Basic analysis for base-band and pass-band transmission system

Ruize Xuan*
Southeast University, Jiangning District, Moling Road, Nanjing, Jiangsu, China
*seuwjx@yeah.net

ABSTRACT

The main type of communication system can be divided into base-band transmission and pass-band transmission. For base-band transmission, polar NRZ coding is utilized among common PCM coding schemes to achieve the best error performance when the SNR stays the same. Matched filter receiver is adopted to carry on error analysis. For pass-band transmission, signal-space is introduced to simplify the procedure with a few examples. Then a new Euclidean space is mentioned, which is sort of different from the former, to directly map sending bits with their actual position on the space. 2D and 3D space, related to the 4-ary system and 8-ary system respectively, is depicted to illustrate the point. At last, a potential application for pixel maps is considered to qualitatively underscore the error impact on the practical system.

Keywords: Base-band transmission, Pass-band transmission, Error analysis, New signal space

1. INTRODUCTION

There are many previous works on a certain part of base-band transmission and pass-band transmission. Base-band signal transmission is often seen in short-distance communication systems such as Local Area Network (LAN). Pass-band transmission is often seen in long-distance systems such as wireless phone services and computer networks. A wide-band digital signal transceiver is mentioned to finish a comparatively complete procedure of digital base-band signal processing[1]. A pass-band equalizer that bears relatively simple structures is introduced to achieve low costs[2]. Also, there are researches on the pass-band ISI effect and error rate[3][4][5][6]. However, less is concerned about the whole transmission procedure with great details combining two kinds of modes.

Moreover, there is an increasing trend, using signal analysis, to calculate related parameters. Finite energy signals are represented by geometry elements. The most common condition is with the AWGN channel. However, the mapping rule is a little indirect. The sending bits are converted to a new coordinate in the space. Sometimes it's difficult to find a corresponding relationship. Therefore, a creative idea for mapping the bits is to be discovered.

In this paper, the basic structures of the base-band and pass-band transmission system are given. Imitating the concept of signal-space analysis, a new Euclidean space is mentioned and illustrated by a few examples with applications. The remainder of this paper is organized as follows. The second part addresses the basic assumptions for the analysis of base-band transmission and pass-band transmission. Part two firstly gives a brief introduction to the PCM coding and receiver structure with error analysis, then signal-constellation for pass-band transmission is given to lay a foundation for the Euclidean space. Simulation results are given in part three with comparisons and analysis for the results in part four. At last, some basic conclusions are given in part five.

2. METHOD

In this paper, a binary transmission system is first considered. Pulse-code modulation, which is robust to channel noise and symbol interference, is utilized to illustrate the point. It's often the case that an analog waveform (message signal) is coded into a PCM wave after sampling, quantizing and encoding. Considering that the first two steps are irrelevant in the analysis, binary digits are assumed already prepared in this circumstance.

In this section, key approaches are illustrated to explain how to select the most appropriate schemes. With different application backgrounds, there are corresponding solutions. As stated before, two types of communication will be considered separately.

2.1. Base-band Transmission

It's not enough to represent information just by binary bits for transmission. To make it appropriate for practical transmission, concrete and intuitive digital waveforms are indispensable. In practice, there are many schemes to choose

International Conference on Neural Networks, Information, and Communication Engineering (NNICE 2022),
edited by Rajeev Tiwari, Proc. of SPIE Vol. 12258, 122581P · © 2022 SPIE
0277-786X · doi: 10.1117/12.2640510

from. To make it general, three common and representative line code methods are adopted to implement this experiment. Assuming we are transmitting 1,0,0,1,1,0,1,0 to test, the corresponding waveform is represented in Figure 1.

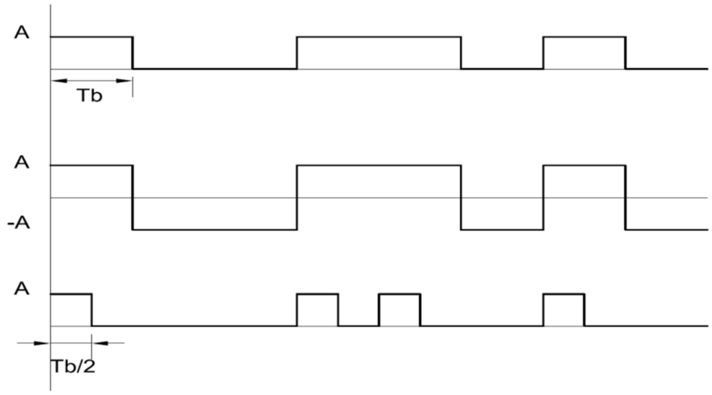

Figure 1. Different coding scheme for 10011010 (a)Unipolar NRZ coding (b)Polar NRZ coding (c)Unipolar RZ coding (from top to bottom)

It's noted that channel noise and intersymbol interference (ISI) are two factors that can cause errors in communication systems. Therefore, two main conditions are separated to be analyzed.

1. In the first condition, the main attention is focused on the channel noise. It often occurs in low-speed rate transmission systems. In this case, the channel can be seen as an ideal one. When a signal is passed through the AWGN channel, the impulse response of the channel can be seen as a unit impulse function. The message signal itself is transmitted without distortion. Noises are added directly to the useful information.

The most commonly used method to cope with channel noise is matched filters. Basic steps for receiving messages contain filtering, sampling and threshold decision. The ultimate goal of matched filters is to maximize the output signal-to-noise ratio. To proceed with the analysis, assuming unipolar NRZ coding is used. Next, a formula for the error rate due to noise is to be obtained.

Model: Symbols 1 and 0 are represented by a rectangular pulse of amplitude A with duration T_b and no pulse respectively. The channel noise is modeled as an additive, white and Gaussian, with zero mean and power spectral density $N_0/2$. Assuming "1" and "0" occur with equal probability. The average transmitted signal energy per bit is defined as E_b.

Deduction: Received signal is denoted as x(t). In a time interval, x(t) can be expressed as follows:

$$x(t) = \begin{cases} +A + n(t), & \text{symbol 1 was transmitted} \\ n(t), & \text{symbol 0 was transmitted} \end{cases} \quad (1)$$

where n(t) denotes the noise component. The structure of the matched filter is shown in Fig. 2.

In the Figure 1, s(t) is the PCM coding waveform obtained from the former part utilizing unipolar NRZ. The matched filter part will generate an impulse response to match the pulse shape of s(t). According to related theory[7], sampling will be added to maximize output SNR. In the end, combined with decision bound and sampling results, the system will return the final choice either 1 or 0.

To reach the ultimate goal of error rate, there are two possible conditions to be considered. An error occurs when 1 was transmitted but decoded as 0 and 0 was transmitted but decoded as 1. Suppose the symbol 1 was transmitted. So the received signal can be expressed as

$$x(t) = A + n(t),\ 0 \le t \le T_b \quad (2)$$

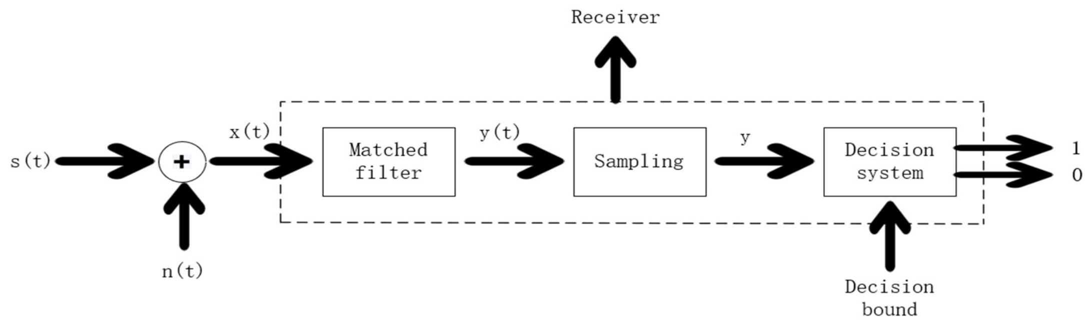

Figure 2. Structure of a matched filter receiver

In matched filters[7], the impulse response is written as

$$h(t) = kg(T - t) \tag{3}$$

In this scheme, g(t) is a rectangular pulse with duration T_b and amplitude $+A$. Before sampling, the received signal was converted to

$$\begin{aligned} y(t) &= x(t) * h(t) \\ &= k\int_0^{T_b} x(\tau)g(T_b - t + \tau)d\tau \\ &= k\int_0^{T_b} g(\tau)g(T_b - t + \tau)d\tau \\ &+k\int_0^{T_b} n(\tau)g(T_b - t + \tau)d\tau \end{aligned} \tag{4}$$

When sampled at t=T_b, y can be expressed by

$$\begin{aligned} y &= y(T_b) \\ &= k\int_0^{T_b} g^2(\tau)d\tau + k\int_0^{T_b} n(\tau)g(\tau)d\tau \\ &= kA^2T_b + kA\int_0^{T_b} n(\tau)d\tau \end{aligned} \tag{5}$$

If kAT_b is set to be unity then we get

$$y = A + \frac{1}{T_b}\int_0^{T_b} n(\tau)d\tau \tag{6}$$

Then y can be seen as a sample of a random variable Y. It follows Gaussian distribution for the sake of noise property.

The variance of Y is

$$\begin{aligned} \sigma_Y^2 &= E[(Y - A)^2] \\ &= E[\frac{1}{T_b^2}\int_0^{T_b}\int_0^{T_b} n(t)n(u)dudt] \\ &= \frac{1}{T_b^2}\int_0^{T_b}\int_0^{T_b} E[n(t)n(u)]dudt \\ &= \frac{1}{T_b^2}\int_0^{T_b}\int_0^{T_b} R_N(t-u)dudt \\ &= \frac{N_0}{2T_b} \end{aligned} \tag{7}$$

Then conditional probability density function of random variable Y can be expressed as

$$f_Y(y|1) = \frac{1}{\sqrt{\pi N_0/T_b}} exp\left(-\frac{(y-A)^2}{N_0/T_b}\right) \tag{8}$$

Accordingly, when 0 was transmitted, conditional probability density function of Y can be expressed as

$$f_Y(y|0) = \frac{1}{\sqrt{\pi N_0/T_b}} exp(-\frac{y^2}{N_0/T_b}) \tag{9}$$

Using the formula of total probability, the error rate considering both cases can be expressed as

$$P_e = p_1 \int_{-\infty}^{b} f_Y(y|1)dy + p_0 \int_{b}^{\infty} f_Y(y|0)dy$$

$$= \frac{1}{4} erfc(\frac{A-b}{\sqrt{N_0/T_b}}) + \frac{1}{4} erfc(\frac{b}{\sqrt{N_0/T_b}}) \tag{10}$$

Where *b* is the decision bound. From the above equation, the probability of error can be seen as a function of *b*. The process of decoding is expressed in Figure 3.

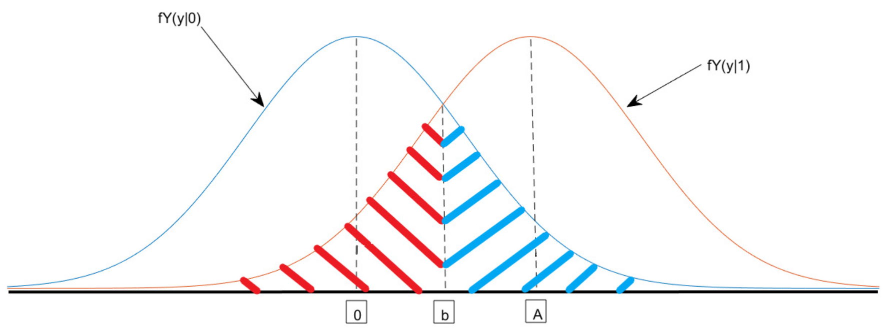

Figure 3. Analysis for error probability of unipolar NRZ coding

Since two symbols are equally sent, *b* can be chosen as *A/2*. To express probability in a more general form, E_b/N_0 is introduced to simplify the result. According to the definition of average energy per bit, E_b can be written as

$$E_b = \frac{A^2 T_b}{2} \tag{11}$$

Combining equation (10) and (11), the ultimate result is

$$P_e = \frac{1}{2} erfc(\sqrt{\frac{E_b}{2N_0}}) \tag{12}$$

For Unipolar RZ coding, it has the same results as above, while polar NRZ coding conforms to the following rule:

$$P_e = \frac{1}{2} erfc(\sqrt{\frac{E_b}{N_0}}) \tag{13}$$

In the second place, ISI becomes the dominant factor. It often happens in high-speed rate transmission. In this circumstance, the channel is dispersal so transmitted symbols are extended which can cause ISI. In a way, high-speed rate transmission indicates a small bit duration period so its spectral property becomes wider and therefore neighboring bits are unavoidably affected by each other. A new model is introduced to solve the problem. To be systematic, transmit filter, channel filter and received filter can be seen as a whole system to consider.

According to the ideal Nyquist Channel theory, it can achieve zero-ISI if it follows a certain rule. However, it's not achievable in actual problems because it is endowed with a sudden transition in the time domain. Thus another useful and highly efficient scheme is introduced, called raised-cosine spectrum which occupies more bandwidth but is achievable. Their differences and characteristics are listed in the following Table 1.

Table 1. Comparison between Ideal Nyquist channel and Raised-cosine spectrum

Item	Channel Bandwidth	Advantages	Disadvantages
Ideal Nyquist channel	$W = \frac{R}{2}$	Minimum bandwidth for transmission	Not achievable
Raised-cosine spectrum	$B_T = W(1+\alpha)$	Achievable	Increased bandwidth

α is called the roll-off factor.

2.2. Pass-band Transmission

In this section, the message PCM wave is modulated onto a higher frequency carrier wave. Only sinusoidal waves are considered. Since it can be trivial and confusing to analyze directly, the concept of signal-space is introduced to effectively simplify the problem.

Similar to the points in Euclidean space, a baseband signal can be expressed by a set of basis functions in the designed space. When a unipolar NRZ coding scheme is adopted in a binary system, the basic function is to be attained. Basis functions are orthonormal to each other. In most cases, only cosine and sine waves are enough to form a signal. As to the digital modulating schemes, coherent BPSK, coherent BFSK and MSK are adopted and compared.

BPSK

For coherent BPSK, there is only one basis function written as

$$\varphi_1(t) = \sqrt{\frac{2}{T_b}} cos(2\pi f_c t), \quad 0 \leq t \leq T_b \tag{14}$$

Then we can represent each of two symbols on the constellation plain. If the modulated wave is denoted as (symbol 1) and (symbol 0), they can be expressed as

$$s_1(t) = \sqrt{E_b}\varphi_1(t), \;\; 0 \leq t \leq T_b \tag{15}$$
$$s_2(t) = -\sqrt{E_b}\varphi_1(t), \;\; 0 \leq t \leq T_b \tag{16}$$

Then, according to the mapping rule, two symbols are represented on the plain in Figure 4.

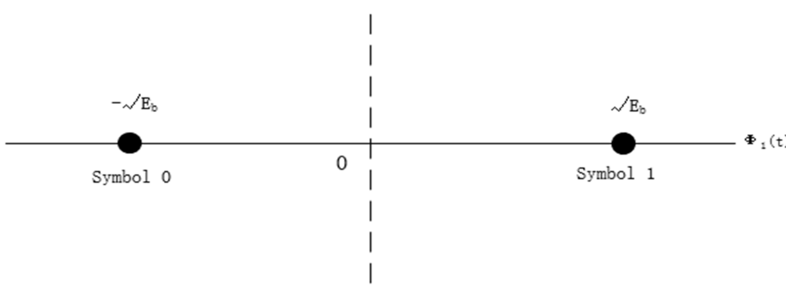

Figure 4. Coherent BPSK constellation

In the coherent frequency-shift keying, there are mainly two categories called Sunde's FSK and MSK. They are both continuous-phase frequency-shift keying schemes.

BFSK

Symbols 0 and 1 are represented by the instant frequency of the carrier wave. Using the most common set of orthonormal basis functions,

$$\varphi_i(t) = \sqrt{\frac{2}{T_b}} cos(2\pi f_i t), \;\; 0 \leq t \leq T_b \tag{17}$$

It follows the rule[9]

$$\Delta f = \frac{1}{T_b} \tag{18}$$

Similarly, modulated waves are represented by two-dimensional vectors in a 2D plain as expressed in Figure 5.

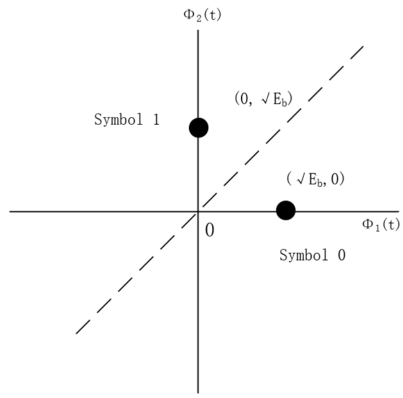

Figure 5. Coherent BFSK constellation

MSK

What distinguish MSK[10] from coherent BFSK is that

$$\Delta f = \frac{1}{2T_b} \tag{19}$$

A new signal space

Following the concept of signal-space[11], a novel idea can be put forward: no matter which type of digital modulation scheme is adopted, an *M*-ary communication system can be turned into an *N*-dimensional Euclidean space where $N=\log_2(M)$. For example, if $M=4$, then there are four possible symbols to be transmitted. It can be expressed in Figure 6(a). And if $M=8$, it can be illustrated in Figure 6(b).

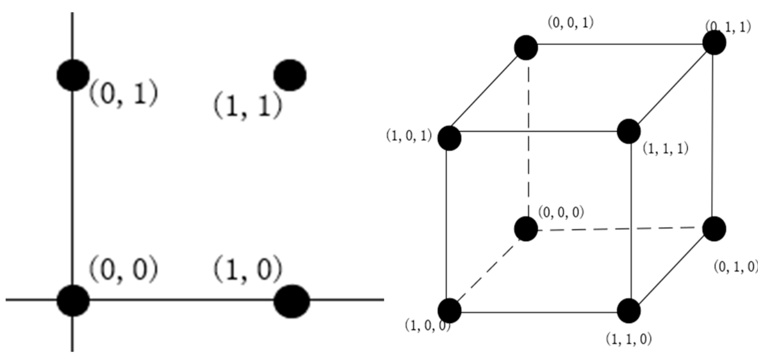

Figure 6. The new Euclidean space (a)4-ary system (b) 8-ary system

The main benefit of this idea is that the actual position of a point can correctly represent corresponding bits without loss and also conforms to the Gray Coding scheme. Thus when the receiver demodulates the signal, it can lower the BER to reach almost the same as SER. In signal space analysis, there are many equivalent kinds of constellation sets under the same message bits due to the invariance of the probability of error to rotation and translation. However, it's not considered in the above idea since it will confuse us about the relation between position and sending bits.

The demodulation part is also similar to the method in the signal-space analysis, using the maximum likelihood scheme. To be more specific, the receiver compares the distance between the received point and transmitted symbols. The closer to a symbol, the more likely it can be estimated as that symbol.

There are some differences between signal-space and this Euclidean space. Firstly, the points in the signal space conform to a more strict rule. Their positions have more complicated meanings such as energy. In general, signal-space is invariant to rotation and translation, and also stays with minimum energy. In the new signal-space, the position of a point always equals its value. Also, their dimension of them varies. In an *M*-ary system, the dimension *N* is always equal or less than *M* in signal-space. While in new signal-space, it always conforms to the rule N=$\log_2$(M).

The general form of a communication system

Next, based on the above discussions, a simplified and practical transmission system model is depicted in Figure 7.

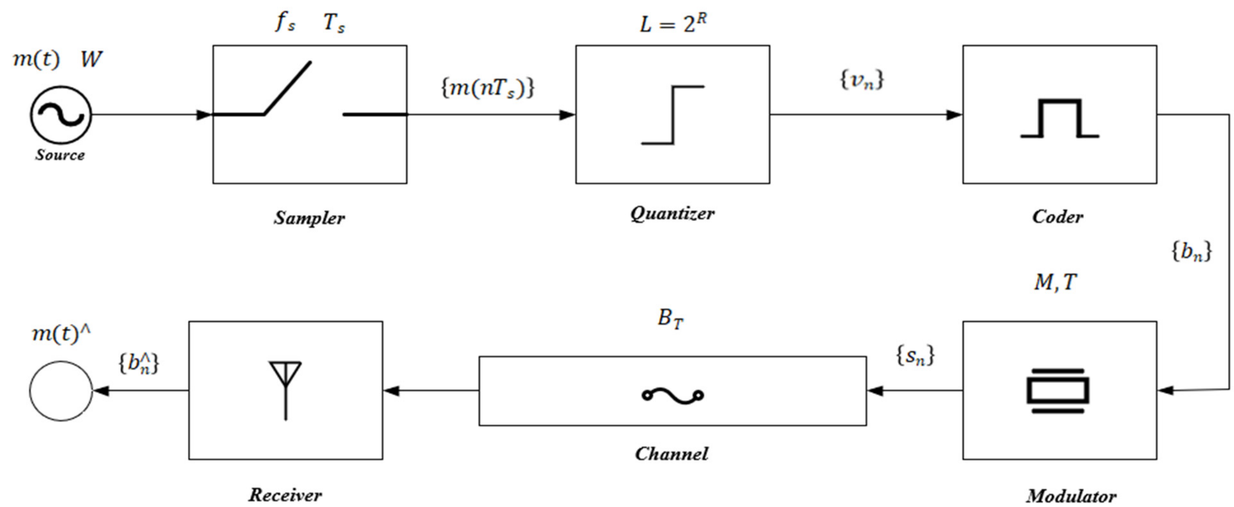

Figure 7. Working flow for a general communication system

The basic working flow can be illustrated as: a message signal $m(t)$ with bandwidth W is to be transmitted. First, it is sent into a sampler with sampling rate f_s (or sampling period T_s). After sampling, the original message becomes a time-discrete and amplitude-continuous series signal $\{m(nT_s)\}$. To make its amplitude discrete, the series signal is sent into a uniform quantizer. The total number of threshold levels $\{v_n\}$ is L, which means each symbol is represented by R bits. The coder mainly achieves two goals: turning threshold levels into bits and then converting bits into waves. As for base-band transmission, the wave is directly sent into the channel. As for *M*-ary pass-band transmission with symbol duration T, the signal is mapped to the signal-space, representing modulated waves $\{s_n\}$. Although the channel part is not directly reflected in the signal space, it's noted that the transmission bandwidth B_T of the channel is to be considered with regard to the message bandwidth W. The minimum bandwidth of the channel conforms to the rule of the ideal Nyquist channel and there are other schemes to be adopted such as raised-cosine spectrum and correlative-level coding. Then the receiver, considered as a whole, recovered the bits according to the maximum likelihood rule. Finally, the recovered message $m(t)^\wedge$ is obtained.

Some related equations and rules are listed as follows:

(1) According to the Nyquist sampling theorem,

$$f_s \geq 2\mathrm{W} \tag{20}$$

(2) The bit rate of the system is $\frac{R}{T_s}$ bit/s or $\frac{\log_2 M}{T}$ bit/s.

(3) Equivalently, the symbol rate can be expressed as $\frac{1}{T}$ sym/s.

(4) To avoid ISI, using the formula of raised cosine spectrum, the transmission bandwidth can be expressed as

$$B_T = (1 + \alpha)W_{Nyquist} \tag{21}$$

where $W_{Nyquist} = \frac{1}{2T}$.

3. RESULT

3.1. Probability of error of coding scheme

According to the result in part three, the probability of error with regard to the SNR can be plotted, which is shown in Figure 8.

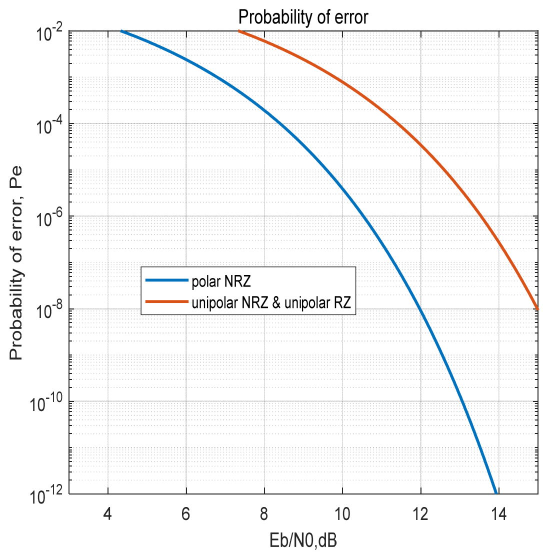

Figure 8. Error performance in different coding schemes

Three systems with different coding schemes all achieve better error performance with increased SNR, which is a general conclusion in digital communication. From Figure 8, as for the same probability of error, polar NRZ gives 3 dB less SNR than unipolar NRZ and unipolar RZ coding scheme, which is more appropriate for base-band transmission. As for the same SNR, polar NRZ performs better in contesting noises. Also, it can be concluded from Figure 8 that although two schemes for coding may be different, they can reach the same error performance. In this experiment, unipolar NRZ and unipolar RZ stay the same result.

3.2. 4-ary system in a 2D space

Similar to the signal-space, in the presence of noise, the received points will gather around the sending symbols. The threshold line is depicted in the plot. Points that stay over the line will be misunderstood and cause errors. For example, when a 4-ary system with an AWGN channel is considered, the noise effect (without ISI) on the message can be directly seen from the following Figure 9.

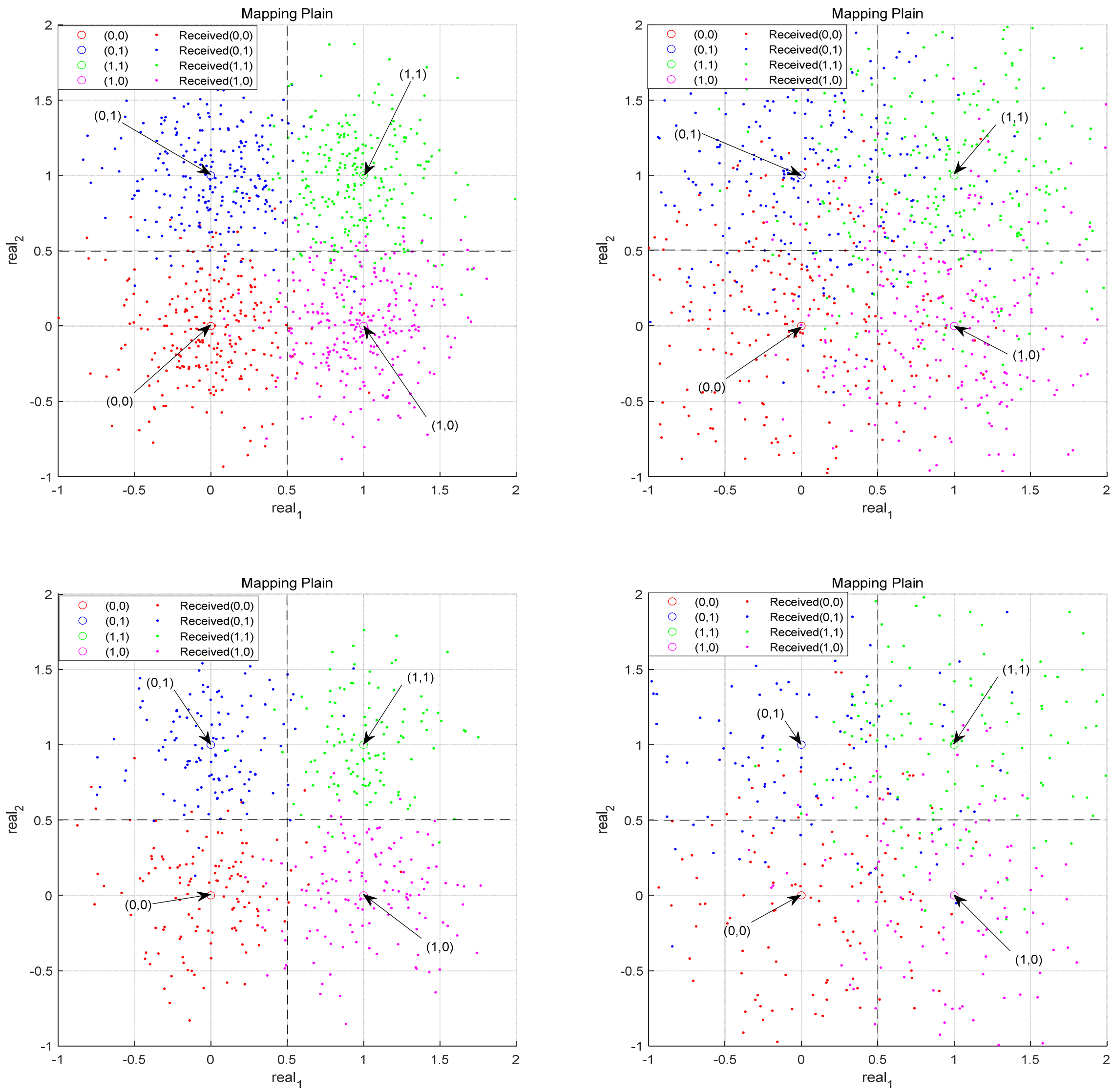

Figure 9. The new signal-space for 4-ary system (a) N=1000, SNR(dB)= 0~10 (b) N=1000, SNR(dB)= 0~5

(c) N=500, SNR(dB)=0~10(d) N=500, SNR(dB)=0~5

(from left to right and from top to bottom)

In Figure 9, with higher SNR, received points are more closely distributed around the corresponding transmitted symbols, which means it's more likely to be demodulated correctly. It is noticeable that the axes of the plot in Figure 9 are described as two real dimensions, which is consistent with a coordinate expressed in a Euclidean space. The point in a classic signal space is represented by a set of basic functions with different coefficients. As a result, the received point in the above picture has non-integer coordinates. It may seem strange since it's not a binary representation. However, it's reasonable under mathematical calculation. For example, when an integer DC signal is added by random noise, the instantaneous value can be non-integer. It goes for the same reason for binary bits. What we are interested in is the final decision given by the demodulator. This received value acts as an intermediate number with no practical meaning.

3.3. 8-ary system in a 3D space

Also, a three-dimensional space with an 8-ary system can be plotted in Figure 10. Higher dimension space can not be plotted directly but the rule still applies.

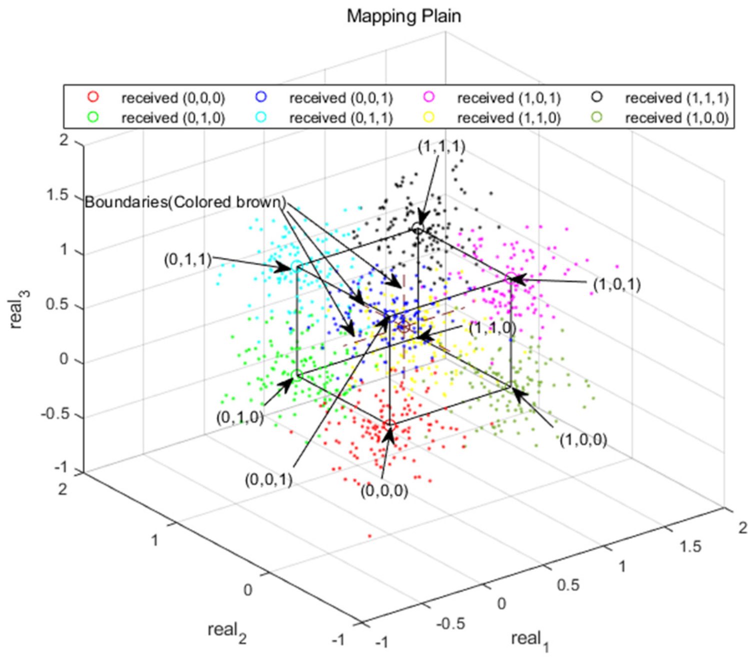

Figure 10. 8-ary system in a 3D space (N=1000, SNR(dB)=0~20)

As can be seen in Figure 10, sending symbols, forming eight vertices of a cube, are surrounded by corresponding received points. According to the maximum likelihood rule, decision boundaries are colored brown which divides the whole space into eight segments. The higher the dimension is, the more difficult for the demodulator to be designed. According to the mapping rule, more bits lead to more boundaries. For example, if the received point in Figure 10 happens to fall on boundaries, it can cause problems for the receiver to detect the signal. An extreme case is that if a point falls on the brown-circular symbol in Figure 10, it's likely to be demodulated as each of eight symbols as a result. Therefore, other judging methods are needed. Sending symbols in Figure 10 is considered equally transmitted, which is a general case. On other special occasions, they are weighted differently. So one possible solution is to consider all these problem points together to achieve the least sum of the distance between them and the most important symbol.

3.4. Pixel maps

Error performance can hugely affect a system. To demonstrate this, for example, it can be clearly illustrated by an example of pixel encoding and demodulation. Assuming that a screen CRT displayer system uses RGB to compose pixels. When there are occurrences of errors, there are many pixel cells misunderstood. As a result, it can cause bad performance. For a 4-ary communication system, the original sending messages and recovered symbols are compared in Figure 11.

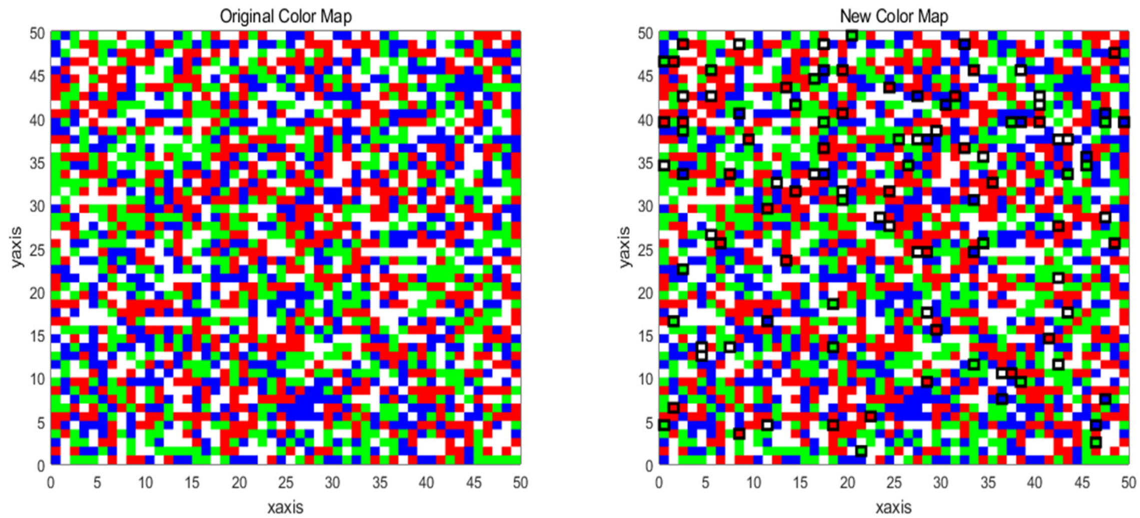

Figure 11. (a)Transmitting pixels in AWGN channel in a 4-ary system

(b)Estimated pixels in AWGN channel in a 4-ary system

Figure 11 just represents a tiny part of the screen. The method adopted above is to use the new Euclidean space with the AWGN channel. Error pixels are shown in the above picture with black frames. White pixels represent lost signals. As can be seen from the results, the error performance is not so satisfying in a 4-ary system. This is an extreme weakness of the unmodified scheme. Other improving details need to be considered. A number of errors or lost pixels can contribute to distinct distortion which is hugely overcome in real products.

4. DISCUSSION

In the paper, two main types of transmission modes are discussed in detail. For the base-band transmission systems, among unipolar NRZ coding, polar NRZ coding and unipolar RZ coding, polar NRZ coding is chosen since it can guarantee the lowest error rate when $\frac{E_b}{N_0}$ is the same. When a 4-ary system is to be considered, it can be seen from the corresponding Euclidean space that with higher SNR, the received points are more close to the sending symbols. To make it more general, a 8-ary system is also considered. As a result, the received points distribute around the original symbols and form a spherical shape. At last, practical use of this method is mentioned in the pixel display transmission system.

The meaning of the new signal-space gives a potential method for signal analysis. Many other properties remain to be discovered. Combining graphical and visualization tools proves to be a powerful and useful method in signal analysis. However, as for designing communication circuits, the signal-space analysis may not present advantages and designers still have to consider complicated details.

As illustrated in Figure 7, a general form and its related parameters in a digital communication system are given. It's highly efficient in analyzing basic components of a communication system both qualitatively and quantitatively. To achieve more advanced structures, many sections remain to be updated.

5. CONCLUSION

A complete procedure for base-band and pass-band signal analysis is given in the paper. An effective coding scheme is identified, in which the error performance proves to be the best. The probability of error, with AWGN channel and matched filter receiver, for base-band signal is obtained quantitatively. In a high-speed transmission system, corresponding to a pass-band signal, the method of signal space analysis is used. With the maximum likelihood detection, it's likely to attain a new Euclidean space for calculating error rates. The most dominant property of this new space is the direct mapping between position and actual sending bits. In future work, the effectiveness and convenience of the new mapping rule could be shown and therefore it can be applied in the pass-band signal analysis.

REFERENCES

[1] C. -m. Liu, H. -m. Gao and Z. -j. Zhao, "Research on Baseband Signal Processing Transceiver Module," 2020 IEEE 3rd International Conference on Electronics Technology (ICET), 2020, pp. 607-611, doi: 10.1109/ICET49382.2020.9119688.

[2] K. Mueller and J. Werner, "A Hardware Efficient Passband Equalizer Structure for Data Transmission," in IEEE Transactions on Communications, vol. 30, no. 3, pp. 538-541, March 1982, doi: 10.1109/TCOM.1982.1095492.

[3] G. Forney, "Lower Bounds on Error Probability in the Presence of Large Intersymbol Interference," in IEEE Transactions on Communications, vol. 20, no. 1, pp. 76-77, February 1972, doi: 10.1109/TCOM.1972.1091101.

[4] S. Asano and T. S. Usuda, "A simple approximation of minimum error probability using trace distance," 2014 International Symposium on Information Theory and its Applications, 2014, pp. 254-258.

[5] E. A. Candreva and G. E. Corazza, "A tighter upper bound on the error probability of signals in white Gaussian noise," 2010 5th Advanced Satellite Multimedia Systems Conference and the 11th Signal Processing for Space Communications Workshop, 2010, pp. 448-448, doi: 10.1109/ASMS-SPSC.2010.5586853.

[6] C. D. Frank and M. B. Pursley, "Bounds on the performance of maximum-likelihood and reduced-state decoding of trellis codes on channels with intersymbol interference," [1992 Proceedings] Vehicular Technology Society 42nd VTS Conference - Frontiers of Technology, 1992, pp. 267-270 vol.1, doi: 10.1109/VETEC.1992.245425.

[7] L. A. Zadeh and J. R. Ragazzini, "Optimum Filters for the Detection of Signals in Noise," in Proceedings of the IRE, vol. 40, no. 10, pp. 1223-1231, Oct. 1952, doi: 10.1109/JRPROC.1952.274117.

[8] Arthur A. Giordano; Allen H. Levesque, "Digital Communications BER Performance in AWGN (FSK and MSK)," in Modeling of Digital Communication Systems Using SIMULINK , Wiley, 2015, pp.101-117, doi: 10.1002/9781119009511.ch5.

[9] Mehta, Sudhanshu & Sharma, Surbhi & Khanna, Rajesh. (2011). SMT-8036 Based Implementation of Secured Software Defined Radio System for Adaptive Modulation Technique. 192. 205-212. 10.1007/978-3-642-22720-2_20.

[10] C. Yun, Y. Jian, L. Jian, S. Dawei and S. Binbin, "Research on Phase Differential Demodulation Method of Narrow Pulse MSK Signal," 2018 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC), 2018, pp. 1-6, doi: 10.1109/CSQRWC.2018.8455446.

[11] Simon Haykin, Communication Systems, Wiley, 4th edition, 2003