Hardware Implementation of an SK Spectrometer

Dale E. Gary[*][†]

New Jersey Institute of Technology
E-mail: dgary@njit.edu

Zhiwei Liu

New Jersey Institute of Technology
E-mail: zwliu@adm.njit.edu

Gelu M. Nita

New Jersey Institute of Technology
E-mail: gnita@adm.njit.edu

A wideband spectrometer of 500 MHz instantaneous bandwidth is described that includes an implementation of the Spectral Kurtosis (SK) algorithm for automatic radio frequency interference (RFI) detection. An SK spectrometer accumulates both power and power-squared, which are then used to develop the SK estimator. The SK spectrometer of the Korean Solar Radio Burst Locator (KSRBL) has a 2 GHz instantaneous bandwidth that is time multiplexed over the 0.24-18 GHz radio frequency range, for the study of solar bursts. We use a novel presentation of the data from the spectrometers to classify types of RFI. It is found that the algorithm is especially useful for excising highly intermittent RFI but is less successful for RFI due to digital signals. The method of multiscale SK is presented that addresses the known blindness of Kurtosis-based estimators to 50% duty-cycle RFI.

---

[*]Speaker.

1. Introduction

The Korean Solar Radio Burst Locator (KSRBL) [1] is the first spectrometer to employ the technique of spectral kurtosis (SK), which has been described in detail by Nita et al. [8] and Nita & Gary [6]. Gary et al. [5] have demonstrated the effectiveness of the SK algorithm for identifying and removing certain kinds of radio frequency interference (RFI) from radio astronomical data. This paper summarizes the results based on KSRBL observations and modeling. The key aspect of the algorithm is that instantaneous power spectral density (PSD) measurements are squared in hardware, without first accumulating or averaging the power, and the resulting sums of power $S_1 = \sum P_i = M\langle P \rangle$ and power-squared $S_2 = \sum P_i^2 = M\langle P^2 \rangle$ are used either on board or in subsequent software to generate the unbiased estimator [5]

$$\widehat{SK} = \frac{M+1}{M-1}\left(\frac{MS_2}{S_1^2} - 1\right), \tag{1.1}$$

where M is the number of accumulations in the sums. Nita & Gary [6] (see also [7] in this issue) have shown that the first four moments of the probability density function (PDF) of $\widehat{SK}$, representing mean μ, variance σ^2, skew γ_1, and kurtosis excess γ_2 are

$$\mu = 1; \ \sigma^2 \simeq \frac{4}{M}; \ \gamma_1 \simeq \frac{10}{\sqrt{M}}; \ \gamma_2 \simeq \frac{246}{M}. \tag{1.2}$$

Although three of these moments approach the normal distribution behavior as $1/M$, the skewness decreases only as $1/\sqrt{M}$. The significance of equations (1.1) and (1.2) is that purely Gaussian noise expected of an astronomical signal should be distributed according to this PDF while signals contaminated with non-Gaussian noise (i.e. certain kinds of RFI) will form outliers that can be identified and flagged.

Nita & Gary [6] describe the procedure for calculating thresholds, which are highly dependent on M, for determining outliers to the distribution. In normal operation, the KSRBL spectrometer accumulates $M = 6104$ estimates of power and power-squared in ≈ 25 ms. Figure 1 shows the measured distribution of $\widehat{SK}$ for an RFI-free band measured with KSRBL, overlaid with the theoretical PDF derived by Nita & Gary [6] for the case of $M = 6104$. In this particular case the appropriate thresholds equivalent to $\pm 3\sigma$ for a normal distribution are $1 - 0.073$ and $1 + 0.081$, which due to the skew of the distribution are slightly asymmetric around the nominal value of 1. There is essentially perfect agreement between theory and data in this RFI-free case.

2. Board Implementation

Gary et al. [5] give details of the board implementation of the SK spectrometer, while Dou et al. [1] describe its use in the KSRBL system. The implementation is based on the Internet Break-Out Board (IBOB) hardware and firmware tools developed by the Center for Astronomical Signal Processing and Electronics Research (CASPER) [10]. The spectrometer design source files may be downloaded from http://casper.berkeley.edu/wiki/Kurtosis_Spectrometer. The hardware includes a CASPER-designed analog/digital converter (ADC) fed by an upstream analog system that employs automatic leveling control. This is important, because variations in power level to
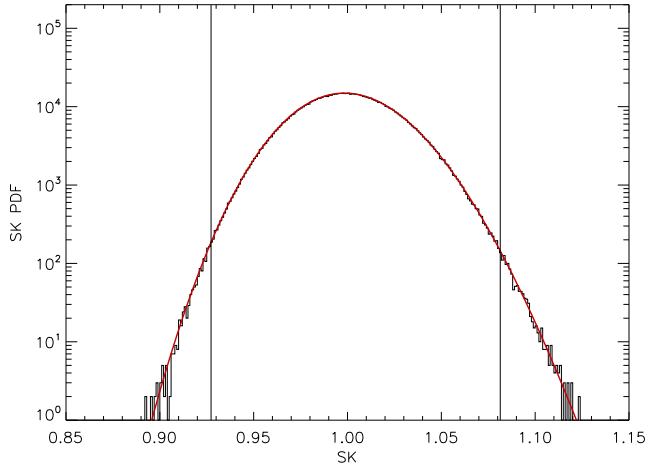
Figure 1: Histogram of spectral kurtosis values for KSRBL data when pointed at the Sun, in the frequency band 7500-8000 MHz (no RFI). The black line shows the histogram of the data, while the red line shows the theoretical PDF for the case of $M = 6104$. The vertical lines show the thresholds given by Nita & Gary [6], corresponding to the $\pm 0.135\%$ tail probability equivalent to $\pm 3\sigma$ for a normal distribution.

the digitizer during an accumulation can alter the statistics of RFI-free data (see [8] for further discussion). Each of the four SK spectrometer boards digitizes a 500 MHz bandwidth at 8-bits using a 1 GHz clock, and uses an FFT-based filterbank to divide the signal into 2048 frequency channels. Power and power-squared are accumulated from the instantaneous PSD measurements on each channel, and dumped for off-board (although real-time) calculation of $\widehat{SK}$ in the recording computer system.

Due to resource and data-volume constraints, the power and power-squared accumulations are done on limited bit-width data. The design includes several scaling and bit-select registers that keep these accumulations in an optimum range, without which the statistics of the SK estimator are adversely affected. Once the spectrometer is properly scaled for a given system, however, the scaling parameters may be left fixed. Gary et al. [5] give further details of scaling and the adverse effects of improperly set bit registers.

3. Classifying RFI Using the $\widehat{SK}$ *versus* S_1 Diagram

To understand the behavior of the SK estimator for non-Gaussian signals (i.e. RFI), Nita et al. [8] compared the theoretical response of $\widehat{SK}$ to a continuous-wave (CW) signal of constant amplitude, which can be used to simulate transient RFI by considering its presence or absence with some duty cycle d. Consider M contiguous PSD estimates of which R are contaminated by RFI at signal to noise ratio (S/N), η_k, where k is the spectral channel index. The duty cycle is therefore defined as $d = R/M$. For this case, the SK estimator expectation value is

$$\widehat{SK} = \frac{M+1}{M-1}\left[1 + \frac{(1/d-2)\eta_k^2}{(1/d+\eta_k)^2}\right], \tag{3.1}$$

which has the following properties: for $d = \frac{1}{2}$ (50% duty cycle), the estimator is always 1; for $d < \frac{1}{2}$ (highly intermittent RFI), the estimator is above 1; for $d > \frac{1}{2}$ (more continuous RFI), the estimator is below 1.

These characteristics are clearly visible in the example shown in Figures 2 and 3, covering the 500-1000 MHz range with considerable RFI. The RFI in this band consists of both continuous and
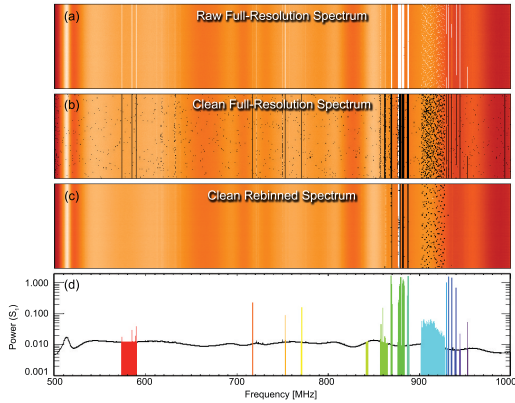
Figure 2: (*a*) The 150 full-resolution snapshot spectra (2048 frequency subchannels), shown as a dynamic spectrum. (*b*) The clean dynamic spectrum, again with 2048 frequency subchannels, after applying the SK flags. Bins where the SK estimator exceeds the threshold are black. (*c*) The same as in (*b*), but now rebinned to 512 subchannels by averaging 4 adjacent frequency bins. (*d*) Average spectrum formed by summing over the 150 snapshot spectra.
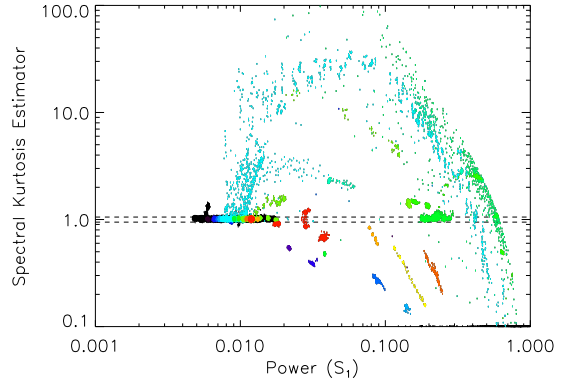
Figure 3: Plot of SK estimator vs. S_1, for the 150 accumulated spectra in Fig. 2. The horizontal dashed lines indicate the thresholds used in Fig. 1. Each dot represents a single frequency-time bin in Fig. 2, and the dots for selected frequency ranges are color-coded with the same colors as in Fig. 2*d*. The colors correspond to the marked frequency ranges in Fig. 2*d*, whether or not RFI is present.

intermittent RFI, and the different types can be traced in the $\widehat{SK}$ *versus* S_1 plot of Fig. 3 by means of the color coding in Fig. 2*d*. The intermittent RFI (cyan color) forms a fountain of points in Fig. 3, lying mostly above $\widehat{SK} = 1$ due to its less than 50% duty cycle. The frequencies of continuous RFI, in contrast, produce points that lie primarily below $\widehat{SK} = 1$ due to their 100% duty cycle. Note, however, the green points in the range $0.2 < S_1 < 0.3$ that correspond to frequencies near 877 MHz in Fig. 2. It is clear that the characteristics of this RFI somehow mimic Gaussian noise statistics so that the SK estimator is closely equal to 1. For that reason, the algorithm does not detect the RFI and it appears unflagged in Fig. 2*c*. These results generally show that the SK algorithm is effective in removing RFI, especially the highly-intermittent type, but that some types of RFI do somehow escape detection.

We have investigated other bands (see [5] for examples) and have found other instances where the RFI escapes detection. In these cases, the source of RFI could be positively identified as due to digital transmissions (XM and Sirius digital radio and digital data links). We therefore have simulated [5] a typical method of encoding information by digital transmitters, viz. Quadrature Phase-Shift Keying [11], which encodes symbols by manipulating their phase and amplitude in a finite set of discrete values. Using such a simulation, we find that indeed the SK estimator PDF mimics that for Gaussian noise almost precisely. We therefore conclude that the SK algorithm is not well suited to identify digital transmissions. Nevertheless, such digital transmissions are likely to be at fixed frequencies that can be identified by other means. The SK algorithm remains extremely useful for identifying the particularly problematic, intermittent RFI that appears randomly and briefly at a given frequency, as well as most sources of continuous, analog RFI.

4. Multiscale SK

One of the advantages of retrieving both S_1 and S_2 information from the hardware (as opposed to calculating SK in hardware and retrieving only the flags) is that these sums can be combined in alternative ways to address the algorithm's blindness to RFI with a duty-cycle near 50%—c.f. equation (3.1). Consider a frequency-time bin b_{ik} in a dynamic spectrum at full resolution that contains RFI at duty cycle $d_{ik} = 0.5$, and an adjacent time bin that contains RFI at some other duty cycle $d_{i+1,k}$. Because of the linearity of S_1 and S_2, new values $S_1 = S_{1,ik} + S_{1,i+1,k}$ and $S_2 = S_{2,ik} + S_{2,i+1,k}$ can be formed, and a new SK can be calculated for the combined bins, whose duty cycle is now $d = (d_{ik} + d_{i+1,k})/2$. If $d_{i+1,k} = 0$ (no RFI) or $d_{i+1,k} = 1$ (continuous RFI), then the combined duty cycle is 0.25 or 0.75, respectively. In either case, the combined SK is likely to exceed the thresholds and could be used to flag bin b_{ik}. This is likely to be a good strategy in the case of highly intermittent RFI such as that shown in Figs. 2 and 3. Alternatively, for more continuous narrow-band RFI one might use instead an adjacent frequency channel, i.e. $S_1 = S_{1,ik} + S_{1,i,k+1}$ and $S_2 = S_{2,ik} + S_{2,i,k+1}$. In general, we define multiscale SK for bin i,k as the determination of alternative values of SK over multiple "macro-bins" of size $m \times n$, i.e.

$$S_{1,ik}^{mn} = \sum_{j=0}^{m} \sum_{l=0}^{n} S_{1,i+j,k+l}, \quad S_{2,ik}^{mn} = \sum_{j=0}^{m} \sum_{l=0}^{n} S_{2,i+j,k+l}, \tag{4.1}$$

each of which will have duty cycle

$$d_{ik}^{mn} = \frac{1}{(m+1)(n+1)} \sum_{j=0}^{m} \sum_{l=0}^{n} d_{i+j,k+l}. \tag{4.2}$$

These multiple values of SK, which we refer to as multiscale SK, can be combined in various ways to flag problem RFI. The KSRBL analysis software routinely applies multiscale SK with $m = n = 1$. See Gary et al. [5] for examples of the performance of the SK algorithm for different choices of m and n.

5. Conclusion

As new radio astronomy instruments explore ever wider radio frequency bandwidths, RFI has come to the fore as one of the key problems to be addressed [e.g. 2, 3, 4, 9]. No single approach will work in all cases, and it is necessary to employ layered RFI avoidance [12] and excision algorithms. The work of Nita et al. [8] and Nita & Gary [6, 7] has placed the SK algorithm on a firm theoretical footing, while Gary et al. [5], briefly summarized here, shows the practical application of the algorithm in hardware. Although the algorithm does not work optimally for all cases, notably for digitally encoded transmissions, it is especially powerful for one problematic type of RFI, highly intermittent RFI in frequency and time. This includes both fixed and mobile transmitters, especially aircraft DME (distance measuring equipment), as shown but not discussed by Nita et al. [8] (their figures 9 and 10). Its ability to identify and flag such RFI in real time makes it an important first line of defense after avoidance and front-end filters.

In this paper we have shown that the hardware implementation we have designed maintains the requisite precision to recover the $\widehat{SK}$ PDF to high accuracy when compared to the theoretical

PDF determined by Nita & Gary [6]. This permits use of the precise theoretical thresholds for identifying outliers due to RFI, and offers excellent detection while minimizing the false alarm rate. By introducing the $\widehat{SK}$ *versus* S_1 plot, we have provided a useful method of characterizing RFI based on its duty cycle. Further examples are given in Gary et al. [5].

The relative simplicity and ease of implementation of the SK algorithm in high-speed digital systems makes it highly suitable for inclusion in new radio spectrometer designs based on either time-domain or spectral-domain filtering. It can easily be used in concert with other RFI excision techniques, and we recommend its use.

References

[1] Dou, Y., Gary, D. E., Liu, Z., Nita, G. M., Bong, S.-C., Cho, K.-S., Park, Y.-D., & Moon, Y. J., *The Korean Solar Radio Burst Locator (KSRBL)*, PASP, 121, 512

[2] Fridman, P. A., *RFI excision using a higher order statistics analysis of the power spectrum*, Astronomy & Astrophysics, 368, 369

[3] Fridman, P. A., & Baan, W. A., *RFI mitigation methods in radio astronomy*, Astronomy & Astrophysics, 378, 327

[4] Fridman, P. A., *Statistically Stable Estimates of Variance in Radio-Astronomy Observations as Tools for Radio-Frequency Interference Mitigation*, AJ, 135, 1810

[5] Gary, D. E., Liu, Z., & Nita, G. M., *A Wideband Spectrometer with RFI Detection*, PASP, 122, 560

[6] Nita, G. M., Gary, D. E., *Statistics of the Spectral Kurtosis Estimator*, PASP, 122, 595

[7] Nita, G. M., Gary, D. E., *Statististics of the SK estimator*, in this proceedings, `PoS(RFI2010)019`

[8] Nita, G. M., Gary, D. E., Liu, Z., Hurford, G. J., & White, S. M., *Radio Frequency Interference Excision Using Spectral Domain Statistics*, PASP, 119, 805

[9] Offringa, A. R., de Bruyn, A. G., Biehl, M., Zaroubi, S., Bernardi, G., & Pandey, V. N., *Post-correlation radio frequency interference classification methods*, MNRAS, 405, 155

[10] Parsons, A., et al., *A Scalable Correlator Architecture Based on Modular FPGA Hardware, Reuseable Gateware, and Data Packetization*, PASP, 120, 1207

[11] Proakis, J. G. 2001, *Digital Communications, 4th Ed.*, McGraw-Hill (NY)

[12] Tzioumis, A., *Radio Quiet Zones (RQZ) - Working with national communication administrations*, 37th COSPAR Scientific Assembly, 37, 3247