# Bispectra of simulated GPS data for potential RFI mitigation

V. van Tonder*[1][2], L. Schwardt[1], A. Faustmann[2] and J. Gilmore[2]

(1) South African Radio Astronomy Observatory, Cape Town, South Africa, www.sarao.ac.za

(2) Department of Electrical and Electronic Engineering, Stellenbosch University, Stellenbosch, South Africa, www.sun.ac.za

## Abstract

Radio telescopes are becoming more powerful, with increased sensitivity to both astronomical and interfering sources. Various interference mitigation techniques exist, however they could be harmful to the signal of interest, thus it is important to develop new techniques. Majority of radio astronomical sources are considered to have a Gaussian distribution which are fully characterised using second-order power statistics. Higher-order statistics is a tool that can suppress Gaussian noise, therefore if interference have energy in higher domains it could form the basis for a new mitigation technique. This work uses a third-order statistic, the bispectrum, to conduct analysis on global positioning system (GPS) signals. It is found that one of the in-phase components of GPS signals has energy in the bispectrum.

## 1 Introduction

Radio astronomy is currently conducted using Earth-based radio telescopes and therefore radio frequency interference (RFI) from terrestrial sources are inevitable. There is a drive to construct radio telescopes in space due to the detrimental effects of RFI along with the ionospheric reflections at low frequencies [1]. For Earth-based telescopes, other strategies exist to mitigate the effects of RFI. This includes legislation to protect the areas in which radio telescopes are constructed [2]. Certain radio frequency bands are reserved specifically for radio astronomy [3]. Additionally, the RFI environment at the site of a telescope is characterised as was done in [4]. In the advent of constructing the world's largest radio telescope, the Square Kilometre Array (SKA), the problem of RFI will continue to hinder scientific observations. This is exacerbated by the advancement of the communications industry and an increase in population. Therefore, research into RFI excision techniques are important.

Signals of scientific interest originate from distant sources whereas RFI is generated locally. Therefore, differentiating between astronomical and terrestrial sources are conducted using power statistics by setting a threshold in a time-frequency plot to identify outliers [5]. In practice, the mean and standard deviation are also used to generate an RFI mask which can be applied to the data, as is done by the PRESTO software [6]. Astronomical signals disperse as

they travel through the interstellar medium (ISM) and scientists use dispersion measure (DM) to estimate distances to the source [7]. A zero-DM filtering technique can therefore be used as an RFI mitigation technique as illustrated by [8]. In [9] the cyclostationarity characteristic of RFI is used to derive two RFI mitigation techniques. Another standard tool for RFI mitigation is the spectral kurtosis (SK) estimator. This tool utilises the fact that, by definition, the excess kurtosis of a Gaussian distribution is zero. The SK estimator makes use of the fourth-order moment which forms part of higher-order statistics (HOS) [10]. Another HOS is the bispectrum, which is used for this analysis.

This work investigates whether GPS signals inherently have energy within their bispectrum. GPS signals are used because it is a prominent source of RFI within the L-band radio spectrum. While most radio-astronomical signals can be characterised as Gaussian noise, pulsar signals have exhibited non-Gaussian components [11]. Furthermore, in [12] it is shown that the folded profile of Pulsar J0437-4715 has energy in the bispectrum. Therefore, if GPS signals have energy in the bispectrum and the signature thereof is different to that of astronomical transient sources, it could potentially be used as an RFI mitigation technique. Section 2 gives an overview of HOS and details the estimation of the bispectrum which is used for this study. Section 3 discuses how the signals are simulated and subsequently processed. Conclusions are drawn in Section 4, which also discusses how this analysis will form the basis of future work.

## 2 Higher-Order Statistics

Traditionally, received random processes are characterised using the power spectrum, which is the Fourier transform of the autocorrelation of the signal. This relationship is known as the Wiener-Khintchine theorem [13]. It is also known as second-order statistics, because the autocorrelation involves the multiplication of the signal with itself. One can also calculate a triple or quadruple product and take the Fourier transform in order to obtain the bispectrum and trispectrum, respectively. Statistics are classified as higher-order when the order is greater than second-order. When working with a zero-mean signal, the moments and cumulants for the first, second, and third-order are the same. An advantage of working with higher-order cumulants is that

they are zero for Gaussian processes, which is not always the case with higher-order moments [14].

There are multiple reasons to use HOS, as outlined in [15, 16]. When the signal of interest (SOI) has a non-Gaussian distribution and is corrupted by Gaussian noise, one can use HOS to suppress the Gaussian noise. This can be seen from definition (19) as given by [16]. When a signal consists of two non-Gaussian components, one can get rid of Gaussian noise by going to higher domains in order to investigate whether the two signals have different signatures in the bispectrum. In addition to Gaussian noise, symmetrically distributed signals also have zero higher-order cumulants.

## 2.1 Estimating the bispectrum

For a zero-mean process, the third-order cumulant is given by

$$c_3(\tau_1, \tau_2) = \mathbb{E}\{x(n)x(n-\tau_1)x(n-\tau_2)\} \quad , \quad (1)$$

where $x(n)$ is a discrete-time random process and $\mathbb{E}$ is the expectation operator. The bispectrum is calculated as the 2-D Fourier transform of Equation 1 as follows:

$$C_3(\omega_1, \omega_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} c_3(\tau_1, \tau_2)e^{-j(\omega_1\tau_1 + \omega_2\tau_2)} \quad (2)$$

$$= X(\omega_1)X(\omega_2)X^*(\omega_1 + \omega_2).$$

The bispectrum can be estimated directly as outlined in algorithm 1, adapted from [17]. Although the literature also documents an indirect method, only the direct method was used in this analysis and is therefore repeated here. Note that one can smooth the bispectrum by applying a window before the averaging in step 5. The bispectrum is computationally expensive and symmetry for real signals are usually exploited for optimisation. A 12-fold symmetry holds for real signals as is illustrated in Figure 1.

---

**Algorithm 1** Direct method

---

0. $x(n)$ is N data samples, $n = 1...N$
1. Reshape into M segments of K samples, where N=MK
2. Subtract the mean of each segment
3. Compute DFT of the $i$th segment, $x_i(n)$
$X_i(k) = \sum_{n=0}^{K-1} x_i(n)e^{-j\frac{2\pi}{K}nk}, k = 0...K-1, i = 1...M$
4. Compute bispectrum of each segment
$\hat{C}_3^i(k_1, k_2) = \frac{1}{K}X_i(k_1)X_i(k_2)X_i^*(k_1+k_2), i = 1...M$
5. Average bispectrum over all segments
$\hat{C}_3(k_1, k_2) = \frac{1}{M}\sum_{i=1}^{M}\hat{C}_3^i(k_1, k_2), k_1 = \frac{K}{2\pi}\omega_1, k_2 = \frac{K}{2\pi}\omega_2$

---

# 3 Bispectra of simulated GPS signals

## 3.1 Simulation methodology

Software provided by [18] is used to simulate GPS waveforms based on the interface standard as documented in [19]. Baseband signals are generated for both legacy
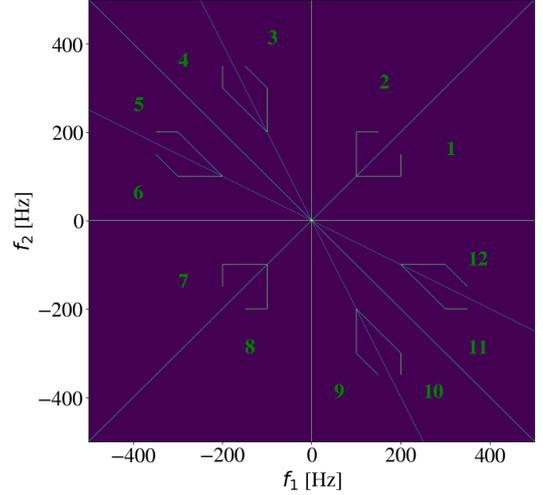


**Figure 1.** Bispectrum symmetry regions for real signals

navigation data (LNAV), denoted by $D$, and civil navigation data (CNAV), denoted by $D_c$. LNAV data frames are generated at 50 bps. An LNAV frame consists of five subframes, where each subframe is 300 bits long. The CNAV data consists of messages which are 300 bits long, generated at 25 bps. CNAV data is passed through a half-rate convolutional encoder to produce an output at 50 bps. All GPS satellites broadcast LNAV data, whereas CNAV data are being broadcasted only by a subset.

The navigation data for each satellite is exclusive-ored with a unique pseudorandom noise (PRN) code, which can also be generated by the receiver. The following three PRN codes exist:

1. $P(Y)$: Precision with anti-spoofing
2. $C/A$: Coarse/Acquisition
3. $L2CM/L2CL$: Civil Moderate / Civil Long

The $P(Y)$ code is generated from a base frequency of $f_b$ = 10.23 MHz, whereas the $C/A$ and $L2CM/L2CL$ are 10 times lower at $\frac{f_b}{10}$ = 1.023MHz. Table 1 provides a summary of PRN codes for the in-phase (I) and quadrature phase (Q) components of the simulated GPS signals. The PRN codes spread the data over a wider bandwidth and is unique to each transmitter. This allows all satellites to transmit on the same frequency. The product of the XOR operation is modulated with a carrier frequency using the bi-phase shift key (BPSK) modulation technique. The carrier frequencies for GPS L1 and L2 bands have frequencies centered at $f_{L1}$=1575.42 MHz and $f_{L2}$= 1227.60 MHz, respectively [18, 19]. For this study the baseband output of the XOR operation is used.

**Table 1.** PRN codes

| I | Q |
|---|---|
| $C/A \oplus D$ | $C/A \oplus D$ |
| $P(Y)$ | $C/A$ |
| $P(Y) \oplus D$ | $L2CM/L2CL \oplus D_c$ |

## 3.2 Bispectral analysis

The simulated signals are processed to compute the estimated bispectrum as per algorithm 1. A 1024-point FFT is used to estimate $X_i(k)$. During the analysis, both the I and Q data streams are treated as real stationary signals. Therefore, when calculating the bispectrum, the 12-fold symmetry displayed in Figure 1 holds, and was used to obtain the full bispectrum.

The results are presented in Figures 2 and 3 for the I branch data, and Figures 4 and 5 for the Q branch data. Initially one bit worth of simulated data was used before comparison with 300 bits of data. This was done to investigate whether additional features would emerge by using more samples.

The bispectrum of the I component for $C/A \oplus D$ data is depicted in Figure 2, and it clearly shows energy within the bispectrum. The result is sustained with a 300-fold increase in the number of data points used in the estimator. It is noted that the bispectral signature of the signal is different to that of Pulsar J0437-4715 shown in [12]. The $P(Y)$ and $P(Y) \oplus D$ bispectra are similar, and therefore only the latter is displayed in Figure 3. No emergent features arise as one increases the number of samples in the experiment.

For the Q branch, the bispectrum of the $C/A \oplus D$ data is depicted in Figure 4, which is equivalent to that of the $C/A$ data and therefore only one is shown. Figure 5 shows the bispectra for the $L2CM/L2CL \oplus D_c$ code. No significant energy is evident from the images using either one or 300 bits of data. The Q branch data consists of $L2CM/L2CL$ and $C/A$ codes, which are generated at a lower bitrate than the $P(Y)$ codes, thereby yielding less data in the same amount of time. However, using a 300-fold increase in data did not yield additional structures in the bispectra. Therefore, the energy seen in the bispectra of the Q components is a result of estimator errors. From this result, one can deduce that all Q components of the simulated signals have a symmetric or Gaussian distribution, as all energy is suppressed when calculating the bispectrum.

## 4 Conclusion

RFI continues to hinder scientific observations and the necessity of mitigation techniques are evident. Various RFI mitigation techniques already exist as was discussed in the introduction. The SK estimator stems from HOS and is an established RFI mitigation technique. Therefore, this work investigated whether GPS signals contained any energy in the bispectrum. The I and Q branch data for all combinations of PRN codes were simulated using a GPS simulator. The data was passed through a bispectrum estimator. Using one bit of data, it was found that bispectral energy within the I branch data for the $C/A \oplus D$ PRN code is present. This bispectrum is different to that of Pulsar J0437-4715, which is a promising result.
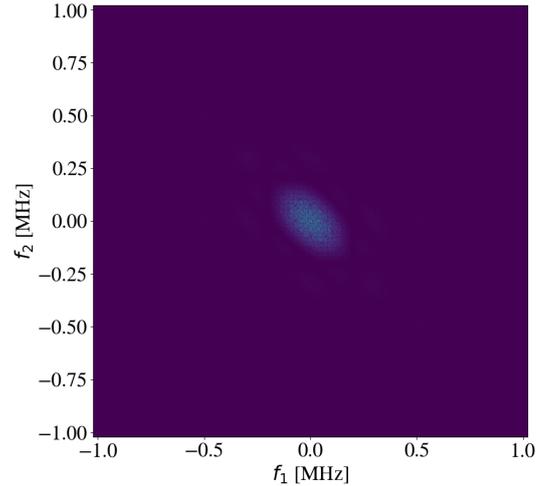


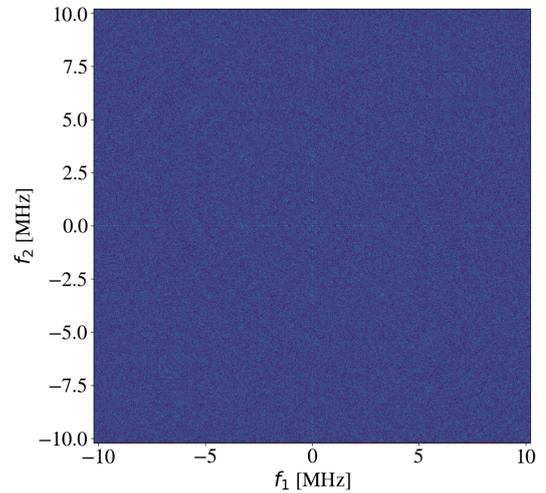**Figure 2.** Bispectrum of I branch $C/A \oplus D$ GPS data



**Figure 3.** Bispectrum of I branch $P(Y) \oplus D$ GPS data
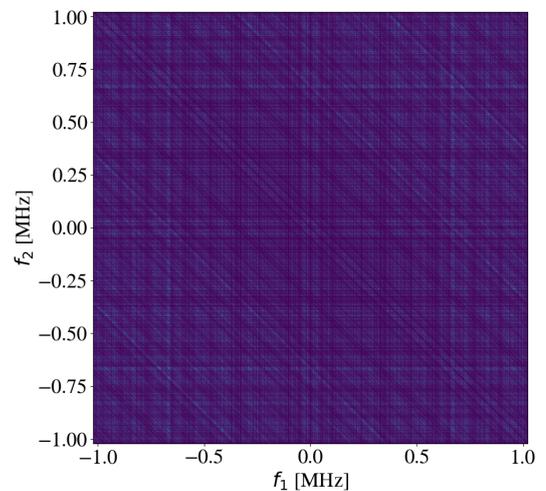


**Figure 4.** Bispectrum of Q branch $C/A \oplus D$ GPS data
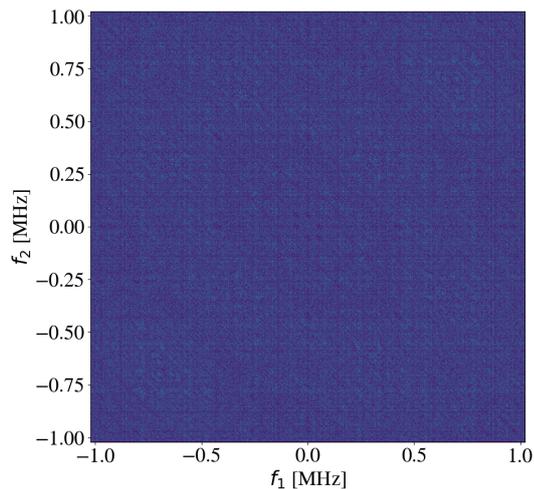
**Figure 5.** Bispectrum of Q branch $L2CM/L2CL \oplus D_c$ GPS data

Future work includes efforts to derive an RFI mitigation technique based on bispectrum statistics. Characterisation needs to be done as to whether the bispectral signatures of all non-Gaussian astronomical sources differs from that of RFI. It also needs to be investigated whether the results are consistent using data from a radio telescope. Further analysis could be performed using the bicoherency index and the trispectrum.

## Acknowledgment

## References

[1] M. Bentum *et al.*, "A Roadmap towards a Space-based Radio Telescope for Ultra-Low Frequency Radio Astronomy," *Advances in Space Research*, vol. 65, September 2019.

[2] *The Astronomy Geographic Advantage Act 21 of 2007*. Republic of South Africa, April 2009.

[3] *Preferred frequency bands for radio astronomical measurements*. Rec. RA.314-10, International Telecommunication Union, June 2003.

[4] A. J. Otto, R. P. Millenaar, and P. S. van der Merwe, "Characterising RFI for SKA phase 1," in *2016 Radio Frequency Interference (RFI)*, pp. 81–84, 2016.

[5] E. van Heerden, A. Karastergiou, and S. J. Roberts, "A Framework for Assessing the Performance of Pulsar Search Pipelines," *Monthly Notices of the Royal Astronomical Society*, pp. 1661–1677, Nov 2016.

[6] S. Ransom, "PRESTO: PulsaR Exploration and Search TOolkit." Astrophysics Source Code Library, ascl:1107.017, July 2011.

[7] D. R. Lorimer and M. Kramer, *Handbook of Pulsar Astronomy*. Cambridge University Press, 1st ed., 2004.

[8] R. P. Eatough, E. F. Keane, and A. G. Lyne, "An interference removal technique for radio pulsar searches," *Monthly Notices of the Royal Astronomical Society*, vol. 395, pp. 410–415, April 2009.

[9] R. Feliachi, C. Dumez-Viou, R. Weber, and A.-J. Boonstra, "RFI mitigation: cyclostationary criterion," in *Proceedings of Wide Field Astronomy & Technology for the Square Kilometre Array — PoS(SKADS 2009)*, vol. 132, p. 033, 2011.

[10] G. M. Nita and D. E. Gary, "The generalized spectral kurtosis estimator," *Monthly Notices of the Royal Astronomical Society*, vol. 406, pp. L60–L64, July 2010.

[11] J. M. Smits, B. W. Stappers, J. P. Macquart, R. Ramanchandran, and J. Kuijpers, "On the search for coherent radiation from radio pulsars," *Astronomy and Astrophysics*, vol. 405, no. 3, pp. 795–801, 2003.

[12] A. Faustmann, J. Gilmore, V. van Tonder, and M. Serylak, "A Bispectral Analysis of the Radio Emissions of Pulsar J0437-4715," *The Journal of Astronomical Instrumentation*, vol. 09, no. 04, p. 2050018, 2020.

[13] P. Z. Peebles and B. Shi, *Probability, Random Variables and Random Signal Principles*. McGraw Hill, 4th ed., 2015.

[14] S. Haykin, *Adaptive Filter Theory*. Pearson, 5th ed., 2014.

[15] C. Nikias and M. Raghuveer, "Bispectrum estimation: A digital signal processing framework," *Proceedings of the IEEE*, vol. 75, no. 7, pp. 869–891, 1987.

[16] C. L. Nikias and J. M. Mendel, "Signal processing with higher order spectra," *IEEE Signal Processing Magazine*, vol. 93, pp. 10–37, 1993.

[17] A. Petropulu, *The Digital Signal Processing Handbook, Second Edition*, ch. Higher-order spectral analysis, pp. 18–1–18–21. CRC Press, Jan. 2009.

[18] MATLAB and Satellite Communications Toolbox Release R2021b, "GPS Waveform Generation." The MathWorks Inc., Natick, Massachusetts, United States.

[19] A. Flores, C. Tiznado, and D. Michael, "NAVSTAR GPS Space Segment/Navigation User Segment Interfaces," Tech. Rep. IS-GPS-200, GPS Enterprise Space & Missile Systems Center, May 2020.