A CROSS-CORRELATION BASED SPECTRAL KURTOSIS RFI DETECTOR

Gelu M. Nita*[1] and Gregory Hellbourg[2]

(1) New Jersey Institute of Technology, Newark, NJ, 08536, USA

(2) California Institute of Technology, Pasadena, CA, USA

URSI GASS 2020, Rome, Italy, 29 August - 5 September 2020

Accurate flagging of Radio Frequency Interference (RFI) is necessary to recover instrumental efficiency and avoid false astronomical detections.

Spectral Kurtosis (SK) is a popular operator in RFI flagging for radio astronomy due to its detection sensitivity to non-Gaussian emissions and its competitive computational cost.

Most SK detection pipelines are applied to single antennas or autocorrelations products.

This paper investigates the application of the SK to antennae cross-correlations and demonstrates an improved detection performance compared to the auto-correlation-based approaches.

ABSTRACT

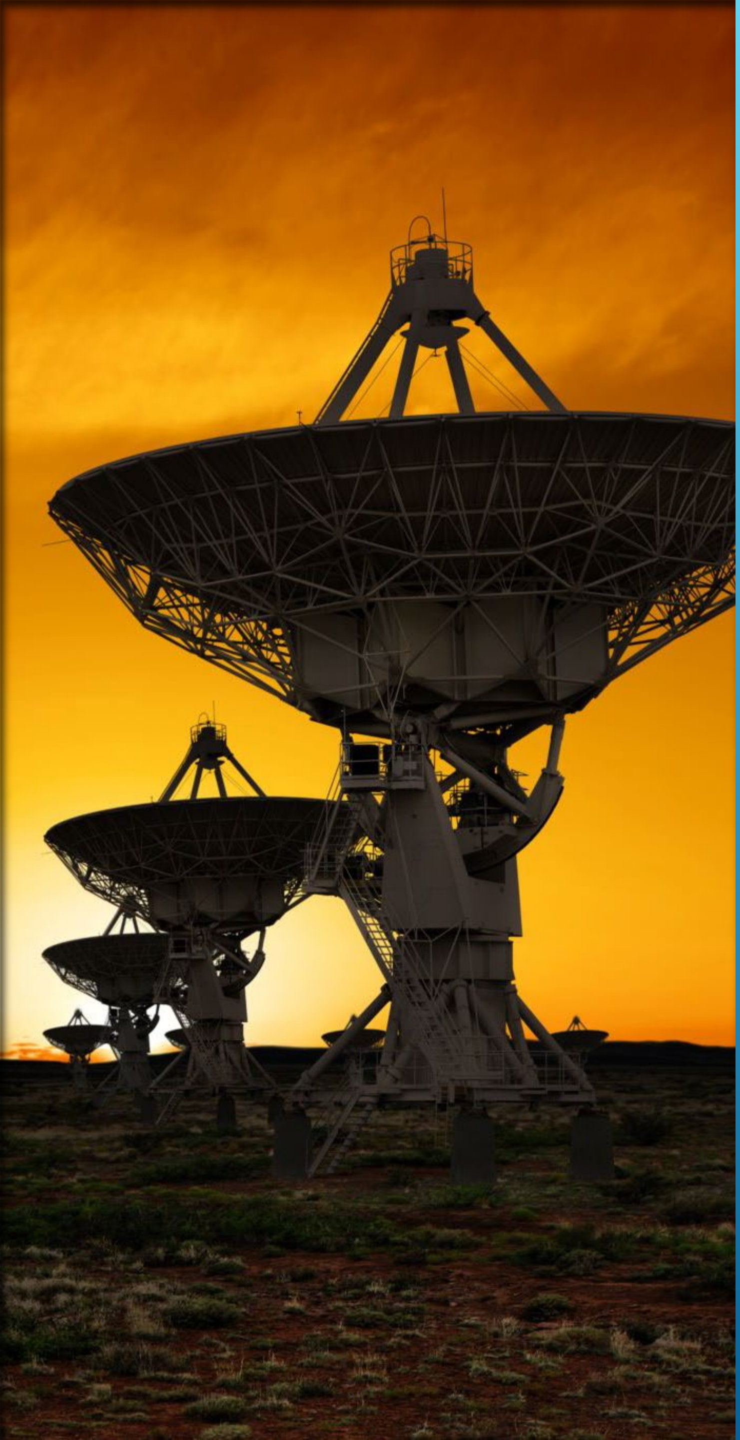
RFI Power

$$\underbrace{\sigma_{i,j}^2 = \sum_{s}^{N_s} a_{s_{i,j}} \sigma_s^2 + \sigma_{n_{i,j}}^2}_{\sigma_{H0}^2} + a_{r_{i,j}} \sigma_r^2 = \begin{cases} \sigma_{H0}^2 & (H0) \\ \sigma_{H0}^2 + a_{r_{i,j}} \sigma_r^2 & (H1) \end{cases}$$

Cross-correlation power

System noise and astronomical contribution

RFI-FREE VS RFI-CONTAMINATED ANTENNA CROSS-CORRELATION

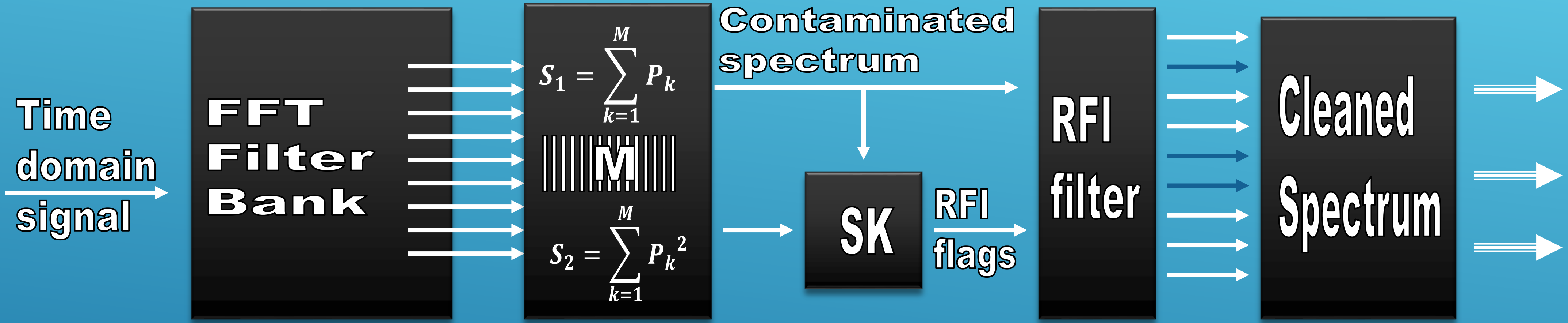


- The Spectral Kurtosis Spectrometer Design
- The Generalized Specral Kurtosis Estimator Theory

SPECTRAL KURTOSIS RFI FLAGING OF SINGLE-ANTENNA AUTO-CORRELATION SPECTRA

THE SPECTRAL KURTOSIS SPECTROMETER

NITA ET AL. 2007 PASP, 119, 805
GARY, LIU & NITA 2010 PASP, 122, 560



The unbiased Spectral Kurtosis Estimator

$$SK \equiv \frac{M+1}{M-1}\left(\frac{MS_2}{S_1^2} - 1\right)$$

$$E(SK) = 1$$

$$\sigma^2(SK) \approx \frac{4}{M}$$

THE GENERALIZED SPECTRAL KURTOSIS ESTIMATOR

Nita & Gary 2010, MNRAS 406 L60-L64

Theorem: Given that, for a particular signal, the set of its power estimates P_k obeys a gamma distribution characterized by the shape parameter ***d***, the infinite series of statistical moments $MS_2/{S_1}^2$, were $S_1 = \sum_{k=1}^{M} P_k$ and $S_2 = \sum_{k=1}^{M} {P_k}^2$, is given by:

$$E\left[\left(\frac{MS_2}{S_1^2}\right)^n\right] = \frac{M^n\Gamma(Md)}{\Gamma(d)^M\Gamma(Md+2n)} \times \frac{\partial^n}{\partial t^n}\left[\sum_{r=0}^{n}\frac{1}{r!}\Gamma(2r+d)t^r\right]^M\Bigg|_{t=0}$$

Corollary: The Generalized Spectral Kurtosis Estimator defined by

$$SK = \frac{Md+1}{M-1}\left(\frac{MS_2}{{S_1}^2} - 1\right)$$

- Has an unbiased unity expectation $E[SK] = 1$, independent of the integrated power S_1
- The infinite series of statistical moments of its PDF are analytically defined only in terms of M and d

The SK estimator is well suited for detecting mixed signals not obeying the same gamma probability distribution: Detection thresholds characterized by analytically defined probabilities of false alarm (PFA)

- OR Operator Flagging (OOF)
- Mean Auto-Correlation Flagging (MACF)
- Cross-Correlation Flagging (CCF)

MULTIPLE ANTENNAE RFI FLAGGING STRATEGIES

OR OPERATOR FLAGGING (OOF)

NITA ET AL. 2016, JAI, 5 (04)

This strategy flags antenna cross-correlations depending on the OR operation between the flags evaluated on the independently computed antenna auto-correlation Spectral Kurtosis RFI flags.

For any given pair of antennae, the probability of false alarm RFI flagging, PFA_{OR}, may be computed in terms of the individual antenna PFA_{auto} as

$$PFA_{OR} = (2 - PFA_{auto})PFA_{auto}$$

where PFA_{auto} may be computed using one of the analytical formulae provided by Nita & Gary 2010, MNRAS, 406, L60-L64

MEAN AUTO-CORRELATION FLAGGING (MACF)

The MACF strategy, proposes by Taylor et al. 2019, JAI, 08 (01) consists in averaging ***N*** individual, antenna-based SK estimators,

$$\langle SK \rangle = \frac{1}{N}\sum_{i=1}^{N} SK_i$$

The $\langle SK \rangle$ estimator has unity expectation and a variance that is ***N*** times smaller than that of the individual antenna-based estimators,

$$\sigma^2_{\langle SK \rangle} = \frac{1}{N} SK_i \approx 2\left(1 + \frac{1}{d}\right)\frac{1}{N \times M} + O\left[\frac{1}{(N \times M)^2}\right],$$

where the first order approximation of the series expansion allows one to compute the PFA associated with a chosen pair of RFI flagging thresholds using the simple substitution

$$M \rightarrow N \times M$$

in the analytical formulae provided by Nita & Gary 2010, MNRAS, 406, L60-L64

CROSS-CORRELATION FLAGGING (CCF)

The cross-correlation SK estimator proposed here is based on the definition of Generalized SK Estimator introduced by Nita & Gary 2010, where, for any given pair of antennas, {i, j}, the sums S_1 and S_2 are constructed in terms of M consecutive visibility squared amplitudes, already accumulated over N consecutive voltage samples,

$$S_1 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} x_i(n) x_j^*(n) \right| ; \quad S_2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} x_i(n) x_j^*(n) \right|^2 ,$$

under the assumption that the true statistical distribution of the inner cross-correlation accumulations $\left|\sum_{n=1}^{N} x_i(n) x_j^*(n)\right|$ may be approximated by a Gamma distribution having an instrumental shape factor **d** that can be empirically determined through a calibration process.

EMPIRICAL CALIBRATION OF THE CROSS-CORRELATION SK ESTIMATOR

EMPIRICAL CALIBRATION STEPS

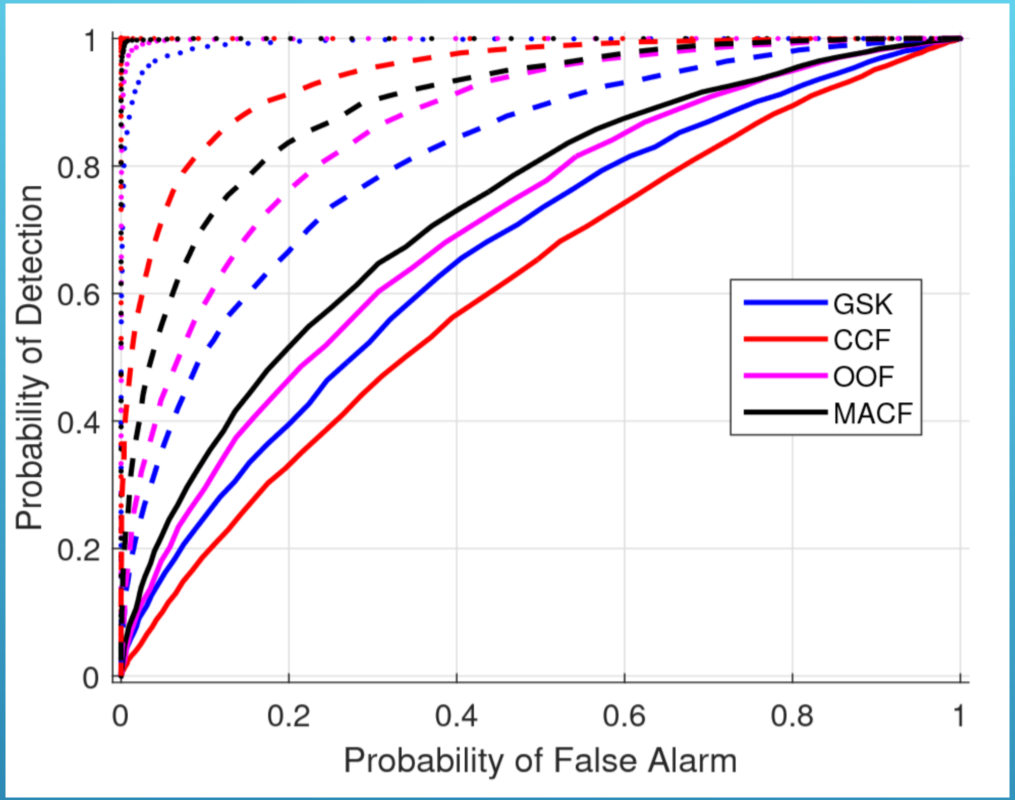
To obtain an empirical calibrated cross-correlation SK estimator having unity expectation, we follows the following steps.

1. Estimate the $(d \equiv 1)$ mean $\mu = \left\langle \frac{M+1}{M-1}\left(\frac{M\,S_2}{S_1^2} - 1\right)\right\rangle$ using a statistically significant set of RFI-free accumulations of squared visibility amplitudes.
2. Estimated the unknown instrumental shape factor need to normalize the SK estimator to unity $d = \frac{M-\mu+1}{\mu M}$
3. Compute the normalized estimator $SK = \frac{Md+1}{M-1}\left(\frac{M\,S_2}{S_1^2} - 1\right)$
4. Compute the RFI detection thresholds corresponding to a desired PFA, using the one of the analytical formulae provided by Nita & Gary 2010, MNRAS, 406, L60-L64 `

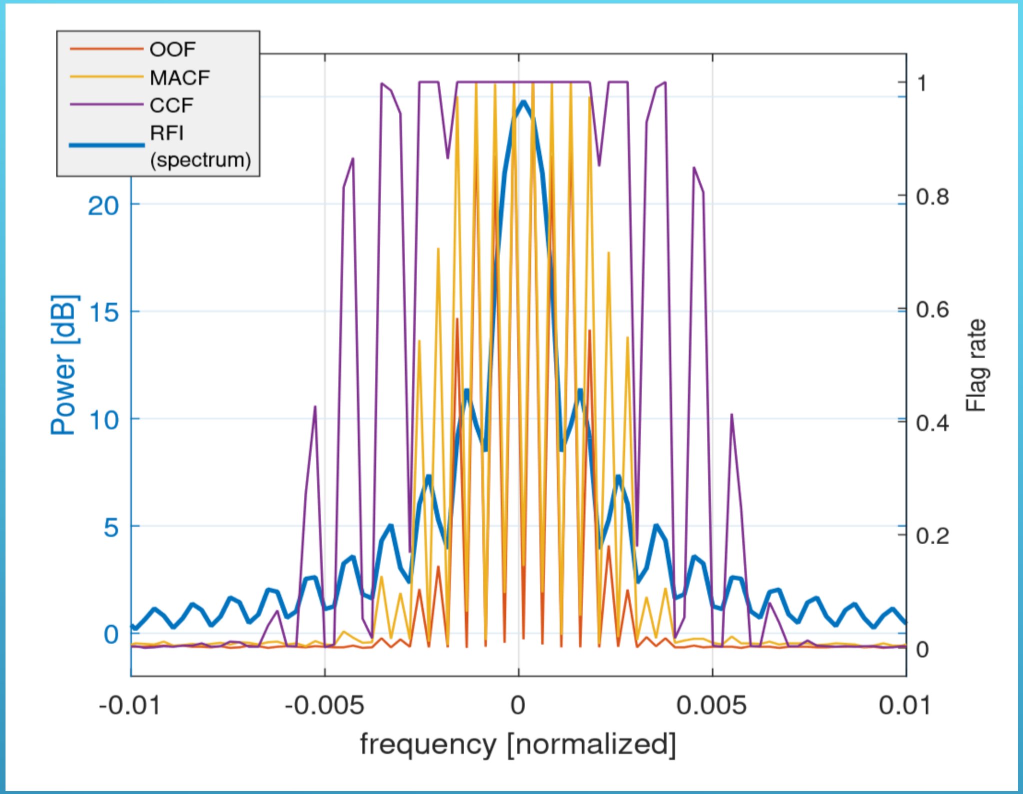
COMPARISON AND PERFORMANCE TESTS

NUMERICAL SIMULATIONS

- To evaluate and compare the performance of the various detection strategies presented here, we conducted a series of Monte-Carlo simulations, to empirically evaluate the probability of detection (PD) and probability of false alarm (PFA).
- For each of the 4096 independent trials performed, sets of n =10,000 data samples were randomly generated under both (H0) and (H1) hypothesizes.
- The correlator and SK engine parameters chosen for this simulation were M = 1,000 and N = 10.



ROC curves for the single antenna SK operator, and 2-antenna CCF, OOF, and MACF detection strategies, for SNR = -10 dB (plain curves), SNR = -8 dB (dashed curves), and SNR = -5 dB (dotted curves)



Comparison of detection sensitivities of the OOF, MACF, and CCF flagging strategies at detecting a BPSK signal with 2 antennas, N=10, M=100, SNR=-8dB PFA=0.013. Statistics evaluated over 1,000 independent realizations.

COMPARISON RESULTS

- Interferometers have good RFI rejection properties due to RFI de-correlation over large baselines, but they are also more sensitive to weak RFI in cross-correlation products over short baselines (shorter than the coherence wavelength of the channelization bandwidth). We propose therefore the use of the SK detector over such data product to improve the detection performance of other SK flagging strategies for array radio telescopes.
- A Monte-Carlo simulation shows that the cross-correlation SK detector outperforms all other detection strategies in a simple data model involving a white Gaussian system noise corrupted by a continuous wave RFI. Moreover, the PFA for the SK detector can be analytically formulated, and accurate detection threshold can be derived under realistic RFI-free data modelling.
- Further improvement in detection performance can be expected by combining the flagging information of multiple short array baselines, depending on the availability of computational resources, similarly to the auto-correlation approaches.
- Further detection performance assessments remain to be conducted on more realistic scenarios, e.g. involving system uncalibration, complex information-bearing sources of RFI, non-negligible astronomical sources, non-iid or non-white system noise.
- The low computational complexity of the SK detector and its detection performance make it an interesting online RFI flagging solution for the next generation radio interferometers with large number of antennas.

CONCLUSION