electronics

MDPI

Article

Covert Channel Based on Quasi-Orthogonal Coding

Krystian Grzesiak *, Zbigniew Piotrowski and Jan M. Kelner

Institute of Communications Systems, Faculty of Electronics, Military University of Technology, 00-908 Warsaw, Poland; zbigniew.piotrowski@wat.edu.pl (Z.P.); jan.kelner@wat.edu.pl (J.M.K.)
* Correspondence: krystian.grzesiak@wat.edu.pl; Tel.: +48-261-885-509

Abstract: The paper presents a method of creating a hidden channel using a signals' superposition. According to this idea, a transmitter simultaneously sends overt and covert (secret) signals, whereby the overt signal is a carrier for the covert one. Due to the need to ensure a low probability of detection for covert communication, the covert signal should have low power. This implies a number of problems relating to its correct reception. This is similar to non-orthogonal multiple access (NOMA) systems, where the collective signal is a superposition of signals with different powers dedicated to different users. In this case, the successive interference cancellation (SIC) process is used in the receiver for the separation of the component signals. SIC requires accurate channel estimation. Even a small channel estimation error causes a significant increase in bit error rate (BER), performance degradation, or connection loss for covert transmission. This is due to the residual signal, i.e., the remnant of the cover signal after an imperfect SIC operation. The paper proposes a method of transforming (i.e., encoding) the applied hidden signal in such a way that the residual signal in the receiver is quasi-orthogonal to the hidden signal. The proposed model is based on appropriate sorting and, compared to methods with fixed constellation points, provides the covert channel with a low BER while maintaining high protection against detection as measured by the Kolmogorov–Smirnov distance. The proposed solution was tested using the USRP-2920 software-defined radio platform.

Keywords: security; steganalysis; covert channel; steganography; undetectability

Citation: Grzesiak, K.; Piotrowski, Z.; Kelner, J.M. Covert Channel Based on Quasi-Orthogonal Coding. *Electronics* **2023**, *12*, 2249. https://doi.org/10.3390/electronics12102249

Academic Editor: Athanasios D. Panagopoulos

Received: 30 March 2023
Revised: 5 May 2023
Accepted: 11 May 2023
Published: 15 May 2023

1. Introduction

Wireless transmission, in its physical layer, is susceptible to all kinds of manipulation, which can be used to create covert channels [1]. It is currently assumed that any method of communication used to illegally transmit information, which violates the system security policy, is a covert channel. Steganography of the physical layer of a radio signal essentially boils down to subtle changes of the parameters of the modulated cover signal. Such parameters can be the carrier frequency and, in the case of an orthogonal frequency-division multiplexing (OFDM) signal, the training sequence of the cover signal [2]. The vast majority of physical layer steganography methods consist of manipulating the position of the constellation points of the in-phase (I) and quadrature (Q) components of the cover signal. Small changes in position correspond to transmitted classified (covert) information, and an uninformed receiver (with no knowledge or ability to ascertain additional transmission) treats them as (channel or hardware) noise.

Examples of such solutions are presented in the literature:

(a) In [3], quadrature amplitude modulation (QAM) covert information constellation points are distributed around the QAM cover constellation points, forming the so-called dirty constellation. Thus, an additional constellation is formed based on the cover constellation.
(b) The authors [4] propose a similar solution, with the dirty constellation being formed using phase drift [5]. The solution can be applied to phase-shift keying (PSK) and QAM modulation.

(c) Hiding the data by moving the constellation points by a given angle (right or left) is shown in [1]. In this case, binary PSK (BPSK) modulation was used for practical implementation.
(d) In [6], the use of 8 frequency-shift keying (8FSK) modulation to embed information in QAM constellation points is presented. In this case, artificial neural networks were used to extract hidden data.
(e) The authors [7] noted that PSK modulations do not use the channel fully in terms of Shannon capacity. Therefore, the so-called residual capacity can be used to hide information. In order to hide this emission from potential observers, pseudo-noise asymmetric shift keying (PN-ASK) modulation is proposed.
(f) An extension to the pseudo randomness element of the [7] method is presented in [8]. The solution, called SteaLTE or Stegano LTE, is a steganographic technique for transmitting hidden data over Long-Term Evolution (LTE) radio networks. The developed method is resistant to steganalysis.
(g) The approach described in [9] is based on the [3] method with elements of additional randomness, which involves an additional shift in the phase of the constellation points of the covert signal. In addition, the authors [9] propose using polar codes to reduce bit error rate (BER).
(h) The transmission of stealth information in the form of noise on a QAM basis is presented in [10]. In this case, the cover's signal is not used to carry information.

The aforementioned literature on the covert channel ignores the issue of channel estimation error [11] and, directly related to this, non-perfect (non-ideal) successive interference cancellation (SIC) [12–15]. The assumption of being able to easily separate the cover signal from the very low-power covert signal is fundamentally difficult to implement. Hence, in this paper, the authors propose an original and novel approach, which is to transmit the covert signal in such a way that it is quasi-orthogonal to the cover signal at the receiver. This is accomplished by sorting, that is, by appropriately ranking the IQ samples of the covert signal against the cover signal over time. In the receiver this results in mutual orthogonalization, thus, easier frequency separation of the signals. The method used can be used for both amplitude–phase and frequency modulations. In the paper, however, FSK modulation is indicated as the optimal solution in terms of transmission capabilities (calculated by transmission speed) as well as protection against steganalysis. The proposed approach and the considerations presented here are a continuation of the work presented in [16].

The work is organized as follows. In Chapter 2, the basics of creating a covert channel in a radio channel are presented. Chapter three describes how the proposed transceiver with sorting circuit works. The results of the computer simulations are included in Chapter 4. The rationale for using FSK modulation to create a covert channel is placed in Chapter 5. Chapter 6 contains the results of the tests conducted based on the universal software radio peripheral (USRP). A summary is included at the end.

2. Radio Physical Layer Steganography

2.1. Creating a Covert Channel

In mathematical terms, the process of creating a covert channel in the physical layer of a radio channel can be described as a superposition of a cover signal (cover) and a covert signal, which can be represented by the formula:

$$s(t) = \sqrt{P_1}x_1(t) + \sqrt{P_2}x_2(t) \quad (1)$$

where x_1 and x_2 are the cover signal and the covert signal, respectively, P ($P = P_1 + P_2$) is the transmitter power.

In order to reduce the probability of detecting a covert channel, the following conditions should be met:

(a) the covert signal power should be significantly less than the cover signal ($P_1 \gg P_2$).

(b) constellation points of the covert signal should have a pseudo-random (noise) characteristic.

The receiver input signal is represented by the following simplified formula:

$$y(t) = h\sum_{i=1}^{2}\sqrt{P_i}x_i(t) + w(t) \tag{2}$$

where h ($h \sim CN(0,1)$) is a channel gain and w is a Gaussian noise.

In an ideal case, when $w(t) = 0$ and the value of the parameter h is known in the receiver, the recovery process of the signal $x_2(t)$ would proceed as follows (ideal SIC):

$$x_2(t) = \frac{y(t) - \sqrt{P_1}x_1(t)}{\sqrt{P_2}} + w(t) \tag{3}$$

Transmission and reception of the cover and covert signal was presented in Figure 1.

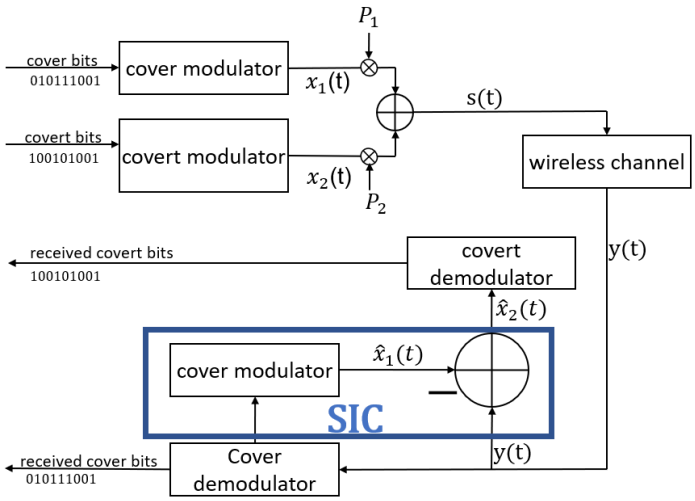

Figure 1. Transmission and reception of a covert signal as a superposition of cover and covert signals.

2.2. Channel Estimation Error

In a real-world situation, the receiver does not have full information about the wireless channel (2), but only estimates the parameter h with a certain error value ε ($\varepsilon \sim CN(0,\sigma_\varepsilon^2)$):

$$\hat{h} = h + \varepsilon, \tag{4}$$

where $\hat{h}$ is a channel gain estimation.

Taking into account the channel estimation error (4), the recovered covert signal $\hat{x}_2(t)$ is distorted by the residual cover signal according to the following formula (non-ideal SIC):

$$\hat{x}_2(t) = h\sqrt{P_2}x_2(t) + \varepsilon\sqrt{P_1}\hat{x}_1(t) + w(t) \tag{5}$$

The error introduced by the SIC propagates and affects the demodulation of the covert signal, i.e., the interference that arose during the recovery of the cover signal (stronger signal) propagates during the recovery of the covert signal (weaker signal) [17–19]. Errors arising during SIC mostly depend on the channel coefficients and the power allocation coefficient [20].

According to Formula (5), the channel estimation error variance value σ_ε^2 has a fundamental effect in the recovery of the covert signal. In this paper, we consider a situation where $\sigma_\varepsilon^2 \neq 0$ as a non-ideal SIC. In such a case, the best way to distinguish signals $x_1(t)$ from $x_2(t)$ is to ensure their mutual orthogonality or (if this is not possible) quasi-orthogonality.

2.3. Quasi-Orthogonality

We speak of orthogonality when the inner product of the $x_1(t)$ and $x_2(t)$ signals is zero, according to the formula:

$$\int_{-\infty}^{\infty} x_1(t) \cdot x_2^*(t)dt = 0 \tag{6}$$

The correlation of two time-limited signals defined over the time interval $0 < t < \tau$ is defined as:

$$r(t) = \int_{-\infty}^{\infty} x_1(\rho) \cdot x_2^*(\rho - t)d\rho, \tag{7}$$

where ρ is a dummy variable.

For two orthogonal signals for each t shift, a zero correlation is obtained if these signals are disjoint in frequency domain. Hence, separated band-limited signals in frequency are the main method of obtaining orthogonal signals. For signals that occupy the same frequency band, there is no possibility of zero cross-correlation for any t time shift.

Quasi-orthogonality [21] refers to signals (waveforms) that exhibit low cross-correlation. Two waveforms are quasi-orthogonal if:

$$\left| \int_{-\infty}^{\infty} x_1(\rho) \cdot x_2^*(\rho - t)d\rho \right| < \varepsilon \tag{8}$$

where $\varepsilon \ll 1$ for any t, and x_1 and x_2 are normalized to a unit energy value.

3. Proposed Model

3.1. Basic Assumptions

In the presented model, a single covert symbol is superimposed on several cover signal symbols. Such a solution stems from the need to reduce the distortion/interference of the cover signal (in order to preserve the energy per bit, variance of the covert signal is reduced), and thus the detection by outsiders of the fact that a covert channel exists. In the analyzed model, as in previous solutions [2,3,7,10], we assume that the cover is QAM amplitude–phase modulation in the form of IQ samples. In order to facilitate the process of demodulation of covert information from Figure 1, it is proposed to cross-orthogonalize the cover and covert signals by sorting. It is assumed that the coherence time of the radio channel, and therefore the channel estimation error, remains constant at least for a single data block for which the sorting operation is performed. Knowing that a typical channel coherence time is between approximately 10 ms to 200 ms [22–24], this assumption will usually remain.

Sorting occurs in both the transmitter and receiver as shown in Figure 2, whereby:

- At the transmitter, samples of the covert signal $x_2(t)$ are sorted based on a given sequence of the cover (QAM modulation) signal $x_1(t)$. In this way, the sorted signal x_{2_p} has a pseudo-random (noise) form.
- At the receiver, the $\hat{x}_2(t)$ signal re-sorting is performed after the SIC operation. Sorting in the receiver aims to:
  - (a) restore the original sample order of the covert signal after the SIC operation $\hat{x}_{2_p}$ to the original order (in an ideal case $\hat{x}_{2_i} = x_2$)
  - (b) restoring the original order of $\hat{x}_{2_p}$ is followed by a simultaneous change in the sample order of the residual signal associated with x_1 Thus, the residual signal becomes orthogonal (quasi-orthogonal) to the covert signal.
  - (c) The covert signal $\hat{x}_{2_i}$ is fed to the input of the covert channel demodulator

Sorting involves dividing covert information into blocks. The block length depends on the covert information modulation used, the value of the cover modulation and the number of IQ samples per signal. Every covert symbol consists of several IQ points imposed on the several cover symbols. In Section 4 are presented simulation results for blocks equalling 16

or 64 cover symbols with (imposed) 1 to 4 covert symbols, respectively. From the proposed principle of sorting, in order to correctly reproduce the covert signal, it is necessary to correctly (without error) reproduce the unclassified (cover) signal (the equation $\hat{x}_1 = x_1$ must be true) because based on the cover signal, the reverse operation of sorting in the transmitter is reproduced (in the covert samples reorder system). Any error in the reception of the cover data in the block results in an error in the covert signal. That means, for example, that if the block has a length of 16 cover symbols with one imposed covert symbol, we lose one covert symbol in case of any cover error. And similarly, for a larger number of covert symbols per block, we lose all covert symbols. Therefore, the block length and the number of covert symbols in the block should be taken carefully.

As shown in Figure 2, the proposed solution, compared with traditional SIC (Figure 1), is based on two additional sorting operations: one in the transmitter and one in the receiver. It can be assumed that the complexity of every sorting operation has complexity O(n^2).

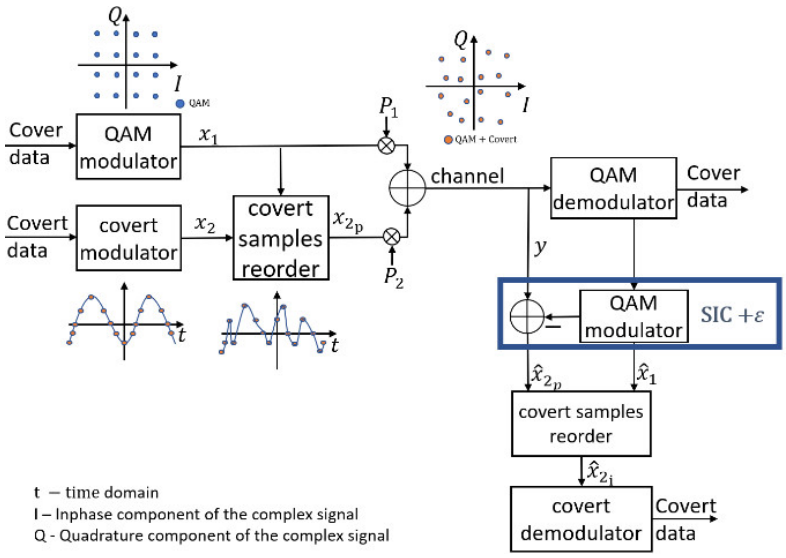

Figure 2. Covert signal transceiver system.

The selection of the optimal modulation of covert information was preceded by simulation tests included in the following chapters. The easiest way to explain orthogonalization (the creation of quasi-orthogonal signals) is to assume that the covert signal is 8FSK modulation. By orthogonalization we mean the mutual transformation of signals in such a way that their spectra are disjointed, i.e., they do not overlap. The 8FSK signal spectrum has eight peaks corresponding to each symbol. The averaged spectrum of the QAM signal is flat over the bandwidth except roll-off, but the instantaneous spectrum calculated for a packet/block of data is characterized by high variability. Hence the concept of using such sorting that will reduce the instantaneous spectrum variation for random IQ values representing QAM symbols. It is reasonable to assume that sorting can, by decoupling the instantaneous spectra of the cover and covert information, reduce the impact of channel estimation error during the recovery of covert information.

3.2. Cover Signal Sorting

In this section, it will be shown on a selected example that by sorting the IQ samples of the cover signal in an appropriate way, its spectral characteristics can be changed, so that the influence of the residual signal on the covert signal x_2 is reduced (Figure 2).

The impact of cover IQ sample sorting will be tested in the frequency domain using fast Fourier transform (FFT) analysis on the example of QAM amplitude–phase modulation. Sorting is performed for a sequence (block) of random IQ samples of signal x_1 consisting of $N = 64$ samples of 64 order QAM modulation ($M = 64$). The probability of each symbol

(from 0 to 63) is equally likely. The averaged FFT spectrum of such a signal does not have a clear main peak. We then sort the samples according to the phase increment defined as the angle $angle(x_1)$. The signal sorted in this way is denoted x_{1_sorted}. Since the phases of the QAM constellation points vary from $-\pi$ to π, a x_{1_sorted} signal close to a sine wave (Figure 3a) with a period equal to N samples will be obtained. The FFT spectrum of the signal thus formed has one strong spectral line for the fundamental frequency (f_0), the normalized value of which is $f_0 = 2/N$. The same will be true if N is a multiple of the modulation order M. By reordering the x_{1_sorted} samples, we can obtain shifted spectral lines on the frequency $f_n = nf_0$, where f_0 is a fundamental frequency, n is a power of 2 and satisfies the condition $-0.5 \leq f_n \leq 0.5$. The above actions can be written using matrix operations. If we have a set of sorted (according to phase increase) samples of the N-QAM x_{1_sorted} signal with indices [1:N], the i indices of the sorted signal with the fundamental frequency k·f_0 (k = 1, 2, 4 ... $\frac{N}{2}$) are obtained according to the formula:

$$i = reordered\_index\_for\ x_{1\_sorted} = \text{parallel\_to\_serial}\left\{\begin{bmatrix} 1:\text{k}:\text{N} \\ 2:\text{k}:\text{N} \\ \vdots \\ \text{k}:\text{k}:\text{N} \end{bmatrix}^{\text{T}}\right\} \tag{9}$$

For example, if we have a sorted (according to the phase increment) set of samples of a 64QAM signal labelled x_{1_sorted} with base frequency f_0 and corresponding indexes from 1 to 64, then the indices of the signal with frequency 16·f_0 and 32·f_0 (which corresponds to the normalized frequency equal to $^1/_4$ and $^1/_2$) are as follows:

$$i = \text{parallel\_to\_serial}\left\{\begin{bmatrix} 1,\ 17, 33, 49 \\ 2,\ 18,\ 34,\ 50 \\ 3,\ 19,\ 35,\ 51 \\ \vdots \\ 16,\ 32,\ 48,\ 64 \end{bmatrix}^{\text{T}}\right\} \Rightarrow \tag{10}$$

and for $k = 32$ (32·f_0)

$$\begin{gathered} i = \text{parallel\_to\_serial}\left\{\begin{bmatrix} 1,\ 33 \\ 2,\ 34 \\ \vdots \\ 32, 64 \end{bmatrix}^{\text{T}}\right\} = [1,\ 33,\ 2,\ 34,\ \ldots.,\ 32,\ 64] \\ \Rightarrow reordered\left(x_{1_{sorted}}\right) = x_{1_i} = \left[x_{1_1},\ x_{1_{33}}, x_{1_2},\ \ldots x_{1_{64}}\right] \end{gathered} \tag{11}$$

In Figure 3, for 64-QAM modulation, for a sequence of 64 random IQ samples, the time and frequency spectrum waveforms are shown for sequences sorted with fundamental frequencies of f_0, 16·f_0 and 32·f_0, respectively.

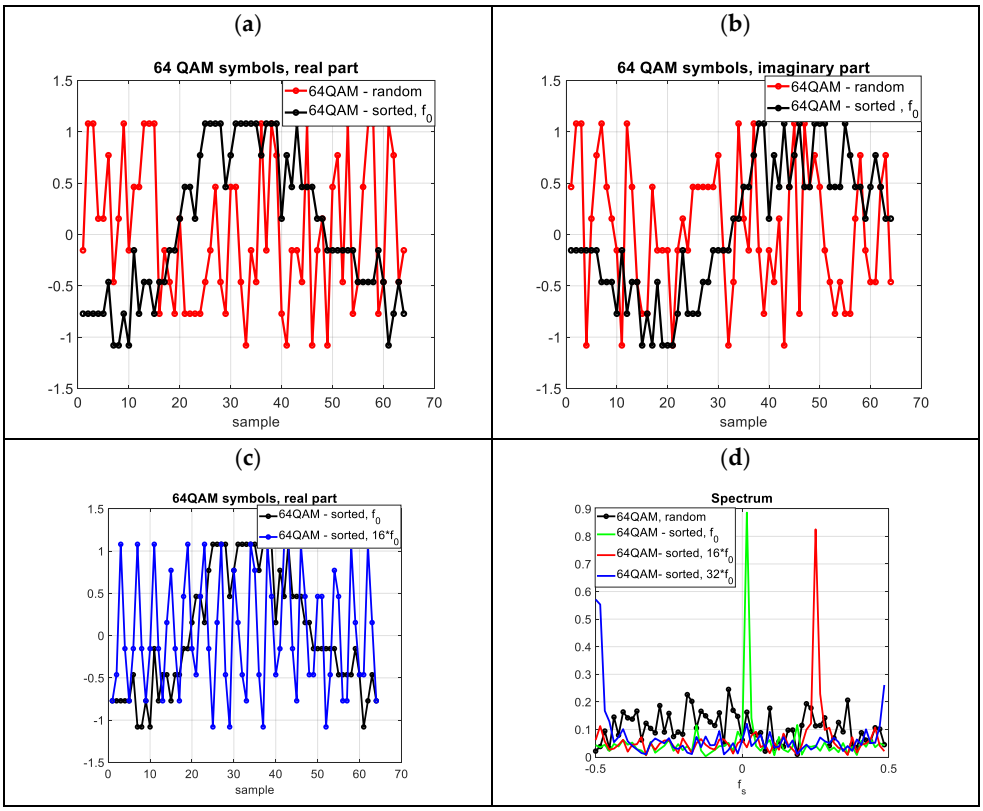

Figure 3. 64QAM signal. (**a**) Random and sorted (x_{1_sorted}, f_0) 64QAM signal-real part of the signal, (**b**) Random and sorted 64QAM signal-imaginary part of the signal (**c**) Sorted 64QAM signal with frequencies of f_0 and $16 \cdot f_0$ (d) Instantaneous FFT spectrum for 64 samples of random and sorted signals.

According to Figure 3, for 64 random 64QAM symbols we can, as a result of sorting, control (to a certain extent) the position of the spectral lines. This enables the parameters of the covert signal to be selected for optimal reception. On the basis of the presented method, the reception of the covert signal in the form of FSK and PSK modulation and the different modulation values of the cover and covert signal will be presented.

4. Simulation Tests

A preliminary evaluation of the feasibility of using the system in Figure 2 to receive a covert signal in the presence of a channel estimation error resulting in a non-perfect SIC was carried out in the MATLAB environment. The following graphs are shown in the figures to better illustrate the phenomena taking place:

(a) Averaged FFT spectrum of cover and covert signal (before and after sorting)
(b) Averaged value and variance of cross-correlation of signals
(c) Probability distribution of cross-correlation of signals estimated using histograms

In order to provide an understanding of the phenomena occurring during the sorting process, the signals were assumed to be unnoisy and of equal power to produce charts containing spectra and cross-correlations (Figure 5a–f). This corresponds to the situation when the $\hat{x}$ signal in Figure 2 consists of a cover and covert signal of equal strengths. This is to show how distinguishable the signals are in the time and frequency domain.

Detection capabilities Figure 5f were investigated in accordance with the diagram in Figure 1 (system without sorting) and Figure 2 (system with sorting), assuming that the signal-to-noise ratio (SNR) of the channel (calculated for the aggregate signal) is 45 dB, and the estimation error of channel ε has a variance σ_ε^2. It was assumed that the covert symbol is transmitted with l IQ samples that are submultiples of the number of cover symbols

for which the sorting operation is performed. The simulation was performed for 10,000 (for each value of σ_ε^2) executions of a random sequence of cover and covert data. It was assumed that the value $\sqrt{P_1} = 1$, $\sqrt{P_2} = 0.005$.

In the case of frequency modulation, it was assumed that the M-FSK covert signal [25–29] in the baseband is defined as:

$$g(t) = \exp(i \cdot \pi \cdot k \cdot \Delta f \cdot t + \theta) \tag{12}$$

$k = \pm 1, \pm 3, \ldots \pm M/2$, and Δf is a frequency deviation, θ random initial phase (random for each symbol).

4.1. Simulation No. 1

The elementary signal from the quadrature transmitter is a composite (superposition) of one symbol of 8FSK (each symbol as 16 samples) and 16 symbols (samples) of 16QAM (Figure 4). The covert signal was sorted in such a way that, at the receiver, the cover fundamental frequency f_0 was increased four times (the normalized frequency of the cover is $4 \cdot f_0 = 0.25$).

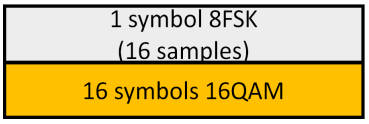

Figure 4. Block of samples subject to sorting.

Figure 4 shows the impact of sorting on the cover spectrum and the cross-correlation between the cover and the covert signal (8FSK) (cross-correlation between two signals). As a result of sorting, the value of the averaged cross-correlation for the 8FSK modulation symbol equal to "5" has increased (Figure 5c), but its variation is eight times smaller (Figure 5e). For the other symbols, the cross-correlation is lower, as is its variance. As a result, the SIC operation runs with a lower probability of BER error.

(**a**) (**b**)

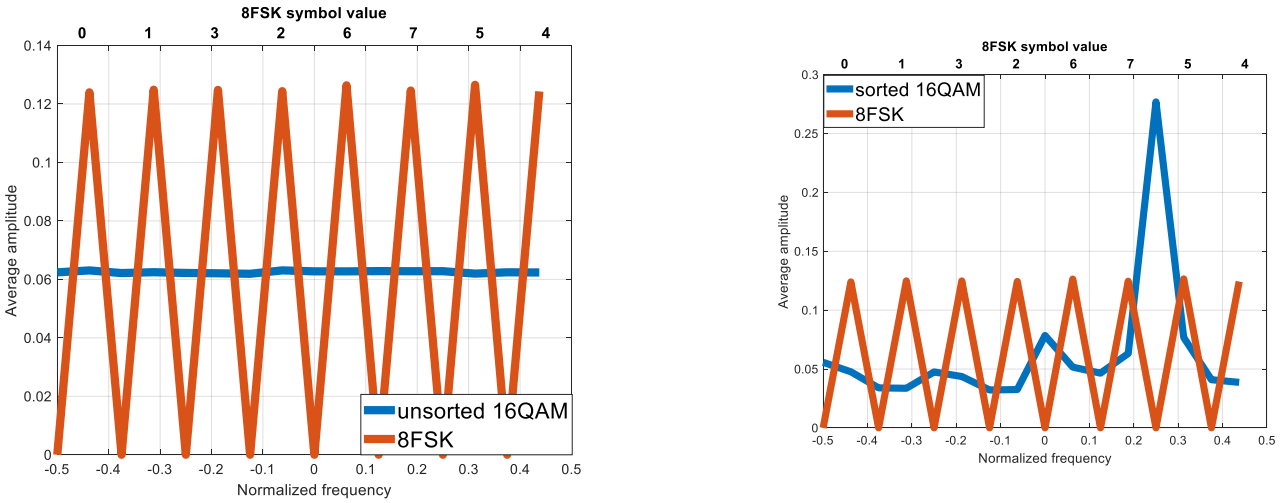

Figure 5. *Cont.*

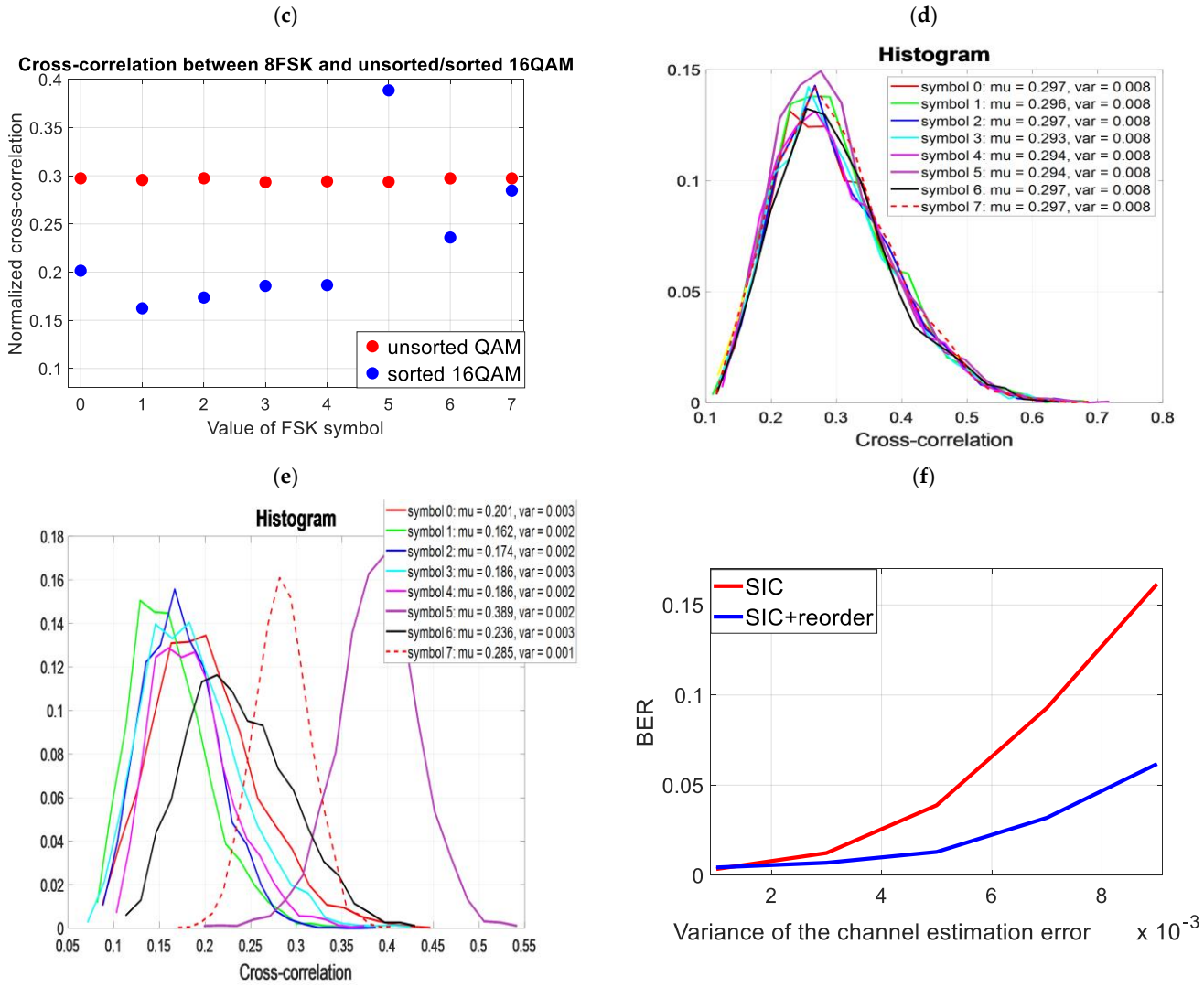

Figure 5. Testing the impact of cover orthogonalization on the non-ideal SIC process (**a**) Averaged FFT spectrum of unsorted cover and 8FSK (**b**) Averaged FFT spectrum of sorted cover and 8FSK (**c**) Averaged cross-correlation value of unsorted cover and 8FSK (**d**) Histogram of the cross-correlation value of the unsorted cover and 8FSK (**e**) Histogram of the cross-correlation value of the sorted cover and 8FSK (**f**) Reception of covert information (8FSK) for the unsorted and sorted cover (non-ideal SIC).

4.2. Simulation No. 2

In order to reduce cross-correlations (from Simulation No. 1), the sorting was changed (the number of covert samples for the block remains the same as in Figure 5). The spectrum in Figure 6b) was obtained by means of two successive repetitions of sorting (according to Formulas (13) and (14)) of the originally sorted x_{1_sorted}.

$$i_1 = \text{paralel\_to\_serial}\left\{\begin{bmatrix}1:2:\mathrm{N}\\ 2:2:\mathrm{N}\end{bmatrix}^{\mathrm{T}}\right\} \tag{13}$$

$$i_2 = f(i_1) = \left[1:\mathrm{k}:\frac{\mathrm{N}}{2},\ \mathrm{N}:-1:\frac{\mathrm{N}}{2}+1\right] \tag{14}$$

As a result of the sorting, the cross-correlation for the symbol "7" is the same before and after the orthogonalization process. Nevertheless, the variance of the cross-correlation is twice as small. Hence, a much lower BER was obtained for the sorted cover than in the previous case (Figure 5f) vs. (Figure 6f).

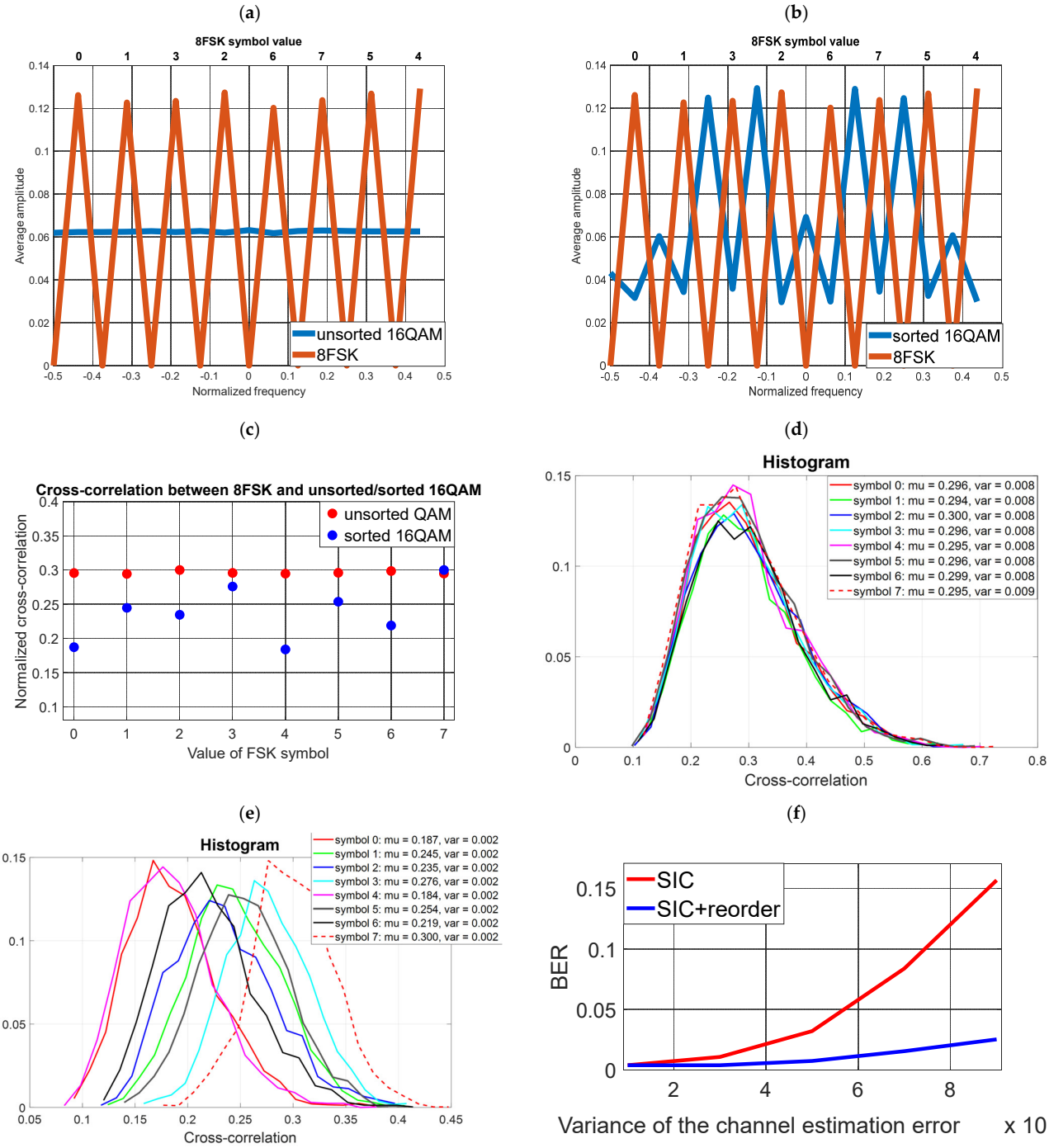

Figure 6. Testing the impact of cover orthogonalization on the non-ideal SIC process (**a**) Averaged FFT spectrum of unsorted cover and 8FSK (**b**) Averaged FFT spectrum of sorted cover and 8FSK (**c**) Averaged cross-correlation value of unsorted cover and 8FSK (**d**) Histogram of the cross-correlation value of the unsorted cover and 8FSK (**e**) Histogram of the cross-correlation value of the sorted cover and 8FSK (**f**) Reception of covert information (8FSK) for the unsorted and sorted cover (non-ideal SIC).

4.3. Simulation No. 3

The cover is 64QAM. For every 64 samples of the cover signal, there are 4 symbols (16 samples each) of the covert signal. To correctly decode the covert data, the sorting operation must be performed sequentially for each block of data shown in Figure 7. The data in the transmitter and receiver are sorted according to the cover signal. The normalized frequency of the sorted cover is $4 \cdot f_0 = 0.5$. The results are presented in Figure 8.

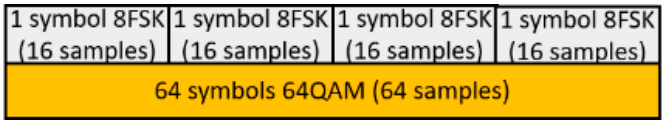

Figure 7. Block of samples subject to sorting.

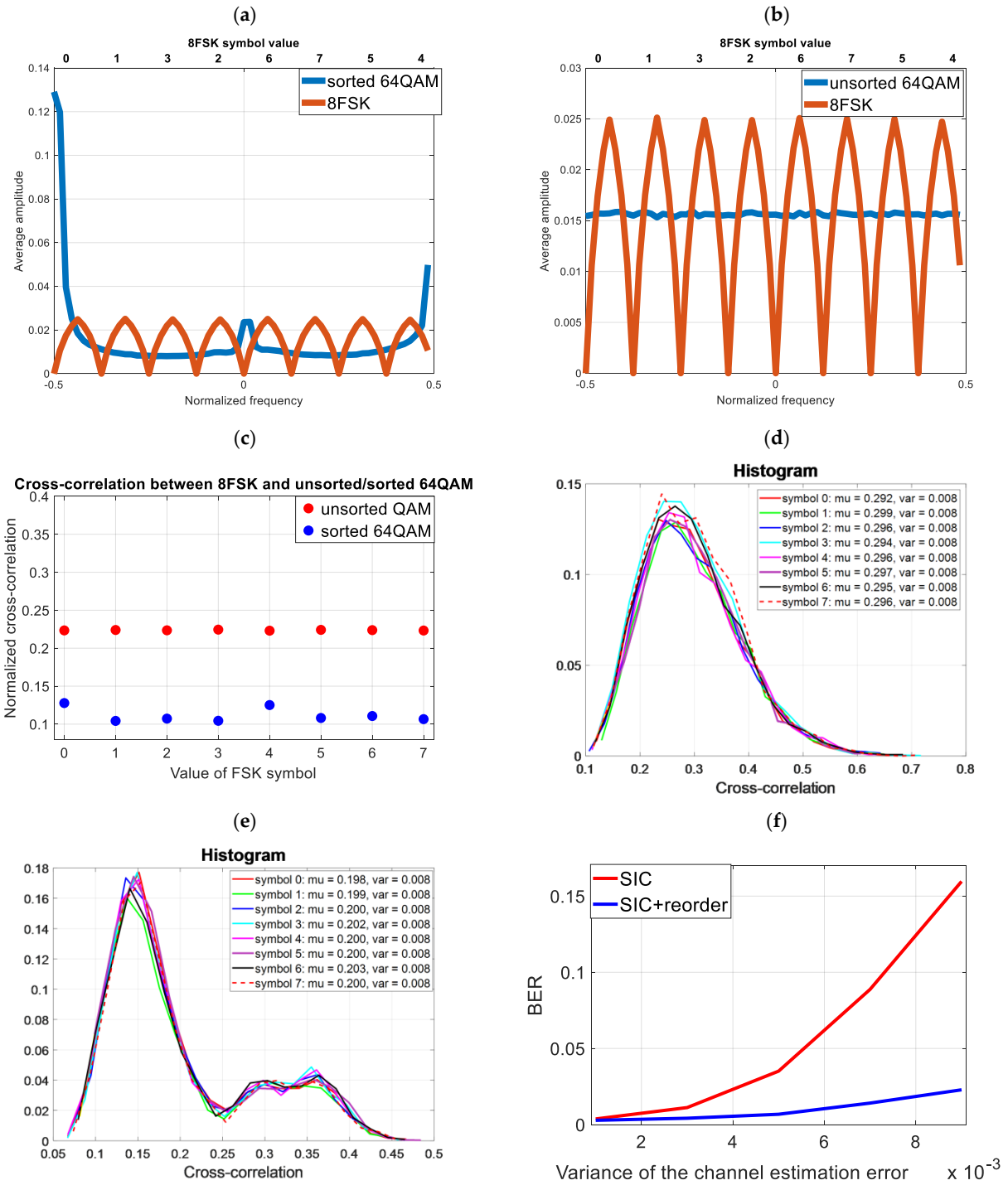

Figure 8. Testing the impact of cover orthogonalization on the non-ideal SIC process (**a**) Averaged FFT spectrum of unsorted cover and 8FSK (**b**) Averaged FFT spectrum of sorted cover and 8FSK (**c**) Averaged cross-correlation value of unsorted cover and 8FSK (**d**) Histogram of the cross-correlation value of the unsorted cover and 8FSK (**e**) Histogram of the cross-correlation value of the sorted cover and 8FSK (**f**) Reception of covert information (8FSK) for the unsorted and sorted cover (non-ideal SIC).

As a result of the sorting, the cross-correlation of the cover and the covert signal was reduced. This has reduced the BER for a non-perfect SIC. The results are compared with Figure 6f).

4.4. Simulation No. 4

In previous simulations (simulation 1 to 3), the modulation of the covert information was 8FSK. However, it should be examined what effect sorting has when the covert signal is the amplitude–phase modulation used in [1,3,7,10]. Let's assume, as in simulations 1 and 2, that one 2PSK covert symbol is transmitted using 16 16QAM cover symbols. The FFT analysis (Figure 9a) of such a signal will show that each covert symbol is a constant value (on the frequency scale it has a non-zero value only for $f = 0$). Orthogonalization will not provide any benefit (Figure 9f) because the block for which we perform orthogonalization is equal to the modulation value of the cover and, at the same time, the number of samples per covert symbol. This is because the average value of the random sorted and unsorted cover in such a case remains constant.

Note that we obtained a relatively low BER for both the sorted and unsorted signal, even for a large estimation error. This is due to the high energy per bit (16 samples represent one bit of data) and, unfortunately, this comes at the expense of reduced resistance to steganalysis (as will be demonstrated in Section 5).

4.5. Simulation No. 5

Simulation conditions are the same as in the previous example, except that we increase the value of modulation of the covert information to 8PSK. The results are presented in Figure 10.

Multi-valued 8PSK modulation requires a higher ratio of energy per bit of information. For this reason (with an assumed SNR = 45 dB), even for zero estimation error, the BER is different from zero (Figure 10f) and there is less immunity to channel estimation errors.

4.6. Simulation No. 6

Sorting was carried out for the parameters as for case no. 3 (sorting a block of data equal to 64, and a covert symbol with a length of 16 samples). The modulation for covert data is 4PSK. The results are presented in Figure 11. The benefits of the sorting are noticeable, although the transmission rate compared with 8FSK is twice smaller.

4.7. Simulation No. 7

Sorting was carried out for the parameters as for case no. 6 (sorting a block of data equal to 64, and a covert symbol with a length of 16 samples). The modulation for covert data is 8PSK. The results are presented in Figure 12.

As expected, sorting yields a lower BER. However, comparing Figure 7 with Figure 12 graphs, it is clear that for the given bit rates and power levels of the covert signal, better results are obtained (regardless of the sorting process) for 8FSK modulation. 8PSK modulation relative to 8FSK requires more energy per bit.

All simulations presented above aimed to show that, for the given waveform of the covert channel, it is possible to find an optimal sorting pattern to minimize imperfect SIC operation in the covert signal demodulator. This seems to be easier for FSK modulation and longer frames. However, it is necessary to keep in mind that the longer the frame, the lower the probability that channel gain is constant, which is the main assumption of this method.

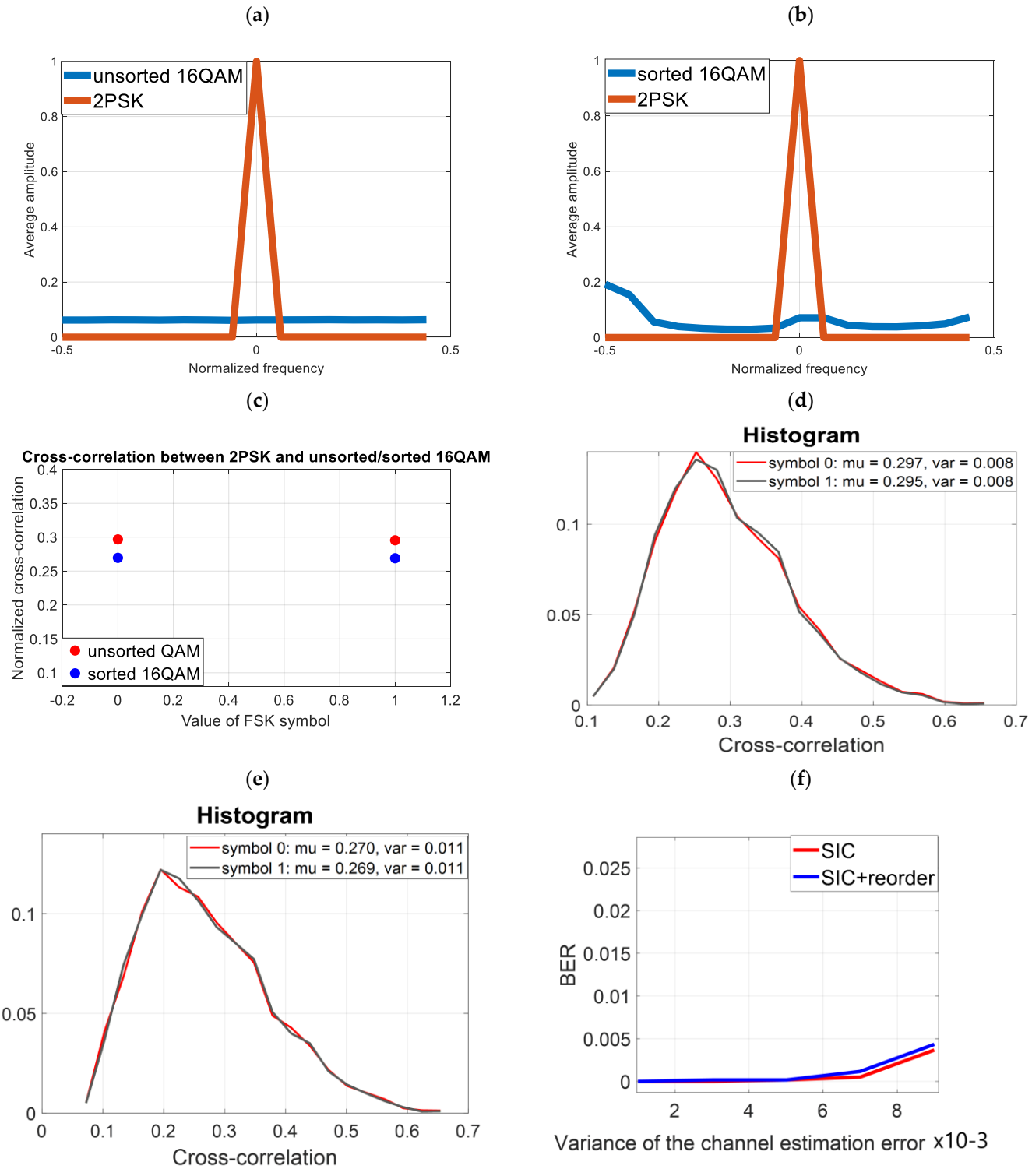

Figure 9. Use of 2-PSK modulation for covert transmission (**a**) Averaged FFT spectrum of unsorted cover and 2PSK (**b**) Averaged FFT spectrum of sorted cover and 2PSK (**c**) Averaged cross-correlation value of unsorted cover and 2PSK (**d**) Histogram of the cross-correlation value of the unsorted cover and 2PSK (**e**) Histogram of the cross-correlation value of the sorted cover and 2PSK (**f**) Reception of covert information (2PSK) for the unsorted and sorted cover (non-ideal SIC).

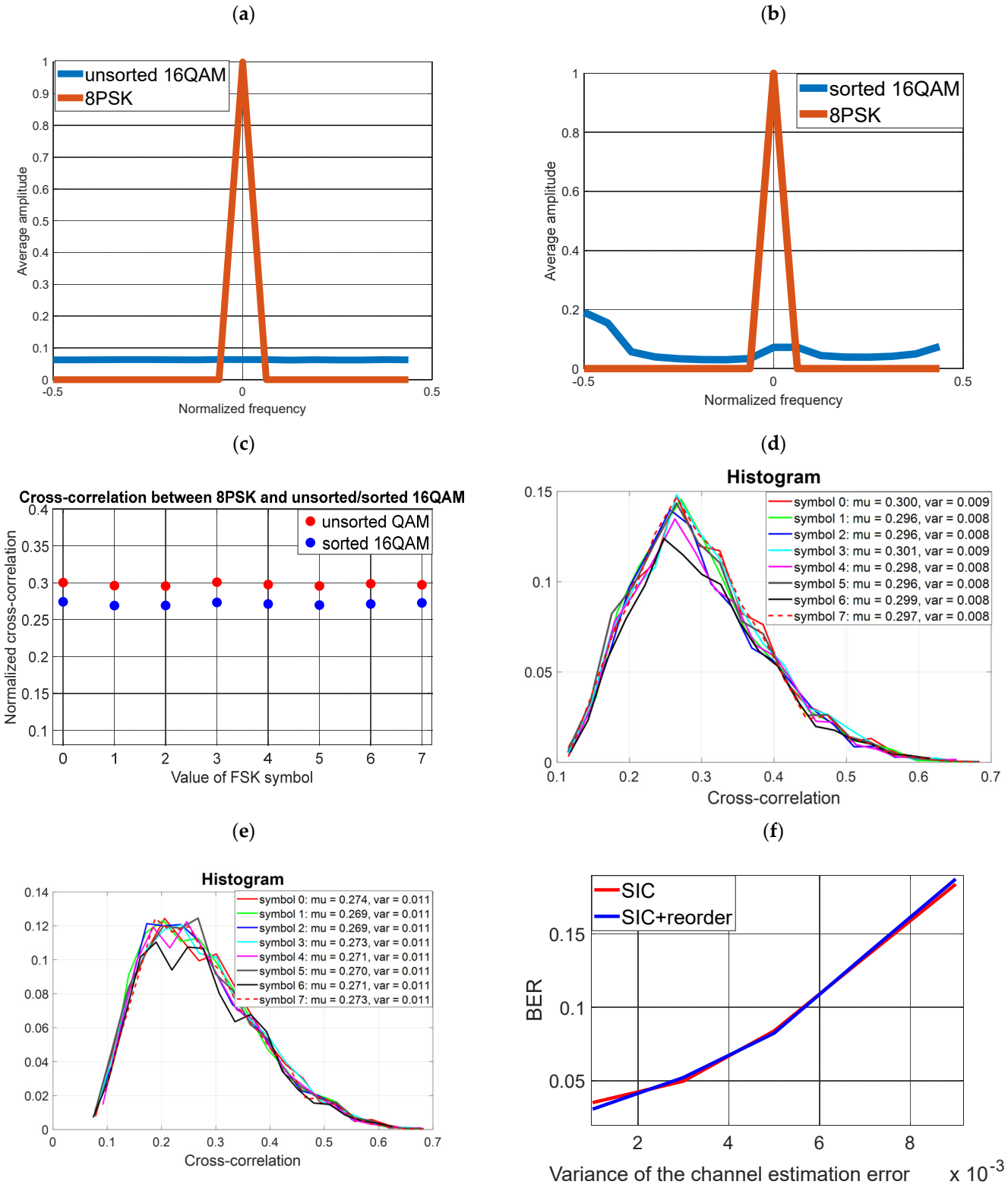

Figure 10. Use of 8-PSK modulation for covert transmission (**a**) Averaged FFT spectrum of unsorted cover and 8PSK (**b**) Averaged FFT spectrum of sorted cover and 8PSK (**c**) Averaged cross-correlation value of unsorted cover and 8PSK (**d**) Histogram of the cross-correlation value of the unsorted cover and 8PSK (**e**) Histogram of the cross-correlation value of the sorted cover and 8PSK (**f**) Reception of covert information (8PSK) for the unsorted and sorted cover (non-ideal SIC).

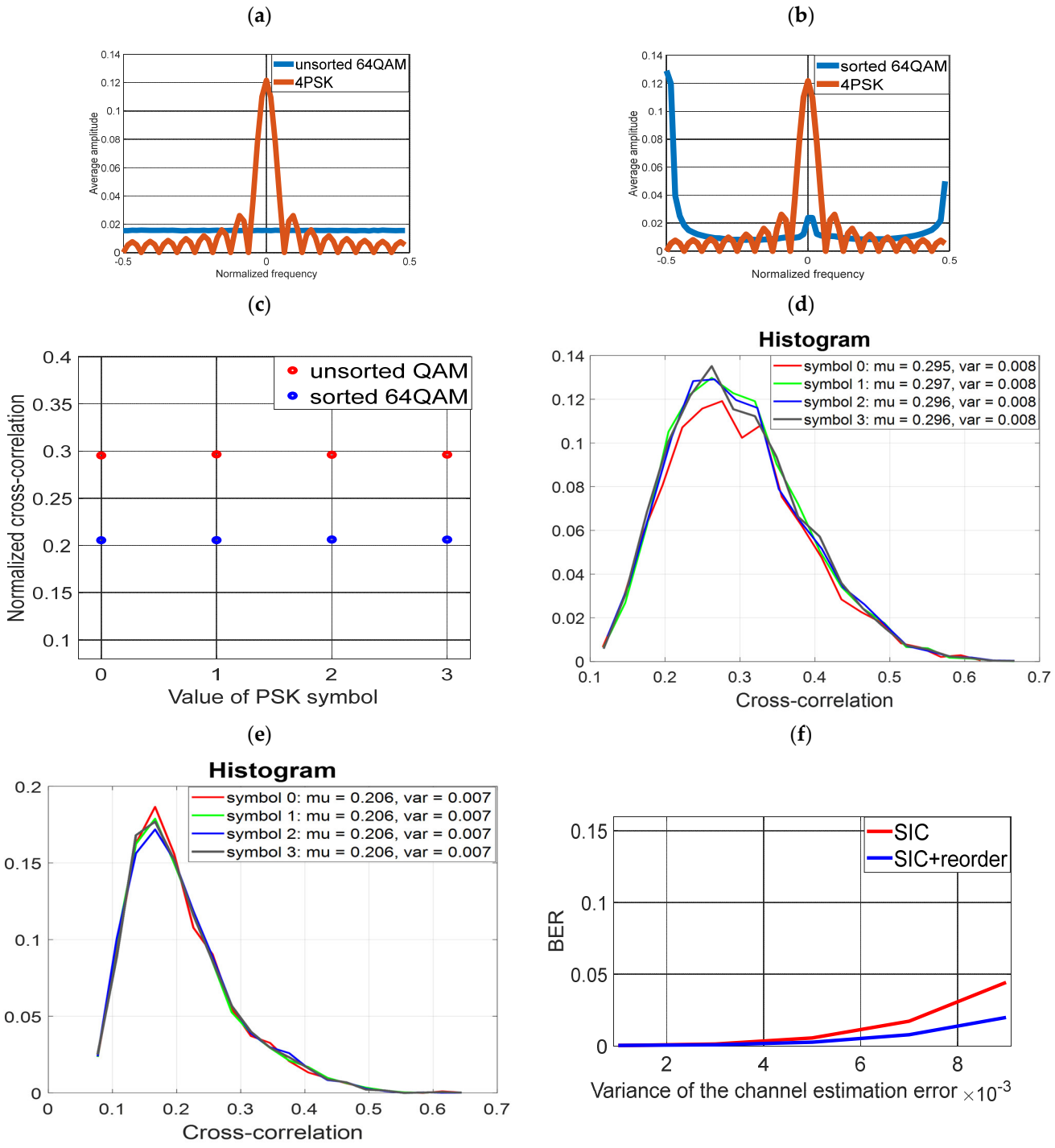

Figure 11. Use of 4-PSK modulation for covert transmission (**a**) Averaged FFT spectrum of unsorted cover and 8PSK (**b**) Averaged FFT spectrum of sorted cover and 4PSK (**c**) Averaged cross-correlation value of unsorted cover and 4PSK (**d**) Histogram of the cross-correlation value of the unsorted cover and 4PSK (**e**) Histogram of the cross-correlation value of the sorted cover and 4PSK (**f**) Reception of covert information (4PSK) for the unsorted and sorted cover (non-ideal SIC).

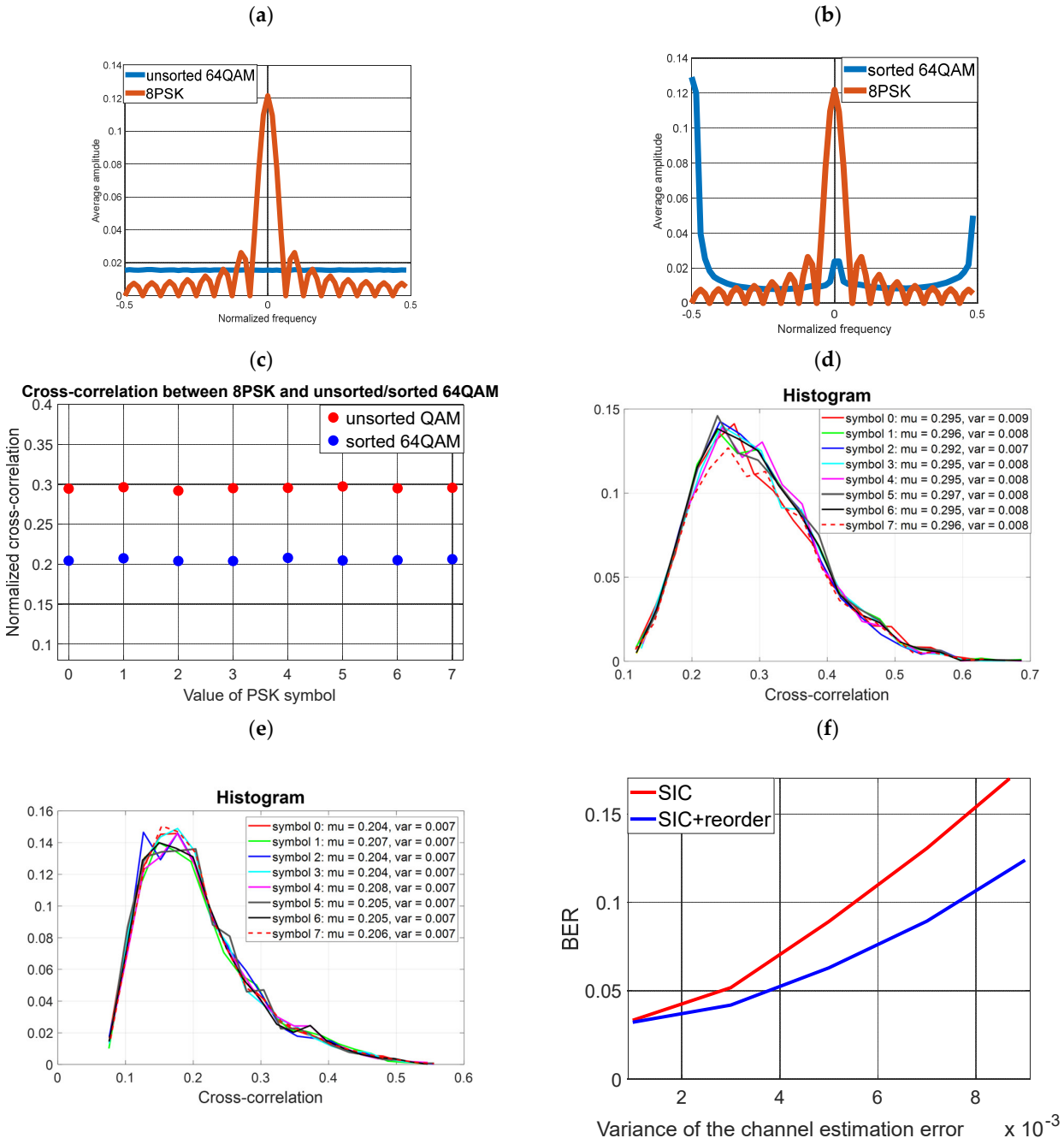

Figure 12. Use of 8-PSK modulation for covert transmission (**a**) Averaged FFT spectrum of unsorted cover and 8PSK (**b**) Averaged FFT spectrum of sorted cover and 8PSK (**c**) Averaged cross-correlation value of unsorted cover and 8PSK (**d**) Histogram of the cross-correlation value of the unsorted cover and 8PSK (**e**) Histogram of the cross-correlation value of the sorted cover and 8PSK (**f**) Reception of covert information (8PSK) for the unsorted and sorted cover (non-ideal SIC).

5. Steganographic Analysis

The choice of FSK modulation as the modulation for covert information is not only due to its good transmission properties and easy orthogonalization (quasi-orthogonalization) process with respect to the cover. The use of FSK modulation provides better properties in terms of low probability of detection (LPD), which is due to increased immunity to steganographic analysis compared to amplitude–phase modulations with constant constellation points.

By steganographic analysis we mean testing of probability density distributions and cumulative distribution function estimated by means of a histogram and cumulative histogram. Quantitatively, a measure of the difference in distributions can be calculated using the Kolmogorov–Smirnov test. To do this, the receiver must have noise information in the radio channel [30–32] and statistics formed from the signal from the SIC system output (we assume that the receiver is able to demodulate the cover information). If we denote the cumulative histogram distribution of the noise and signal after performing the SIC operation by F_w and $F_{\hat{x}_2}$, respectively, the Kolmogorov–Smirnov distance $KSTEST$ is expressed by the formula [10,33]:

$$KSTEST = max\left|F_w - F_{\hat{x}_2}\right| \tag{15}$$

Results of $KSTEST$ calculated on the basis of 200,000 IQ samples for $SNR = 45$ dB and 50 dB conditions relative to the cover in the form of 64QAM, $\sqrt{P_1} = 1$, $\sqrt{P_2} = 0.005$ and zero channel estimation error are shown in the Table 1.

Table 1. KSTEST calculation.

Covert Modulation	*KSTEST*	
	***SNR* = 45 dB**	***SNR* = 50 dB**
2PSK	0.153	0.321
4PSK	0.087	0.214
8PSK	0.082	0.188
2FSK	0.081	0.186
4FSK	0.081	0.186
8FSK	0.080	0.185

Example histograms and cumulative histograms for 2PSK and 2FSK modulations are presented in Figure 13. It was assumed that $SNR = 45$ dB.

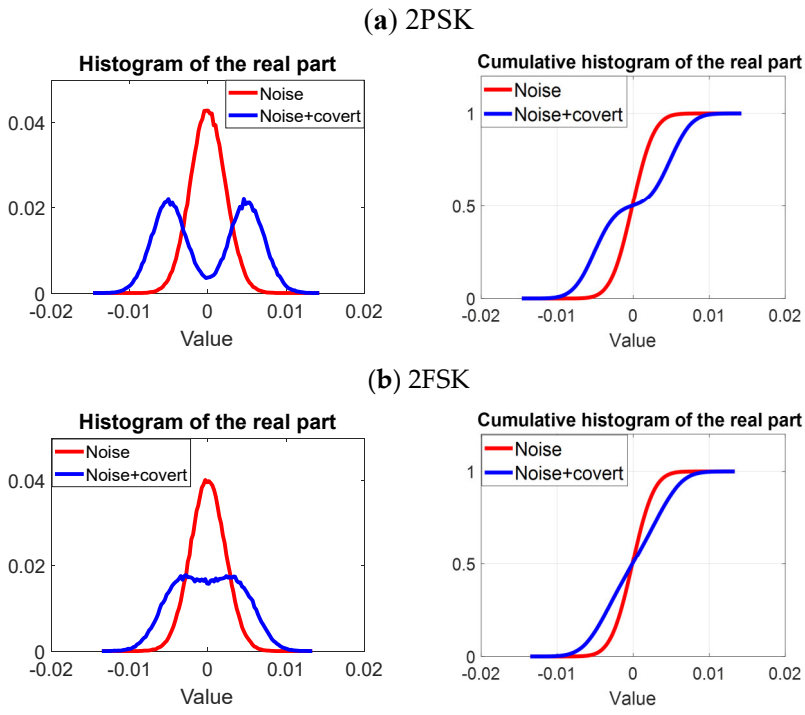

Figure 13. Histograms and cumulative histograms for (**a**) 2PSK and (**b**) 2FSK signals.

The analysis indicates that the 8FSK signal for the additive white Gaussian noise (AWGN) channel and ideal SIC provides the highest bit rate for the covert signal, while providing the best (calculated in K-S distance) steganographic protection. In addition, simulation tests have proven (Section 4) that sorting can be successfully used, which effectively reduces the impact of channel estimation error.

6. Practical Implementation

The concept of a covert channel based on quasi-orthogonal coding was implemented by using the USRP-2920 [34] hardware platform manufactured by National Instruments. USRP is the essential hardware part for generating a radio signal, while the software part is provided by the LabView software (with Matlab scripts) installed on a personal computer (PC). An Ethernet network adapter with a bit rate of 1 Gb/s is used to provide communication between USRP and the PC. Two USRP-2920 were used to implement a test stand (Figure 14) for detectors (in the transmitter–receiver system) connected with the computer by an unmanaged switch. The system was placed in an office room, and the distance between the transmitter and receiver was 5 m. The line-of-sight (LOS) propagation conditions were ensured disturbance only by office equipment such as PCs and monitors.

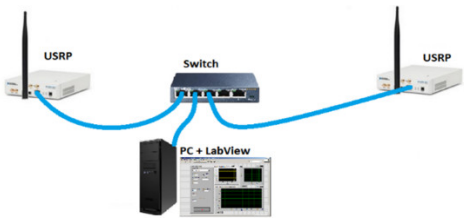

Figure 14. Test system used for examinations.

The data were preceded by a short and long training sequence (Figure 15) as defined in [35]. On this basis, transmission channel parameters were estimated and synchronization and frequency and phase correction were made. For performance analysis, there was no channel encoding during the signal transmission. In order to compare the results obtained, tests were performed for the case of transmission with and without sorting. Sorting was done as in simulation #2 in Section 4. The results obtained are shown in Figure 16. Estimated SNR value refers to cover signal. Cover detection are intended to show that a certain minimum SNR for the cover channel is required to receive the covert channel. The cover signal has to be detected correctly first and then the covert signal can be received. During the test, a low power covert signal was selected deliberately. First, the authors intended to make the signal as difficult to detect as possible, and second, to obtain conditions under which it is more sensitive to channel estimation errors. The test verified the previously assumed and simulation-validated thesis that sorting aimed at mutual quasi-orthogonality of signals can improve the bit error rate. The difference for a signal with sorting versus without sorting for parameters defined in Table 2 is about 3 dB. The proposed algorithm effectively reduces channel estimation error and improves SIC operation. Significant gain was achieved, although the channel parameters were estimated every 64 blocks, which should give small channel estimation error. Improved CSI (lower error) could be achieved by, for example, additional pilot signals and training sequences, however, this would come at a cost of system resources and maximal bit rate.

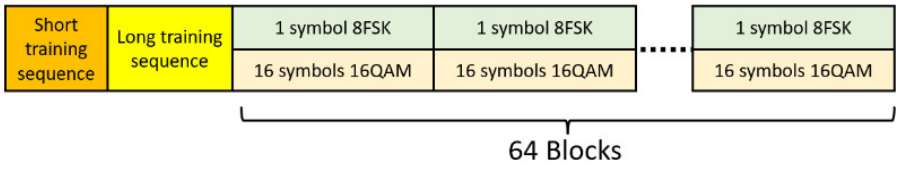

Figure 15. Structure of the transmitted signal.

Table 2. Parameters of radio signal.

Carrier Frequency	**850 MHz**	
Cover (carrier)	Modulation	16QAM
	Bandwidth	8 MHz
	Transmission rate	32 Mb/s
	Block length	16
	Power	P_0
Covert information	Modulation	8FSK
	Number samples per symbol	16
	Number symbols in block	1
	Transmission rate	1.5 Mb/s
	Power	$0.01 \cdot P_0$

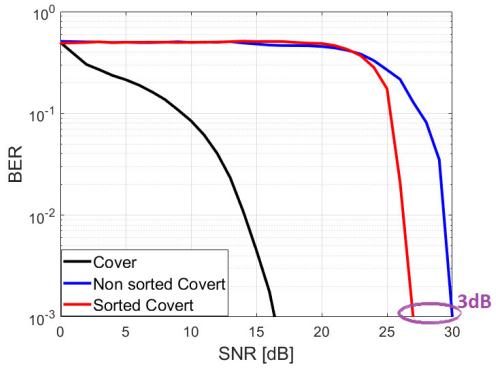

Figure 16. BER versus SNR for covert signal transmission with and without sorting.

7. Summary

Creating a covert channel in the physical layer of wireless communications is an issue that is difficult to implement in practice. The low probability of detecting such a channel and the need to affect the cover's signal as little as possible entails the low power that can be allocated to covert transmission. The natural solution in such a situation is to increase energy per bit by increasing its duration while accepting a lower transmission speed. Such a solution encounters a serious problem, arising from the estimation of channel parameters, which becomes apparent in the inability to extract the covert signal. The purpose of the article was to identify solutions to this type of problem. First, it was noted that higher transmission speed can be achieved by using FSK modulation, which does not require an increase in signal power if covert modulation order is increased, since the energy per bit remains constant, and this is done at the expense of signal bandwidth. However, as noted, FSK modulation is more difficult to receive in non-perfect SIC compared to low-value PSK modulation. The solution in such a case may be the use of sorting, which aims to more easily extract the signal through greater separability of signals in the frequency domain. Although, at the transmitter, the primary FSK signal is converted to a pseudo-noise sequence, the final reception is performed by a traditional FSK demodulator. The proposed solution for creating quasi-orthogonal signals can also be applied to other modulations, which was also simulated in this paper. Importantly, the sorting method is determined by the cover signal, hence there is no need to send additional information between the transmitter and receiver. Although a correct decoding of a block of cover data is required to receive a single or several covert symbols, this is not a major limitation, since a cover signal is a strong signal by its very definition.

Author Contributions: Conceptualization, K.G.; Methodology, Z.P. and J.M.K.; Validation, K.G., Z.P. and J.M.K.; Formal analysis, K.G.; Investigation, Z.P.; Writing-original draft, K.G. and J.M.K.; Writing-review & editing, Z.P. and J.M.K.; Visualization, K.G.; Funding acquisition, J.M.K. 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.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors would like to express their great appreciation to the sensors journal editors and anonymous reviewers for their valuable suggestions, which have improved the manuscript quality.

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

References

1. Chen, O.; Meadows, C.; Trivedi, G. Stealthy Protocols: Metrics and Open Problems. In *Concurrency, Security, and Puzzles. Lecture Notes in Computer Science*; Gibson-Robinson, T., Hopcroft, P., Lazić, R., Eds.; Springer: Cham, Switzerland, 2017; Volume 10160. [CrossRef]
2. Classen, J.; Schulz, M.; Hollick, M. Practical covert channels for wifi systems. In Proceedings of the 2015 IEEE Conference on Communications and Network Security (CNS), Florence, Italy, 28–30 September 2015; pp. 209–217.
3. Dutta, A.; Saha, D.; Grunwald, D.; Sicker, D. Secret Agent Radio: Covert Communication through Dirty Constellations. In *Information Hiding. IH 2012. Lecture Notes in Computer Science*; Kirchner, M., Ghosal, D., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; Volume 7692. [CrossRef]
4. Grzesiak, K.; Piotrowski, Z.; Kelner, J.M. A Wireless Covert Channel Based on Dirty Constellation with Phase Drift. *Electronics* **2021**, *10*, 647. [CrossRef]
5. Piotrowski, Z. Drift Correction Modulation Scheme for Digital Signal Processing. *Math. Comput. Model.* **2013**, *57*, 2660–2670. [CrossRef]
6. Grzesiak, K.; Piotrowski, Z. NN-Based 8FSK Demodulator for the Covert Channel. *Sensors* **2022**, *22*, 7181. [CrossRef] [PubMed]
7. D'Oro, S.; Restuccia, F.; Melodia, T. Hiding Data in Plain Sight: Undetectable Wireless Communications Through Pseudo-Noise Asymmetric Shift Keying. In Proceedings of the IEEE INFOCOM 2019–IEEE Conference on Computer Communications, Paris, France, 29 April–2 May 2019; pp. 1585–1593.
8. Bonati, L.; D'Oro, S.; Restuccia, F.; Basagni, S.; Melodia, T. SteaLTE: Private 5G cellular connectivity as a service with full- stack wireless steganography. In Proceedings of the IEEE INFOCOM, Vancouver, BC, Canada, 10–13 May 2021.
9. Qiao, S.; Liu, G.; Shi, J.; Ji, X.; Liu, W. Wireless Covert Channel with Polarized Dirty Constellation in Backscatter Communication. *Res. Square* **2021**. [CrossRef]
10. Cao, P.; Liu, W.; Liu, G.; Ji, X.; Zhai, J.; Dai, Y. A Wireless Covert Channel Based on Constellation Shaping Modulation. *Secur. Commun. Netw.* **2018**, *2018*, 1214681. [CrossRef]
11. Oyerinde, O.O.; Mneney, S.H. Review of channel estimation for wireless communication systems. *IETE Tech. Rev.* **2012**, *29*, 282–298. [CrossRef]
12. Yue, X.; Liu, Y.; Kang, S.; Nallanathan, A.; Chen, Y. Modeling and Analysis of Two-Way Relay Non-Orthogonal Multiple Access Systems. *IEEE Trans. Commun.* **2018**, *66*, 3784–3796. [CrossRef]
13. Do, D.T.; Nguyen, T.T.T. Impacts of imperfect SIC and imperfect hardware in performance analysis on AF non-orthogonal multiple access network. *Telecommun. Syst.* **2019**, *72*, 579–593. [CrossRef]
14. Yue, X.; Qin, Z.; Liu, Y.; Dai, X.; Chen, Y. Outage Performance of a Unified Non-Orthogonal Multiple Access Framework. In Proceedings of the 2018 IEEE International Conference on Communications (ICC), Kansas City, MO, USA, 20–24 May 2018; pp. 1–6. [CrossRef]
15. Kara, F.; Kaya, H. BER performances of downlink and uplink NOMA in the presence of SIC errors over fading channels. *IET Commun.* **2018**, *12*, 1834–1844. [CrossRef]
16. Grzesiak, K.; Piotrowski, Z. From Constellation Dithering to NOMA Multiple Access: Security in Wireless Systems. *Sensors* **2021**, *21*, 2752. [CrossRef] [PubMed]
17. Ikki, S.; Aissa, S. Two-way amplify-and-forward relaying with Gaussian imperfect channel estimations. *IEEE Commun. Lett.* **2012**, *16*, 956–959. [CrossRef]
18. Wang, C.; Liu, T.-K.; Dong, X. Impact of channel estimation error on the performance of amplify-and-forward two-way relaying. *IEEE Trans. Veh. Technol.* **2012**, *61*, 1197–1207. [CrossRef]
19. Ma, Y.; Jin, J. Effect of channel estimation errors on M-QAM with MRC and EGC in Nakagami fading channels. *IEEE Trans. Veh. Technol.* **2007**, *56*, 1239–1250. [CrossRef]
20. Yang, Z.; Ding, Z.; Fan, P.; Karagiannidis, G.K. On the Performance of Non orthogonal Multiple Access Systems with Partial Channel Information. *IEEE Trans. Commun.* **2016**, *64*, 654–667. [CrossRef]
21. Keel, B.M.; Baden, J.M.; Heath, T.H. A Comprehensive Review of Quasi-Orthogonal Waveforms. In Proceedings of the 2007 IEEE Radar Conference, Waltham, MA, USA, 17–20 April 2007; pp. 122–127. [CrossRef]

22. Strinati, E.C.; Simoens, S.; Boutros, J. New error prediction techniques for turbo-coded OFDM systems and impact on adaptive modulation and coding. In Proceedings of the 2005 IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications, Berlin, Germany, 11–14 September 2005; Volume 2, pp. 1116–1119. [CrossRef]
23. Desset, C.; Ahmed, N.; Dejonghe, A. Energy Savings for Wireless Terminals through Smart Vertical Handover. In Proceedings of the 2009 IEEE International Conference on Communications, Dresden, Germany, 14–18 June 2009; pp. 1–5. [CrossRef]
24. Hamed, E. Practical distributed MIMO for WiFi and LTE. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, 2018.
25. Bob, W. *FSK: Signals and Demodulation*; Watkins-Johnson Company Technotes 7.5: Palo Alto, CA, USA, 1980.
26. Boonrungruedee, T.; Khumsat, P. 27-MHz FSK Wireless System Resilient to In-band Interference for IoT Applications. In Proceedings of the 2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Phuket, Thailand, 24–27 June 2020; pp. 692–695. [CrossRef]
27. Saadeh, W.; Altaf, M.A.B.; Alsuradi, H.; Yoo, J. A Pseudo OFDM With Miniaturized FSK Demodulation Body-Coupled Communication Transceiver for Binaural Hearing Aids in 65 nm CMOS. *IEEE J. Solid-State Circuits* **2017**, *52*, 757–768. [CrossRef]
28. Shang, Z.; Zhao, Y.; Lian, Y. A Low Power Frequency Tunable FSK Receiver Based on the N-Path Filter. *IEEE Trans. Circuits Syst. II Express Briefs* **2019**, *66*, 1708–1712. [CrossRef]
29. Chiu, C.Y.; Zhang, Z.C.; Lin, T.H. Design of a 0.6-V 429-MHz FSK Transceiver Using Q-Enhanced and Direct Power Transfer Techniques in 90-nm CMOS. *IEEE J. Solid-State Circuits* **2020**, *55*, 3024–3035. [CrossRef]
30. Park, D.; Ahn, J.; Choe, C.; Woo, S.; Ahn, S.; Choi, J. A Noise-Shaped Signaling Method for Vehicle-to-Everything Security. *IEEE Access* **2021**, *9*, 75385–75397. [CrossRef]
31. Choi, J.; Park, D.; Kim, S.; Ahn, S. Implementation of a Noise-Shaped Signaling System through Software-Defined Radio. *Appl. Sci.* **2022**, *12*, 641. [CrossRef]
32. Xu, Z.; Jin, W.; Zhou, K.; Hua, J. A Covert Digital Communication System Using Skewed α-Stable Distributions for Internet of Things. *IEEE Access* **2020**, *8*, 113131–113141. [CrossRef]
33. Ahmaderaghi, B.; Kurugollu, F.; Rincon, J.M.D.; Bouridane, A. Blind Image Watermark Detection Algorithm Based on Discrete Shearlet Transform Using Statistical Decision Theory. *IEEE Trans. Comput. Imaging* **2018**, *4*, 46–59. [CrossRef]
34. 20 MHz Bandwidth, 50 MHz to 2.2 GHz USRP Software Defined Radio Device. Available online: https://www.ni.com/pl-pl/support/model.usrp-2920.html (accessed on 5 January 2023).
35. IEEE. 802.11a-1999. *IEEE Standard for Telecommunications and Information Exchange Between Systems—LAN/MAN Specific Requirements—Part 11: Wireless Medium Access Control (MAC) and physical layer (PHY) specifications: High Speed Physical Layer in the 5 GHz band.* 1999. Available online: https://pdos.csail.mit.edu/archive/decouto/papers/802.11a.pdf (accessed on 10 May 2023).