channel, a set of $N \times M/(N/2+1) \simeq 2M$ independent PSD estimates. Therefore, for a given amount of time-domain data, the $\widehat{V_k^2}$ variance of $\sim 4/M$

has to be compared with a $\widehat{K}$ variance of about $\sim 12/M$. Therefore, for a given time-domain data set, the SK estimator relying on DFT based PSD estimates may be more efficient by a factor of 3, in terms of its variance, than a kurtosis estimator based on filter-bank PSD estimates.

4.3 Software Implementation of an RFI Excision Algorithm for the FST Interferometric System

Having established the theoretical properties of the SK estimator, the aim is to suggest a general approach for its practical implementation as an RFI detection and excision tool. A straightforward implementation of an RFI excision algorithm for a 1-channel DFT-based radio spectrograph would require accumulating the sums S_1 and S_2 defined in equation (4.12) over the desired M spectra, each derived from N blocks of time-domain data, and normalized by the band-integrated total power, if necessary. From these sums, the SK estimator, equation (4.14), is calculated and compared with a threshold value based on equation (4.15) (or equation (4.16)) to define a flag array to be applied to S_1 to remove bad channels. Because of the narrow bandwidth of most RFI, one may maximize retention of RFI-free data channels by accumulating the SK statistics at finer frequency resolution than the desired (science-based) bandwidth, and combine frequency channels after RFI excision.

In the case of an interferometric radio array, the logical AND of the RFI flag arrays (assuming channels with RFI are flagged as zero) yield an RFI flag array for each pair of antennas. Note that this approach avoids the theoretical difficulties of defining a statistical estimator for the correlated spectra, which would have to take into account that the data sets recorded by two different antennas are not statistically independent (Fridman 2001).

One of the main goals of the newly developed FASR Subsystem Testbed (FST; Liu et al. 2007) is to provide a testbed for studying methods for RFI excision. The design of this instrument, which records full time-resolution time-domain data at a sampling rate of 1 ns, allows playing back data in simulated real time while experimenting with software-implemented RFI excision algorithms in either time or frequency domains over an instantaneous 500 MHz bandwidth. This 500 MHz band can be tuned anywhere in the 1–9 GHz frequency range. Since the FST system is composed of three independent data channels linked to three of the seven antennas of the Owens Valley Solar Array (OVSA; Gary & Hurford 1999), the performance of the RFI excision algorithms on correlated data also can be investigated. However, this flexibility does not come without cost. Due to the limitation to 2 MB/channel of the on-board memory of the digitizing system (Acqiris DC271), no more than about 2 ms of contiguous time domain data segments can be recorded at a time, with time gaps between data segments due to the time needed to download the recorded data and store it on the host computer hard drive. The duty cycle for data acquisition is roughly 1%. The data used to illustrate the SK algorithm were taken in modes 1 and 3, as defined in Liu et al. (2007). For 3-channel data taken in mode 1, 2 ms of contiguous time-domain data are recorded, with time gaps between segments of the order of 160 ms. In mode 3, 100 μs of data are separated by gaps of 20 ms. As a consequence of these hardware constraints, the contiguous SK statistics cannot be extended beyond an upper limit (for mode 1 data) of $N \times M = 2 \times 10^6$ contiguous data samples, which, depending on the frequency resolution desired, puts an upper limit on the maximum number of PSD estimates, M, that can be obtained from the same contiguous data set. Nevertheless, as will be shown later in the text, this limitation can be overcome by ignoring the time gaps and performing SK

statistics based on more or less evenly distributed, noncontiguous PSD estimates. But one should keep in mind that the efficiency of the estimator using this operating mode may be affected by the temporal behavior of non-stationary RFI signals.

The FST system also provides the ability to investigate the validity and efficiency of the RFI excision algorithms under extreme conditions of rapidly-varying power levels of the gaussian signal, which occur during the impulsive phases of some solar radio bursts recorded by the instrument. As shown in the previous section, it is necessary in such situations to obtain SK statistics based on normalized PSD estimates. However, the FST system does not need special treatment of such situations due to its built in automatic gain control (AGC) loop, which dynamically maintains a fixed power level of 6 dBm over the 500 MHz bandwidth data channel inputs of the digitizer. Therefore, the unscaled PSD estimates obtained directly from the raw data are already normalized by the default behavior of the system. However, it is one of the main goals of this study to check whether or not the SK estimator really behaves as expected under the extreme conditions of fast-varying radio bursts, a situation considered of great practical interest for the design of future solar radio telescopes with real-time RFI excision capability. Moreover, studying the behavior of the SK estimator under radio burst conditions should help one evaluate the extent to which the assumption of an underlying gaussian parent population still holds in the presence of various coherent or incoherent emission mechanisms that might be responsible for the natural radio emission produced by the solar radio bursts. In particular, it is desired to investigate whether or not a dynamic SK-based RFI excision algorithm can be safely used during a radio burst without the risk of confusing them with artificial RFI signals.

4.3.1 Performance of the Algorithm in the Case of Strong RFI

In the 1-9 GHz observational range of the FST instrument, the 1.5-2.0 GHz frequency band is one that is strongly contaminated by RFI. Figure 4.1 displays a dynamic spectrum of the quiet Sun that is heavily contaminated by RFI signals with a variety of time-domain characteristics. The dynamic spectrum has a frequency resolution of $\sim$ 1 MHz and a time resolution of $\sim$ 142 ms, each spectrum being obtained by averaging data obtained from 2 ms segments of contiguous data corresponding to one mode-1 acquisition, separated by 140 ms of dead time. Each segment has been divided into $M = 244$ adjacent blocks of $N = 8192$ data samples each, resulting in 244 PSD estimates in each of 4097 frequency channels, having a frequency resolution of $\sim$ 122 kHz. Then integrating over eight frequency bins (subchannels) and accumulating in time produce a time-averaged spectrum with $\sim$ 1 MHz frequency resolution.

Figure 4.2 illustrates the performance of the RFI excision algorithm in the case of $\widehat{V_k^2}$ statistics based on a single averaged spectrum. A detection threshold of $\sim 1 \pm 3\sigma = 1 \pm 6\sqrt{M}$ has been used to flag the RFI-contaminated channels and remove them prior to subchannel integration. Channels for which subchannels have been removed are proportionally scaled in amplitude. Given the small number ($M = 244$) of PSD estimates available as a result of the low duty cycle of FST, the performance illustrated by Figure 4.2 is considered to be promising. Even in this case of low-number statistics, Figure 4.2 shows that the algorithm successfully removes most of the transient RFI without affecting the continuum background. The removal of the strong, apparently continuous RFI is less effective, perhaps due to a combination of low-number statistics, the time gaps forced by the low duty cycle of the instrument, or residual non-stationary power level, which can reduce the

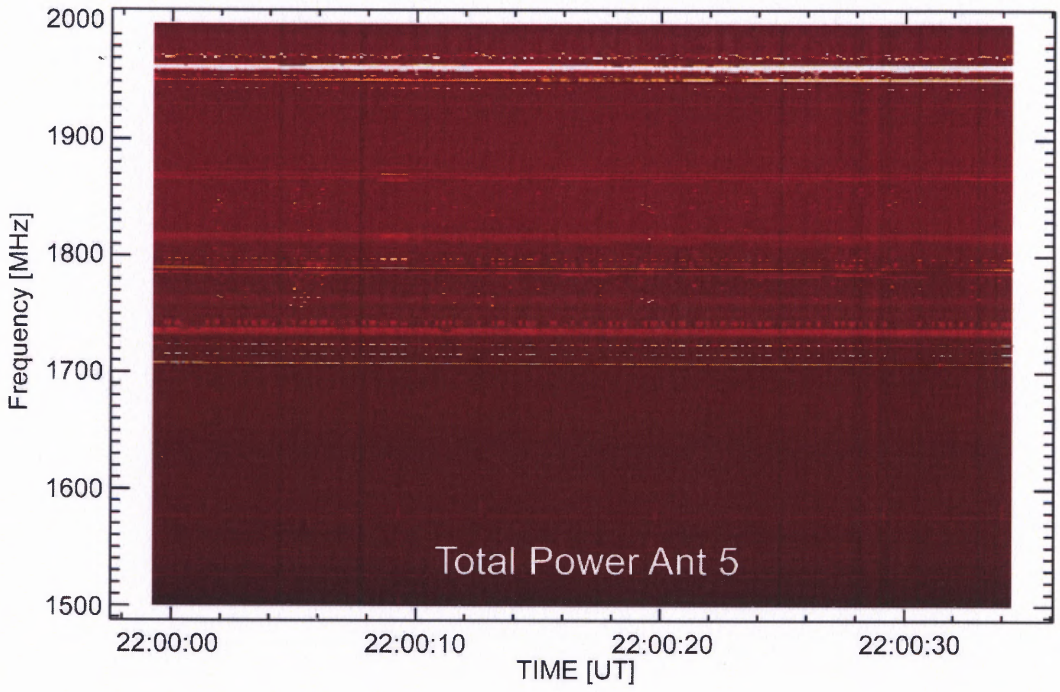


Figure 4.1 RFI-contaminated dynamic spectrum produced from FST radio data obtained during a quiet Sun observation. The time and frequency resolution of the averaged dynamic spectrum are $\sim$ 142 ms and $\sim$ 1 MHz, respectively. Various types of continuous, clustered, or scattered RFI contamination are evident.

effectiveness of the algorithm for continuous RFI. It is anticipated that the algorithm will perform better for continuous RFI in an instrument with 100% duty cycle. Indeed, a detailed inspection of the RFI flags generated by the algorithm reveals that it is sensitive to the continuous RFI. Figure 4.3 zooms in on two areas in Figures 4.1 and 4.2, and shows the number (0 to 8) of surviving (unflagged) subchannels in each frequency and time bin. It is evident that, in the case of continuous RFI components, much of the time at least one out of eight frequency channels are flagged as being contaminated by RFI, and it is anticipated that if better statistics (larger M) were possible the results would have been better. The general salt and pepper aspect of the flag spectrum is explained by normal statistical fluctuations of the SK estimator, expected to be of order $\sim 1\%$ for the 3σ detection threshold used. The total number of bins flagged as RFI represent $\sim 2.5\%$ of the original data, which suggests that $\sim 1.5\%$ of the data has been successfully identified as contaminated with RFI, and removed.

To check the extent to which the performance of the SK estimator would be improved by better statistics, one can reduce the estimator frequency resolution by a factor of 8, yielding $M = 1952$ PSD estimates in each frequency-time bin, to get the results as in Figure 4.4. To allow a better comparison with the results presented in Figure 4.2, in this case there is no frequency integration, so that the RFI flagged bins appear black in the plot. Although the results have improved, please note that the algorithm performance is still not optimal. This is because, by reducing the frequency resolution of the estimator, the signal to noise ratio of narrow band RFI is also reduced due to the wider frequency bins. The total data loss in this case is $\sim 7\%$, which may be partially due to better performance, but is also due to the fact that the narrow band RFI contaminates larger frequency bins.

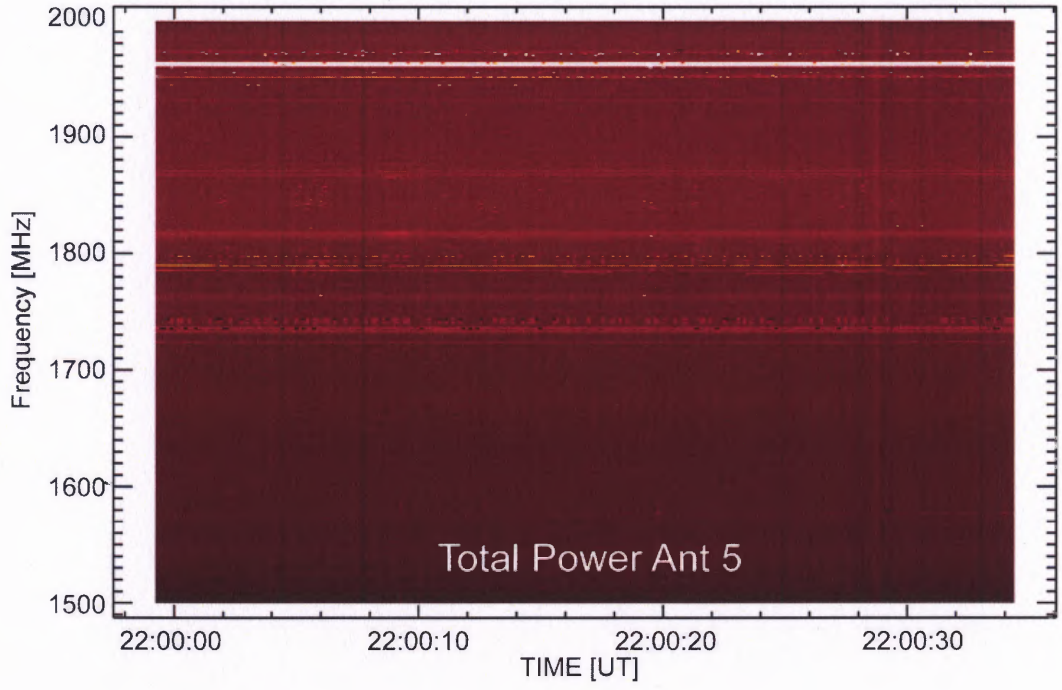


Figure 4.2 The RFI-cleaned dynamic spectrum obtained from the same data displayed in Figure 4.1. For each averaged spectrum, the RFI flags have been generated based on the $\widehat{V_k^2}$ statistics obtained from $M = 244$ contiguous PSD estimates. Most of the RFI spikes have been removed, and the continuous RFI spectral lines are reduced.

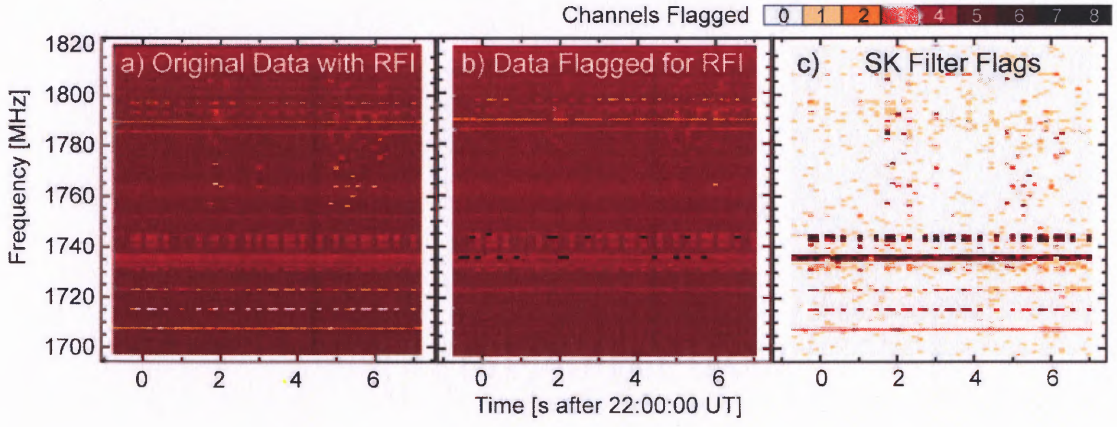


Figure 4.3 (*a*) A magnified portion of Figure 4.1, containing RFI. (*b*) The same portion of Figure 4.2, with some of the RFI removed. (*c*) The number of subchannels flagged for RFI, in dynamic spectrum format, corresponding to the same portion of the dynamic spectrum as in *a* and *b*. The number of subchannels is color-coded according to the scale above the panel *c*.

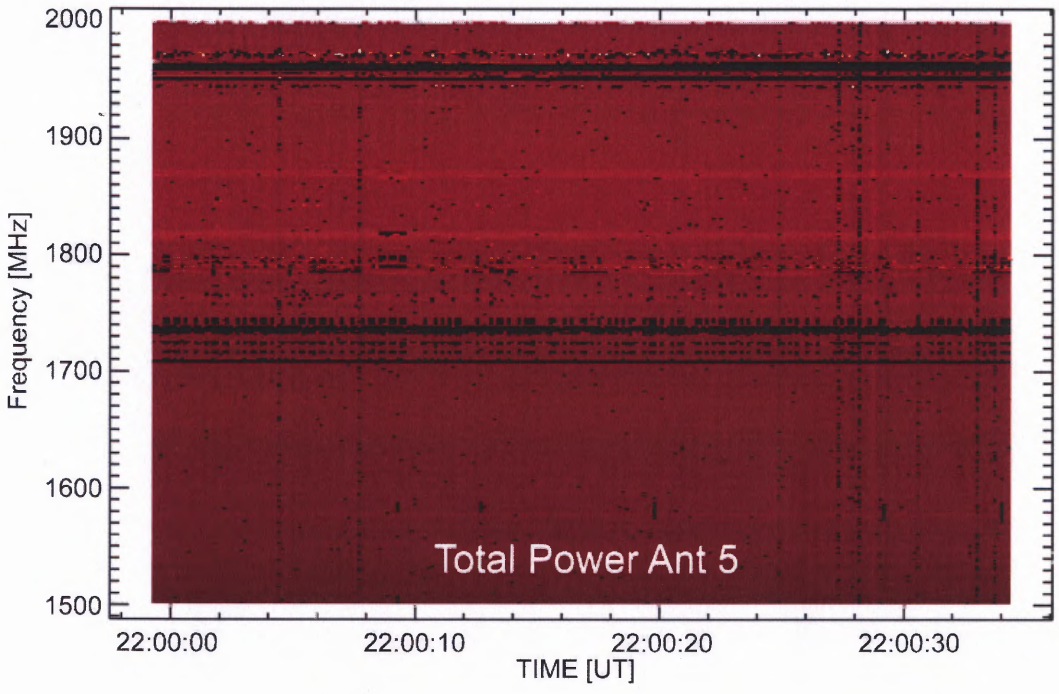


Figure 4.4 The RFI-cleaned dynamic spectrum obtained from the same data displayed in Figure 4.1, but with frequency resolution reduced by a factor of 8 in order to improve the SK statistics ($M = 1952$ contiguous PSD estimates). To allow a better comparison with Figs. 4.1 and 4.2, no subchannel integration has been performed.

4.3.2 Weak Solar Radio Burst Spectrum Contaminated by Moderate RFI

Figure 4.5 displays a dynamic spectrum corresponding to a solar radio burst recorded by the FST instrument on April 05 2006 in the 1.0–1.5 GHz range. The burst consists of several weak emissions, the most intense being observed around 21:06:30 UT. Four on/off periodic RFI signals can be best observed at about 1200, 1360, 1400, and 1440 MHz, as well as a few transient RFI spikes near 1100 MHz. A periodic change in the background intensity every 4 s is due to feed-switching between right-hand and left-hand circular polarization. The spectrum in each time bin results from $N \times M = 10^5$ time samples of 1 ns each ($N = 8192$, $M = 12$, $\Delta f = 0.122$ MHz) separated by 25 ms dead time, while ~ 100 ms time gaps occur between 20 consecutive spectra (vertical black stripes in the plot) due to additional dead time for writing the data to disk.

Although the number of PSD estimates, $M = 12$, is far too small for useful statistics for an individual bin, the efficiency of the SK algorithm may be greatly improved by accumulating the sums S_1 and S_2, on which the SK estimator is based, in running buffers $\{S_1\}$ and $\{S_2\}$ that span a much longer time. Sums over the $M = 12$ PSD estimates for each time bin are added to limited-length arrays holding a predefined number of partial sums. When the length limit of these arrays is reached, the oldest $\{S_1, S_2\}$ pair is discarded to make room for the incoming one. This procedure allows one to generate, based on the total sums $S_1 = \sum S_{1k}$ and $S_2 = \sum S_{2k}$, a statistically significant SK estimator that is used to flag the RFI-affected spectral components of the incoming time bin.

Figure 4.6 presents the cleaned version of the spectrum presented in Figure 4.5, based on running history buffers maintaining 240 partial sums of 12 contiguous PSD estimates. Thus, the statistics are generated from $M = 2880$ PSD estimates. Comparing

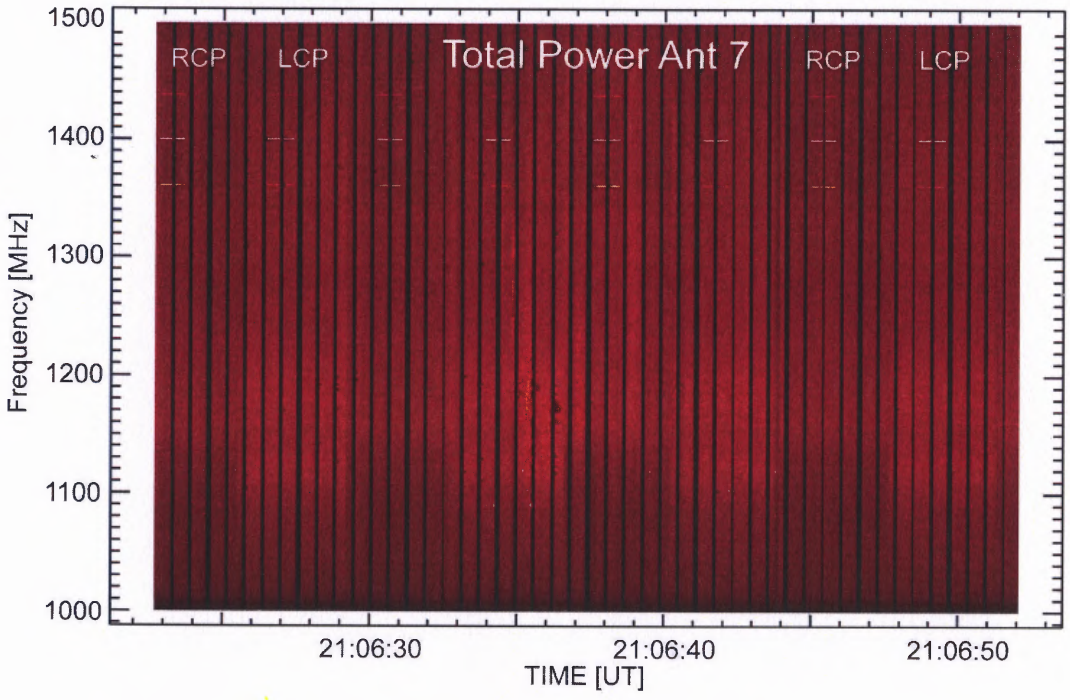


Figure 4.5 Total power dynamic spectrum of a weak solar radio burst recorded on 2006 April 05, around 21:06:30 UT, in the 1-1.5 GHz frequency band. The time resolution of the dynamic spectrum is 25 ms and the frequency resolution is 1 MHz. The vertical black stripes in the plot represent variable time gaps needed to dump each separate acquisition on the storage media. Four on/off periodic RFI signals can be seen at 1190, 1360, 1400, and 1440 MHz, as well as a few narrowband features scattered below 1200 MHz, most likely due to aircraft distance measuring equipment. RCP and LCP labels indicate a few of the periodic changes in the background intensity, every about 4 s, due to periodic switching of the observed circular polarization, which is the default mode of operation of OVSA.

Figures 4.6 and 4.5, one may conclude that the approach successfully removes most of the RFI spectral components without affecting the spectrum of the solar radio burst. Note that the algorithm fails to detect the RFI in the first few seconds, before a statistically significant number of PSD estimates is accumulated in the running SK buffers. While the "learning curve" of the estimator should not be of concern given the fact that this is a time-limited "start-up" effect, the high sensitivity of the estimator to transient RFI spikes might result in an undesired complete blanking of some spectral regions that are only transiently affected by RFI, as shown by the missing channel near 1090 MHz. RFI spikes in a few time bins can result in removal of all data for the affected frequency until the contribution to the sum "ages off" and leaves the buffer. A balance must be struck between increasing buffer size for improved statistics and the corresponding loss of data due to over-long response to transient RFI.

4.3.3 Strong Solar Radio Burst with Fine Time-Frequency Structures

One of the solar radio bursts recorded by FST, which occurred on 06 December 2006, displayed extremely strong, narrow-band (few MHz), short-duration (< 20 ms) radio spikes that nominally look a great deal like RFI. Figure 4.7 displays ~ 7 s of the dynamic spectrum, where the highly right-circularly polarized (RCP) spike bursts are apparent during the first ~ 4 s, when the system was measuring RCP, but are nearly absent when the system was switched to left-circular polarization (LCP) in the last ~ 3 s. At the same time, RFI is weak during the first ~ 4 s, but more readily visible in the last ~ 3 s. As in the previous burst (Figure 4.5), each time bin of the spectrum is obtained from $N \times M = 10^5$ time samples of 1 ns each ($N = 8192$, $M = 12$, $\Delta f = 0.122$ MHz) separated in this case by 20 ms dead time.

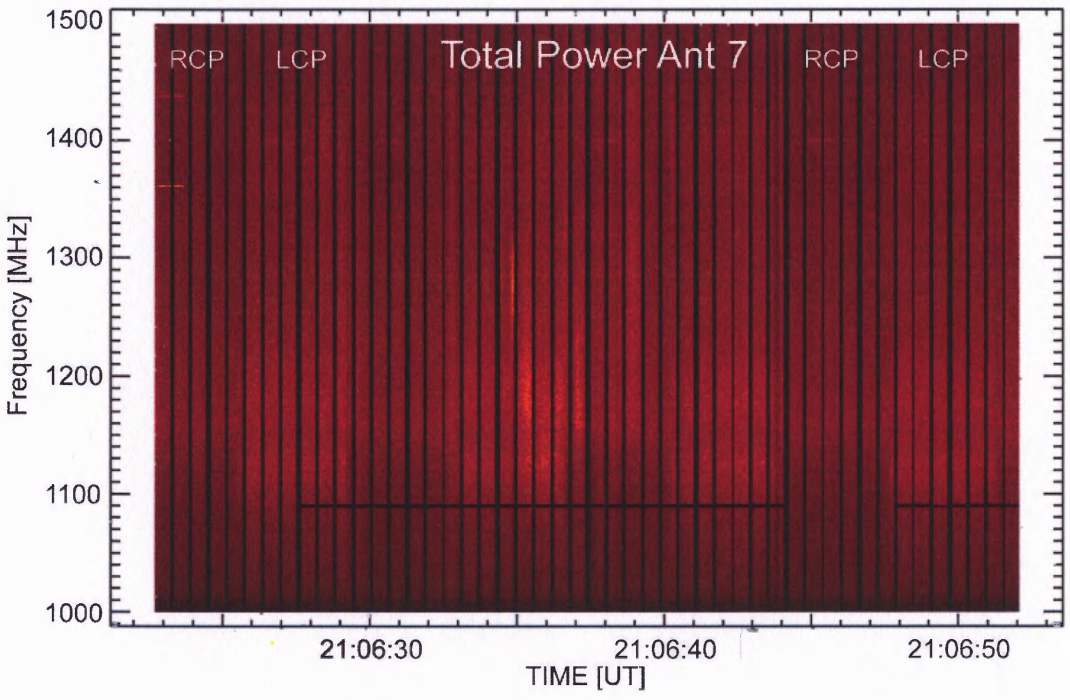


Figure 4.6 RFI-cleaned version of the spectrum displayed in Figure 4.5, using a running buffer as explained in the text. $M = 2880$ PSD estimates have been used to evaluate the SK estimator and generate the RFI flags. In comparison with the dynamic spectrum displayed in Figure 4.5, the periodic fixed frequency RFI signals, and nearly all of the scattered RFI spikes, have been removed while the spectrum of the solar radio burst is not affected by the excision algorithm. The algorithm fails to detect the RFI spectral lines at the earliest times, until a statistically significant number of PSD estimates is accumulated in the running buffers.

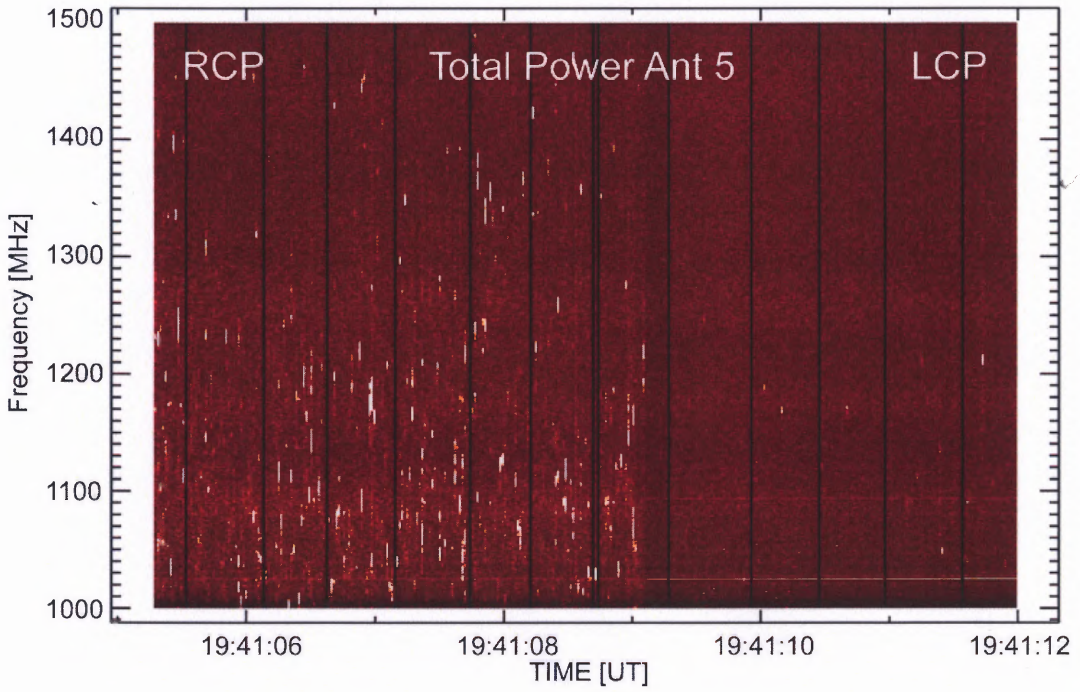


Figure 4.7 Portion of the dynamic spectrum of a solar radio burst recorded by FST on 2006 Dec 06, 19:41:00 UT, when strong radio spikes where observed. The frequency and time resolutions of the dynamic spectrum are ~ 0.122 MHz and 25 ms, respectively. However, each averaged spectrum corresponding to a 25 ms time bin was obtained only from 0.1 ms of contiguous data samples, resulting in $M = 12$ consecutive PSD estimates. The first 4 s of data, taken with a right hand polarized (RCP) feed, reveal intense radio spikes, with only weak RFI contamination. The rest of the data, taken with a left hand polarized (LCP) feed, reveal just a few radio spikes, and clear RFI contamination. However, some of the spiky structures present in both polarizations might also be due to low-duty-cycle RFI.

A legitimate question arising from the displayed spectral structures in Figure 4.7 is whether or not the $\widehat{V_k^2}$ statistics will remain insensitive to the natural spiky structures, while still discriminating the true RFI contamination. The answer to this question is important from both practical and theoretical points of views. From a practical standpoint of a real-time RFI excision algorithm, it would be highly undesirable for the natural solar emission to be irreversibly removed due to its being mistaken for low-duty-cycle RFI contamination, to which Nita et al. (2007) proved the high sensitivity of the SK estimator. From a theoretical point of view, the $\widehat{V_k^2}$ analysis of this data set might reveal whether or not the intensity of such natural radio spikes, generally believed to be produced by the electron cyclotron maser emission mechanism (see Fleishman et al. 2003, and references therein), nevertheless obeys a gaussian statistical distribution.

The number $M = 12$ of contiguous PSD estimates within a single time bin is too small to be statistically significant, and when one applies the running buffer approach of the previous section the algorithm does indeed incorrectly flag the spikes as RFI. However, as mentioned before, one can increase M by decreasing N (i.e. decreasing spectral resolution), which is likely to be fruitful since the spikes have an intrinsic spectral width of several MHz. Figure 4.8 displays the dynamic spectrum obtained for the same data set but with frequency resolution of 3.9 MHz. As expected, most of the individual spikes are still resolved, while some of the weaker RFI spectral lines present in Figure4.7 are no longer visible due to the fact that their SNR has been decreased by increasing the width of the frequency bins. While the temporal resolution of the dynamic spectrum remains the same, each of the individual spectrograms is now obtained from $M = 390$ contiguous PSD estimates, which, according to Nita et al. (2007), should be sufficient to discriminate

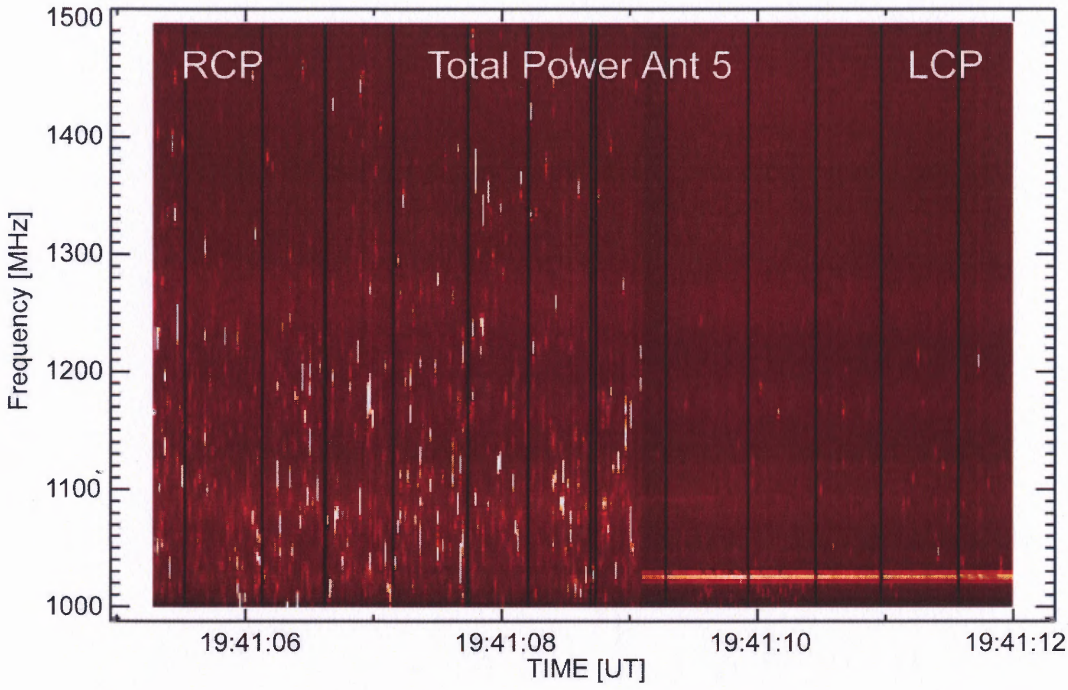


Figure 4.8 The same dynamic spectrum displayed in Figure 4.7, but with much lower frequency resolution. Each of the contiguous 0.1 ms data segments have been divided in $M = 390$ adjacent blocks of $N = 256$ data samples, for ~ 3.9 MHz frequency resolution. No integration in the frequency direction has been performed. Most of the spike clusters are still partially resolved in the RCP segment of the spectrum, while only one RFI spectral line at about 1020 MHz remains clearly visible.

continuous RFI with a signal to noise ratio as low as ~ 1.3.

Figure 4.9 presents the results of the RFI excision algorithm applied to the data displayed in Figure4.8, with no frequency integration. Remarkably, almost all of the natural radio spikes survived the application of the $\widehat{V_k^2}$ discriminator. This indicates that, in the limit of the 390 PSD estimates available, a gaussian distribution is at least a good approximation for the probability distribution function of spike intensity. On the other hand, two of the three adjacent RFI spectral lines have been completely removed, as well as a few other bright features, believed to be RFI. The low-frequency edge of the main RFI-affected bins is not removed, possibly due to too low SNR, or a duty cycle near 50%. The random

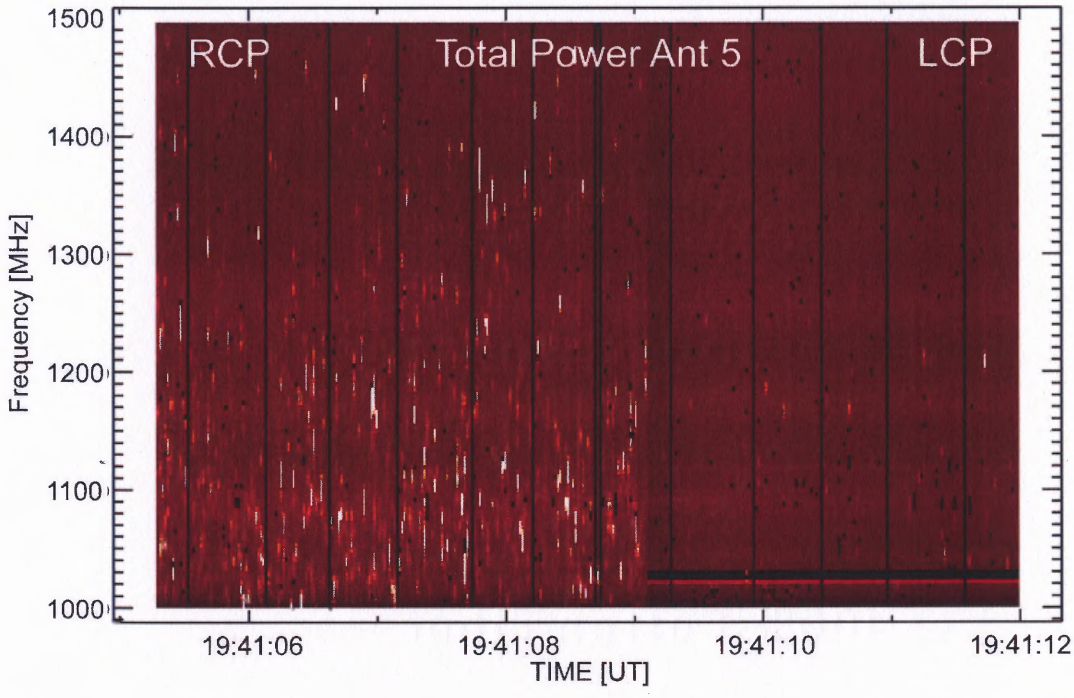


Figure 4.9 The results of the RFI excision algorithm applied to the data displayed in Figure 4.8. Since no frequency integration has been performed, the RFI-flagged spectrogram bins appear black in the plot. Remarkably, the SK estimator proves to be insensitive to the natural spiky emission, while it successfully detects most of the continuous RFI spectral line at about 1020 MHz. The spectrogram also reveals a few scattered bins flagged by the algorithm, most likely to be false RFI alarms. However, the total data losses, $\sim 0.7\%$ in the RCP section, are close to the $\sim 1\%$ losses due to statistical fluctuations expected for the $\pm 3\sigma$ RFI detection threshold used.

removal of spectral bins is about equally numerous in both the RCP and LCP sections, and is consistent with the 1% removal expected for a 3σ discriminator.

Therefore, one may conclude that the implementation of an SK-based RFI excision algorithm in a real time instrument, which can be designed to assure sufficient statistics in the 20 ms lifetime of a spike, would be unlikely to trigger more than statistically expected false RFI alarms even in the presence of intense and highly fluctuating natural radio emission.

4.4 Comparison with Time Domain Kurtosis Estimator

Since the newly constructed FST system saved the digital data with full resolution, one can play back the same set of raw, full resolution recorded data to examine the performance of the time domain and spectral kurtosis algorithms on an equal basis. Nita et al. (2007) already has evaluated the performance of spectral kurtosis for strong RFI, weak solar radio burst spectrum contaminated by moderate RFI, and strong solar radio burst with fine time-frequency structures, respectively. For simplicity, the spectral kurtosis estimator $\widehat{SK}$ is compared with the time domain kurtosis estimator $\widehat{K}$ only for the data of quiet Sun contaminated by strong RFI signals (Figure 4.1). One reason for chosing this example is that the strong RFI data has the longest contiguous time domain data (2 ms), which gives it better statistics.

As discussed in section 4.2 the variance of a time-domain kurtosis estimator (from a filter-bank spectrometer) is about three times larger than the variance of the SK estimator (from an DFT-based spectrometer). Implementation of the time domain kurtosis algorithm involves the filter bank design, various design windows (such as Blackman-Harris, Hamming, Hanning, gaussian, Kaiser, Equiripple, and Chebyshev), number of coefficients, and the number of bits used per coefficient (Ruf et al. 2006). Since this comparison is implemented in software, it is not necessary to be distracted by the filter design details from the underlying real algorithm performance comparison. Therefore, the IDL function "digital_filter" is used to generate a near-perfect filter with a very large number of coefficient (51201), even if that means a long computer calculation time.

Figure 4.4 shows the performance of the spectral kurtosis algorithm to excise the RFI signals from the quiet Sun. In order to compare with the time domain kurtosis algo-

rithm under the same conditions the same data of 2 ms of contiguous sampled signal (2 million samples) need to be channelized to 512 channels by passing the data through a filter bank, giving $N = 3906$ data samples for each subband. Figure 4.10 shows the response function of one of the 512 channels filters. The out-of-band rejection is better than 130 dB, which make it essentially an ideal filter. The process for each subband is the convolution of

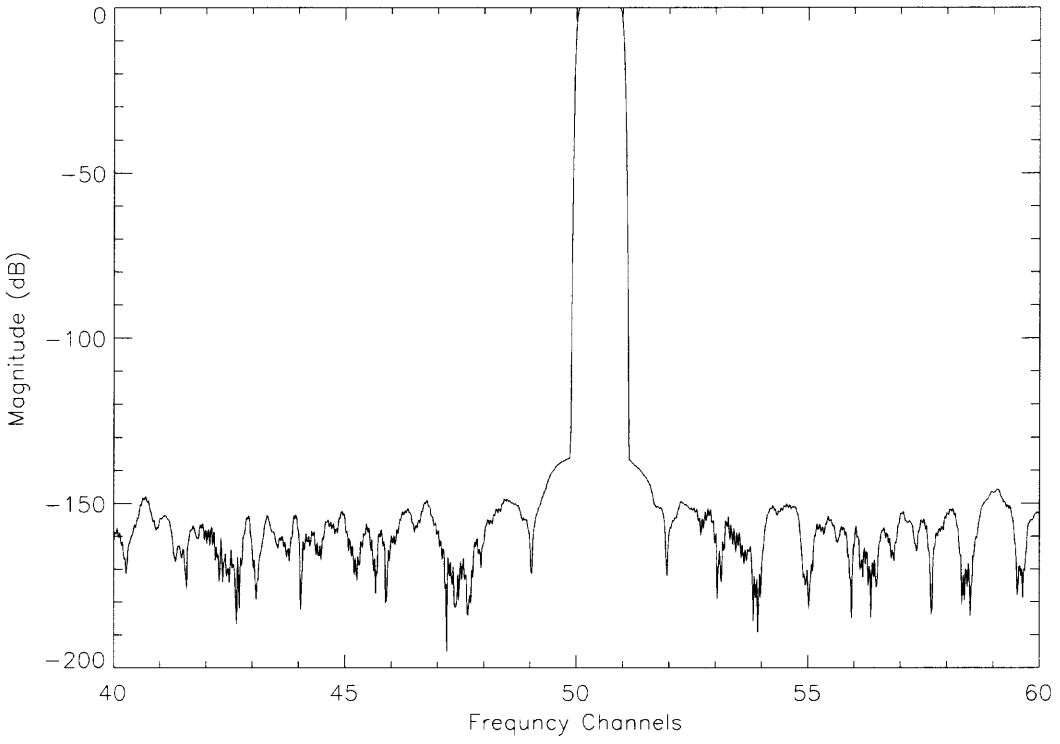
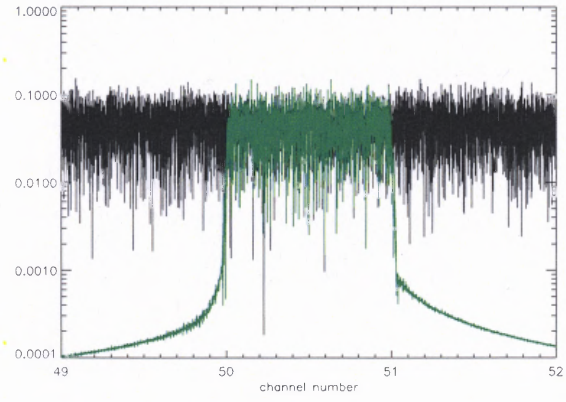
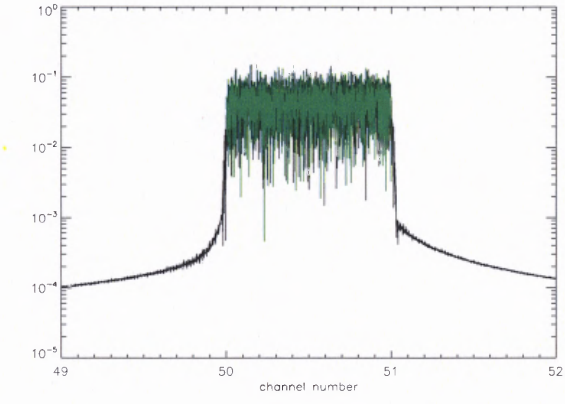


Figure 4.10 One of the sub-band filter in the filter bank to channelize the digitized digital data to multi-band time domain signal. .

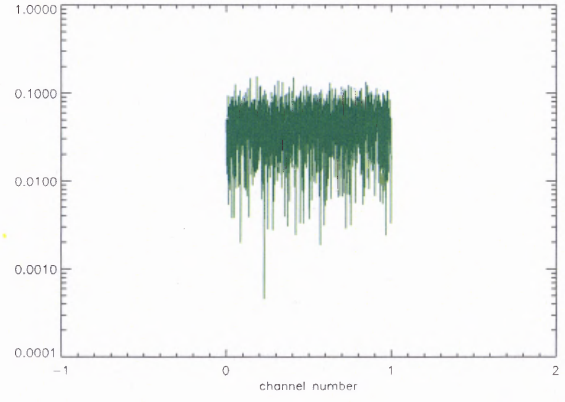
2 million sampled data points with 51201 filter coefficients. The resulted 2 million point of data can be decimated by a factor of 512 since the bandwidth of the subband is only 1/512 of the full bandwidth. The decimation aliases the particular subband signal to the base band with the same spectrum information. Figure 4.11 shows this procedure for a single channel (channel 50). The data in Figure 4.11(c) are integrated to provide the total power



(a)



(b)



(c)

Figure 4.11 The filter and decimation process. (a)the original spectrum (black) overplotted with the filtered spectrum (green) (b) the filtered spectrum (black) overplotted with the spectrum after decimation, the decimation spectrum has been shifted back to the proper channel from the baseband for better comparison. (c) the spectrum of decimated data in baseband.

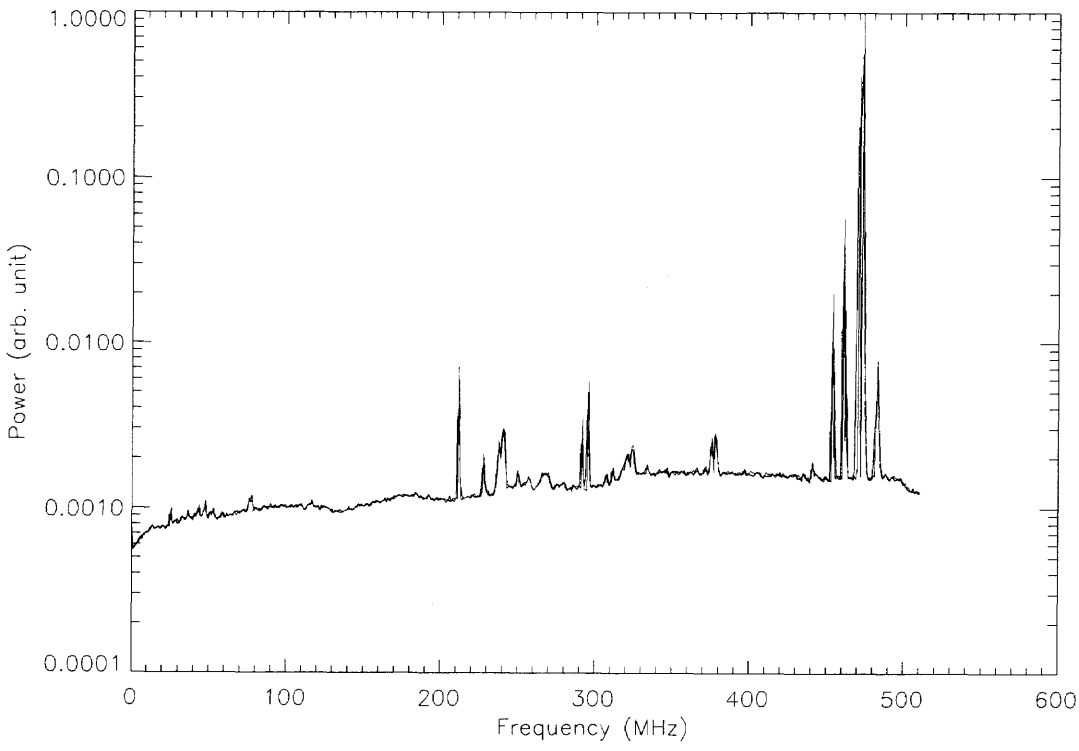
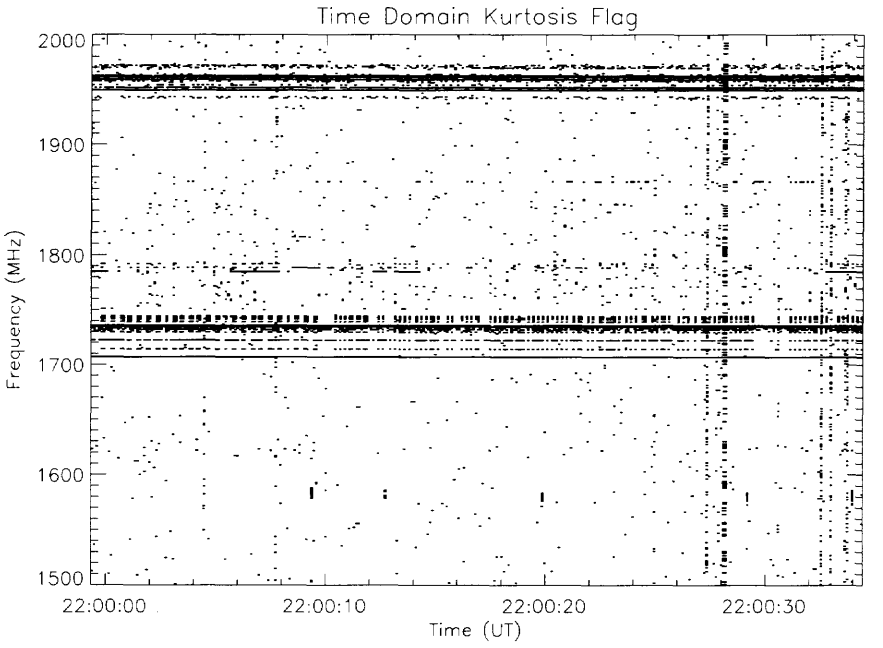


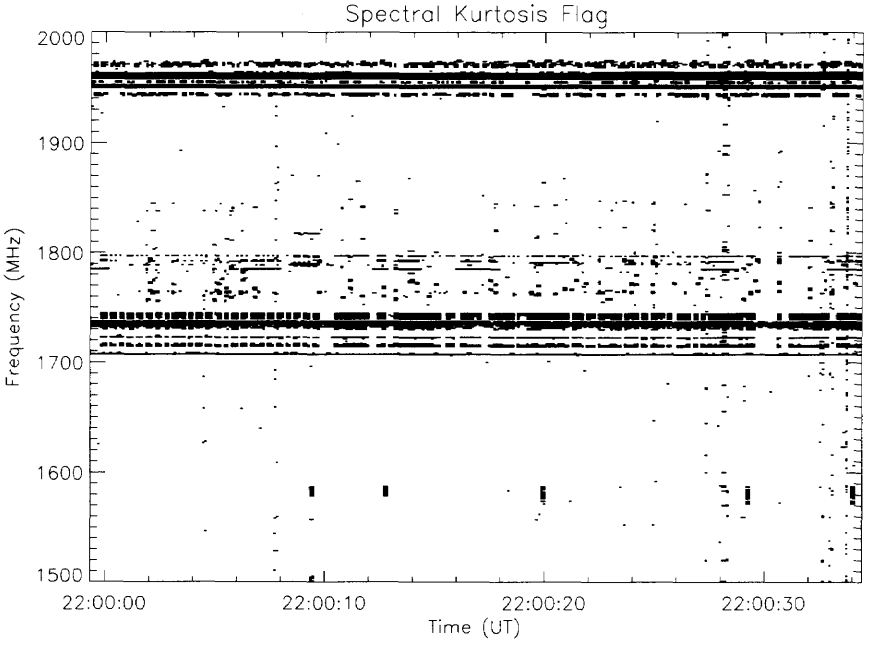
Figure 4.12 The power spectrum at one time, the black curve was obtained by taking DFT from one block of 1024 time domain signal points and integrated by 1953 blocks, green curve was obtained by passing the data through a filter bank and integrated the signal power for each sub-band. The two curves are identical expect that the black spectral lines are wider, due to the leakage of non-window effect of DFT.

for channel 50, and this procedure is repeated for all 512 channels. Figure 4.12 compares the power spectrum based on DFT with that from the filter bank method. The results are essentially identical. This shows that the data obtained after the filter bank channelization still has the same power distribution, the spectral properties have not been altered.

Once the channelized time domain signals are obtained, the time domain kurtosis can be easily calculated by using IDL function "kurtosis" which evaluates equation 4.1, but subtracts three from the raw kurtosis value since three is the kurtosis for gaussian distribution.

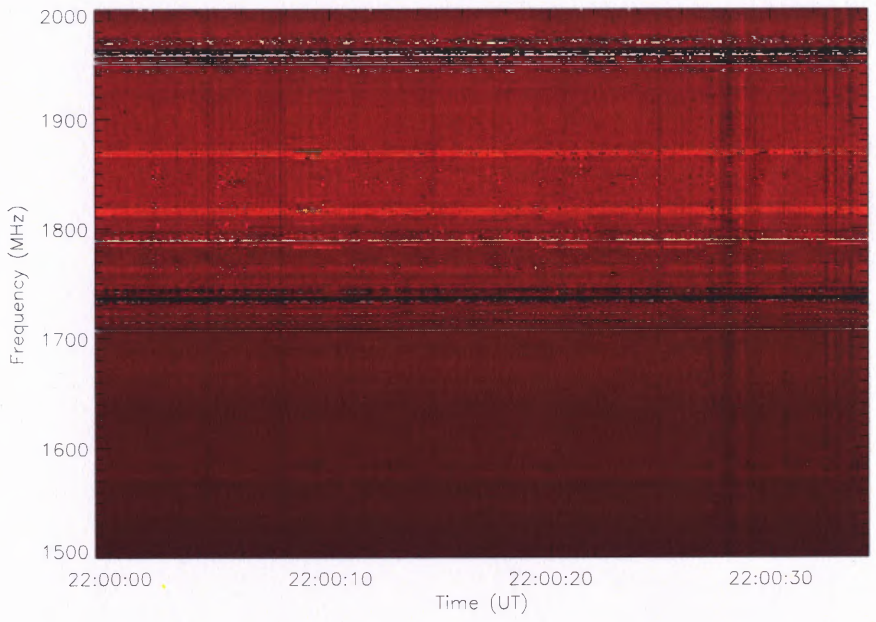


(a)

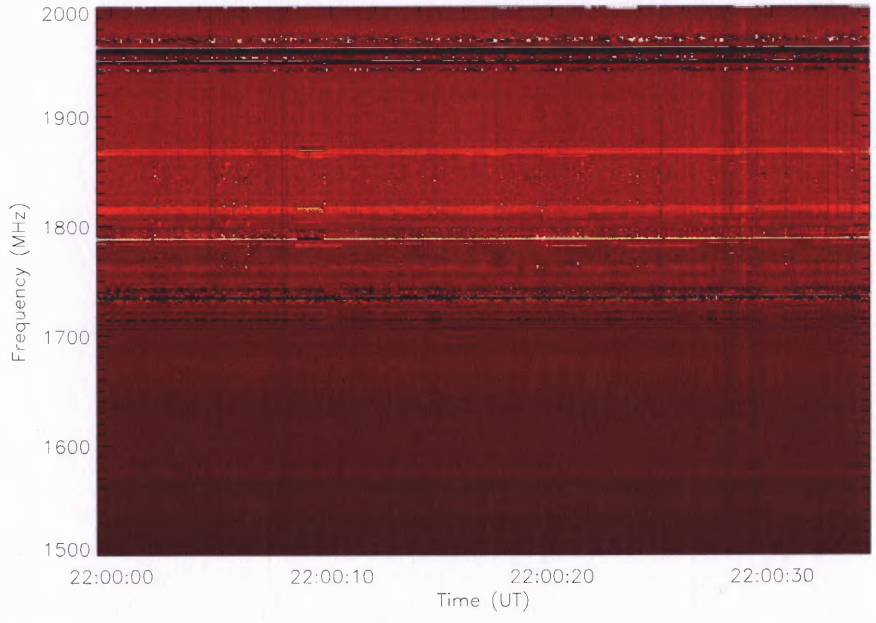


(b)

Figure 4.13 RFI flags for time-domain and spectral kurtosis estimators. (a) time domain kurtosis flag using 3σ as the detection threshold. (b)spectral kurtosis flag also using 3σ as the detection threshold.



(a)



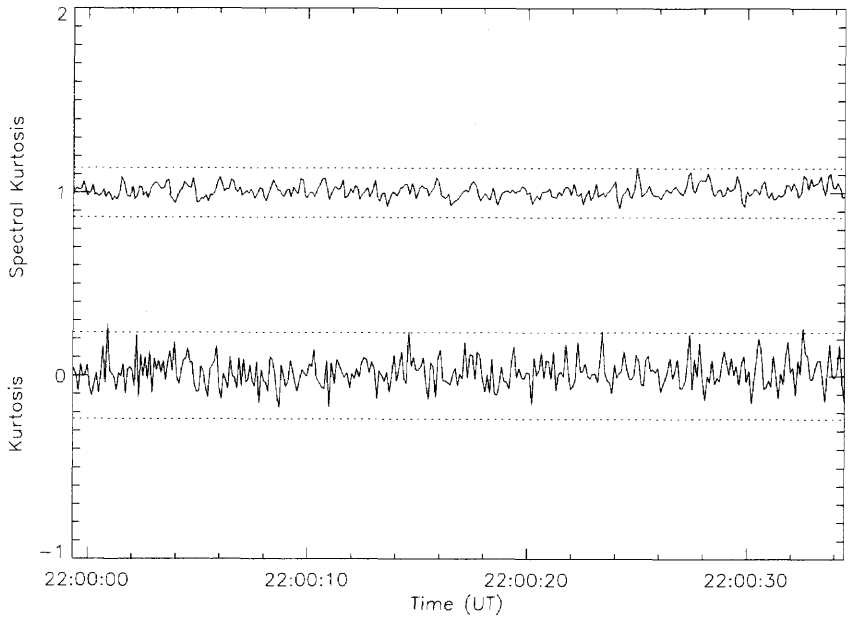
(b)

Figure 4.14 (a)RFI-cleaned dynamic spectrum obtained by the time domain kurtosis estimator flag. (b) RFI-cleaned dynamic spectrum obtained by the spectral kurtosis estimator flag.

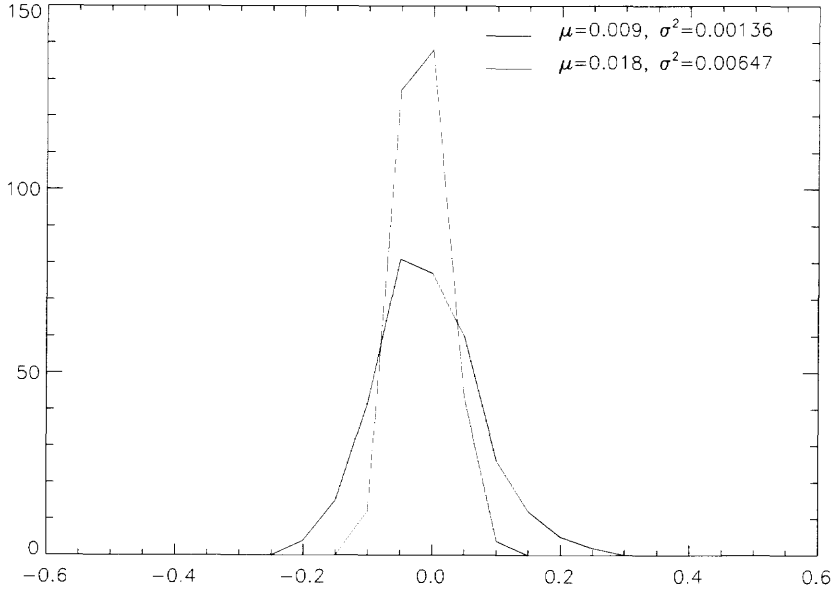
Figure 4.13(a) illustrates the RFI flag for the time domain kurtosis algorithm, obtained using 3σ as the detection threshold. Apparently this flag is more noisy compared with the spectral kurtosis flag (Figure 4.13(b)). No doubt, this reflects the fact that the spectral kurtosis estimator is more efficient and one can use the data to investigate whether the variance differ by a factor of three as predicted from theory. Figure 4.13(a) and 4.14(b) illustrate the flagged spectrum using the time domain kurtosis flag and spectral kurtosis flag, respectively.

Rather than comparing the flag pattern by eye, one can further quantitatively compare their variances from the RFI contaminated solar data. Figure 4.15(a) shows the kurtosis and spectral kurtosis for one particular channel (channel 150, corresponding to 16.47 MHz) without RFI contamination. The horizontal dotted lines represent the $\pm 3\sigma$ detection thresholds for $M = 1953$ points (SK) and $N = 3906$ points (kurtosis), respectively. The kurtosis estimator value has had the value 3 subtracted from the raw value, so it varies around zero. The spectral kurtosis estimator value is around 1 correspondingly. Their histograms are illustrated in Figure 4.15(b). The red and green gaussian-like curves are the histograms of spectral kurtosis and time domain kurtosis estimator for channel 150, respectively.

The spectral kurtosis estimator was reduced by 1 to align the bell shaped SK estimator distribution to zero value for better comparison with kurtosis estimator. Their variance for SK and time domain kurtosis estimators are found to be 0.00136 and 0.00647, respectively. These measurements can be compared to the theoretical values predicted by equation 4.16 and 4.3 by using $M = 1953$ and $N = 3906$, which are 0.00204 and 0.00614. Thus, for this channel the SK algorithm seems to be better than expected, while the time-domain algorithm is worse. The ratio between variances of time domain kurtosis and SK estimator



(a)



(b)

Figure 4.15 (a) The kurtosis and spectral kurtosis for one particular channel (channel 150, 1647 MHz) which is free of RFI. (b) The distribution of kurtosis (green) and spectral kurtosis (red).

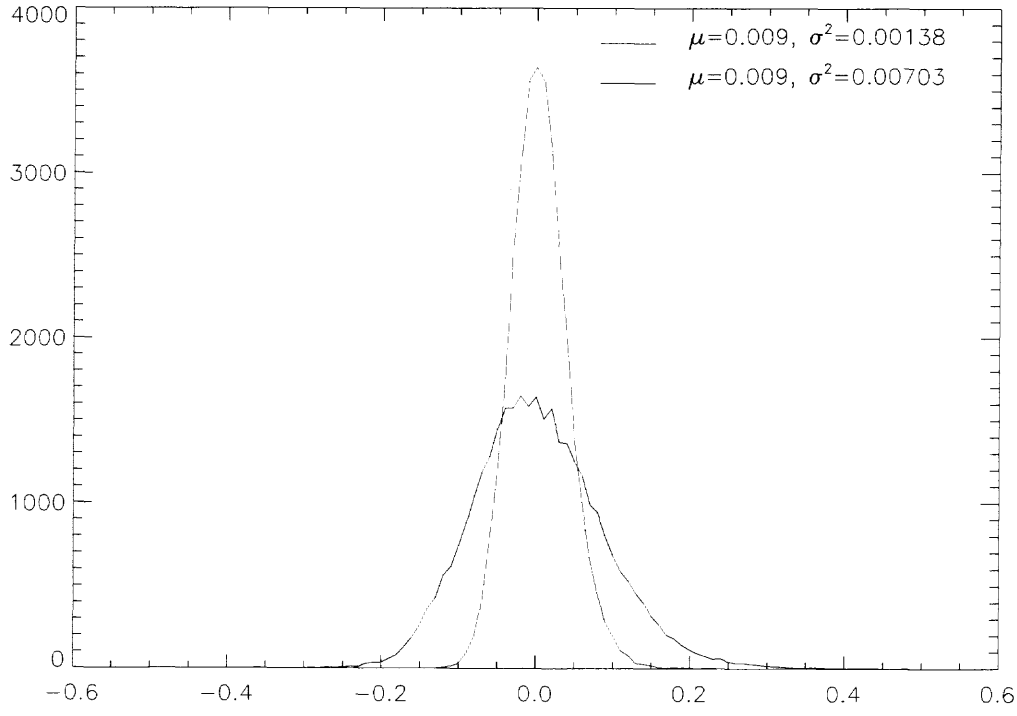


Figure 4.16 The distribution of kurtosis (green) and spectral kurtosis (red) of 100 channels (channel 100–200, 1.6–1.7 GHz) for a better statistics.

is 4.75 which is almost 60% larger than the predicted theoretical value 3.

The histograms for multiple channels from 100 to 200 (1.6-1.7 GHz) were combined in order to improve the statistics of this comparision (Figure 4.16). Very similar variance values, 0.00138 and 0.00703, were obtained, compared with theoretical values. The ratio between the variance of $\widehat{SK}$ and $\widehat{K}$ is now 5.09, even larger than the expected value of 3. This comparision has shown that the difference in variance is real and significant, and that the SK algorithm performs better than the time-domain algorithm for the same data set. The cause of the discrepancy in the relative variance with respect to the theoretical one is unknown, but may be due to a slight non-gaussian behavior, perhaps instrumental in nature.

4.5 Conclusion

The statistical properties of the PSD estimates obtained in the spectral domain (e.g., via FFT) have been analyzed, and an RFI excision algorithm suitable for real–time implementation in radio telescopes has been presented. The algorithm is based on spectral kurtosis obtained from DFT-based spectrograph data. The value of the SK estimator is unity in most spectral bins (see eq. 4.9), with a variance given by equation (4.15), which for practical purposed reduces to

$$var\left(\widehat{V_k^2}\right) \simeq \begin{cases} 24/M, & k = 0, N/2 \\ 4/M, & k = 1, \ldots, (N/2-1). \end{cases} \tag{4.17}$$

The variance in the DC and Nyquist bins ($k = 0, N/2$) is the well-known value $24/M$ obtained from time-domain analysis (c.f., Ruf et al. 2006), which unifies the time-domain and spectral-domain approaches.

The spectral kurtosis estimator, however, is expected to be more efficient than the time-domain kurtosis in terms of statistical variance by a factor of 3. The choice of which algorithm to be used is driven by whether spectral channels are derived from an DFT-based approach (where SK would be used) or from a finite impulse response (FIR) or polyphase filter approach (where time-domain kurtosis would be used).

The above results have been applied to the problem of RFI identification and excision using full-resolution time-domain data recorded by the FST instrument (Liu et al. 2007) during RFI surveys and solar observations. Taken into account the limitation on number of PSD estimates allowed by the FST instrument, the results are considered to be a successful experimental validation of the use of the SK estimator for RFI excision. More-

over, the algorithm appears to be most effective in the case of transient, low-duty-cycle RFI. If it proves less effective for channels with continuous RFI, such channels are easy to detect from simple power level considerations and it is always possible to flag them using a pre-defined, static frequency mask.

The comparison between classical time-domain kurtosis and the proposed spectral kurtosis estimator has been carried out not only qualitatively but quantitatively. The comparison further proved the superiority of the spectral kurtosis algorithm relative to time domain kurtosis. Although the spectral kurtosis estimator, $\widehat{SK}$, has a smaller variance by a factor of 3 than time domain kurtosis, $\widehat{K}$, according to the theory, the author obtained a larger factor, 5.09, by using real data in a direct head-to-head comparision.

The author agrees with Fridman & Baan (2001) that no single RFI excision technique will be 100% successful in identifying all possible artificial signals, but the SK estimator is a powerful tool in combination with other techniques, and is superior to the time-domain kurtosis estimator.

CHAPTER 5

CONCLUSION

The Sun, as the main source of energy to the Earth, is closely-linked to human life. Despite thousands of years of pondering the mysterious Sun, humanity had little physical understanding of the Sun until the last three centuries with the development of new technologies and the scientific method. Especially in recent decade, a numbers of earth-based observatories were built, and solar-dedicated spacecraft were launched, to observe and gather the data from the sun at all wavelengths. Among these emissions from the Sun, radio emission is considered very important since it provides the unique local physical information—especially magnetic fields—and the opportunity to study the energy release process in the atmosphere of the Sun. Observations of radio emission associated with solar flares provide diagnostics for flare physics, CME origin, particle acceleration and propagation, and coronal heating. But no existing radio facility has the ability to obtain both imaging and spectroscopy simultaneously over a large bandwidth with angular, time and spectral resolution corresponding to the properties intrinsic to solar radio emissions. From a recognition of this need, the concept of Frequency Agile Solar Radiotelescope was born. The FASR will be the first instrument that can obtain the precision information with 4-dimension (two spatial, spectral, and temporal) resolution simultaneously. In order to accomplish this goal, the detailed prototyping and characterization of the broadband RF transmission system used for FASR is worth careful study.

The FASR Subsystem Testbed (Liu et al. 2007) has been developed to prototype and characterize the broadband RF transmission system and digital signal processing to be

used for FASR. As a prototype for FASR, FST also provides the opportunity to study the design, calibration and interference-mitigation requirements of FASR. Three antennas in Owens Valley Solar Array were utilized, and upgraded to the new broadband RF transmission system. The design and construction of the instrument described in Chapter 2 utilizes new technology that has only recently become available, such as the image-rejecting spectral line downconverter and the four-channel high-speed digitizer. The data rate out to digitizer (> 40 MB $s-1$), and the data volume (300 GB/day), can only be handled due to recent advances in computing hardware. Taking advantage of these recent new capabilities makes FST unique and state-of-the-art. All observations and tests show that FST performs its expected function extremely well. Tracking GPS and geosynchronous satellite made it possible to calibrate and test the performance for both total power and interferometry. The solar bursts captured by the FST system nicely demonstrate the promise of FST, and hence that of the future FASR. In particular, Since the correlated spectra are derived through offline software from full-resolution data, FST provides the flexibility to process the data multiple times with alternate algorithms and parameters, to simulate various hardware designs using real data. It is this unique capability that makes FST so valuable as a testbed instrument, and this was key to the head-to-head comparison of RFI mitigation algorithm included in this thesis. FST observations provide an excellent database for testing RFI mitigation algorithms in both total power and correlated data, under real-world conditions such as one might expect for FASR. FST will continue to provide a platform for testing more advanced RFI mitigation algorithm, as well as detailed testing of digital signal processing designs for FASR.

By the end of year 2006, 12 events during six months daily observation have been

successfully captured. Many of the events captured are decimetric type III burst, as expected when planning the observations, but it was surprising to see several strong events with type IV continuum, including the intermediate drift fine structure. Especially surprising was the record-setting burst of 2006 December 6, which produced the largest solar radio flux ever measured, around one million solar flux units. This event had a severe impact on the reception of GPS signals. This brought more attention to the impact of solar radio energy release on the satellite and wireless communication systems at Earth. The 2006 December 6 burst shows the fine set of fibers bursts that drift from high to low frequencies over 6-10 s across the whole observing frequency band ($1.0 \sim 1.5$GHz) with frequency drift rate $-70 \sim -25 MHz/s$. With this three element interferometer, their source locations was determined relative to the background noise storm emission. In order to obtain the absolution location information, phase calibration across the 500MHz bandwidth is needed. This is a problem that remains to be solved for FST, because of the lack of point calibration sources bright enough to be measured with only three 2-m antennas. One solution is to use satellite for calibration, although in Chapter 2 it was pointed out that better satellite coordinates are needed to make this possible. Calibration will not be a problem for FASR, with of order 100 of these antennas. Despite the uncertainty in absolute position of the fibers, however, the work described in this thesis shows the promise of FST, and even more so for FASR, to provide detailed physical information on the density and magnetic field structure of flaring loops through the tracking of fiber bursts and other drifting bursts.

The quality of radio astronomy scientific data may be greatly affected by radio frequency interference (RFI) contaminating the natural radio emission produced by astrophysical objects. With the increasing demand for wireless communication systems oper-

ating in frequency bands of scientific interest, while astronomers seek to observe over an ever broader part of the radio spectrum, the RFI environment becomes ever more hostile.

The techniques in RFI mitigation may be broadly categorized according to whether the excision is based on time-domain or frequency-domain analysis, and on whether a reference signal is available to provide adaptive noise cancellation (Fridman & Baan 2001). Among these methods, RFI mitigation algorithms based on higher order statistics in both the time domain (Ruf et al. 2006) and the frequency domain (Fridman 2001; Fridman & Baan 2001) have become increasingly popular in recent years due to the feasibility of their implementation using field programable gate arrays (FPGA), with which many modern digital instruments are equipped. Based on the fact that system and radio source noise is thermal noise which has a gaussian distribution and a zero mean, whereas signal sources other than thermal noise have non-gaussian distribution, high order of statistics can be efficient in discriminating gaussian and non-gaussian signals. Time domain kurtosis algorithm has been applied to the RFI mitigation first in the remote sensing field (Ruf et al. 2006). For lack of a clear and unifying picture of the subject, and lack of instruments technically able to perform broadband spectral analysis with the temporal and frequency resolution needed for efficient implementation and testing of such RFI excision algorithms, The frequency domain statistical analysis (spectral kurtosis algorithm) has not previously been developed as a RFI detection tool in astrophysical radio instrument despite its general applicability. From this perspective, the newly designed FST offers a unique testbed capability. Its special design allows the direct comparison of various RFI excision algorithms by playing back in a simulated real-time mode.

Nita et al. (2007) analyzed the statistical properties of the PSD estimates obtained

in the spectral domain (e.g., via FFT) and presented an RFI excision algorithm suitable for real–time implementation in radio telescopes. The algorithm is based on spectral kurtosis obtained from DFT-based spectrograph data. The value of the SK estimator is unity with a variance given by

$$var\left(\widehat{V_k^2}\right) \simeq \begin{cases} 24/M, & k = 0, N/2 \\ 4/M, & k = 1, \ldots, (N/2-1). \end{cases} \tag{5.1}$$

The spectral kurtosis estimator, however, is expected to be more efficient than the time-domain kurtosis in terms of statistical variance by a factor of 3. The choice of which algorithm to be used is driven primarily by whether spectral channels are derived from an DFT-based approach (where SK would be used) or from a finite impulse response (FIR) or polyphase filter approach (where time-domain kurtosis would be used). This thesis provides the first direct, head-to-head comparison of these two kurtosis-based algorithms, in order to quantitatively examine their relative efficiencies. The comparison confirms the superiority of the spectral kurtosis algorithm relative to time domain kurtosis. Although the spectral kurtosis estimator, $\widehat{SK}$, has a smaller theoretical variance by a factor of 3 than time domain kurtosis, $\widehat{K}$, according to the theory, the direct, head-to-head comparison yielded a larger factor, 5.09, for reasons not fully understood. The speculation is that the real-world FST data may not obey perfectly gaussian statistics.

There is no single RFI excision technique will be 100% successful in identifying all possible artificial signals. Indeed, both the spectral and time-domain kurtosis estimators share on weakness, a "blindness" to RFI with a 50% duty cycle (Nita et al. 2007). But the SK estimator is a powerful tool in combination with other techniques, and as this thesis

demonstrates, it performs better for a given quantity of data than the time-domain kurtosis estimator.

The work described in this thesis has made unique, original contributions in several areas of instrument development, algorithm development for RFI excision, and scientific understanding of the Sun. Studies of FST data will continue to provide insight into solar radio emission processes and physical conditions in the solar atmosphere. At the same time, FST will remain unique and valuable for the design of FASR, and will continue to contribute to efforts to improve data quality of radio instruments in the face of ever increasing radio frequency interference.

REFERENCES

Allaart, M. A. F., van Nieuwkoop, J., Slottje, C., & Sondaar, L. H. 1990, Sol. Phys., 130, 183

Amblard, P. O., Gaeta, M., & Lacoume, J. L. 1996a, Signal Process., 53, 1

—. 1996b, Signal Process., 53, 15

Antoni, J. 2006, Mechanical Systems and Signal Processing, 20, 282

—. 2007, Mechanical Systems and Signal Processing, 21, 108

Aurass, H., Rausche, G., Mann, G., & Hofmann, A. 2005, Astron. Astrophys., 435, 1137

Bastian, T. S. in , Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 4853, Innovative telescopes and instrumentation for solar astrophysics. edited by stephen l. keil, sergey v. avakyan . Proceedings of the SPIE, ed. S. L. KeilS. V. Avakyan, 98–110

Bastian, T. S., Benz, A. O., & Gary, D. E. 1998, Annu. Rev. Astron. Astrophys., 36, 131

Bastian, T. S., Bradley, R., White, S., & Mastrantonio, E. 2005, AGU Spring Meeting Abstracts, A16+

Benz, A. O. Solar and Space Weather Radiophysics, ed. , D. E. GaryC. U. Keller (Kluwer Academic Publishers)

Benz, A. O., Magun, A., Stehling, W., & Su, H. 1992, Sol. Phys., 141, 335

Benz, A. O. & Mann, G. 1998, Astron. Astrophys., 333, 1034, provided by the Smithsonian/NASA Astrophysics Data System

Benz, A. O., Saint-Hilaire, P., & Vilmer, N. 2002, Astron. Astrophys., 383, 678

Bock, D. C.-J., Bolatto, A. D., Hawkins, D. W., Kemball, A. J., Lamb, J. W., Plambeck, R. L., Pound, M. W., Scott, S. L., Woody, D. P., & Wright, M. C. H. 2006, in Ground-based and airborne telescopes. edited by Stepp, Larry M. Proceedings of the SPIE, Vol. 6267, 626713

Bogod, V. M., Garaimov, V. I., Komar, N. P., & Korzhavin, A. N. 1999, in Esa sp-448: magnetic fields and solar processes. Edited by Wilson, A. et al., 1253–+

Chan, J. C. & Navarro, D. 2001, in ESA SP-473: Space Debris, ed. H. Sawaya-Lacoste, 771–779

DeBoer, D. R., Welch, W. J., Dreher, J., Tarter, J., Blitz, L., Davis, M., Fleming, M., Bock, D., Bower, G., Lugten, J., Girmay-Keleta, G., D'Addario, L. R., Harp, G. R., Ackermann, R., Weinreb, S., Engargiola, G., Thornton, D., & Wadefalk, N. 2004, in Ground-based Telescopes. Edited by Oschmann, Jacobus M., Jr. Proceedings of the SPIE, Volume 5489, pp. 1021-1028 (2004)., ed. J. M. Oschmann, Jr., 1021–1028

Dulk, G. A. & McLean, D. J. 1978, Sol. Phys., 57, 279

Dwyer, R. F. 1983, International Conference on Acoustic, Speech, and Signal Processing, Boston, 607

Elgaroy, O., ed. 1973, Proc. 3rd Meeting CESRA (Brodeaux Floirac)

Elgaroy, O. 1982, Intermediate Drift Bursts, Tech. rep.

Fleishman, G. D., Gary, D. E., & Nita, G. M. 2003, Astrophys. J., 593, 571

Fridman, P. A. 2001, Astron. Astrophys., 368, 369

Fridman, P. A. & Baan, W. A. 2001, Astron. Astrophys., 378, 327

Gary, D. E. & Hurford, G. J. Proceedings of the nobeyama symposium, held in kiyosato, japan, oct. 27-30, 1998, eds.: t. s. bastian, n. gopalswamy and k. shibasaki, nro report no. 479., pp. 429-432, ed. , T. S. BastianN. Gopalswamy & K. Shibasaki

Gary, D. E., Hurford, G. J., Liu, Z., Nita, G. M., & White, S. M. 2006, in AAS/Solar Physics Division Meeting

Gary, D. E. & Keller, C. U. 2004, Solar and Space Weather Radiophysics - Current Status and Future Developments (ASTROPHYSICS AND SPACE SCIENCE LIBRARY Vol. 314 Kluwer Academic Publishers, Dordrecht)

Hartley, R. 1928, U.S.Patent 1666 206

Hurford, G. J., Read, R. B., & Zirin, H. 1984, Sol. Phys., 94, 413

Karpman, V. I. & Washimi, H. 1977, Journal of Plasma Physics, 18, 173

Kendall, M. G. & Stuart, A. 1958, The Advanced Theory of Statistics Vol. 1, ed. Griffin

Kenney, J. F. & Keeping, E. S. 1962, Mathematics of Statistics, 3rd ed. Princeton, NJ, Van Nostrand

Kerdraon, A. & Delouis, J.-M. 1997, in Lecture Notes in Physics, Berlin Springer Verlag, Vol. 483, Coronal Physics from Radio and Space Observations, ed. G. Trottet, 192–+

Kuijpers, J. 1975, Sol. Phys., 44, 173

Kuijpers, J. & Slottje, C. 1976, Sol. Phys., 46, 247

Lang, K. R. 1980, Astrophysical Formulae (Springer-Verlag, Second and enlarge Edition)

Lecacheux, A. 2000, in Radio astronomy at long wavelengths. Edited by Stone, R. G., Weiler, K. W., Goldstein, M. L., Bougeret, J.-L., 321–+

Liu, Z., Gary, D. E., Nita, G. M., White, S. M., & Hurford, G. J. 2007, PASP, 119, 303

Mann, G. 1985, Journal of Plasma Physics, 33, 21

McLean, D. J. & Labrum, N. R. 1985, Solar radiophysics: Studies of emission from the sun at metre wavelengths (Cambridge University Press)

Meléndez, J. L., Sawant, H. S., Fernandes, F. C. R., & Benz, A. O. 1999, Sol. Phys., 187, 77

Messmer, P., Benz, A. O., & Monstein, C. 1999, Sol. Phys., 187, 335

Nakajima, H., Nishio, M., Enome, S., Shibasaki, K., Takano, T., Hanaoka, Y., Torii, C., Sekiguchi, H., Bushimata, T., Kawashima, S., Shinohara, N., Irimajiri, Y., Koshi-ishi, H., Kosugi, T., Shiomi, Y., Sawa, M., & Kai, K. 1994, Proceedings of The IEEE, 82, 705

Napier, P. J., Thompson, A. R., & Ekers, R. D. 1983, Proceedings of the IEEE, 71, 1295

Newkirk, G. J. 1961, Astrophys. J., 133, 983

Nita, G. M., Gary, D. E., Lanzerotti, L. J., & Thomson, D. J. 2002, Astrophys. J., 570, 423

Nita, G. M., Gary, D. E., Liu, Z., Hurford, G. J., & White, S. M. 2007, PASP, 119

Paesold, G., Benz, A. O., Klein, K.-L., & Vilmer, N. 2001, Astron. Astrophys., 371, 333

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. 1989, Numerical Recipies in Pascal (Cambridge University Press)

Prestage, N. P., Luckhurst, R. G., Paterson, B. R., Bevins, C. S., & Yuile, C. G. 1994, Sol. Phys., 150, 393

Ruf, C. S., Gross, S. M., & Misra, S. 2006, IEEE Transactions On Geoscience and Remote Sensing, 44

Slottje, C. 1972, Sol. Phys., 25, 210

Spatschek, K. H., Shukla, P. K., Yu, M. Y., & Karpman, V. I. 1979, Physics of Fluids, 22, 576

Staehli, M. & Benz, A. O. 1987, Astron. Astrophys., 175, 271

Suzuki, S. & Dulk, G. A. Solar Radiophysics, ed. , D. J. McLeanN. R. Labrum (Camebridge University Press), 289

Thompson, A. R., Moran, J. M., & Swenson, Jr., G. W. 2001, Interferometry and Synthesis in Radio Astronomy, 2nd Edition

Treumann, R. & Bernold, T. E. X. 1981, Phys. Rev. Lett., 47, 1455

Treumann, R. A., Guedel, M., & Benz, A. O. 1990, Astron. Astrophys., 236, 242

Trottet, G., Pick, M., House, L., Illing, R., Sawyer, C., & Wagner, W. 1982, Astron. Astrophys., 111, 306

Vrabie, V., Granjon, P., Maroni, C.-S., & Leprettre, B. 2004, in 5th International Conference on Acoustical and Vibratory Surveillance, Methods and Diagnostic Techniques, Senlis, France

Vrabie, V. D., Granjon, P., & Servière, C. 2003, IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing

Young, C. W., Spencer, C. L., Moreton, G. E., & Roberts, J. A. 1961, Astrophys. J., 133, 243

Zirin, H., Baumert, B. M., & Hurford, G. J. 1991, Astrophys. J., 370, 779