



Article

Counter-Interception and Counter-Exploitation Features of Noise Radar Technology

Gaspare Galati ^{1,2} , Gabriele Pavan ^{1,2,*} , Kubilay Savci ^{3,4} and Christoph Wasserzier ⁵

¹ Department of Electronic Engineering, Tor Vergata University of Rome and CNIT-Consortium for Telecommunications, Research Unit of Tor Vergata University of Rome, via del Politecnico 1, I-00133 Rome, Italy; gaspare.galati@uniroma2.it

² CNIT-Consortium for Telecommunications, Research Unit of Tor Vergata University of Rome, via del Politecnico 1, I-00133 Rome, Italy

³ Turkish Naval Research Center Command, Istanbul 34890, Turkey; ksavci@ieee.org

⁴ Department of Electrical and Electronics Engineering, Koc University, Istanbul 34450, Turkey

⁵ Fraunhofer Institute for High Frequency Physics and Radar Techniques-FHR, 53343 Wachtberg, Germany; christoph.wasserzier@fhr.fraunhofer.de

* Correspondence: gabriele.pavan@uniroma2.it

Abstract: In defense applications, the main features of radars are the Low Probability of Intercept (LPI) and the Low Probability of Exploitation (LPE). The counterpart uses more and more capable intercept receivers and signal processors thanks to the ongoing technological progress. Noise Radar Technology (NRT) is probably a very effective answer to the increasing demand for operational LPI/LPE radars. The design and selection of the radiated waveforms, while respecting the prescribed spectrum occupancy, has to comply with the contrasting requirements of LPI/LPE and of a favorable shape of the ambiguity function. Information theory seems to be a “technologically agnostic” tool to attempt to quantify the LPI/LPE capability of noise waveforms with little, or absent, a priori knowledge of the means and the strategies used by the counterpart. An information theoretical analysis can lead to practical results in the design and selection of NRT waveforms.

Keywords: noise radar; low probability of intercept; pseudorandom numbers; random processes; electronic warfare; electronic support measures; radar signal processing; signal analysis



Citation: Galati, G.; Pavan, G.; Savci, K.; Wasserzier, C. Counter-Interception and Counter-Exploitation Features of Noise Radar Technology. *Remote Sens.* **2021**, *13*, 4509. <https://doi.org/10.3390/rs13224509>

Academic Editors: Vito Pascazio, Konstantin Lukin, Krzysztof Kulpa and Piotr Samczynski

Received: 23 September 2021

Accepted: 3 November 2021

Published: 9 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The most relevant features of Noise Radar systems in defence applications are tightly related to modern Electronic Warfare (EW) systems, whose *intercept*, *identification* and *jamming* capabilities are quickly evolving following the evolution of radar threats. Both EW and radar systems are boasting more and more “intelligence” thanks to the tighter and tighter convergence of computer science, communications, signal processing and big data analytic means. The history of anti-interference radars is old: more or less “clever” anti-jamming techniques have been proven for over half a century in the radar context [1]. They include (to name only two of them) the Adaptive Frequency Selection with which the radar automatically selects the least jammed operating frequency, and the Adaptive Antenna System to counteract sidelobe jamming [2,3]. The normal follow-on has been a generation of more and more adaptive radars, arriving in this century to the concept of *cognitive radar* [4,5] with some (more or less partial) implementations [6,7] of it. Modern adaptive radars [8] may change their operating modes and the radiated waveforms almost instantaneously: these *radar threats* are adaptable and reprogrammable, creating for EW engineers an urgent request to characterise them correctly.

At the same time, the increasing scarcity of the electromagnetic spectrum, which is, of course, the main resource for radio communications, radio navigation and radar, has generated a lot of interest in research and development activities, both academic and industrial. Hot topics today are the Communication and Radar Spectrum Sharing (CRSS), and the

related usage of Artificial Intelligence with its Machine Learning/Deep Learning tools [9], with new applications to Machine Learning of existing mathematical tools such as tensors. *Tensor* is a multidimensional array providing a natural representation for multivariate, high-dimensional data. Hence, tensor decomposition has rapidly found many applications in signal processing and machine learning problems. The most well-known and widely-used tensor decompositions are CANDECOMP/PARAFAC (CP) decomposition and Tucker decomposition which are extensions, for tensors, of the widely used singular value decomposition, SVD. Tensors, generalising the concept of a matrix over the classical two dimensions, are a powerful tool in automated and rapid analysis of huge quantities of complex data, [10].

In general, there are two main research directions in CRSS: (i) Radar-communication coexistence (RCC) and (ii) Dual-Functional Radar-Communication (DFRC) system design. By considering the coexistence of individual radars and communication systems, the first category of research aims for developing efficient interference management techniques, so that both systems can operate without unduly interfering with each other [11]. On the other hand, DFRC techniques focus on designing joint systems that can simultaneously perform wireless communication and remote sensing [12]. Doing so enables a paradigm change, where previously competing transmissions can be jointly optimised. It is worth pointing out that these explorations have gone far beyond their original motivation of realising spectrum sharing between radar and communication systems and were extended to numerous novel applications including vehicular network, indoor positioning and secure communications.

The term LPI (Low Probability of Intercept) is a property of radar that, because of its low power, wide bandwidth, frequency variability, or other design attributes, makes it difficult to be detected by passive intercept devices such as electronic support (ES), radar warning receivers (RWRs), or electronics intelligence (ELINT) receivers. The term LPI in any case refers to a certain application and is strongly related to the ratio of the detection range of the LPI radar and the intercept range by a particular class of EW receivers. The success of an LPI radar is measured by how hard it is to extract the radar emission parameters by the EW system. Since LPI radars typically use wideband CW signals that are difficult to intercept and/or identify, modern intercept receivers must resort to more sophisticated signal processing systems to extract the waveform parameters necessary to create a proper jamming response.

The main methods for intercepting LPI radars include: (a) filter bank processing with higher-order statistics [13,14]; (b) methods based on time–frequency transforms, e.g., Short-Time Fourier transform (STFT), Wigner–Ville transform [15,16]; (c) quadrature mirror filter banks [17] and (d) cyclostationary processing [18]. A novel method, based on the use of two receivers on board a fast-moving platform (e.g., an airplane or a satellite), is described in [19,20]. Recently, a new waveform recognition technique (WRT), based on a convolutional neural network (CNN), was proposed in [21]. The input parameters of the CNN (number of filters, filter size, number of neurons) are designed for various waveforms to guarantee the best classification performance and a reduction of the computational cost.

The increasing capability of intercept receivers to detect, locate and identify a radar set (which can quickly lead to an electronic reaction, i.e., soft kill technique or physical reaction, i.e., hard kill technique by an Anti-Radiation Missile or by guided ammunition) are stressing the LPI requirements more and more. The obvious counter measure of interrupting the radar transmission is difficult, or even not allowed, in some safety-critical applications such as the marine navigation radar and the airborne radar altimeter, as well as in tactical surveillance. Hence, LPI features have ever-increasing importance in modern military radar. The authors of [22] state the need for radar engineers to consider not only the radar detection and measurements but also its survivability. The following LPI features are summarised: (i) CW operation, with bistatic architecture because of the antennas coupling problem, (ii) wide beam transmission, (iii) narrow beam reception with a multi-beam, quickly scanning, low sidelobes antenna, and finally, (iv) large time-bandwidth product for

the transmitted signal [23,24]. LPI radar techniques are receiving more and more attention in the open literature for over three decades, see for instance [25–29].

Nowadays, the above-mentioned evolution is changing the classical paradigms of the EW, which has to cope with typical battlefield spectral environments with thousands of radio emitters (and billions of radar pulses per second). The possible answers to the problems are related to (a) the analysis of the electromagnetic spectrum, (b) the characterisation of the sources, and (c) the exploitation of the spectrum for Electronic Attack (EA) and Electronic Defense (ED) purposes. Transition is ongoing towards the full automation of most EW/ED functions, with the human component operating at higher and higher supervision levels. EW pursues the way to cognitive systems, after radar. Cognitive EW uses modern machine learning techniques for tasks such as target recognition, decision-making, and autonomous learning. For instance, adaptive radars make it difficult for an EW system to analyse the radar emission on a pulse basis by standard emitter libraries and, thus, to define and actuate an adequate response. This issue causes EW technology to turn to Machine Learning and to create cognitive EW systems with software-defined capabilities, resulting in more operational flexibility, quicker upgrades, and greater affordability. The mostly automated, quick interactions between EA, ED and radar functions [30] also call for new analysis and design tools, also based on Game Theory, [31].

Modern Electronic Warfare uses technology advancements such as high-performance Digital Signal Processing (DSP) and Field Programmable Gate Arrays (FPGA) [32] systems to achieve appropriate values of resolution and dynamic range and to sustain the ever-increasing algorithm complexity. Powerful algorithms have added the Specific Emitter Identification (SEI, [33]) feature to some intercept receivers, permitting the identification of a particular source, e.g., the long-range radar of a given platform (i.e., not limiting themselves to identify the type of source) thanks to a detailed analysis of the signal modulation [34–36] aiming at the identification of characteristics and therefore identifications of individual radar sets. Modern agile radars demand paradigm changes in EW systems evolving from pulse analysis to threat-based classifications. Thus, the design, development and test/qualification phases of these *cognitive* systems require a set of *agnostic* (i.e., not limited to specific types of equipment by particular suppliers) threat models, embedded in an electromagnetic environment generator able to support multiple equipment types and technologies, [37].

Summing up, as the battlefield electromagnetic spectrum becomes more and more crowded, both cognitive radars and modern EW systems are trying to understand in real-time the intent of every system using the spectrum, rather than relying on a priori assumptions and modelling for the operational scenario. In such a frame, EW systems and radars can be programmed on the fly to change waveforms and create unique signatures in real-time: these systems are moving from merely being adaptive to using Machine Learning capabilities to analyse any change in spectrum use. For what concerns Noise Radar, this technology, whenever correctly applied (e.g., granting the radar signals a “good enough” randomness, as discussed below) appears to promise better resilience to modern and future EW interceptors as compared with adaptive and cognitive radars using “deterministic”, although varying, waveforms. Note that for many decades radar parameters may vary during radar operation—such as the carrier frequency (either due to Automated Frequency Selector or to programmed “frequency hopping/agility” [26,38]) and/or the pulse repetition frequency (PRF stagger, [39]) and/or the transmitted code. However, these variations still aim for a certain radar task which limits their degrees of freedom and allows the counterpart to catch information to be exploited for threat-based signal classification. For example, a radar waveform optimised for target tracking is selected by the radar for a good reason and may be of a higher threat level to an airborne platform than is a different emission that is clearly associated with a surveillance task.

Thus, present-day operational radars, also agile ones, may be “more easily” intercepted to feed the “libraries” of the emitters of tactical and strategic interest, while the

Noise Radar pseudorandom radiated signals may only be statistically analysed, a matter to be discussed in the remaining parts of this paper.

As an example, let us consider (among many) two highly-automated EW systems: the Autonomous decoys and the Electronic deception means to confuse enemy's intelligence, surveillance, and reconnaissance (ISR) systems. A common element of these applications is the Digital Radio Frequency Memory, DRFM, system shown in Figure 1.

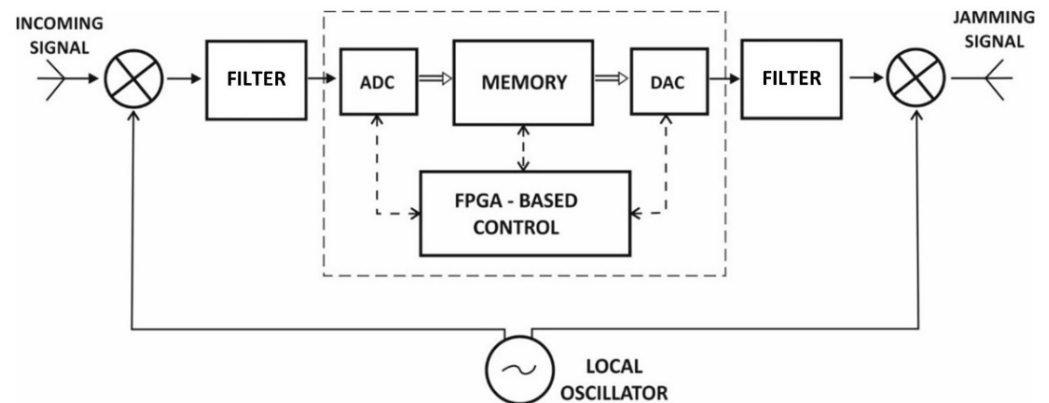


Figure 1. General schematic of DRFM systems.

The key elements of the DRFM system are a fast Analog-to-Digital Converter, ADC, a fast Digital-to-Analog Converter, DAC, and a fast dual memory or fast memory, [40]. Using them, a “copy” of the radar signal is acquired, appropriately delayed and transmitted a number of times to jam the radar receiver. When the radar changes its signal, this particular jammer is only effective at radar ranges greater than the range of the platform carrying the DRFM, but if the radar signal is predictable, or even transmitted unchanged many times, all radar ranges may be jammed. When the EW system combines the DRFM technology with waveform analyses in the ES domain that have the aforementioned *intelligent* features the jammer becomes “smart” with deceptive features, as shown in Figure 2.

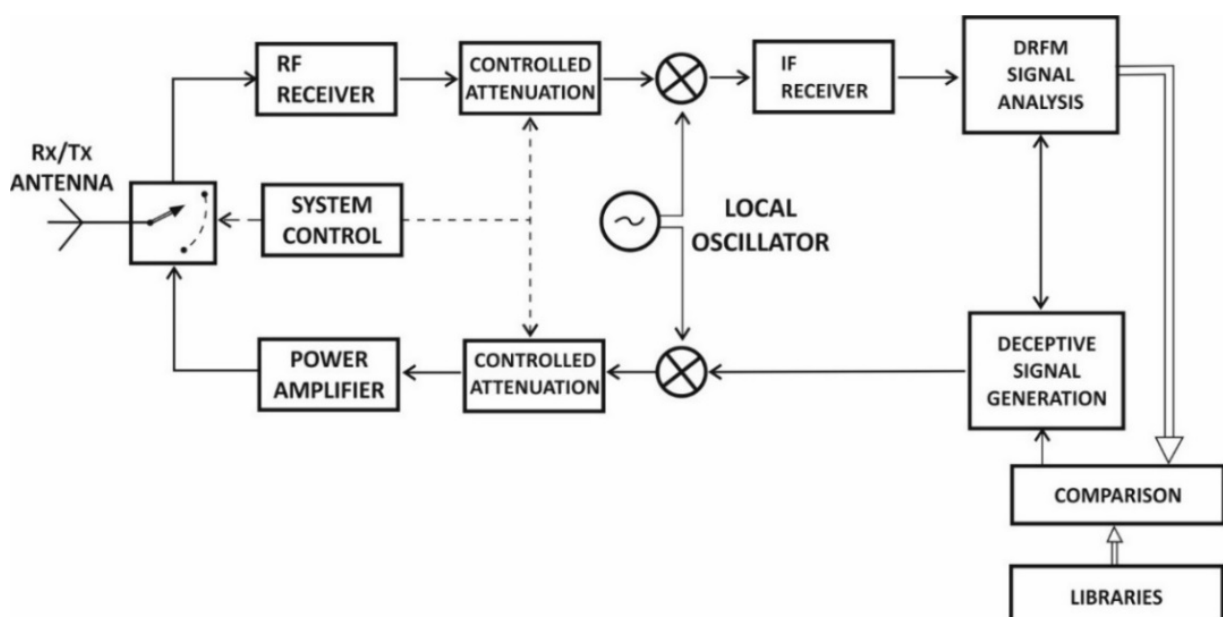


Figure 2. General block diagram of a deceptive/smart radar jammer using Digital Radiofrequency (RF) Memory (DRFM). Remark: The block “Libraries” includes both the “signal” and the “threat” level.

In this case the jamming signal results from the aforementioned steps (a), (b) and (c) with the Machine Learning/Deep Learning analysis of the signals emitted by radar (see for instance [41]) according to their “signatures” and to their statistical features, and the comparison of the result with “a priori” information stored in the ad hoc “libraries”.

In this general frame, the remaining part of this paper is dedicated to the robustness of Noise Radar waveforms; for example, in Section 4.2 of [22] dedicated to Waveforms, it is claimed that a pure random-phase coded signal is the best waveform for CW LPI radar. This solution was studied and tested [42–44] by the Liu Guosui group, which is one of the forerunners of NRT. In fact, the acquisition/recording of pseudorandom radar signals (not necessarily phase-only modulated) is of little advantage to the counterpart and their limited information content, in general, does not allow the enemy to feed any “signal library”; more generally, the analysis of this content allows us to quantify the Low Probability of Exploitation (LPE) property of NRT.

In [45] a novel system concept is introduced combining active sensing by a noise radar (NR) with electronic warfare (EW) tasks. This joint NR/EW sensor concept provides simultaneous operations of spectral sensing, jamming and radar target detection.

2. Pseudo-Random Numbers Generators and Cryptography Security

The *entropy* concept, in its broad sense, can be used to describe the *disorder* and *lack of predictability* of a system (Appendix A). Computation units are *deterministic* systems, hence they cannot generate entropy, but they only may collect it from outside sources when needed. Strictly speaking, computers may not generate a random sequence, but only a sequence of *pseudo-random numbers* (PRN), which is not strictly random being the result of an algorithm: the (pseudo) randomness indicates that these numbers “appear random” to an *external observer*, i.e., they may pass some statistical tests. It is well-known that a PRN generation (PRNG) algorithm has a starting point, called “seed”, which defines the whole sequence till a repetition point. In a finite-state machine such as a digital unit, repetition cannot be avoided, but pushed away and away exploiting the digital resources, registers and memory above all. The widely used “Mersenne Twister” [46] is the collective name of a family of PRNGs based on F_2 -linear maps, whose period, for 32-bit integers, is the huge $2^{19937} - 1$. Today the Mersenne Twister is the default generator in C compilers, the Python language, the Maple mathematical computation system, and in many other environments. In practice, repetition is a minor problem with respect to the low statistical quality of some widely used PRNGs [47].

In principle, an unpredictable *truly-random numbers generator* (TRNG), also said hardware-based RNG, is a nondeterministic system driven by a noise source (i.e., a physical process) governed by quantum mechanical laws, such as reflections of photons by a beam-splitter, radioactive decay and many others. However, practical methods to derive the requested numbers from such sources (i.e., to extract and exploit their entropy compliant with the *recommendation for the entropy sources used for random bit generation* [48]) do not always supply acceptable results, as important problems of *accuracy* and *dynamic range* arise in most cases. For example, one could try to get random numbers by measuring radioactive decay times and taking the time interval X between two successive decays. It is well known that for a fixed *decay-rate* λ the probability distribution function of the time interval (the inter-arrival time in traffic theory) X is an *exponential* type random variable with distribution: $F_X(x) = 1 - \exp(-\lambda x)$, $x \geq 0$. Hence, taking the output quantity x and transforming it into $u = 1 - \exp(-\lambda x)$, one should obtain (in theory) a random numbers generator whose output U has a *uniform* distribution between 0 and 1. In practice, problems arise. (a) Particle detectors have an efficiency less than unity. (b) The clock, needed to measure X , has a finite resolution Δx , i.e., the resulting random number is zero when two successive different decay times differ by less than Δx . (c) A real clock has drifts and higher-order errors. (d) The decay-rate λ is a priori unknown and has to be measured, leading to some possible errors in the distribution of U . For these reasons, the desired *uniform* distribution cannot be exactly achieved in practice.

In general, methods to extract, from physical sources, entropy/randomness are costly, hard to be implemented and only used in particular, highly sensitive applications. This is why computer-based PRNGs are referred to as “physical” generators and are so widely used. An exception may be low-quality methods in consumer computer applications (e.g., electronic games) to create “entropy pools” based on easily usable phenomena such as random movements of the mouse, least significant digits of the clock at a given event and so on.

However, recently a PRNG was developed based on the use of *beta* radiations enabled by integrated circuits (ICs) suitably designed to detect the very low energy of these radiations [49,50]. This generator, although with a low number of bits, has shown a relatively simple structure, low-cost and small volume, passing the National Institute of Standards and Technology (NIST) test [48].

Cryptograph [51] is a driving factor for research on PRNGs, which are used in various cryptographic steps, with an overall security level mostly depending on the quality of the used pseudo-random number generator. A PRNG suitable for cryptographic use is called a *Cryptographically Secure Pseudo-Random Number Generator* (CSPRNG). The strength of a cryptographic system heavily depends on the properties of these CSPRNGs. Depending on the particular application, a CSPRNG might generate sequences with some of these features: (i) to appear random to an external observer, (ii) to be unpredictable in advance, (iii) cannot be reproduced using affordable means. A perfect CSPRNG is one that produces a sequence of independent, equally distributed numbers, i.e., if X_i is the generated integer number (between 1 and N) at the i -th step, the probability of X_i equals $1/N$ independently on all the outputs of the other steps, i.e., all X_k , with $k \neq i$. In other words, knowing the past or future numbers does not help to predict the current number.

A perfect CSPRNG would permit the implementation of an unbreakable cryptographic system, i.e., robust to any attack even using unlimited computation power. It is the celebrated *one-time pad*, in which each bit of the message is coded by addition (*XOR* operation) to the corresponding bit of the key, which is the output of the CSPRNG and decoded with the same operation by the legitimate recipient knowing the key. In practice, this method has the important limitation of the key to be kept secret, to be used only once, in addition to the well-known fundamental problem of the distribution of the cryptographic keys.

The features of CSPRNGs make them suitable to generate waveforms to be emitted by Noise Radars, which have to be secure against statistical analysis and reproduction for jamming and spoofing purposes. The main difference with respect to the aforementioned cryptographic applications is the inherent randomness of the physical medium (including receiver noise, environmental noises and channel/target fluctuations), [52–54]. This topic is also studied in recent research on Physical Layer secure communications and the Internet of Things [55,56]. In order to benefit the legitimate receiver (the own radar receiver) while denying the operation of a counterpart receiver, one shall exploit the difference between the channel to the legitimate receiver and the one to an eavesdropper (an intercept receiver in radar/EW applications) to securely transmit confidential messages (even without using an encryption key).

A possible example could be transmitting one communication code with a main, narrow-beam antenna and a different code, for deception purposes, by an auxiliary antenna (similar to the Interrogator Side Lobe Suppression—ISLS—one in Secondary Surveillance Radar—Selective Mode—SSR Mode S) whose pattern is higher than the sidelobes structure of the main antenna [57,58], thus masking its sidelobes signals.

3. Information Content of Radar Signals

When studying the properties of noise radar waveforms, the main question is: “How much information about the signals emitted from a particular radar may be obtained by analysing more and more samples from the radar emission?” Likely, the answer mainly depends on the operation (and performance) of the counterpart system and on the operational theatre. However, a rather general answer may be searched in terms of Information

Theory [59], in which a measure of the information contained in a signal is related to the *entropy* concept [60] from which the *mutual information* is derived [61] (see Appendix A).

To introduce the concept of mutual information, we start considering a generic measurement system as sketched in Figure 3 [61]. Generally speaking, X is a vector whose components define the parameters of the “object” we wish to measure. The “measurement mechanism” maps X into a random vector Y (“observer”) introducing an inherent inaccuracy due to the measurement errors and to disturbing effects. We denote $I(X; Y)$ the *mutual information* between X and Y , i.e., the amount of information that the measurement (Y) provides about the parameters of the “object” (X). More and more information can be obtained about X when $I(X; Y)$ increases. In communication systems, it is desirable to choose, among all transmissible signals, the ones that maximise the mutual information. Conversely, a low mutual information level implies a high difficulty in identifying the “object”, as in the case of an LPI radar system.

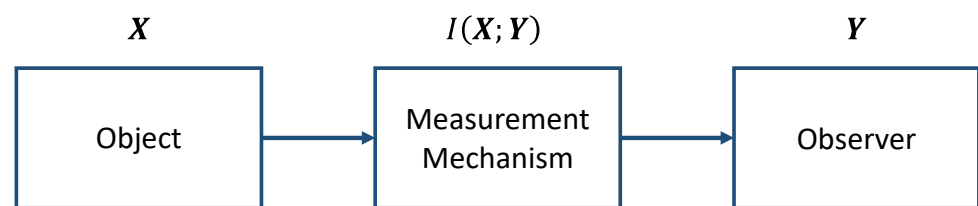


Figure 3. Block diagram of a general measurement system.

In Electronic Support (ES) measurements, the “object” is a particular radar (able to transmit some types of waveform) and the components of the vector X are the (relatively few) parameters of the waveform as obtained by the EW system, for example, TOA (time of arrival), duration (PW, Pulse Width), bandwidth B , codes—called MOP (Modulation on Pulse), samples of the power spectrum and more. Adding goniometry to the ESM an additional component is available, i.e., the DOA (direction of arrival). With the classification function or even the Specific Emitter Identification (SEI) feature, the result is a vector of integers that enumerate the radar sources in the *library*. In such a case, the information of X is not only a set of parameters (*estimation problem*) but it includes *identification (decision problem* on many hypotheses) or, as a subcase, *classification* of the radar source.

The “observer” is the vector Y , whose components represent the samples of the signal as intercepted by the ESM. The vector Y is affected by noise, multipath, and measurement errors. Hence, $I(X; Y)$ is a measure of the (partial and corrupted) information that Y provides about X . The aim of LPI radar designs is to design and select a set of radar waveforms that *minimise* the information $I(X; Y)$ for best, or optimal, LPI features. However, this set of waveforms has to guarantee the requirements for optimal target detection, i.e., a high peak-sidelobe level in its auto-correlation function (and, considering the Doppler shift, in the ambiguity function) to mitigate the masking effect due to strong targets. The constraints include a limited bandwidth to meet the frequency regulations, the required range-resolution and, finally, the efficiency in the transmitted power. Concerning the latter, noise radars show the possibility to control the peak-to-average power ratio (*PAPR*) of the radiated waveform. For a given noise sequence $g[k]$ of length n , the *PAPR* is defined as:

$$PAPR = \frac{\max_k |g[k]|^2}{\frac{1}{n} \sum_{k=1}^n |g[k]|^2} \quad (1)$$

To maximise the transmitted power, deterministic waveforms (e.g., chirp, Barker, . . . , [62]) are normally Phase/Frequency only modulated with unitary *PAPR*. Unless the amplitude is saturated, a noise waveform has a *natural PAPR* around 10 (or greater) with reduced transmitted energy, which causes a loss, compared to the *PAPR* of unity, in signal-to-noise ratio (SNR) equal to:

$$Loss[\text{dB}] = -10 \cdot \log_{10}(PAPR) \quad (2)$$

i.e., about 10 dB or greater for *natural PAPR*. To ensure a loss lower than 2 dB, the *PAPR* must be less than 1.6. Of course, when a hard limiter is applied to the noise waveform, the loss is zero. However, in this case, the noise waveform has a reduced number of degrees of freedom (DoF) equal to the number (BT) of phase values versus of 2BT degrees of full freedom, i.e., amplitude and phase pairs, or equivalently, I and Q (in-phase and quadrature) pairs. More generally, a reduction in *PAPR* causes a decrease in the equivalent number of DoF and therefore potentially a greater probability of interception of the waveform.

When the intercepted (observed) signal Y is modelled as a complex discrete-time random process, it is natural to arrive at the concept of the *Mutual Information Rate (MIR)* for the real part and the imaginary part of Y as a measure of the rate of growth of the common information versus the time, as explained in paragraph 3.1.

3.1. Mutual Information of a Random Process

For a real discrete-time random process represented by a sequence of n equally-distributed random variables $\{X_1, X_2, \dots, X_n\}$, with *marginal* and the *joint entropies* $h(X_i)$, $i = 1, 2, \dots, n$, and $h(X_1, X_2, \dots, X_n)$ respectively, the *mutual information* (not to be confused with the $I(X; Y)$ of Figure 3) can be evaluated (see Appendix A) as [59]:

$$I(X_1, X_2, \dots, X_n) = \sum_{i=1}^n h(X_i) - h(X_1, X_2, \dots, X_n) \quad (3)$$

If $\{X_1, X_2, \dots, X_n\}$ are *independent*, the mutual information is zero as $h(X_1, X_2, \dots, X_n) = \sum_{i=1}^n h(X_i)$. As a measure of the rate of growth of the common information versus the time, we introduce the *mutual information rate (MIR)*:

$$MIR = I(X_1, X_2, \dots, X_n) - I(X_1, X_2, \dots, X_{n-1}) \quad (4)$$

Substituting Equation (3) in Equation (4) and using the relation between the *joint* and the *conditional* entropy (see Appendix A for details), the *MIR* can be written as:

$$MIR = h(X_n) - h(X_n | X_{n-1}, X_{n-2}, \dots, X_1) \quad (5)$$

Therefore the *MIR* represents the *entropy* of a single sample $h(X_n)$, reduced by the *knowledge of its past*, i.e., by the *conditional entropy* $h(X_n | X_{n-1}, X_{n-2}, \dots, X_1)$. If the process is *stationary* in a *wide sense (WSS)*, and n tends to the infinity, Equation (5) becomes (for details see Appendix A):

$$MIR = h(X) - h_r(X_1, X_2, \dots, X_n) \quad (6)$$

with $h(X_n) = h(X)$ for each n , and $h_r(X_1, X_2, \dots, X_n)$ the *entropy rate*, i.e., the measure of the *average information* carried by each sample in a random sequence of n consecutive samples. In many cases, the evaluation of $h(x)$ and $h_r(X_1, X_2, \dots, X_n)$ are computationally hard, a well-known exception being one of the WSS Gaussian processes, where:

$$h(X) = \ln(\sqrt{2\pi e}) + \frac{1}{2} \ln \left[\frac{1}{2\pi} \int_{-\pi}^{+\pi} S(\omega) d\omega \right] \quad (7)$$

$$h_r(X_1, X_2, \dots, X_n) = \ln(\sqrt{2\pi e}) + \frac{1}{4\pi} \int_{-\pi}^{+\pi} \ln[S(\omega)] d\omega \quad (8)$$

with $S(\omega)$ denoting the power spectrum density of the Gaussian process. Therefore:

$$MIR = \frac{1}{2} \left\{ \ln \left[\frac{1}{2\pi} \int_{-\pi}^{+\pi} S(\omega) d\omega \right] - \frac{1}{2\pi} \int_{-\pi}^{+\pi} \ln[S(\omega)] d\omega \right\} \quad (9)$$

The *MIR* is *non-negative* and equal to zero if and only if $S(\omega)$ is a constant, i.e., for a *white* Gaussian process. By Equation (9), we can introduce the *Spectral Flatness Measure* (*SFM*) for the Gaussian process:

$$SFM = \exp(-2 \cdot MIR) = \frac{\exp\left[\frac{1}{2\pi} \int_{-\pi}^{+\pi} \ln[S(\omega)] d\omega\right]}{\frac{1}{2\pi} \int_{-\pi}^{+\pi} S(\omega) d\omega} \quad (10)$$

The *SFM* is a well-known and widely accepted method for evaluation of the “*whiteness*” (or “*compressibility*” in audio or imaging applications) of a signal. From the non-negativity of *MIR*, it is easy to show that $0 < SFM \leq 1$. Values of *SFM* close to zero ($MIR \gg 1$) correspond to a more *structured* (or less *random*) signal; an *SFM* = 1 ($MIR = 0$) corresponds to a *random, unpredictable* signal.

Table 1 shows the theoretical *MIR* and *SFM* evaluated using Equations (9) and (10) for three different power spectrum densities: Uniform, Hamming and Blackman-Nuttall (see Figure 4). In Section 4, they will be used to generate noise waveforms with the assigned spectrum. The Blackman-Nuttall spectrum will produce a more structured signal in contrast to the others.

Table 1. Theoretical *MIR* and *SFM*.

Spectrum	<i>MIR</i>	<i>SFM</i>
Uniform	0.00	1.00
Hamming	0.272	0.580
Blackman-Nuttall	1.258	0.081

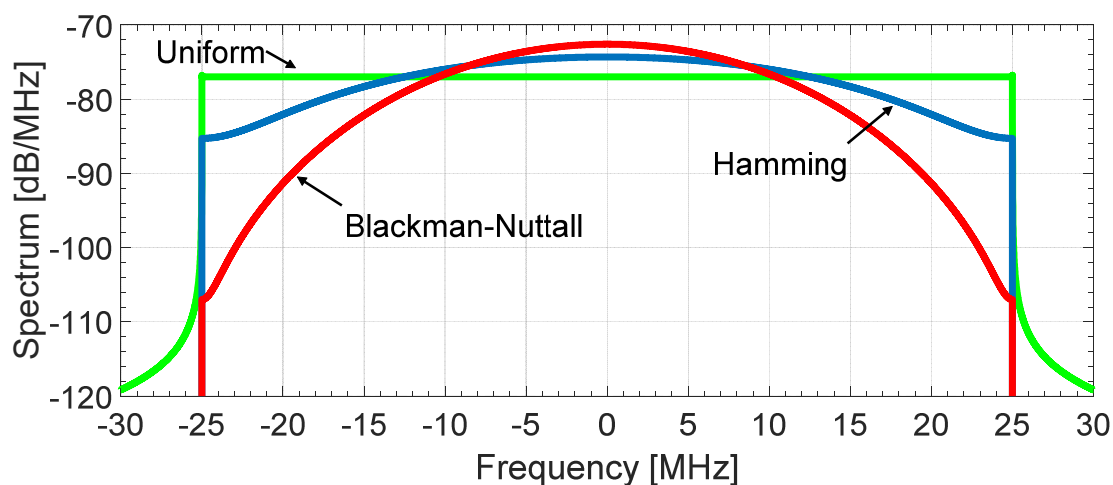


Figure 4. Uniform, Hamming and Blackman-Nuttall spectrum.

In Appendix B the definition of *MIR* is extended to a complex process [63–65]. If the process $\{Z_n\}$ is a *second-order circular doubly white* Gaussian (see Appendix B for the definition), the marginal entropy $h(Z)$ is the sum of the entropy of the real and the one of the imaginary part. Additionally, the entropy rate results equal to the sum of the entropy rate of the real and the imaginary parts. Hence:

$$MIR = \ln\left[\frac{1}{2\pi} \int_{-\pi}^{+\pi} S(\omega) d\omega\right] - \frac{1}{2\pi} \int_{-\pi}^{+\pi} \ln[S(\omega)] d\omega \quad (11)$$

3.2. On the Significance of *MIR* and *SFM* in Radar

The *MIR* and the *SFM* are used to characterise the information content of various signals, e.g., the ones from musical instruments [66] and, sometimes, radar signals [67].

The *MIR* analysis of radar signals, e.g., to characterise their LPI/LPE features, would prove really useful when one takes into account the full information chain of the counterpart: reception/interception, analysis/extraction up to data exploitation, which also includes some a priori information stored in tactical databases or “libraries”. Such an analysis is beyond the scope of this paper. Only, we wish to emphasise the importance of the a priori information by the ideal, simple example which follows.

Let the counterpart know the radar’s operating band, with bandwidth B (e.g., $B = 50$ MHz), of the victim radar, and be able to sample its signals at the Nyquist rate. Let the intercept receiver operate at a constant SNR measured on the band B before processing. The exemplary radar, here, is assumed to have two basic options:

- (a) Linear Frequency Modulation (LFM) pulse with duration T such that $B \cdot T \gg 1$.
- (b) Noise Radar operation emitting noise waveforms with three cases:
 - (b₁) “natural” Peak-to-Average Power Ratio, $PAPR \geq 10$ (Gaussian process for the components I and Q);
 - (b₂) “low $PAPR$ ”, e.g., $PAPR = 1.5$ (non-Gaussian process for I and Q);
 - (b₃) “unimodular” waveform, $PAPR = 1$ (non-Gaussian process for I and Q).

In case (a), when the counterpart knows T (does not matter whether by intelligence or measurement; remember that B is supposed to be known), the signal is fully determined (apart from an immaterial constant phase). Hence, in principle, the “new” information from the waveform’s samples is nearly zero, which is consistent with the “flat” signal spectrum of a chirp signal with $SFM = 1$, hence with a nearly null *MIR*. Similar considerations apply to any signal defined by a finite number of parameters. These parameters determine the values of the signal samples or, vice versa, in principle, may be determined by a suited, limited number of these samples.

Some preferred signal analysis methods against LPI radar use time–frequency transforms. Their main aim is to classify the emitter (e.g., to define which signals are transmitted, e.g., choose among Barker (nested Barker), polyphase, Frank, Golay, and other codes [62]). In short, they are aimed at classification, a role in which *MIR* is less important. Moreover, the ESM “sees” a signal that consists of *radar* emissions plus *noise* (from its receiver) so Equation (11) has only a theoretical value.

Different considerations apply to the case of interception of NRT signals, case (b), where the number of (real) parameters defining the (pseudo-random) signals reaches the—possibly huge—value $2BT$. Hence, even when the parameters B and T are known to the intercepting part, a number of real parameters slightly less than $2BT$ remain unknown, and the *MIR* analysis may be meaningful, and sometimes interesting, at least to compare different pseudorandom waveforms. In case (b₁) the process is Gaussian and Equation (11) can be applied, conversely, in cases (b₂) and (b₃) the low $PAPR$ makes the noise process no longer Gaussian and Equation (11) is not applicable anymore. For non-Gaussian statistics, it is necessary to evaluate the distance from the Gaussian one. A possible approach is based on the use of the *negentropy* (see Appendix A), which will be considered in Section 4.2.

4. Estimation of Entropy and Mutual Information by Simulation

Three exemplary noise waveforms $Z(k) = X(k) + jY(k)$ of N samples, duration $T = N \cdot T_s = 100 \mu\text{s}$ (T_s is the sampling time), with spectrum inside $\left(-\frac{B}{2}, +\frac{B}{2}\right)$ defined by three spectral windows (Uniform, Hamming and Blackman–Nuttall) with $B = 50$ MHz, are generated as described in [68] with a detailed functional description of the blocks of Figure 5.

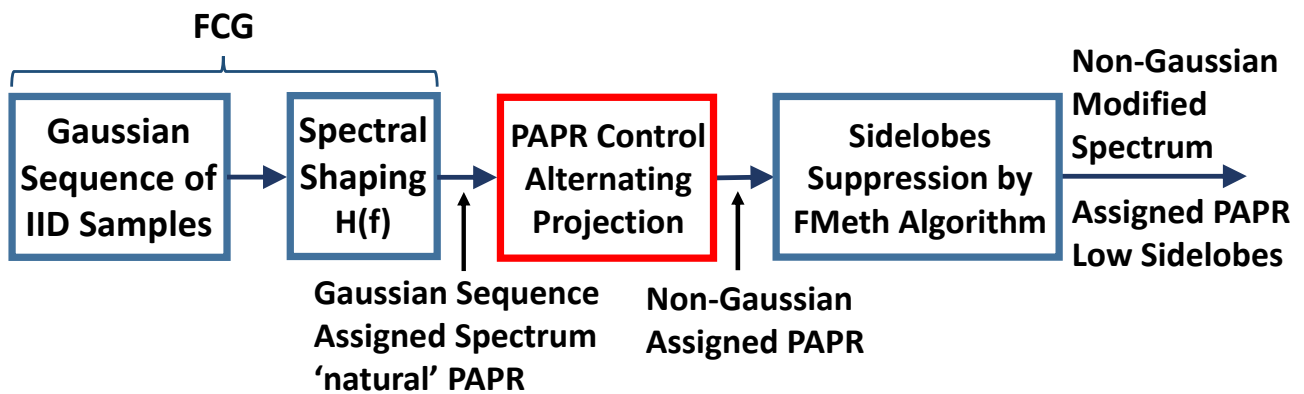


Figure 5. FCG (Fast Convolution Generator) noise waveforms with *PAPR* reduction by Alternating Projection algorithm and Sidelobes Suppression (FMeth).

We assume that the sampling frequency $F_s = \frac{1}{T_s}$ is set to B , hence the number of complex samples is BT corresponding to $2BT$ real samples (or $2BT$ degrees of freedom). The *PAPR* is varied from the “natural” value (ten or greater) to a value of one (minimum loss on Signal-to-Noise Ratio). In this frame, the three waveforms are compared in terms of marginal entropy, joint entropy and mutual information.

4.1. “Natural” *PAPR* (Gaussian Process)

A “natural” *PAPR* makes $X(k)$ and $Y(k)$ zero-mean Gaussian-distributed random variables of the same variance $\sigma_X^2 = \sigma_Y^2 = \sigma^2$. The latter, here, is set to 0.5 for the given waveforms with unit power. Posing $[X_i(1) X_i(2) \dots X_i(N)]$ where i denotes the trial i -th ($i = 1, 2, \dots, N_{trial}$ with N_{trial} set to 10^5), the analysis is limited here to the first and at the second probability hierarchy, i.e., it is performed only on a couple of extracted random variables $X_1 = [X_i(1)]$ and $X_2 = [X_i(2)]$; the same is carried out for Y_1 and Y_2 . In the following Montecarlo simulation the marginal entropy $h(X_1)$, $h(X_2)$, the joint entropy $h(X_1, X_2)$ and the mutual information $I(X_1, X_2)$ are estimated following two different approaches.

In the first approach, the marginal entropy is estimated by the analytical expression, valid for a Gaussian model, using the sample variance $S_{1(2)}^2$ (The sample variance is the estimator $S^2 = \frac{1}{N-1} \sum_{i=1}^N [X(i) - \bar{X}]^2$ with $\bar{X} = \frac{1}{N} \sum_{i=1}^N X(i)$ the sample mean. For the Gaussian population S^2 is unbiased and consistent with $Var[S^2] = \frac{2\sigma^4}{N-1}$ [60]. For $N = 10^5$ and $\sigma^2 = 0.5$, the standard deviation is 2.24×10^{-3}):

$$\hat{h}(X_{1(2)}) = \frac{1}{2} \ln(2\pi e S_{1(2)}^2) \quad (12)$$

In Table 2 for $N = 10^5$, the value of S_1^2 and S_2^2 are shown for the real part of the signal (X). Similar results are obtained for the imaginary part (Y). The maximum deviation of $S_{1(2)}^2$ from 0.5 (theoretical value) is around $3 \cdot 10^{-3}$ comparable with the standard deviation of the estimator equal to 2.24×10^{-3} .

Table 2. Estimated marginal entropy of the real part of two contiguous samples with natural *PAPR*.

Spectrum	$\sigma_{1(2)}^2$	$h(X)$	S_1^2	S_2^2	$\hat{h}(X_1)$ by S_1^2	$\hat{h}(X_2)$ by S_2^2	$\hat{h}(X_1)$ by histogram	$\hat{h}(X_2)$ by histogram
Uniform	0.5	1.0724	0.5004	0.5013	1.0727	1.0737	1.0723	1.0732
Hamming	0.5	1.0724	0.5002	0.4994	1.0725	1.0718	1.0722	1.0718
Black-Nuttall	0.5	1.0724	0.4974	0.4970	1.0697	1.0703	1.0692	1.0697

The joint entropy is estimated using Equation (13):

$$\hat{h}(X_1, X_2) = \frac{1}{2} \ln \left\{ (2\pi e)^2 |\hat{\mathbf{K}}| \right\} = \frac{1}{2} \ln \left\{ (2\pi e)^2 S_1^2 S_2^2 (1 - \hat{\rho}_{12}^2) \right\} \tag{13}$$

with $|\hat{\mathbf{K}}|$ the determinant of the estimated covariance matrix $\hat{\mathbf{K}} = \begin{vmatrix} S_1^2 & S_1 S_2 \hat{\rho}_{12} \\ S_1 S_2 \hat{\rho}_{12} & S_2^2 \end{vmatrix}$; $\hat{\rho}_{12}$

is the sample correlation coefficient: $\hat{\rho}_{12} = \frac{\sum_{i=1}^N [X_1(i) - \bar{X}_1][X_2(i) - \bar{X}_2]}{\sqrt{\sum_{i=1}^N [X_1(i) - \bar{X}_1]^2 \sum_{i=1}^N [X_2(i) - \bar{X}_2]^2}}$. Introducing the

Fisher’s auxiliary variable $W = \frac{1}{2} \ln \left(\frac{1 + \hat{\rho}_{12}}{1 - \hat{\rho}_{12}} \right)$, for large N , the distribution of W is Gaussian with mean $\eta_W \cong \frac{1}{2} \ln \left(\frac{1 + \hat{\rho}_{12}}{1 - \hat{\rho}_{12}} \right)$ and variance $\sigma_W^2 \cong \frac{1}{N-3}$ [60]; setting $N = 10^5$ the standard deviation of W is $\sigma_W \cong 3.16 \times 10^{-3}$.

For Uniform spectrum at lag $T_s = \frac{1}{B}$ the random variable $\{X_1, X_2\}$ are uncorrelated ($\rho_{12} \cong 0$), while for Hamming and Blackman-Nuttall the width of the main lobe of the autocorrelation increases and ρ_{12} assumes the value reported in Table 3. The maximum deviation of $\hat{\rho}_{12}$ from its theoretical value as shown in Table 3, after Fisher’s transformation into W , is comparable with the standard deviation of W .

Table 3. Estimated joint entropy and information of the real part of two contiguous samples with natural PAPR.

Spectrum	ρ_{12}	$\hat{\rho}_{12}$	$h(X_1, X_2)$	$\hat{h}(X_1, X_2)$ by $\hat{\rho}_{12}$	$\hat{h}(X_1, X_2)$ by 2D hist.	$I(X_1, X_2)$	$\hat{I}(X_1, X_2)$ by $\hat{\rho}_{12}$	$\hat{I}(X_1, X_2)$ by 2D hist.
Uniform	0	$5.3 \cdot 10^{-5}$	2.1447	2.1447	2.1345	0	$1.4 \cdot 10^{-9}$	0.0111
Hamming	0.4258	0.4267	2.0447	2.0443	2.0338	0.10002	0.1005	0.1099
Black-Nuttall	0.6727	0.6681	1.8435	1.8491	1.8366	0.3013	0.2956	0.3024

Finally, by Equation (3) with $n = 2$, the mutual information is estimated as:

$$\hat{I}(X_1, X_2) = \hat{h}(X_1) + \hat{h}(X_2) - \hat{h}(X_1, X_2) = -\frac{1}{2} \log \left(1 - \hat{\rho}_{12}^2 \right) \tag{14}$$

It results that, having considered a pair of successive samples, the mutual information depends only on the correlation coefficient and, for $\rho_{12} \cong 0$, it is very close to zero. If $|\rho_{12}|$ tends to one (X_1, X_2 are becoming perfectly correlated) the mutual information is going to infinity.

The estimates of $\hat{h}(X_1)$, $\hat{h}(X_2)$, $\hat{h}(X_1, X_2)$ and $\hat{I}(X_1, X_2)$, using the sample parameters, are shown in Tables 2 and 3.

A second approach to estimate the marginal and the joint entropy is carried out using the 1D and 2D histograms as an approximation of the probability density and joint density function (details about the entropy estimation by histogram are reported in Appendix C). Then, the mutual information is estimated using $\hat{h}(X_1)$, $\hat{h}(X_2)$ and $\hat{h}(X_1, X_2)$ in Equation (3) as in the previous approach. Figure 6 shows the projections of the 2D histograms on the plane (X_1, X_2) . A uniform spectrum shows a circular symmetry due to the independence of (X_1, X_2) . With Hamming and Blackman–Nuttall spectrum, after $\frac{1}{B}$ the theoretical correlation coefficient is 0.4258 and 0.6727 respectively (the level curves of the joint density function are ellipses with the major axis sloping).

The estimated entropy using histogram and the corresponding estimation of the mutual information, which are shown in Table 2, are in good agreement with those estimated by the sample parameters ($S_1^2, S_2^2, \hat{\rho}_{12}$) and with the theoretical ones.

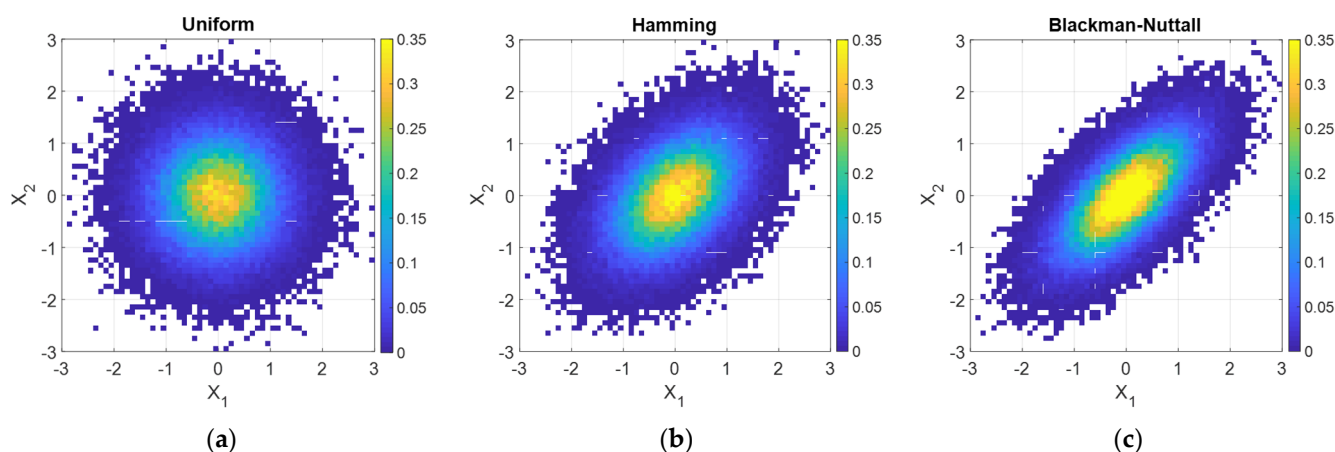


Figure 6. Projection of the 2D histogram on the plane (X_1, X_2) . (a) Uniform spectrum $\rho_{12} \cong 0$; (b) Hamming spectrum $\rho_{12} \cong 0.4258$; (c) Blackman-Nuttall spectrum $\rho_{12} \cong 0.6727$. Parameters: $B = 50$ MHz, $F_s = \frac{1}{B}$, $\sigma_1^2 = \sigma_2^2 = 0.5$.

4.2. Controlled PAPR (Non-Gaussian Process)

A reduction of the PAPR from the natural value by a non-linear transformation (Alternating Projection algorithm in Figure 5) distorts the random process reducing or deleting its Gaussianity. For a real random variable X , the negentropy $J(X)$ measures the loss of Gaussianity and is defined as:

$$J(X) = h_G(X) - h(X) \tag{15}$$

where $h_G(X) = \frac{1}{2} \ln(2\pi e \sigma^2)$ is the entropy of a Gaussian random variable with the same variance σ^2 as the random variable X . As well known, see for instance [60], the Gaussian distribution has the maximal entropy among all distributions with the same variance, hence a positive $J(X)$ measures the distance of X from the Gaussian model. The main problem using negentropy is its evaluation, very difficult when the probability density function of X is unknown.

Figure 7 shows the estimated negentropy of the real part of a single waveform generated by FCG (with Uniform, Hamming and Blackman–Nuttall spectrum, $B = 50$ MHz, $\sigma^2 = 0.5$) varying the PAPR, where $h_G(X) = \frac{1}{2} \ln(2\pi e \sigma^2) = \ln(\sqrt{\pi e}) \cong 1.0724$ and $h(X)$ are estimated using the histogram (Appendix C). The negentropy results are independent of the spectrum shape. When the PAPR is less than $4\sigma \cong 3$ the process begins to be definitely non-Gaussian. The same trend is obtained for the imaginary part. According to Figure 7, for $PAPR > 3$ the joint entropy and the joint information are the ones of a Gaussian process (see Figure 8a,b). For $PAPR < 3$ the non-linear transformation de-correlates the samples reducing both the joint entropy (Figure 8a) and the mutual information (Figure 8b).

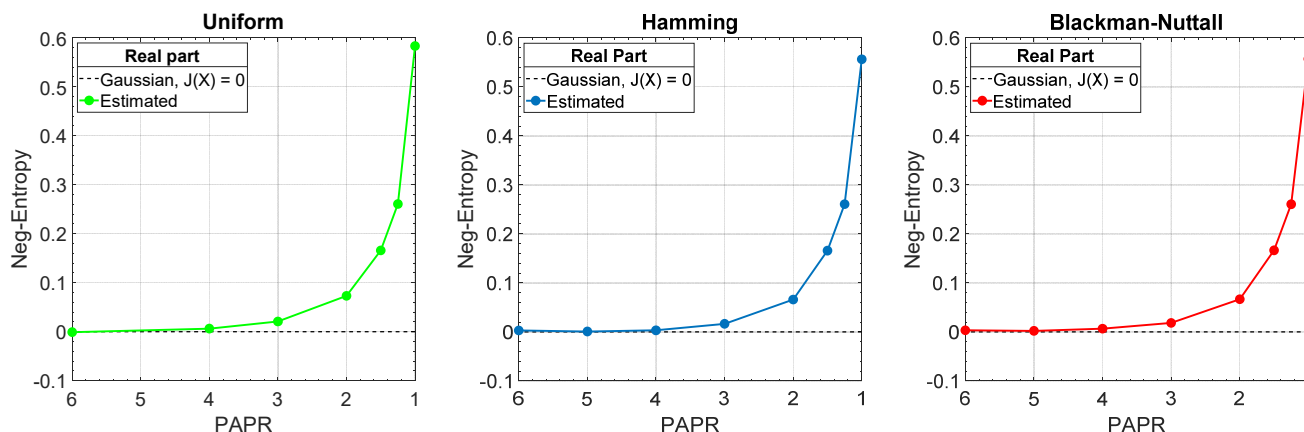


Figure 7. Neg-Entropy vs. PAPR of the real part of a single noise waveform with unit power generated using FCG.

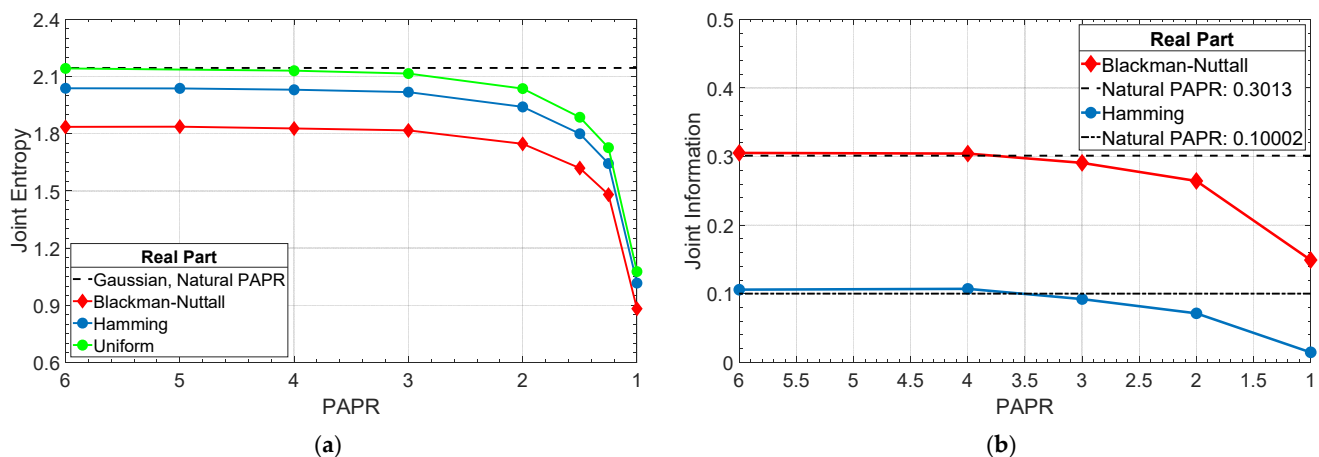


Figure 8. (a) Joint entropy (b) joint information vs. $PAPR$ of the real part of a single noise waveform with unit power.

5. Conclusions, Recommendations and Perspectives for Future Research

The transmitted waveform is a key performance element in every radar design. Using Noise Radar Technology, the pseudo-random waveforms shall be suitably “tailored” to satisfy contrasting requirements in terms of power efficiency (calling for a “low” and often nearly unitary $PAPR$) and of the information available to any intercepting counterpart (calling for a “high” $PAPR$ equal or close to that, order of 9–10, of a Gaussian process). In addition, low sidelobes of the autocorrelation function are requested by many applications, calling for a significant spectral weighting (e.g., Blackman–Nuttall or Taylor).

From a preliminary information-theoretic analysis, using the *negentropy* concept, the effect of the selected $PAPR$ on the information content that is available to EW receivers was analysed in its main trends. It results that for a $PAPR$ value above an approximate threshold of three, the Gaussian approximation for the entropy holds. On the other hand, when the $PAPR$ goes from about three towards the unity, there is a steep increase in the negentropy and a significant decrease in the joint information of a pair of successive samples due to their progressive decorrelation. This decorrelation increases the quantity of information available to the counterpart and might limit the LPI properties of the radar by supporting a detection or interception of the noise radar emission. Thus, the trade-off in noise radar waveform design between longer detection ranges (demanding for higher effective transmitter powers often implemented efficiently by lower $PAPR$ values) and LPI features remains. This paper aims to support such design decisions by providing some quantitative analysis of the relationship between the $PAPR$ and the detectability or exploitability of the radar waveform.

Author Contributions: G.G. defined the main content, including a synthesis of the modern EW systems, and the overall organisation of the paper. G.P. developed the waveform analysis tools and obtained the related results. K.S. and C.W. studied, checked, and critically reviewed the waveform generation methods and the related analysis of their information content. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1. Entropy and Negentropy

From an energetic point of view, in a transformation process, maximum efficiency is obtained when the sum of all involved energies after the process has occurred has an ability to produce work equal to their sum before. The II Law of Thermodynamics, the well-known “increase of entropy principle”, states that maximum efficiency is attained in energy transformations for an (ideal) reversible process, i.e., an “infinitely slow” one, in which the entropy production is zero, hence, the order is saved. The extent of the entropy creation due to irreversibility is a measure of the “lack of ideality” of a process. In energy transformations, we get the maximum efficiency using processes in which the entropy creation due to irreversibility is minimal.

The mechanical–statistical character of the entropy concept and its logarithmic connection to the probability is due to Ludwig E. Boltzmann (20 February 1844–5 September 1906) and his kinetic gas theory. The tombstone of Boltzmann in the Viennese Zentralfriedhof shows his celebrated formula for the entropy (S) of a system with W possible microstates: $S = k_B \cdot \log(W)$, where k_B is the Boltzmann’s constant and $\log(\cdot)$ the natural logarithm.

In the Information Theory context [59], entropy is generally seen as a measure of the uncertainty (unpredictability) of an event, and for a random variable X , of the “distance” of its realisations from the predictable ones.

Given a discrete random variable X with probability mass function $p(x_i) = P\{X = x_i\}$, the entropy is defined as (being $E[\cdot]$ the expected value operator):

$$H(X) = -E[\log\{p(x_i)\}] = -\sum_i p(x_i) \cdot \log\{p(x_i)\} \quad (\text{A1})$$

For a continuous random variable X with probability density function $f(x)$, the differential entropy (denoted with the lowercase letter) is defined as:

$$h(X) = -E[\log\{f(x)\}] = -\int f(x) \cdot \log\{f(x)\} dx \quad (\text{A2})$$

In the present discussion, the natural logarithm will be used. Note that when dealing with discrete random variables the equivalent formulation with base-2 logarithm (\log_2) is more widely used, and leads to the well-known information unit called bit, while the natural logarithm, more widely used in signal processing, leads to the (equivalent) *nat* unit ($1 \text{ nat} = \frac{1}{\ln(2)} \cong 1.443 \text{ bit}$).

A tutorial example of entropy evaluation is the coin-tossing experiment with outcomes head (with probability $p_1 = P\{\text{head}\}$) and tail (with probability $p_2 = P\{\text{tail}\} = 1 - p_1$), which generates a Bernoulli-type variable. Its entropy is: $H_{\text{Bernoulli}} = -[p_1 \log(p_1) + p_2 \log(p_2)]$, which is equal to 1, i.e., one bit of information, when \log_2 is used and the coin is unbiased ($p_1 = 0.5$), i.e., the result is minimally predictable (the disorder is maximum). For a biased coin, predictability increases and the entropy is smaller: for $p_1 = 0.25$, it equals 0.8114 bit (0.5623 nat), and for p_1 going to zero, or to one, the entropy goes to zero, i.e., the result becomes predictable.

When the random variable X is normalised (i.e., with zero mean and unit variance) its entropy only depends on the particular shape of its probability density function $f(x)$ or probability mass function $p(x_i)$. For real continuous random variables the interesting result holds that, among all normalised random variables, the Gaussian (Normal) variable has the maximum entropy: $h_G(X) = \frac{1}{2} \ln(2\pi e) \cong 1.42 \text{ nat}$.

Hence, any non-linear transformation of a Gaussian variable reduces the entropy, or “creates some negentropy”. The term negentropy is first found in Schrödinger’s book [69], which discusses the complexity of life and its evolution towards more and more “organized”, or somehow, “ordered”, living species, a situation which seems to “create order” and to “negate the entropy”, apparently against the II Law of Thermodynamics.

For a random variable X , the negentropy $J(X)$ is defined as:

$$J(X) = h_G(X) - H(X) \quad (\text{A3})$$

where $h_G(X) = \frac{1}{2} \ln(2\pi e \sigma^2)$ is the entropy of a Gaussian random variable with the same variance σ^2 of X . As said, the Gaussian distribution has the maximal entropy among all distributions of the same variance, hence $J(x) \geq 0$.

The main problem using negentropy is its evaluation, very difficult when the probability density function of X is unknown and it has to be estimated. Therefore, approximations of negentropy use the higher-order (up to four) moments [70]:

$$J(X) \approx \frac{1}{12} E[X^3]^2 + \frac{1}{48} \left\{ E[X^4] - 3E[X^2]^2 \right\}^2 \quad (\text{A4})$$

Appendix A.2. Self-Information and Mutual-Information

Posing $I(X) = -\log\{p(x_i)\}$ in Equation (A1) or $I(X) = -\log\{f(x)\}$ in Equation (A2), the entropy of X equals the expectation of the random variable $I(X)$. The latter is said *self-information* of X and measures the information obtained by observing X , or equivalently, it measures the *a priori uncertainty* in the outcome of the event $\{X = x_i\}$ in the discrete case and $\{x < X < x + dx\}$ for a continuous random variable.

A real random process $\{X(t)\}$, after sampling, can be described by n samples, i.e., by n real random variables X_1, X_2, \dots, X_n (or equivalently in the row vector notation $\mathbf{X}_n \equiv [X_1, X_2, \dots, X_n]$) with *marginal* probability density function $f(x_i)$ for $i = 1, 2, \dots, n$, and *joint* probability density function $f(\mathbf{X}_n) \triangleq f(x_1, x_2, \dots, x_n)$. The *marginal entropy* for each X_i and the *joint entropy* of \mathbf{X}_n are:

$$h(X_i) = -E[\ln\{f(x_i)\}] \quad \forall i \quad (\text{A5})$$

$$h(\mathbf{X}_n) \triangleq h(X_1, X_2, \dots, X_n) = - \int_{\mathbb{R}^n} f(x_1, x_2, \dots, x_n) \ln\{f(x_1, x_2, \dots, x_n)\} dx_1 dx_2 \dots dx_n \quad (\text{A6})$$

The *mutual information* of \mathbf{X}_n is defined as [59]:

$$I(\mathbf{X}_n) \triangleq I(X_1, X_2, \dots, X_n) = \int_{\mathbb{R}^n} f(x_1, x_2, \dots, x_n) \ln \left[\frac{f(x_1, x_2, \dots, x_n)}{f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_n)} \right] dx_1 \dots dx_n \quad (\text{A7})$$

Developing the integrand function of Equation (A7), we obtain:

$$I(X_1, X_2, \dots, X_n) = \sum_{i=1}^n h(X_i) - h(X_1, X_2, \dots, X_n) \quad (\text{A8})$$

If $\{X_1, X_2, \dots, X_n\}$ are independent, the mutual information $I(X_1, X_2, \dots, X_n) = 0$ and it results in $h(X_1, X_2, \dots, X_n) = \sum_{i=1}^n h(X_i)$.

For a *multivariate* Gaussian distribution with *covariance* matrix \mathbf{K} , the *joint entropy* is:

$$h(X_1, X_2, \dots, X_n) = \frac{1}{2} \ln\{(2\pi e)^n |\mathbf{K}|\} \quad (\text{A9})$$

where $|\mathbf{K}|$ denotes the determinant of \mathbf{K} . For a *bivariate* Gaussian distribution with covariance matrix $\mathbf{K} = \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix}$ being ρ the correlation coefficient, by Equation (A9), with $h(X_1) = h(X_2) = \frac{1}{2} \ln(2\pi e \sigma^2)$ and $h(X_1, X_2) = \frac{1}{2} \ln\{(2\pi e)^2 |\mathbf{K}|\} = \frac{1}{2} \ln\{(2\pi e)^2 \sigma^4 (1 - \rho^2)\}$, the mutual information depends on only ρ and it results:

$$I(X_1, X_2) = -\frac{1}{2} \ln(1 - \rho^2) \quad (\text{A10})$$

If $\rho = 0$, $\{X_1, X_2\}$ are independent and the mutual information is zero. If $\rho = \pm 1$, $\{X_1, X_2\}$ are perfectly correlated and the mutual information is infinite.

By Equation (A8), we can introduce the mutual information rate (MIR) as a measure of the rate of growth of the common information as a function of the time:

$$MIR = I(X_1, X_2, \dots, X_n) - I(X_1, X_2, \dots, X_{n-1}) \tag{A11}$$

If $\{X_1, X_2, \dots, X_n\}$ are independent, $MIR = 0$. By Equation (A8), MIR can be written as:

$$MIR = h(X_n) - h(X_1, X_2, \dots, X_n) + h(X_1, X_2, \dots, X_{n-1}) \tag{A12}$$

Using the relation between the *joint* and the *conditional* densities for $\{X_1, X_2, \dots, X_n\}$, i.e.,: $f(x_1, x_2, \dots, x_n) = f(x_n|x_{n-1}, \dots, x_1) \cdot f(x_{n-1}|x_{n-2}, \dots, x_1) \cdot \dots \cdot f(x_2|x_1) \cdot f(x_1)$, the *joint entropy* is written as a function of the *conditional entropy*:

$$h(X_1, X_2, \dots, X_n) = h(X_n|X_{n-1}, \dots, X_1) + h(X_{n-1}|X_{n-2}, \dots, X_1) + \dots + h(X_2|X_1) + h(X_1) \tag{A13}$$

and the MIR becomes:

$$MIR = h(X_n) - h(X_n|X_{n-1}, X_{n-2}, \dots, X_1) \tag{A14}$$

For a real *wide sense stationary* (WSS) process Equation (A14) can be written as:

$$MIR = h(X) - h_r(X_1, X_2, \dots, X_n) \tag{A15}$$

being $h(X_n) = h(X)$ for each n and, when $n \rightarrow \infty$, it can be demonstrated that $\lim_{n \rightarrow \infty} \frac{1}{n} h(X_1, X_2, \dots, X_n) = \lim_{n \rightarrow \infty} \frac{1}{n} h(X_n|X_{n-1}, X_{n-2}, \dots, X_1) \triangleq h_r(X_1, X_2, \dots, X_n)$, where the second limit defines the *entropy rate* $h_r(X_1, X_2, \dots, X_n)$, i.e., the measure of the *average information carried by each sample* in a random sequence of n consecutive samples.

The above concepts can be extended to n discrete random variables. For example, considering a process generating, at each step, a random natural number X_i in the interval from 1 to N , (where N in computer applications may be, for example, the largest representable integer number, equal to 2^b where b is the number of bits used in the representation) if the distribution of X is *uniform* and each generation is *independent* of the others, the resulting entropy after M steps and for $N \gg 1$ is $\frac{M}{N} \log(N)$. It seems useful to define $\frac{1}{N} \log(N)$ as the *entropy rate*, or *information rate*, of this process.

Appendix B

For a complex random variable $Z = X + jY$ the probability density function, $f(Z)$, is defined by the *joint* density of the *real* (X) and the *imaginary* (Y) part, i.e., $f(Z) \triangleq f(X, Y)$.

Considering a complex random vector ($Z \in \mathbb{C}^n$), with n complex components $\{Z_1, Z_2, \dots, Z_n\}$, the *joint* probability density function is $f(Z_n) \triangleq f(X_n, Y_n)$ where $X_n, Y_n \in \mathbb{R}^n$ are real vectors denoting the real and the imaginary part of Z_n .

The entropy of Z_n is:

$$h(Z_n) \triangleq h(X_n, Y_n) = -E\{\log[f(X_n, Y_n)]\} = H(Z_1, Z_2, \dots, Z_n) \tag{A16}$$

A complex random process $\{Z(t)\}$, after sampling, can be described by n complex samples, i.e., by a raw vector $Z_n = [Z_1, Z_2, \dots, Z_n]$, whose components are complex random variables $Z_i = X_i + jY_i$ with $i = 1, 2, \dots, n$. Without loss of generality, we suppose Z_n to be *zero-mean*.

In this context of complex processes, to extend the definition of the *entropy rate*, we introduce the *second-order stationary* (SOS) concept. The only *covariance function* $R(i, i + m) = E[Z_{i+m}Z_i^*]$ is not sufficient to entirely describe an SOS process. It is necessary to introduce a second function, said the *pseudo-covariance function* (also called the *relation function* in [65]), defined as $\tilde{R}(i, i + m) = E[Z_{i+m}Z_i]$.

Definition of a Second-Order Stationary (SOS) Process

A complex random process Z_n is said SOS if:

- (i) it is WSS, i.e., $\forall i E[Z_i]$ is constant and the covariance function $R(i, i + m) = E[Z_{i+m}Z_i^*]$ only depends on the index difference m , i.e., $R(i, i + m) = R(m)$;
- (ii) $\forall i$ the pseudo covariance function $\tilde{R}(i, i + m) = E[Z_{i+m}Z_i]$ only depends on the index difference m , i.e., $\tilde{R}(i, i + m) = \tilde{R}(m)$.

For complex random processes, the WSS assumption does not necessarily imply the SOS. In many applications (communications, medical image processing, . . .) signals (realisations of the process), in the most general case, are complex with the real and imaginary parts being possibly *correlated* to each other. In this case the $\tilde{R}(m)$ is complex. However, when the real and imaginary parts are *uncorrelated*, $\tilde{R}(m)$ has real values, and when $\tilde{R}(m) = 0$, the process is said *second-order circular (SOC)* [63]. (Definition: A complex process is a *second-order circular (SOC)* if its second-order statistics are *invariant* in any phase transformation, i.e., considering \mathbf{Z} and $\mathbf{Z} \cdot e^{j\alpha}$, their covariance functions, $R(m)$, are equal for any real number α , but their pseudo-covariance functions, $\tilde{R}(m)$, are equal *if and only if* they are zero).

If the operator $\mathcal{F}\{\cdot\}$ defining the *Fourier Transform* is applied to the covariance function and to the pseudo-covariance function, the *power spectrum* $S(\omega) = \mathcal{F}\{R(m)\}$ and the *pseudo-power spectrum* $\tilde{S}(\omega) = \mathcal{F}\{\tilde{R}(m)\}$ are defined.

For a real process $S(\omega) = \tilde{S}(\omega)$, and uniquely it defines the *power spectrum* of the process, which shows the well-known properties of *symmetry*, *non-negative* and a *finite-value* of the integral over the interval $[-\pi, +\pi]$. Conversely, any function $S(\omega)$ satisfying the previous three properties can be considered to be the Fourier transform of the covariance function of a real process.

For a complex process, given the two functions $S(\omega)$ and $\tilde{S}(\omega)$, they are respectively the *power spectrum* and the *pseudo-power spectrum* of a complex SOS random process if they satisfy the necessary conditions:

$$(i) S(\omega) \geq 0, (ii) \tilde{S}(\omega) = \tilde{S}(-\omega), (iii) |\tilde{S}(\omega)|^2 \leq S(\omega)S(-\omega) \tag{A17}$$

An SOS complex process is said “*white*” if the covariance function $R(m) = R(0) \cdot \delta(m)$, where $\delta(m)$ is the *delta function*, and *no restriction* is imposed on $\tilde{R}(m)$. Instead, an SOS complex process is said “*doubly white*” if it is “*white*”, and $\tilde{R}(m) = \tilde{R}(0) \cdot \delta(m)$.

In general $\tilde{R}(m)$ is complex, but when the real and imaginary parts are *uncorrelated*, it assumes a real value and, when $\tilde{R}(m) = 0$, the process is a *circular white process*.

Now, after the introduction of the above concepts and definitions, we define the *entropy rate* of a complex random process \mathbf{Z}_n . As it occurs in the real case, the *entropy rate* is defined:

$$h_r(\mathbf{Z}_n) = h_r(Z_1, Z_2, \dots, Z_n) = \lim_{n \rightarrow \infty} \frac{1}{n} h(Z_1, Z_2, \dots, Z_n) \tag{A18}$$

if the limit exists, and it results: $h(Z_1, Z_2, \dots, Z_n) \leq \sum_{k=1}^n h(Z_k)$, where the equality occurs if and only if the Z_i are independent.

Hence, the entropy rate can be used to measure the sample dependence and it reaches the upper bound when all samples of the process are independent.

Given a complex SOS Gaussian process \mathbf{Z}_n with *power spectrum* $S(\omega)$ and *pseudo-power spectrum* $\tilde{S}(\omega)$, the entropy rate is [64]:

$$h_r(\mathbf{Z}_n) = \ln(\pi e) + \frac{1}{4\pi} \int_{-\pi}^{+\pi} \ln \left[S(\omega)S(-\omega) - |\tilde{S}(\omega)|^2 \right] d\omega \tag{A19}$$

If \mathbf{Z}_n is a Gaussian *second-order circular* process, $\tilde{S}(\omega) = 0$, and Equation (A19) becomes:

$$h_r^{circ.}(\mathbf{Z}_n) = \ln(\pi e) + \frac{1}{4\pi} \int_{-\pi}^{+\pi} \ln[S(\omega)S(-\omega)] d\omega \tag{A20}$$

hence, being in the most general case $|\tilde{S}(\omega)|^2 \geq 0$, for the second-order *circular* and *non-circular* Gaussian process with the same covariance function $R(m)$, we have:

$$h_r^{non-circ.}(\mathbf{Z}_n) < h_r^{circ.}(\mathbf{Z}_n) \tag{A21}$$

Now we show that for a *doubly white* Gaussian random process with $\tilde{R}(0) \in \mathbb{R}$ (i.e., with the real and imaginary parts *uncorrelated*) the entropy rate, Equation (A19), is the sum of the entropy rate of the real part, $h_r(\mathbf{X}_n)$, and imaginary parts, $h_r(\mathbf{Y}_n)$, of the process. Since \mathbf{X}_n and \mathbf{Y}_n are *uncorrelated*, we can directly use the entropy rate formula for real Gaussian process, and the entropy rate of the complex Gaussian process is:

$$h_r(\mathbf{Z}_n) = h_r(\mathbf{X}_n) + h_r(\mathbf{Y}_n) \tag{A22}$$

Since X_n and Y_n are white, i.e., with constant spectrum: $S_X(\omega) = R_X(0)$ and $S_Y(\omega) = R_Y(0)$, due to the uncorrelation hypothesis for X_n, Y_n : $R(0) = R_X(0) + R_Y(0)$ and $\tilde{R}(0) = R_X(0) - R_Y(0)$, it results $S(\omega) = S(-\omega) = S_X(\omega) + S_Y(\omega)$ and $\tilde{S}(\omega) = S_X(\omega) - S_Y(\omega)$ real. Then:

$$S(\omega)S(-\omega) - |\tilde{S}(\omega)|^2 = 4 \cdot S_X(\omega)S_Y(\omega) \quad (\text{A23})$$

Inserting Equation (A23) in Equation (A19), we obtain Equation (A22) being:

$$h_r(X_n) = \ln(\sqrt{2\pi e}) + \frac{1}{4\pi} \int_{-\pi}^{+\pi} \ln[S_X(\omega)]d\omega \quad (\text{A24})$$

$$h_r(Y_n) = \ln(\sqrt{2\pi e}) + \frac{1}{4\pi} \int_{-\pi}^{+\pi} \ln[S_Y(\omega)]d\omega \quad (\text{A25})$$

If the Gaussian random process is *second-order circular white*, $\tilde{R}(m) = 0$, the entropy rate is simply given by twice the entropy rate in the real domain and $MIR = h(Z) - h_r(Z_1, Z_2, \dots, Z_n)$, can be evaluated considering the marginal entropy $h(Z) = \ln(\pi e) + \ln\left[\frac{1}{2\pi} \int_{-\pi}^{+\pi} S(\omega)d\omega\right]$ (sum of the entropy of the real and the imaginary part) and also the entropy rate as the sum of the entropy rate of the real and the imaginary part: $h_r(Z_1, Z_2, \dots, Z_n) = \ln(\pi e) + \frac{1}{2\pi} \int_{-\pi}^{+\pi} \ln[S(\omega)]d\omega$. Hence:

$$MIR = \ln\left[\frac{1}{2\pi} \int_{-\pi}^{+\pi} S(\omega)d\omega\right] - \frac{1}{2\pi} \int_{-\pi}^{+\pi} \ln[S(\omega)]d\omega \quad (\text{A26})$$

Appendix C

The following Matlab function allow us to estimate the marginal and the joint entropy of the real (imaginary) part of a noise waveforms.

```
function E = EntropyEstimationHist(X1)
h = histogram(X1);
x = h.BinEdges;
BinCount = h.BinCounts;
zero = find(BinCount == 0);
for k = 1:length(zero)
BinCount(zero(k)) = 1e - 18;
end
BinCount = BinCount/sum(BinCount);
step=x(2) - x(1);
E = log(step) - sum(BinCount.*log(BinCount));
end
```

```
function E = EntropyEstimationHist2D(X1,X2)
h = histogram2(X,X2);
x = h.XBinEdges;
y = h.YBinEdges;
stepx = x(2) - x(1);
stepy = y(2) - y(1);
BinCount = h.BinCounts;
zero = find(BinCount == 0);
for k = 1:length(zero)
BinCount(zero(k)) = 1e - 18;
end
BinCount = BinCount/sum(sum(BinCount));
E = log(stepx*stepy) - sum(sum(BinCount.*log(BinCount)));
```

It uses the approximation of the density function and the joint density function by the 1D and 2D histogram.

The performance of the estimation depends on the number of samples (or pairs for 2D histograms) used to evaluate the histogram. Figure A1a shows the estimated marginal entropy, Figure A1b the joint entropy versus the number of samples (or pairs) and their standard deviation (Figure A1c). Generally, this approach presents an underestimate of the entropy for a low number of samples (lack of values in the histogram evaluation). For a large number of samples (or pairs), i.e., $\geq 10^5$, the performance becomes to be appreciable.

This approach conceptually could be extended to estimate the joint density in the multivariate case ($n > 2$), however, the “curse of dimensionality” [71] makes it impossible to realise. This term “curse of dimensionality” was coined by R. E. Bellman (1961) to indicate that the number of samples needed to estimate an arbitrary function with a given level of accuracy grows exponentially with the number of variables that it comprises.

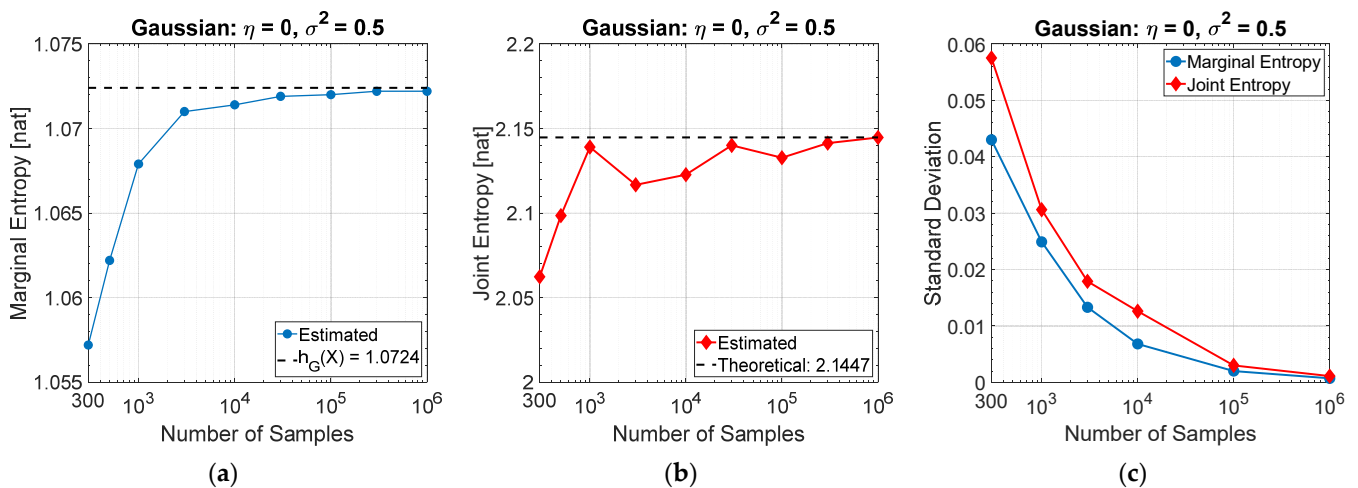


Figure A1. (a) Estimated marginal entropy, (b) estimated joint entropy, (c) standard deviation of the estimators, versus the number of samples used to evaluate in (a) the histogram of the Gaussian random variable ($\sigma^2 = 0.5$, dashed line is the theoretical value: $h_G(X) = \frac{1}{2} \ln(2\pi e \sigma^2) \cong 1.0724$); in (b) the bi-dimensional histogram of the bivariate Gaussian independent variables ($\sigma^2 = 0.5$, dashed line is the theoretical value: $h_G(X_1, X_2) = \frac{1}{2} \ln\{(2\pi e)^2 \sigma^4\} \cong 2.1447$).

References

- Farina, A.; Galati, G. Surveillance radars: State of the art, research and perspectives. In *Optimised Radar Processors*; IET Press: London, UK, 1987; ISBN 9781849191760. [\[CrossRef\]](#)
- Farina, A. Electronic Counter-Counter Measures, Chapter 9. In *Radar Handbook*, 2nd ed.; Skolnik, M.I., Ed.; Mc. Graw Hill: New York, NY, USA, 1990.
- Li, N.-J.; Zhang, Y.-T. A Survey of Radar ECM and ECCM. *IEEE Trans. Aerosp. Electron. Syst.* **1995**, *31*, 1110–1120. [\[CrossRef\]](#)
- Guerci, J.R. Cognitive radar: A knowledge-aided fully adaptive approach. In Proceedings of the 2010 IEEE Radar Conference, Arlington, VA, USA, 10–14 May 2010; pp. 1365–1370. [\[CrossRef\]](#)
- Haykin, S.; Xue, Y.; Setoodeh, P. Cognitive radar: Step toward bridging the gap between neuroscience and engineering. *Proc. IEEE* **2012**, *100*, 3102–3130. [\[CrossRef\]](#)
- Aberman, K.; Aviv, S.; Eldar, Y.C. Adaptive frequency allocation in radar imaging: Towards cognitive SAR. In Proceedings of the 2010 IEEE Radar Conference, Seattle, WA, USA, 8–12 May 2017; pp. 1348–1351. [\[CrossRef\]](#)
- Horne, C.; Ritchie, M.; Griffiths, H. Proposed ontology for cognitive radar systems. *IET Radar Sonar Navig.* **2018**, *12*, 1363–1370. [\[CrossRef\]](#)
- Mitchell, A.E.; Garry, J.L.; Duly, A.J.; Smith, G.E.; Bell, K.L.; Rangaswamy, M. Fully Adaptive Radar for Variable Resolution Imaging. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9810–9819. [\[CrossRef\]](#)
- Lang, Y.-C. *Dynamic Spectrum Management—From Cognitive Radio to Blockchain and Artificial Intelligence*; Springer: Berlin/Heidelberg, Germany, 2020. [\[CrossRef\]](#)
- Sidiropoulos, N.D.; De Lathauwer, L.; Fu, X.; Wang, K.; Papalexakis, E.E. Tensor Decomposition for Signal Processing and Machine Learning. *IEEE Trans. Signal Process.* **2017**, *65*, 3551–3582. [\[CrossRef\]](#)
- Chiriyath, A.R.; Paul, B.; Bliss, D.W. Radar-Communications Convergence: Coexistence, Cooperation, and Co-Design. *IEEE Trans. Cogn. Commun. Netw.* **2017**, *3*. [\[CrossRef\]](#)
- Liu, F.; Zhou, L.; Masouros, C.; Li, A.; Luo, W.; Petropulu, A. Toward Dual-functional Radar-Communication Systems: Optimal Waveform Design. *IEEE Trans. Signal Process.* **2018**, *66*, 4264–4279. [\[CrossRef\]](#)
- Oroian, T.C.; Enache, F.; Ciotirnae, P. Some considerations about third-order statistics for different types of radar signals. In Proceedings of the 10th Intern. Symposium on Advanced Topics in Electrical Engineering (ATEE), Bucharest, Romania, 23–25 March 2017. [\[CrossRef\]](#)
- Aly, O.A.M.; Omar, A.S. Detection and localization of RF radar pulses in noise environments using wavelet packet transform and higher order statistics. *Prog. Electromagn. Res. PIER* **2006**, *58*, 301–317.
- Barbarossa, S. Parameter estimation of undersampled signals by Wigner–Ville analysis. *IEEE Conf. Acoust. Speech Signal Process. ICASSP 91* **2002**, *5*, 3944–3947.
- Gulum, T.O.; Pace, P.E.; Cristi, R. Extraction of Polyphase Radar Modulation Parameters Using a Wigner-Ville Distribution Radon Transform. In Proceedings of the IEEE International Conf. on Acoustics, Speech and Signal Processing, Las Vegas, NV, USA, 31 March–4 April 2008. [\[CrossRef\]](#)
- Copeland, D.B.; Pace, P.E. Detection and analysis of FMCW and P-4 polyphase LPI waveforms using quadrature mirror filter trees. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Orlando, FL, USA, 13–17 May 2002; pp. IV-3960–IV-3963. [\[CrossRef\]](#)

18. Roberts, R.S.; Brown, W.A.; Loomis, H.H. Computationally efficient algorithms for cyclic spectral analysis. *IEEE Signal Process. Mag.* **1991**, *8*, 38–49. [[CrossRef](#)]
19. Hejazi Kookamari, F.; Norouzi, Y.; Nayebi, M.M. Using a moving aerial platform to detect and localise a low probability of intercept radar. *IET Radar Sonar Navig.* **2017**, *11*, 1062–1069. [[CrossRef](#)]
20. Hejazi Kookamari, F.; Norouzi, Y.; Kashani, E.S.; Nayebi, M.M. A Novel Method to Detect and Localize LPI Radars. *IEEE Trans. Aerosp. Electron. Syst.* **2019**, *55*, 2327–2336. [[CrossRef](#)]
21. Kong, S.H.; Kim, M.; Hoang, L.M.; Kim, E. Automatic LPI radar waveform recognition using CNN. *IEEE Access* **2018**, *6*, 4207–4219. [[CrossRef](#)]
22. Liu, G.S.; Gu, H.; Su, W.M.; Sun, H.B. The analysis and design of modern Low Probability of Intercept radar. In Proceedings of the 2001 CIE International Conference on Radar Proceedings (Cat No.01TH8559), Beijing, China, 15–18 October 2001; pp. 120–124. [[CrossRef](#)]
23. Wirth, W.D. Omni directional low probability of intercept radar. In Proceedings of the International Conference on Radar 89, Paris, France, 24–28 April 1989; Volume 1 (A90-40951 18-32), pp. 25–30.
24. Wirth, W.D. Long term coherent integration for a floodlight radar. In Proceedings of the IEEE 1995 International Radar Conference, Alexandria, VA, USA, 8–11 May 1995; pp. 698–703. [[CrossRef](#)]
25. Schleher, D.C. Low probability of intercept radar. In Proceedings of the International Radar Conference, Arlington, VA, USA, 6–9 May 1985; Record (A86-32576 14-32). Institute of Electrical and Electronics Engineers, Inc.: New York, NY, USA, 1985; pp. 346–349.
26. Burgos-Garcia, M.; Sanmartin-Jara, J. A LPI tracking radar system based on frequency hopping. In Proceedings of the International Radar Symposium, Munich, Germany, 15–17 September 1998; pp. 151–159.
27. Gross, F.B.; Chen, K. Comparison of detectability of traditional pulsed and spread spectrum radar waveforms in classic passive receivers. *IEEE Trans. Aerosp. Electron. Syst.* **2005**, *41*, 746–751. [[CrossRef](#)]
28. Wang, W. Potential transmit beamforming schemes for active LPI radars. *IEEE Aerosp. Electron. Syst. Mag.* **2017**, *32*, 46–52. [[CrossRef](#)]
29. Pace, P.E. *Detecting and Classifying Low Probability of Intercept Radar*, 2nd ed.; Artech House Remote Sensing Library: Norwood, MA, USA, 2008.
30. Gao, L.; Liu, L.; Cao, Y.; Wang, S.; You, S. Performance analysis of one-step prediction-based cognitive jamming in jammer-radar countermeasure model. *J. Eng.-IET* **2019**, *21*, 7958–7961. [[CrossRef](#)]
31. Bachmann, D.J.; Evans, R.J.; Moran, B. Game theoretic analysis of adaptive radar jamming. *IEEE Trans. Aerosp. Electron. Syst.* **2011**, *47*, 1081–1100. [[CrossRef](#)]
32. Zhou, H.; Guo, L. Self-adaptive frequency agility realized with FPGA. In Proceedings of the International Conference on Image Analysis and Signal Processing, Taizhou, China, 11–12 April 2009; pp. 419–422. [[CrossRef](#)]
33. Talbot, K.I.; Duley, P.R.; Hyatt, M.H. Specific Emitter Identification and Verification. *Technol. Rev. J.* 2003, pp. 113–133. Available online: http://jmfriedt.org/phase_digital/03SS_KTalbot.pdf (accessed on 22 September 2021).
34. Anjaneyulu, L.; Sarma, N.V.; Murthy, N.S. Identification of LPI radar signals by higher order spectra and neural network techniques. *Int. J. Inf. Commun. Technol.* **2009**, *2*, 142–155. [[CrossRef](#)]
35. Kawalec, A.; Owczarek, R. Specific emitter identification using intrapulse data. In Proceedings of the First European Radar Conference, EURAD, Amsterdam, The Netherlands, 14–15 October 2004; pp. 249–252.
36. D’Agostino, S.; Foglia, G.; Pistoia, D. Specific Emitter Identification: Analysis on real radar signal data. In Proceedings of the European Radar Conference (EuRAD), Rome, Italy, 30 September–2 October 2009; pp. 242–245.
37. NEWEG-Electronic Warfare Signal Environment by Naval Air Systems Command (US Navy)—EW Simulation and Stimulation. Available online: <https://www.navair.navy.mil/nawctsd/sites/g/files/jejdrs596/files/2019-07/2016-neweg.pdf> (accessed on 5 November 2021).
38. Vankka, J. Digital frequency synthesizer/modulator for continuous-phase modulation with slow frequency hopping. *IEEE Trans. Veh. Technol.* **1997**, *46*, 933–940. [[CrossRef](#)]
39. Modarres-Hashemi, M.; Nayebi, M.M. LPD feature improvement in random PRF radar signals. *IEE Proc. Radar Sonar Navig.* **2004**, *151*, 225–230. [[CrossRef](#)]
40. De Martino, A. *Introduction to Modern EW Systems*, 2nd ed.; Artech House Inc.: Norwood, MA, USA, 2018.
41. Zhi, Z.M.; Li, H.A.; Huang, G. LPI Radar Waveform Recognition Based on Features from Multiple Images. *Sensors* **2020**, *20*, 526. [[CrossRef](#)]
42. Liu, G.; Gu, H.; Su, W. The development of random signal radar. *IEEE Trans. Aerosp. Electron. Syst.* **1999**, *35*, 770–777.
43. Liu, G.; Shi, X.; Lu, J.; Yang, G.; Song, Y. Design of noise FM CW radar and its implementation. *IEE Proc. Radar Sonar Navig.* **1991**, *138*, 420–426.
44. Hong, G.; Guosui, L.; Xiaohua, Z. The study of the random binary phase coded CW radar system. *Acta Electron. Sin.* **1995**, *23*, 71–74.
45. Wasserzler, C.; Worms, J.G.; O’Hagan, D.W. How Noise Radar Technology Brings Together Active Sensing and Modern Electronic Warfare Techniques in a Combined Sensor Concept. In Proceedings of the Sensor Signal Processing for Defence Conference, Brighton, UK, 9–10 May 2019. [[CrossRef](#)]

46. Matsumoto, M.; Nishimura, T. Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator. *ACM Trans. Model. Comput. Simul.* **1998**, *8*, 3–30. [CrossRef]
47. Vigna, S. It is high time we let go of the Mersenne Twister. *Comput. Sci. Data Struct. Algorithms* **2019**, arXiv:1910.06437. Available online: <https://arxiv.org/pdf/1910.06437> (accessed on 5 November 2021).
48. NIST Special Publication (SP) 800-90B. *Recommendation for the Entropy Sources Used for Random Bit Generation*; 100 Bureau Drive: Gaithersburg, MD, USA, 2018. [CrossRef]
49. Park, S.; Choi, B.G.; Kang, T.; Park, K.; Kwon, Y.; Kim, J. Efficient hardware implementation and analysis of true random-number generator based on beta source. *ETRI J. Spec. Issue SoC AI Process.* **2020**, *42*, 518–526. [CrossRef]
50. Park, K.; Park, S.; Choi, B.G.; Kang, T.; Kim, J.; Kim, Y.H.; Jin, H.Z. A lightweight true random number generator using beta radiation for IoT applications. *ETRI J.* **2020**, *42*, 951–964. [CrossRef]
51. Ferguson, N.; Schneier, B. *Practical Cryptography*; Wiley & Sons, Inc.: Hoboken, NJ, USA, 2003; ISBN 978-0-471-22357-3.
52. Gopala, P.; Lai, L.; El Gamal, H. On the Secrecy Capacity of Fading Channels. *IEEE Trans. Inf. Theory* **2008**, *54*, 4687–4698. [CrossRef]
53. Negi, R.; Goel, S. Guaranteeing secrecy using artificial noise. *IEEE Trans. Wirel. Commun.* **2008**, *7*, 2180–2189. [CrossRef]
54. Liang, Y.; Poor, V.; Shamai, S. Secure Communication Over Fading Channels. *IEEE Trans. Inf. Theory* **2008**, *54*, 2470–2492. [CrossRef]
55. Atzori, L.; Ferrari, G. Internet of Things: Technologies, Challenges and Impact; CNIT Technical Report-05; Texmat. 2020. ISBN Print Version 9788894982381, Digital Version 9788894982398. Available online: <https://www.texmat.it/collana-cnit.html> (accessed on 22 September 2021).
56. Suo, H.; Wan, J.; Zou, C.; Liu, J. Security in the Internet of Things: A Review. In Proceedings of the 2012 International Conference on Computer Science and Electronics Engineering, Hangzhou, China, 23–25 March 2012; pp. 648–651. [CrossRef]
57. Available online: <http://web.mit.edu/6.933/www/Fall2000/mode-s/sidelobe.html> (accessed on 5 November 2021).
58. Smoll, A.E. Radar Beacon System with Side Lobe Suppression. U.S. Patent 2,966,675, 23 October 1957.
59. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 2nd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2006.
60. Papoulis, A.; Pillai, S.U. *Probability, Random Variables and Stochastic Processes*, 4th ed.; McGraw-Hill: New York, NY, USA, 2002; Chapter 14.
61. Bell, M.R. Information Theory and Radar Waveform Design. *IEEE Trans. Inf. Theory* **1993**, *39*, 1578–1597. [CrossRef]
62. Levanon, N.; Mozeson, E. *Radar Signals*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2004.
63. Neeser, F.D.; Massey, J.L. Proper Complex Random Processes with Applications to Information Theory. *IEEE Trans. Inf. Theory* **1993**, *39*, 1293–1302. [CrossRef]
64. Picinbono, B.; Bondon, P. Second-Order Statistics of Complex Signals. *IEEE Trans. Signal Process.* **1997**, *45*, 411–420. [CrossRef]
65. Xiong, W.; Li, H.; Adali, T.; Li, Y.O.; Calhoun, V.D. On Entropy Rate for the Complex Domain and Its Application to i.i.d. Sampling. *IEEE Trans. Signal Process.* **2010**, *58*, 2409–2414. [CrossRef] [PubMed]
66. Dubnov, S. Generalization of Spectral Flatness Measure for Non-Gaussian Linear Processes. *IEEE Signal Process. Lett.* **2004**, *11*, 698–701. [CrossRef]
67. Tohidi, E.; Nazari Majd, M.; Bahadori, M.; Jariani, H.H.; Nayebi, M.M. Periodicity in Contrast with Sidelobe Suppression in Random Signal Radars. In Proceedings of the IEEE CIE International Conference on Radar, Chengdu, China, 24–27 October 2011; pp. 442–445.
68. De Palo, F.; Galati, G.; Pavan, G.; Wasserzier, C.; Savci, K. Introduction to Noise Radar and its Waveforms. *Sensors* **2020**, *20*, 5187. [CrossRef] [PubMed]
69. Schrödinger, E. *What is Life—The Physical Aspect of the Living Cell*; Cambridge University Press: Cambridge, UK, 1944.
70. Hyvarinen, A. New Approximations of Differential Entropy for Independent Component Analysis and Projection Pursuit. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 1998; pp. 273–279. Available online: <http://papers.nips.cc/paper/1408-new-approximations-of-differential-entropy-for-independent-component-analysis-and-projection-pursuit.pdf> (accessed on 5 November 2021).
71. Bellman, R.E. *Adaptive Control Processes*; Princeton University Press: Princeton, NJ, USA, 1961.