Fluctuation and Noise Letters
Vol. 6, No. 2 (2006) L117–L125

AN EMPIRICAL STUDY OF THE PROBABILITY DENSITY FUNCTION OF HF NOISE

J. GIESBRECHT*,†,‡, R. CLARKE‡ and D. ABBOTT*,†

*Centre for Biomedical Engineering, The University of Adelaide, SA 5005, Australia
†School of Electrical & Electronic Engineering, The University of Adelaide, SA 5005, Australia
‡Ebor Computing, 147 Henley Beach Rd., Mile End, SA 5031, Australia

Received 31 January 2006
Revised 2 March 2006
Accepted 3 March 2006
Communicated by Boris M. Grafov

To many, high-frequency (HF) radio communications is obsolete in this age of long-distance satellite communications and undersea optical fiber. Yet despite this, the HF band is used by defense agencies for backup communications and spectrum surveillance, and is monitored by spectrum management organizations to enforce licensing. Such activity usually requires systems capable of locating distant transmitters, separating valid signals from interference and noise, and recognizing signal modulation. Our research targets the latter issue. The ultimate aim is to develop robust algorithms for automatic modulation recognition of *real* HF signals. By *real*, we mean signals propagating by multiple ionospheric modes with co-channel signals and non-Gaussian noise. However, many researchers adopt Gaussian noise for their modulation recognition algorithms for the sake of convenience at the cost of accuracy. Furthermore, literature describing the probability density function (PDF) of HF noise does not abound. So we describe a simple empirical technique, not found in the literature, that supports our work by showing that the probability density function (PDF) for HF noise is generally not Gaussian. In fact, the probability density function varies with the time of day, electro-magnetic environment, and state of the ionosphere.

Keywords: HF; HF noise; Gaussian; Bi-Kappa.

1. Introduction

Long distance radio-communications became practical in the early-1900s with the experiments and inventions of Guglielmo Marconi [1]. Marconi was able to demonstrate long-distance wireless communications because ionized gases in the ionosphere cause refraction of high-frequency (HF) signals (nominally 2 MHz to 30 MHz). Such refraction enables signals to propagate beyond the horizon to distant receivers unable to be reached by higher frequency (VHF and above) signals. This fact makes the HF band attractive for private and commercial interests as well as for defense forces spread across the globe. For example, spectrum management agencies monitor the HF band for unlicensed operators and military agencies use the HF band for communications.

A receiver for conventional HF communications consists of filters, oscillators, mixers, usually at least one low-noise amplifier (LNA), and many discrete components. Its function is to down-convert (in frequency) signals acquired by an antenna to baseband and to then demodulate the signal to extract the information content. A particularly versatile HF receiver may even have more than one demodulator so that signals of various types can be handled.

Today, there are so many different digital and analog modulation techniques that even the "versatile" receiver cannot be used for more than a few modulations. Many of these are of the family of space-time layered signals including spread spectrum, time-domain multiplexing (TDM), frequency-domain multiplexing (FDM), and parallel transmission of data through multiple antennas and/or frequencies. Traditional HF receivers cannot easily handle these signals, but the advent of software radio is helping to alleviate this problem.

Software radio aims to replicate hardware functions in software running on a generic platform. In so doing many of the problems associated with receiver implementations are avoided. In addition, receiver chains can easily be changed to accommodate various modulation schemes. The idea is to digitize the incoming radio-frequency (RF) signal directly and then to perform down-conversion and demodulation in digital hardware and/or software. This is now possible with current high-resolution Analog-to-Digital Converters (ADCs) and Digital Down-Converters (DDCs). Modulation recognition[1] algorithms are then used to automatically choose the correct demodulator.

However, the modulation recognition algorithms and models that many researchers [2–9] use rely on an assumption that background noise is Gaussian in character. This assumption is not generally valid for HF communications. HF noise consists of: galactic noise, atmospheric noise, and man-made noise [10]. Galactic noise is produced by cosmic radiators such as the sun and other stars. Atmospheric noise changes with time of year and is largely influenced by lightning activity; while man-made noise is produced by electro-magnetic radiating devices such as electric machinery, power lines and transformers, and computers.

Note that this definition of noise does not include interfering or co-channel signals. Such signals, as well as the actual HF noise, are greatly affected by the state of the ionosphere. So we briefly digress into a short discussion about the ionosphere. Budden [11] and Davies [12] provide more detailed information on the ionosphere. In a simplistic model, the ionosphere has three main refraction regions: the D-, E-, and F- regions. These regions consist of ionized gases and free electrons or plasma. The concentration of free electrons is greatly affected by ultra-violet and x-ray radiation from the sun.

The D-region is weakly ionized and exists from about 50–80 km above the earth. It allows communication at very low frequencies (VLF) because the D-region and earth form a type of waveguide but, is highly absorptive for signals with frequencies

[1]Modulation recognition is the process of automatically determining the modulation type of a signal with no foreknowledge of the signal modulation.

between about 100 kHz–2 MHz[2]. As a result, effective daytime communications in this frequency range is limited to about 300 km. Attenuation by the D-region of signals above 2–5 MHz decreases as the signal frequency increases. After sunset, the D-region disappears because ultraviolet light from the sun no longer ionizes the gases in the lower ionosphere and consequently electrons and ions quickly recombine. The impact on communications is that, at night, signals passing through what was the D-region during the day are not absorbed but are propagated via the E and F regions[3].

The E-region covers altitudes of approximately 80–150 km above the earth. Like the D-region after sunset, electrons and ions in the E-region also quickly recombine but, the E-region does not disappear entirely. The E-region permits limited medium-range (about 500–2500 km) communications at low HF frequencies (typically less than 15 MHz). Sometimes areas of enhanced ionization in the E-region occur during the day or night. These enhanced regions are called Sporadic-E.

The F-region exists from about 150–1000 km above the earth. During the day, the F-region consists of the F1 layer (150–200 km) and the F2 layer (above 250 km). Only one F layer is generally present at night from about 270–320 km above the earth. Maximum ionization in the F-region occurs typically between mid-day and mid-afternoon but, unlike the other regions, the F-region typically remains weakly ionized throughout the night. Since the F-region is the highest of the regions, a signal propagating by a single-hop of the F2 layer can reach from 3000 km to about 5000 km; multiple hops are required for greater distances. The maximum usable frequency (MUF) that the F-region will refract over a defined ground range depends on the degree of sunspot activity (and time of day and season in equatorial regions). During sunspot maximum, the F-region can refract signals up to $\approx$ 100 MHz, while at sunspot minimum the MUF might only reach 10 MHz. The sun is currently at sunspot minimum.

With this in mind, it is clear that the HF noise distribution is affected by the time of day, electro-magnetic environment, and ionosphere. For example, lightning activity at a distant location can, by various ionospheric modes, contribute to the local HF noise distribution.

Separating the true HF noise from all interfering signals is difficult. Therefore this paper, which continues work presented earlier [13–16], uses an empirical method to show that the probability density function (PDF) of HF noise is generally not Gaussian. The resulting implication is that common assumptions for theoretical communication models do not necessarily apply to the HF band.

2. Research Platform

An advanced HF receiver is under development at Ebor Computing. This receiver consists of an array of antennas, signal conditioning electronics, multiple digital receivers, a data processing system, and an independent research platform (see Fig. 1).

The Analog Pre-Processing system consists of antennas, amplifiers, attenuators, and filters. Outputs of this system are fed to a rack of digital receivers that directly

[2]Depending on the sunspot cycle the upper limit can be as high as 5–8 MHz
[3]This is why an amplitude modulation (AM) radio station can be received at locations far distant from the transmitter during the night.

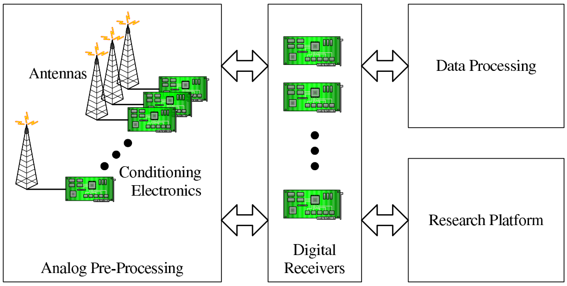

Fig. 1. Architecture of the advanced HF receiver.

sample RF signals up to 50 MHz and down-convert the HF channels to baseband. The digitization of the RF signals is achieved by high-speed ADCs, while down-conversion is accomplished with DDCs. The Data Processing block collects and analyzes the baseband information. The Research Platform supports the present investigation and operates independently of the Data Processing system. In this research, active antennas (Rohde & Schwarz HE011) that are vertically-polarized are used to sense the vertical component of the electric field of the HF environment. These antennas are sensitive to local groundwave propagation as well as skywave propagation from distant emitters (greater than a few hundred kilometers away).

3. Measurement Method

Data to determine the PDF of HF noise was gleaned from the output of a pre-emptive null-steering algorithm [17]. This algorithm was used for purposes entirely different from the present investigation and therefore will not be discussed further. However, the algorithm did provide the raw output of a DDC (bandwidth of ≈ 4 kHz and sampling rate of ≈ 10 kHz[4]) that was swept across the HF band at 125 kHz/s from 2 MHz to 25 MHz every 15 minutes. It is this output that forms the basis for the present work.

Sweeps were made at an HF site with relatively high environmental noise over a 48-hour time span (February 16–17, 2005) in Adelaide, Australia. Only data sets created shortly before sunrise[5] and shortly after sunrise are considered because at sunrise the D-region in the ionosphere begins to form and to absorb distant man-made and natural noise propagating by ionospheric modes. Moreover, the level of local man-made noise is relatively low because people are still asleep but, as people wake and attend to their daily activities the local man-made noise increases significantly. The "best" estimation of the distribution of the HF noise is therefore during the sunrise period.

Data from each DDC sweep is segmented into bands, BW, (see Fig. 2) within which the noise statistics are assumed stationary in the wide sense. For a sweep rate, $F_{\rm swp}$, and a sampling rate, $F_{\rm s}$, the bandwidth covered by the sweep at each sampling instant is denoted by $f_{\rm n} = \frac{F_{\rm swp}}{F_{\rm s}}$. The number of samples in a given bandwidth, BW,

[4]The exact sampling rate was $\frac{31250}{3}$ Hz.
[5]On February 17, 2005 sunrise was at 6:51 am in Adelaide, Australia.

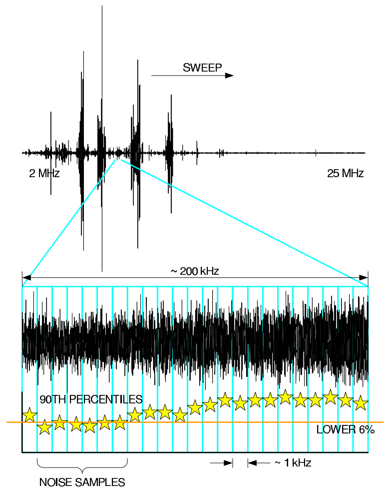

Fig. 2. The method of determining noise samples requires a sweep in time and frequency across much of the HF band. Data collected from the sweep is sub-divided into $\approx$ 200 kHz bands and further sub-divided into $\approx$ 1 kHz bands. The 1 kHz bands having the lowest 90^{th} percentiles are those that are used in the computation of the PDF for HF noise.

is $N_{\text{samp}} = \frac{\text{BW}}{f_{\text{n}}}$. Within BW, M sub-channels of bandwidth, BW_{c}, are chosen such that $M = \frac{\text{BW}}{\text{BW}_{\text{c}}}$ and where the number of samples in each sub-channel are $L = \frac{\text{BW}_{\text{c}}}{f_{\text{n}}}$. Typical values for BW, BW_{c}, and N_{samp} are $\approx$ 200 kHz, $\approx$ 1 kHz, and $\approx$ 16000 respectively.

The 90^{th} percentile of the absolute value of the samples in each sub-channel is taken as a measure of the signal level in each sub-channel. From the list of signal levels, those sub-channels having their 90^{th} percentile below a threshold are chosen as signal-free sub-channels. All the time-series samples that correspond to the sub-channels falling below an L_p percent threshold (in this case 6%) are considered noise samples. Normalizing the histogram of these samples by the number of samples for all chosen sub-channels yields the PDF of the of HF noise. Then, by mapping each sample to a voltage at the output of the antenna and accounting for the antenna factor, the noise probability can be plotted as a function of electric field strength. This mapping incorporates the dominant gains and losses in the receiver.

Well known research [18] by the International Telecommunication Union (ITU) shows that the electric field strength of HF noise can be described by

$$E_n = F_a - 95.5 + 20\log_{10}(F_{\text{MHz}}) + 10\log_{10}(b)\,, \tag{1}$$

where E_n is the root-mean-square (RMS) electric field strength of the noise in units of dB with respect to a μV/m (denoted by $\text{dB}_{\mu\text{V/m}}$) for a bandwidth of b Hz, and where F_a is the effective noise figure (dB) of the antenna at a center frequency of F_{MHz} MHz. Moreover, the antenna factor, f_a, is

$$f_a = \frac{p_n}{kT_o b}\,, \tag{2}$$

where p_n is the noise power (W) available from an equivalent loss-free antenna, k is Boltzmann's constant, and T_o is the reference temperature (288 K as suggested by the ITU). Calculating $10\log_{10}(f_a)$ and inserting into (1) yields

$$E_n = P_n + 20\log_{10}(F_{\mathrm{MHz}}) + 108.5\,, \tag{3}$$

where E_n has units of $\mathrm{dB}_{\mu\mathrm{V/m}}$, and P_n has units of dB with respect to a Watt ($\mathrm{dB_W}$). To determine P_n we compute $20\log_{10}(|y[k]|) - 149.7$ where $y[k]$ represents the sequence of samples output from the digital receiver and where the constant, 149.7, accounts for all gains and losses in the receiving system. This produces a noise power, $P_n[k]$, at each sampling instant, k. Inserting $P_n[k]$ into (3), taking the inverse-logarithm of the result, and accounting for the sign of $y[k]$ renders the electric field strength ($\mu V/m$) at each sampling instant, $e_n[k] = 10^{E_n[k]/20} \times \mathrm{sign}\,(y[k])$.

4. Results

Consider Fig. 3 which shows the PDF of HF noise between 6.1 MHz and 6.3 MHz for various times shortly after sunrise. The sub-channel bandwidth, $\mathrm{BW_c}$, is 1.2 kHz and the bandwidth over which the noise statistics are assumed stationary in the wide sense, BW, is 228 kHz (these choices of bandwidth are reasonable for the purpose of this demonstration). From 7:00 am to 8:00 am the PDFs are decidedly non-Gaussian because most noise emitting sources have not yet been turned on (people are just waking). During this time the D-region is becoming more ionized and is absorbing sky-wave transmissions. After 8:00 am the PDF becomes more and more Gaussian[6]. This is indicative of nearby emitters of all types[7] contributing to the noise background as people begin enabling their devices. Consequently, man-made noise through ground-wave transmissions are likely affecting the PDF after sunrise due to the fact that the D-region is absorbing most distant man-made and natural sources of noise.

Similar results are achieved for the PDF at frequencies between 11.6 MHz and 11.8 MHz, as well as between 13.6 MHz and 13.8 MHz. At the "noisy" measurement site, the PDF at frequencies between 5.4 MHz and 5.6 MHz is markedly Gaussian at all times. This band has high signal levels throughout the day.

All of this suggests that the assumption of Gaussian noise can lead to inaccurate measures of algorithm or model performance. Consequently any HF communications model must account for the surrounding HF environment (*i.e.* mixture and strength of noise and interference components), the frequency band, and the time of day. However, given the enormous difficulties involved in developing robust modulation recognition algorithms, it is not unreasonable to begin with an assumption of Gaussian noise and then to progress to a more realistic form of noise.

So, if the PDF of HF noise is not Gaussian what distribution does it resemble? It actually resembles the Bi-Kappa distribution presented by Leubner and Vörös [19] to describe the probability distribution of the solar wind where,

$$p\,(x;\kappa) \;=\; \begin{cases} \frac{1}{2\sqrt{\pi}\sigma}\left[1+\frac{x^2}{\kappa\sigma^2}\right]^{-\kappa} & \kappa > 0 \\ \frac{1}{2\sqrt{\pi}\sigma} & \kappa = 0 \\ \frac{1}{2\sqrt{\pi}\sigma}\left[1-\frac{x^2}{\kappa\sigma^2}\right]^{\kappa} & \kappa < 0 \end{cases} \tag{4}$$

[6] This appears to be a result of the central limit theorem.
[7] These transmitters could be genuine HF transceivers, or welders and car ignitions for example.

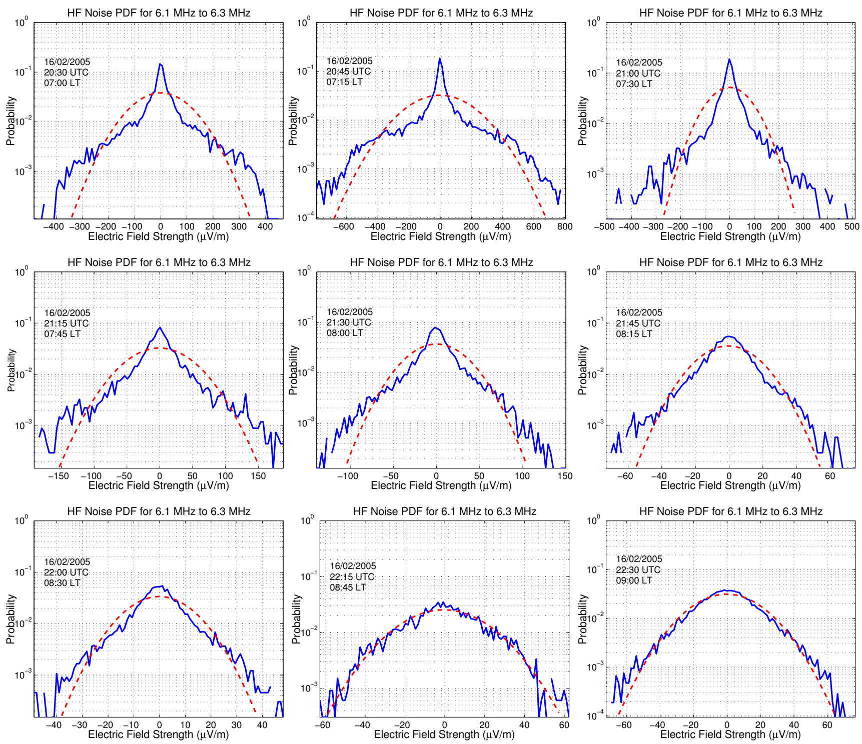

Fig. 3. Probability versus electric field strength (vertical component) of HF noise in a 4 kHz band for various times shortly after sunrise, 6:51 am local time (LT), from 6.1 MHz to 6.3 MHz. A Gaussian PDF is superimposed (dashed curves) to show how poorly it fits the PDF of HF noise. Notice that as the day progresses the PDF becomes more and more Gaussian. This is indicative of transmitters of all types contributing to the noise background as people awake and begin enabling their devices. The central-limit theorem seems to be playing a role here.

and where x is the speed of the solar wind, σ is the standard deviation of the speed of the solar wind and κ is a tuning parameter. In our case, x corresponds to the electric field strength (μV/m) of the noise and σ is its standard deviation. This distribution is very similar to that of HF noise described by Johnson [10]. Figure 4 fits the Bi-Kappa distribution to two PDFs of Fig. 3, clearly showing that HF noise is not Gaussian.

5. Conclusion

With a simple empirical method it is shown that the probability-density-function (PDF) of HF noise is not necessarily Gaussian and, in fact, depends on geographic location, time of day, and surroundings (among other things). Just before dawn HF noise is dominated by atmospheric noise propagating via skywaves. Shortly after dawn the skywave component of HF noise is much weaker, due to the absorptive properties of the D-region, and therefore local noise propagating by groundwaves becomes the dominant component. A Gaussian distribution is sufficient for devel-

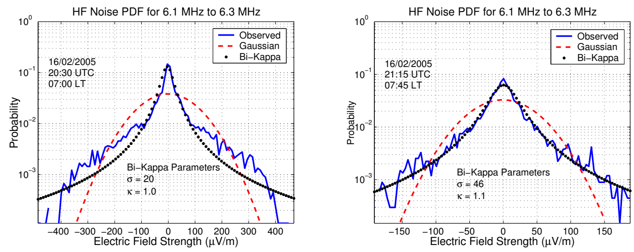

Fig. 4. Fit of a Bi-Kappa distribution (dotted curve) and Gaussian distribution (dashed curve) to the PDF of HF noise in a 4 kHz band at 7:00am (*left*) and 7:45am (*right*). For the 7:00am observation the measured standard deviation is 100 μV/m, while $\sigma = 20$ μV/m for the Bi-Kappa distribution. In this case the Bi-Kappa distribution fits the observed PDF in the range of -100 μV/m to 100 μV/m. At 7:45am, the Bi-Kappa distribution fits much better. Here $\sigma = 46$ μV/m for the Bi-Kappa distribution, and is the same as the measured standard deviation (46 μV/m).

opment of preliminary models and algorithms, but progression to a more accurate distribution is necessary before making conclusions about the behavior of such models and algorithms.

Acknowledgments

Data for noise analysis is provided by Nigel Brine (School of Electrical & Electronic Engineering, The University of Adelaide). Financial support for this work is provided by the The University of Adelaide and Ebor Computing.

References

[1] *Nobel Lectures – Physics 1901-1921* (Elsevier Publishing Company, Amsterdam, 1967).

[2] K. E. Nolan, L. Doyle, P. Mackenzie and D. O'Mahony, Modulation scheme classification for 4G software radio wireless networks, in *Proc. IASTED Int. Conf. Signal Processing, Pattern Recognition, and Applications* (*SPPRA 2002*) (Crete, Greece, 2002), pp. 25–31.

[3] A. O. Hero III and H. Hadinejad-Mahram, Digital modulation classification using power moment matrices, in *Proc. IEEE 1998 Int. Conf. Acoustics, Speech, and Signal Processing*, Vol. 6 (Seattle, 1998), pp. 3285–3288.

[4] M. L. D. Wong and A. K. Nandi, Automatic digital modulation recognition using spectral and statistical features with multi-layer perceptrons, in *Proc. IEEE Int. Symp. Signal Processing and Its Applications*, Vol. 2 (Kuala Lumpur, Malaysia, 2001), pp. 390–393.

[5] M. Goris, Categories of Radio-Frequency Interference. Technical report, National Foundation for Radio Astronomy, Holland (February 1998).

[6] W. Akmouche, Detection of multicarrier modulations using 4th-order cumulants, in *Proc. MILCOM 1999 Military Communications Conference*, Vol. 1 (Atlantic City, NJ, USA) (November 1999), pp. 432–436.

[7] S.-Z. Hsue and S. S. Soliman, Automatic modulation classification using zero crossing, *IEE Proceedings* **137**(6) (1990) 459–464.

[8] A. K. Nandi and E. E. Azzouz, Modulation recognition using artificial neural networks, *Signal Processing* **56** (1997) 165–175.

[9] A. K. Nandi and E. E. Azzouz, Algorithms for automatic modulation recognition of communication signals, *IEEE Transactions on Communications* **46**(4) (1998) 431–435.

[10] E. E. Johnson, R. I. Desourdis, Jr., G. D. Earle, S. C. Cook and J. C. Ostergaard, *Advanced High-Frequency Radio Communications* (Artech House Inc., Norwood MA, USA, 1997).

[11] K. G. Budden, *The Propagation of Radio Waves*: *The Theory of Radio Waves of Low Power in the Ionosphere and Magnetosphere* (Cambridge University Press, 1988).

[12] K. Davies, *Ionospheric Radio* (Institution of Electrical Engineers, 1990).

[13] J. E. Giesbrecht, R. Clarke and D. Abbott, Monitoring the HF spectrum in the presence of noise, in *Proc. SPIE Second Int. Symp. Fluctuations and Noise*, Vol. 5473 (Maspalomas, Gran Canaria, Spain, 2004), pp. 76–84.

[14] J. E. Giesbrecht, R. Clarke and D. Abbott, Modulation recognition for HF signals, in *Proc. SPIE Int. Symp. Microelectronics, MEMS, and Nanotechnology*, Vol. 5649 (Sydney, Australia, 2005), pp. 501–512.

[15] J. E. Giesbrecht, R. Clarke and D. Abbott, Coherence as a feature of real HF signals, in *Proc. SPIE Third Int. Symp. Fluctuations and Noise*, Vol. 5847 (Austin TX, USA, 2005), pp. 188–198.

[16] J. E. Giesbrecht, R. Clarke and D. Abbott, Modulation recognition for real HF signals, in *Proc. SPIE Int. Symp. Microelectronics, MEMS, and Nanotechnology*, Vol. 6035 (Brisbane, Australia, 2006), pp. 458–469.

[17] N. L. Brine, C. C. Lim, A. D. Massie and W. Marwood, A pre-emptive null-steering ionosonde, in *Proc. Fourth Symp. Radiolocation and Direction Finding* (Southwest Research Institute, San Antonio TX, USA, 2002).

[18] Characteristics and applications of atmospheric radio noise data. Technical Report 322-3, International Telecommunication Union, Geneva (1986).

[19] M. P. Leubner and Z. Vörös, A nonextensive entropy path to probability distributions in solar wind turbulence, *Nonlinear Processes in Geophysics* **12** (2005) 171–180.