

Research on the Technology of Breaking the Shannon Channel Capacity and Shannon Limit

Dequn Liang (✉ ldqlgn@163.com)

Dalian Maritime University

Xinyu Dou

Dalian Maritime University

Article

Keywords: Shannon channel capacity, signal uncertainty principle, time-delay overlapping technique, non-orthogonal multi-carrier modulation

Posted Date: July 12th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1643836/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Research on the Technology of Breaking the Shannon Channel Capacity and Shannon Limit

Dequn Liang* and Xinyu Dou

Information Science and Technology College, Dalian Maritime University, Dalian, China

Corresponding author: Dequn Liang. Email: ldqlgn@163.com

Abstract

In this paper, firstly, the Shannon channel capacity is briefly stated, and the relationship between it and the signal uncertainty principle is analyzed in order to prepare for deriving the formula which can break through the Shannon channel capacity. Then, as practical examples of breaking the Shannon channel capacity, the time-shift non-orthogonal multicarrier modulation technology (TS-NMT) and the time frequency pulse time delay multicarrier modulation technology (TFP-TDMC) are introduced. After more than twenty years of development, this technique is proved to be a practical modulation technique for digital communication. At the end of the paper, a hypothesis about the physical nature of information is proposed.

Keywords: Shannon channel capacity, signal uncertainty principle, time-delay overlapping technique, non-orthogonal multi-carrier modulation

1. Introduction

At present, the fifth-generation mobile communication technology (5G) has been listed. Through mobile communication technology, the integration of intelligent devices and the Internet is becoming more and more in-depth, which indicates that a new industrial revolution is coming. Due to the unprecedented demand for flexible, efficient and reliable communication between people, between people and objects and between objects, wireless communication technology must provide higher data transmission rate, higher spectral efficiency and lower delay than before [1-3]. Therefore, the industry has proposed many technical solutions, such as developing spectrum resources and various new multi-user access technologies, new signal processing technologies and channel coding methods [4-16]. In addition, large-scale Multiple-Input-Multiple-Output (Massive MIMO) will be deployed in 5G. By significantly increasing the number of antennas, the data transmission rate can be further improved [3,4,17-21]. However, the above-mentioned new methods to improve communication performance are mainly focused on mining new resources and improving the original technology. Therefore, some people in the industry are confused about the development future of communication technology and products, because in theory, the transmission rate has been limited by Shannon limit! Although it has been proved that MIMO can break the Shannon limit of Single-Input-Single-Output (SISO) [22], its essence is to multiplex multiple spatial physical channels by using multiple antennas, and the rate of a single physical channel still does not break the Shannon limit. In 1948, Shannon gave three theorems of source coding, channel coding and channel capacity limit with detailed mathematical derivation in his paper *mathematical theory of communication* [23], which laid a theoretical foundation for signal processing and communication technology. Since then, academia has done a lot of effective work in source

coding and channel coding, but there has been no way to break the Shannon limit in terms of channel capacity.

In the second section of this paper, the Shannon channel capacity formula is briefly introduced, and the theoretical completeness of the formula is analyzed according to the principle of signal uncertainty. Then, according to Shannon's information entropy theory, a theoretical method to break through Shannon's limit is introduced. This method is called delay overlap method. For more than 20 years, we have constructed a digital modulation method that breaks through the Shannon limit by using the idea of delay overlap method. Through a large number of computer simulation, physical experiments and theoretical analysis, it is proved that this digital modulation method has reached the practical stage. It will show strong vitality in the field of 5G wireless communication. In the process of studying this theory, we realize that Shannon theory is not only of great significance in communication, but also plays an important role in economics, sociology and other disciplines. As Thomas M. Cover pointed out in *elements of information theory*: "... therefore, some people think that information theory is a subset of communication theory. We think it's more. In fact, it makes important contributions to statistical physics (thermodynamics), Computer Science (Kolmogorov complexity or algorithm complexity), statistical inference (Okam razor: "the simplest explanation is the best ") and probability and Statistics (error index used to test and estimate the best hypothesis)." [24]. Further, we think that Shannon theory only gives the mathematical quantitative analysis of information, but never involved the physical essence of information. Therefore, a hypothesis about the physical properties of information are discussed at the end of this paper.

2. The Shannon Limit and the Method of Breaking the Shannon Limit

2.1 Shannon Limit

Fig. 1 shows a basic process model of signal transmission in a communication system. In the system model, $x(t)$ is a signal carrying information, $n(t)$ is the noise (including various other interferences), and $x(t)$ becomes $y(t) = x(t) + n(t)$ after passing through the channel. If we refer to $x(t)$ as the event to be sent, $y(t)$ is the sent event received. $y(t)$ is not equal to $x(t)$ due to the existence of noise. We call it as "information loss". Therefore, the problem of information measurement is proposed.

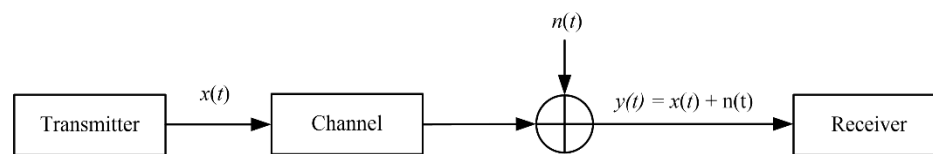


Fig. 1 The model structure of a communication system

A simple answer is that events convey messages, and messages carry information. Common sense tells us that the information provided by the event with small probability can attract more attention of the receiver than the event with high probability. That is to say, the amount of information of the event with small probability is large. In order to explain this phenomenon quantitatively, Hartley defined the information content of a discrete message as

$$I(x_i) = \log \frac{1}{p(x_i)} = -\log p(x_i), \quad (1)$$

where $p(x_i)$ is the probability density function [25]. Shannon introduced the **entropy** of thermodynamics into the communication system. The statistical average value of information is used as a measurement of amount of information for discrete sources

$$H(X_i) = -\sum_{i=1}^M p(x_i) \log p(x_i), \quad (2)$$

where X_i refers to the random variable form of the transmitted signal and $H(X_i)$ is called the **entropy** of a discrete information source.

Further, let $\Delta x_i = x_{i+1} - x_i$, $i = -M, \dots, -1, 1, \dots, M$, and there is $p(x) = \lim_{\Delta x \rightarrow 0} p(x_i) \Delta x_i$. So, there is

$$H(X) = -\int_{-\infty}^{\infty} p(x) \log p(x) dx. \quad (3)$$

This is the entropy of the continuous information source, where $p(x)$ is the probability density function. According to the sampling theorem $x(t)$ can be expressed to its discrete form $x(t_j)$, $j = 1, 2, \dots, 2W$, and W is the bandwidth of the signal $x(t)$. So that the discrete form of continuous entropy can be written as follows

$$H(X_j) = -\sum_{j=1}^{2W} p(x_j) \log p(x_j). \quad (4)$$

Shannon defined condition entropy $H(Y_j/X_j)$, and the association relationship between $y(t_j)$ and $x(t_j)$ was established in the sense of entropy. Due to $y(t_j) = x(t_j) + n(t_j)$, there is $H(Y_j/X_j) = H((X_j+N_j)/X_j) = H(N_j)$. It can be understood that when X_j is known, the uncertainty of Y_j is only determined by N_j (which is the part of Y_j more than X_j). Then there is the average amount of the mutual information of a single sampling point

$$I(Y_i, X_i) = H(Y_i) - H(Y_i / X_i) = H(Y_i) - H(N_i). \quad (5)$$

As a result, the transmission rate on the continuous channel is

$$R = 2WT \cdot I(Y_i, X_i) / T = 2W \{H(Y_i) - H(Y_i / X_i)\} = 2W \{H(Y_i) - H(N_i)\}. \quad (6)$$

Shannon has proved that when $n(t)$ is additive Gaussian white noise and the power is $P = \sigma^2$, and when $p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/(2\sigma^2)}$ and $p(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-y^2/(2\sigma^2)}$ (called Optimal probability density function function), there are the maximum entropy as $H(X) = \log_2 \sqrt{2\pi eP}$, $H(Y) = \log_2 \sqrt{2\pi e(P+N)}$ and $H(Y/X) = \log_2 \sqrt{2\pi eN}$.

Finally, Shannon limit be represented as

$$C = R_{\max} = \max \{I(X;Y) \cdot 2W\} = W \log_2 \frac{P+N}{N} = W \log_2 \left(1 + \frac{P}{N}\right) = W \log_2 (1 + \text{SNR}), \quad (7)$$

where $N = n_0 W$, and n_0 is the unilateral power spectral density of noise. It can be seen that C increases with W , but at the same time the noise also increases, and finally the maximum transmission rate reaches a limit

$$\lim_{W \rightarrow \infty} C = \lim_{W \rightarrow \infty} W \log_2 \left(1 + \frac{P}{n_0 W}\right) = \frac{P}{n_0} \lim_{W \rightarrow \infty} \frac{n_0 W}{P} \log_2 \left(1 + \frac{P}{n_0 W}\right) = \frac{P}{n_0} \log_2 e = 1.44 \frac{P}{n_0}. \quad (8)$$

Shannon then proposed a method to approximate this rate limit. Firstly, the statistical property of the transmitted signal must approximate a white noise. Secondly, the least R.M.S error theory can be used to implement the demodulation, i.e. by comparing the received signal with all the possible $M=2^s$ samples of the transmitted signal, the sample which has the least discrepancy is chosen as the demodulated signal. This method can be represented as

$$\lim_{\varepsilon \rightarrow 0} \lim_{T \rightarrow \infty} \frac{\log M(\varepsilon, T)}{T} = W \log \frac{P+N}{N}, \quad (9)$$

where ε and T represent the frequency of errors and the duration of the signal, respectively. This means that no matter how small the ε is, we can make the transmission rate approximate to $W \log(1+P/N)$ as possible as we can through sufficiently large T . What's more, $\varepsilon \rightarrow 0$ and $t \rightarrow \infty$ must be satisfied at the same time to make (9) more approximate to (7).. This just fits **signal uncertainty principle**. Specifically, only when $T \rightarrow \infty$ the bandwidth W of the signal is finite, and the sampling theorem is valid. The signal uncertainty principle is the representation of the uncertainty principle of quantum mechanics in signal processing domain. In an article "Quantum Interpretation of Signal's Uncertainty Principle", Wang Peng and Li Jian-ping indicated that according to the uncertainty principle of the quantum mechanics, the signal in some sense can be treated as a quantum-like system with wave-particle duality, and just due to the wave-particle duality, we cannot simultaneously determine both the duration and the bandwidth of the signal^[26]. Here, we propose a "time-frequency finite duality" conception. In addition, some papers demonstrated the signal uncertainty principle from the Fourier transform point of view [27,28].

Although formula (11) theoretically solves the completeness problem of approximate realization of Shannon limit, in practical communication services, it is impossible to continuously transmit information in infinite time, so it is difficult to realize a limited bandwidth in frequency domain. Fortunately, in practical engineering, under the condition of given error frequency ε , the bandwidth can be limited in a certain range, so the industry has defined various bandwidth, such as 3dB bandwidth, spectrum zero bandwidth, minimum fading bandwidth, etc. We will also face another problem: how to guarantee $\varepsilon \rightarrow 0$. Encouragingly, Shannon gave the coding theorem. If the estimated rate of source v_1 is R_1 , the output of the

source is encoded and transmitted to the channel with capacity C . As long as $R_1 < C$, the fidelity is as close to v_1 as possible. What's more, $R_1 \geq C$ is impossible.

According to the method given by Shannon to approximate the maximum transmission rate [29], we designed the numerical simulation program and the simulation results are shown in Fig. 2. To save the space of this paper, the detailed program does not presented and just show the main parameters. Let n_0 and n_s be the power spectral density (PSD) of the narrow-band Gaussian white noise and the transmitted signal which approximates the white noise, respectively, and n_s/n_0 represents the SNR . Different SNR can be obtained by changing n_s while keeping n_0 unchanged. . At the same time, both the bandwidth of the transmitted signal and the noise are set to be W and there are $P_0 = \sigma_0^2 = Wn_0$ and $P_{si} = \sigma_{si}^2 = Wn_{si}$. In addition, according to the sampling theory, in the simulation the transmitted signal $x(t)$, the noise $n(t)$ and the received signal $y(t) = x(t) + n(t)$ are represented as $x(t_i)$, $n(t_i)$ and $y(t_i) = x(t_i) + n(t_i)$ represently, where $i=1, \dots, 2TW$ and T is the duration of the signal. The demodulation progress is shown in Fig. 2(a). Obviously, in the simulation T is impossible to be very long, then the program must be repeated many times to ensure the simulation results approximate the theoretical formula.

Researching comprehensively formula (1) and the coding theorem given by Shannon, we can say that Shannon established a complete information communication theory with mathematics and physics for us 70 years ago! Today, 70 years later, looking back on Shannon's contribution, how can we not admire its greatness! From quantum physicists, we can see Shannon's greatness in another way.

The entropy was firstly employed to describe the change of the temperature under thermal injection by German Physicist R. Clausius in 1865 [30], and the statistical physical expression was proposed by L. Boltzmann in 1877 [31]. Until it was used to measure the information, people really realized that the entropy is a measurement of the disorder of the system. For this reason, E. Jaynes regarded the entropy of the thermodynamics an application of Shannon's information theory [32]. With the development of astrophysics and micro-physics, uncertainty principle proposed by W. Heisenberg in 1927 has become more and more universal [33]. The information theory is therefore no longer limited to the engineering theory but gradually becomes a basic theory, which has a solid mathematical support and can be used in many disciplines.

We should not only praise Shannon's contribution, but also enrich and develop his thoughts and methods. This is also the ultimate goal of this paper

2.2 Theoretical Derivation of Breaking through the Shannon Limit

We use the delay overlapping method as the basic means to break through the Shannon limit. The implementation of this method is actually very simple. It is only necessary to copy the discrete sampling points of continuous signal on the basis of Shannon limit formula deduced in the previous section, and then insert the newly obtained $2TW$ sampling points into the interval of the original $2TW$ sampling points to form a composite signal composed of $4TW$ sampling points, as shown in Fig. 2.

The following formula can explain the principle of this method more clearly. Let two continuous signals be $\hat{x}_1(t_{2j+1})$ and $\hat{x}_2(t_{2j} - \tau)$, where $j=0, 1, \dots, 2W-1$ and $\tau = \frac{1}{2} \frac{1}{2W}$. This indicates that the sampling points of the second signal delays half of the sampling interval comparing with the first signal. When the two signals are added, there is

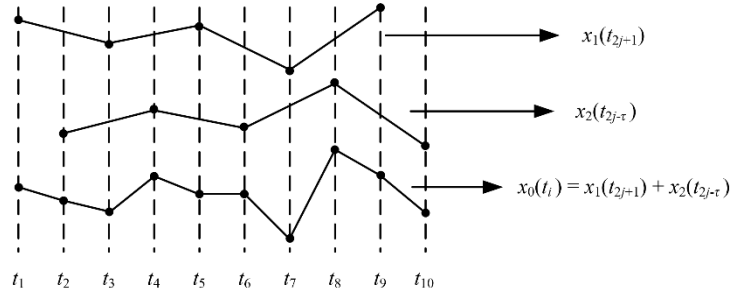


Fig. 2 Diagram of the time-delay overlapping method.

$$\hat{x}_0(t_i) = \hat{x}_1(t_{2j+1}) + \hat{x}_2(t_{2j} - \tau), \quad (10)$$

where $i=1, \dots, 4TW$, $j=0, 1, \dots, 2TW-1$ and $\hat{x}_1(t_{4TW}) = \hat{x}_2(t_0) = 0$.

This means that the sampling points of the second signal are inserted into intervals of sampling points of the first signal in turn. It is specified that the two signals have the same characteristics in terms of bandwidth, PSD and average power, and the sampling points of the two signals are independent of each other¹. The corresponding random variables are denoted as $\hat{X}_1$ and $\hat{X}_2$, and the entropies are $H(\hat{X}_1)$ and $H(\hat{X}_2)$. It should be noted that according to the linear characteristics of the Fourier transform, the bandwidth of the synthetic wave $\hat{x}_0(t_i)$ is same with the bandwidth of the wave composed alone by $\hat{x}_1(t_{2j+1})$ or $\hat{x}_2(t_{2j} - \tau)$. That is to say, the bandwidth of $\hat{x}_0(t_i)$, $\hat{x}_1(t_{2j+1})$ and $\hat{x}_2(t_{2j} - \tau)$ is equal.

In the receiver, the signal is expressed as

$$\hat{y}_0(t_i) = \hat{x}_0(t_i) + \hat{n}_0(t_i) = \hat{x}_1(t_{2j+1}) + \hat{x}_2(t_{2j} - \tau) + \hat{n}_0(t_i), \quad (11)$$

where $\hat{n}_0(t_i) = \begin{cases} \hat{n}_1(t_{2i+1}) \\ \hat{n}_2(t_{2i}) \end{cases}$, $i = 1, \dots, 4TW$. The mean power of the noise is

¹ When the discrete sampling points in the waveform pass through the channel, as long as the "ideal" channel estimation and equalization are used, it can be considered that the adjacent sampling points will not form mutual interference, and independent sampling points can be obtained in the receiver. Of course, this is an ideal state. There will still be some interference during the implementation of the project, but it can be ignored.

$$\hat{N}_0 = \left\{ \left[\hat{n}_1(t_{2i+1}) \right]^2 + \left[\hat{n}_2(t_{2i}) \right]^2 \right\} / 4TW = \left[\hat{n}_1(t_{2i+1}) \right]^2 / 2TW = N. \quad (12)$$

This means that the time-delay overlapping method only changes the synthetic wave and does not change the noise. Then there are

$$I(\hat{Y}_i, \hat{X}_i) = H(\hat{Y}_i) - H(\hat{Y}_i / \hat{X}_i) = H(\hat{Y}_i) - H(\hat{N}_i), \quad (13)$$

$$I(\hat{X}_0, \hat{Y}_0) = H(\hat{X}_0) - H\left(\frac{\hat{X}_0}{\hat{Y}_0}\right) = H(\hat{Y}_0) - H\left(\frac{\hat{Y}_0}{\hat{X}_0}\right) = H(\hat{Y}_0) - H(\hat{N}), \quad (14)$$

and

$$R_0 = I(\hat{X}_0, \hat{Y}_0) \cdot \frac{4TW}{T} = \left[H(\hat{X}_0) - H\left(\frac{\hat{X}_0}{\hat{Y}_0}\right) \right] \cdot 4W = \left[H(\hat{Y}_0) - H\left(\frac{\hat{Y}_0}{\hat{X}_0}\right) \right] \cdot 4W. \quad (15)$$

Thus, the channel capacity of the time-delay overlapping method is

$$C_0 = R_{0\max} = I(\hat{X}_0, \hat{Y}_0) \cdot 4W. \quad (16)$$

Due to $p(\hat{x}_0) = p(x)$, $p(\hat{y}_0) = p(y)$, we get

$$C_0 = \max \left[H(\hat{Y}_0) - H\left(\frac{\hat{Y}_0}{\hat{X}_0}\right) \right] \cdot 4W = 2W \log \frac{\alpha P + N_0}{N_0}. \quad (17)$$

The comparison between formula (17) and (7) should be carried out under the same condition of signal power and noise power. Due to the nonlinear characteristics of the signals and the time-delay structure between the signals, their amplitudes can be canceled each other in many time periods. Therefore, the power of the synthetic wave will not linearly increase with the overlapping of the signals. The power can be expressed as $\hat{P}_0 = \alpha P < \hat{P}_1 + \hat{P}_2$. The value of α can be determined by numerical simulation, and its value range is $0 \leq \alpha \leq 2$. We call this phenomenon power efficiency and will demonstrate it in the next section. This phenomenon means that the power of the synthesized signal can be reduced by using the delay overlap method when the number of bits carried by the subcarriers of the symbol remains unchanged.

Moreover, we have proved $\hat{N}_0 = N$. For this reason, we can rewrite the formula (17) as

$$C_0 = \max \left[H(\hat{Y}_0) - H\left(\frac{\hat{Y}_0}{\hat{X}_0}\right) \right] \cdot 4W = 2W \log \frac{\alpha P + N}{N} = 2W \log(1 + \alpha P / N). \quad (18)$$

Obviously $C_0 > C$. This indicates that the Shannon Limit has been broken. The above derivation can be easily extended to the case of L signals with the same probability density function, as shown below

$$C_L = LW \log \frac{P_L + N}{N} = LW \log \frac{\beta P + N}{N} = LW \log(1 + \beta P / N), \quad (19)$$

where $P_L = \sum_{l=1}^L P_l$ and $0 < \beta \leq L$.

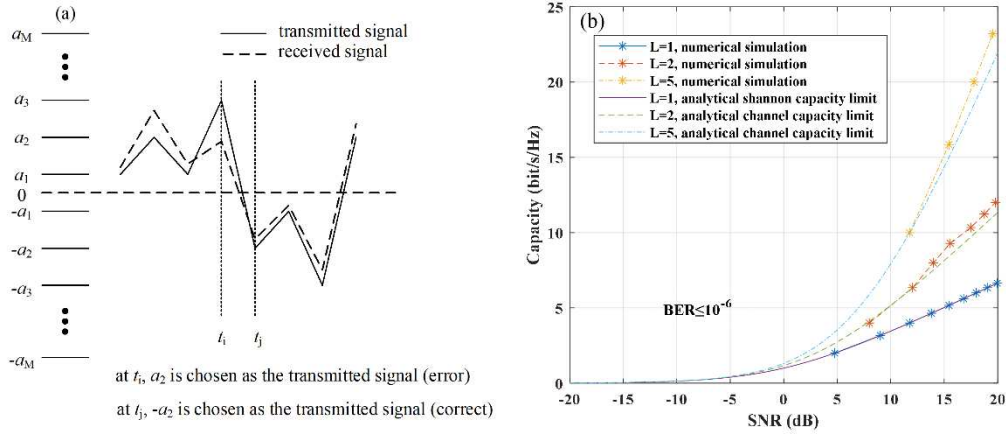


Fig. 3 Channel capacities based on the time-delay overlapping method. (a) Demodulation process proposed by Shannon. (b) Theoretical curve of the Shannon Limit (The purple curve represents the limit curve given by Shannon in another article^[29]) and the numerical simulation results.

Above process describes a method to break through the Shannon Limit. In order to distinguish, we call the channel capacities described by Eq. (7) and Eq. (19) the Shannon Capacity (ShC) and the Time-delay overlapping Shannon Capacity (TDSHC), respectively. The simulation results of the channel capacity corresponding to Eq. (19) are shown in Fig. 3 under three different conditions, i.e. $L = 1, 2$, and 5 . We can draw some conclusions as follows. Firstly, the TDSHC is larger than ShC when L is bigger than 1 , which means that the Shannon Limit has been broken through. Secondly, when L increases, the TDSHC is obviously enhanced, which indicates that large L will bring a higher channel capacity.

A question will be asked: will L be increased infinitely? The time-delay overlap method to construct waveform means to increase the sampling rate, which is equivalent to increasing the bandwidth. However, the signal bandwidth has been limited, and increasing the bandwidth results in the increase of white noise. According to Eq. (8), when the bandwidth is infinite, the Shannon limit of a single wave is $1.44 P / N$, so formula (24) becomes

$$\lim_{W \rightarrow \infty} C_L = 1.44L \cdot \frac{\beta P}{n_0} \quad (20)$$

3. An Example to Break Through the Shannon Limit -- Digital communication technology based on the time-delay overlapping modulation method

In 1997, authors started to explore a more efficient modulation method for digital communication. In 2000, the patent “Phase-shift overlap codes and decode method” (grant no. 1151641) is authorized. Then, authors implemented lots of theoretical studies, numerical simulations and physical experiments, and acquired many new patents. Among those patents, the patent “Time-frequency-phase mixed multicarrier modulation method” (grant no. 101662437) reflect more comprehensively the idea of constructing high-efficient digital modulation technique through time-delay overlapping method. After that, we call this

modulation technique the time-shifting non-orthogonal modulation technology (TS-NMT). During the research of TS-NMT, we gradually realized that the waveform construction method based on time-delay overlapping in TS-NMT is a practical method to break through the Shannon Limit. In recent years, on the foundation of TS-NMT, we propose a time-frequency pulse shaping based delay multicarrier modulation and demodulation method (TFP-TDMC), so that the transmission efficiency of the channel is further greatly improved. Computer simulation shows that TFP-TDMC can achieve 5G rate on 4G channel.

3.1 The Waveform Structure of TS-NMT

Firstly, the waveform of the TS-NMT symbol constructed by the time-delay overlapping modulation technique is shown in Fig. 4. A single sinusoidal or cosine wave is called the base sub-carrier, and multiple subcarriers with identical frequency form a sub-channel on the frequency domain. It is similar with OFDM belong to the multi-carrier category. There are many forms of TS-NMT and they can be uniformly expressed as follows

$$\left. \begin{aligned} g_t(t) &= \sum_{h=1}^H a_h \bar{g}_t \omega_h(t - \tau_h) \\ g_r(t) &= \sum_{h=1}^H a_h \bar{g}_r \omega_h(t - \tau_h) \end{aligned} \right\} t \in T, \tau_h \in T_{dh} \quad (21)$$

where subscripts t and r represent the transmitter and the receiver, respectively. T is the period of the symbol of the synthetic wave, and $T_{dh}=T_{d(h+1)}=T_d$ represents the duration of each subcarrier and is called the lifetime. $g_t(t)$ and $g_r(t)$ are the synthetic wave of the transmitted and received symbol, respectively. $a_h \bar{g}_t \omega_h(t - \tau_h)$ and $a_h \bar{g}_r \omega_h(t - \tau_h)$ are subcarriers, in which a_h is the modulation amplitude, $\bar{g}_t \omega_h(t - \tau_h)$ and $\bar{g}_r \omega_h(t - \tau_h)$ are base-subcarriers. Base-subcarriers can be sinusoidal, cosine or other waveforms with no DC component, and ω_h is the angle frequency. It is worth noting that, $\omega_h = \gamma \omega_{h+1}$ and γ can be any rational number, which means that orthogonality between the sub-channels is not needed. τ_h is the time delay, and the starting points of all the subcarriers are in the life-time of the first subcarrier. Obviously, $\tau_0=0$. We can see that there are three differences between TS-NMT and OFDM:

- i. There are time delays among each sub-carrier;
- ii. Orthogonality in the frequency domain is not required;
- iii. The number of the sub-carriers in a single frequency sub-channel is not limited to 2.

Specifically, there is no upper limit of the number of the sub-carriers.

In addition, due to the interference of the channel fading and the noise, $\bar{g}_r \omega_h(t - \tau_h) \neq \bar{g}_t \omega_h(t - \tau_h)$ under normal circumstances. However, $\bar{g}_r \omega_h(t - \tau_h) \approx \bar{g}_t \omega_h(t - \tau_h)$ can be realized through channel estimation and equalization.

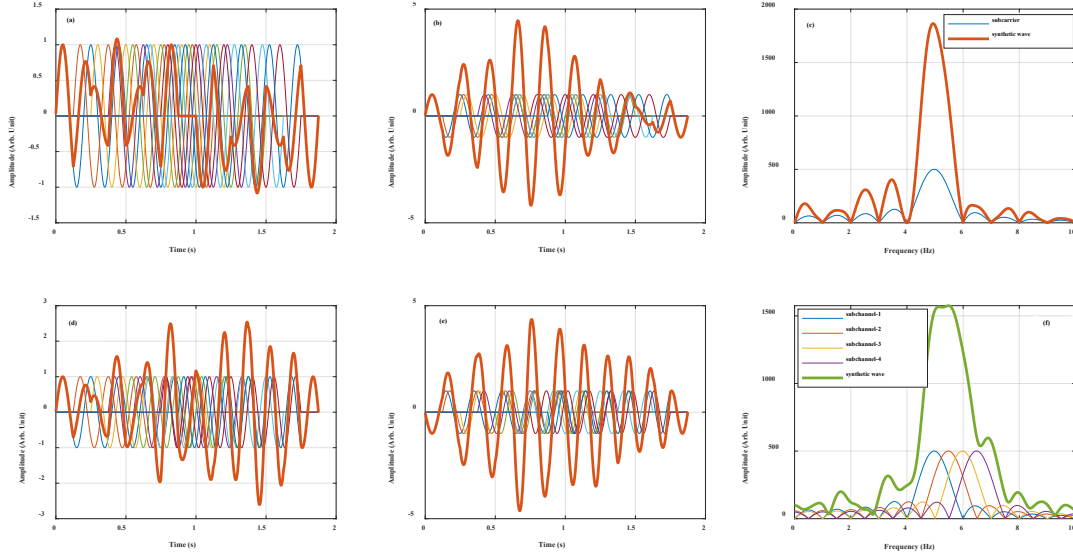


Fig. 4. Waveform structure of TS-NMT. (a) TS-NMT/S constructed by X sub-carriers with same frequency and modulation amplitude (b) TS-NMT/S constructed by Y sub-carriers with same frequency but the modulation amplitude are ± 1 alternately. (c) Spectrum of the synthetic wave of TS-NMT/S and a one subcarrier. (d) TS-NMT/F constructed by the subcarriers in the sub-channels with 4 frequencies, each sub-channel containing two sub-carriers of which amplitudes are +1 (e) TS-NMT/F symbol. All the parameters are same with that in (d) except for the modulation amplitude of each base subcarrier is ± 1 alternately. (f) Spectrum of each frequency sub-channel and the synthetic wave.

3.2 Demodulation Methods of TS-NMT

The demodulation of TS-NMT can be realized through a specific method of solving the set of linear equations. The diagram of TS-NMT symbol with 4 subcarriers (the synthetic wave is not drawn) and the simple demonstration of the demodulation process are shown in Fig. 5. The goal of demodulation is to solve the modulation amplitude of each base-subcarrier from the periodic symbol waveform in turn. The modulation amplitude of a base-subcarrier can be obtained through the integral operation expressed as $\frac{1}{2}a_h = \int_{T_{dh}} a_h \sin \omega_h(t - \tau_h) \sin \omega_h(t - \tau_h) dt$, which is defined as the coherent operation. Due to the overlapping of all the subcarriers in T_{d1} , although it is impossible to distinguish the amplitudes of all the subcarriers through only one coherent operation, we can construct and solve an set of equations to get all the modulation amplitudes.

In the linear set of equation, $\mathbf{R}$ is the demodulation matrix, in which the elements are obtained through the coherent operations on each base-subcarrier, i.e. $\mathbf{R} = [r_{hk}, h, k = 1, \dots, H]$, $r_{hk} = \int_{T_{dhk}} \bar{g}_r \omega_h(t - \tau_h) \bar{g}_r \omega_k(t - \tau_k) dt$. $\mathbf{A}$ is the amplitude vector, and it is consisted of the quantified amplitude as $\mathbf{A} = [a_h, h = 1, \dots, H]$, also is the result of the linear equation set. It can be expressed as

$$\mathbf{A} = \mathbf{R}^{-1} \mathbf{B} \quad (22)$$

where $\mathbf{B}$ is the coherent vector and consists of the results of coherent operations, which is defined as $B_h = \int_{T_{dh}} g_r(t) \bar{g}_r \omega_h(t - \tau_h) dt, h = 1, \dots, H$.

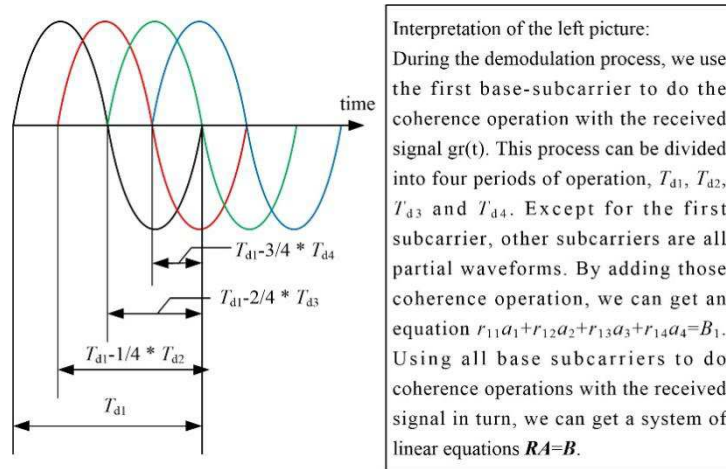


Fig. 5. Diagram of the demodulation of TS-NMT.

3.3 Transmission Rate and Performance Analysis of TS-NMT

In reference [34,35], the performances comparison between TS-NMT and OFDM are analyzed in detail, and the corresponding numerical simulations are given. It is demonstrated that TS-NMT has the advantages of high spectrum utilization, low peak to average power ratio, strong frequency shift resistance and low power consumption compared with OFDM.

Next, we will analyze the feasibility of TS-NMT as a practical method to break through the Shannon Limit. Four aspects are considered: power, transmission rate and efficiency, ill-condition set of the equations and the sampling frequency of the system.

Firstly, it can be seen from Fig. 4 intuitively that the synthetic wave of TS-NMT contains superposition of positive and negative amplitudes during some time sections, which makes the composite amplitude very small or even approximate zero. The numerical simulation results also indicate that the power of the whole synthetic wave doesn't increase proportionally with the number of subcarriers, which means that higher power efficiency can be achieved by TS-NMT. From Fig. 6(a), comparing with OFDM, the increment of the TS-NMT symbol power along with the number of subcarriers is much lower.

In the aspect of data transmission rate, the comparison of the spectral-efficiency between TS-NMT and the Shannon Limit is given in Fig. 6(b). It can see that the spectral-efficiency of TS-NMT exceeds the Shannon Limit, especially in the higher SNR domain. It is worth noting that, for TS-NMT/S, the frequency of all the subcarriers are the same, which means that the TS-NMT/S symbol occupies only one sub-channel. Meanwhile, for TS-NMT/F, although multiple sub-channels are occupied, due to the non-orthogonal between frequency sub-channels, the spectral-efficiency is much higher than any existed multi-carrier modulation technologies, such as OFDM.

Thirdly, as we all know that, there is a problem of ill-condition in the numerical solution of the set of equations, i.e. the perturbation of the equations will bring large deviation between the accurate solution and the numerical solution. For TS-NMT, we proposed a unique method through adding additional waves to solve the ill-condition problem. From Fig. 6(c) and (d), it can see that when there exist additional waves the channel capacity can be further increased and the frequency of error is further decreased.

Last but not least, TS-NMT requires a higher sampling frequency than that specified in the sampling theorem, and the transmission rate is directly proportional to the sampling frequency. Therefore, some researchers believe that the sampling theorem used by Shannon is not perfect. This paper does not draw this conclusion. In fact, according to equation (7) and signal uncertainty principle, under the framework of signal uncertainty principle, $t \rightarrow \infty$ can ensure the completeness of sampling theorem for continuous signals.

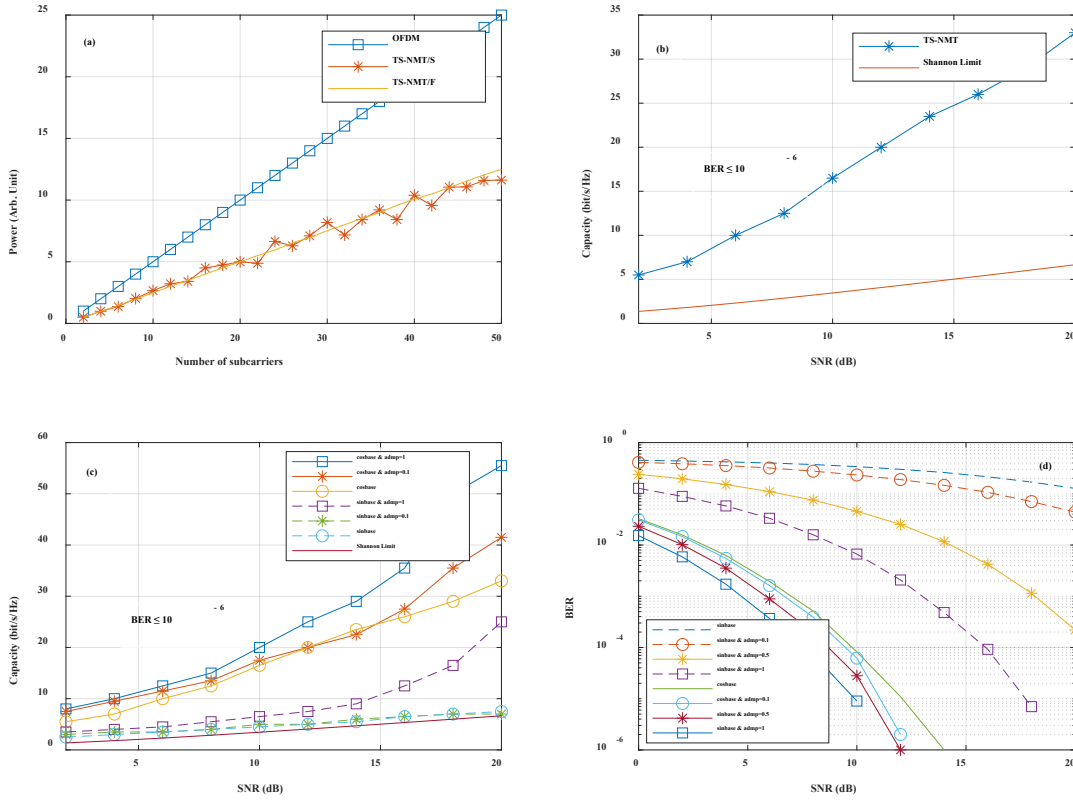


Fig. 6. Simulation results of the performance of TS-NMT. (a) Comparison of the symbol power of OFDM, TS-NMT/S and TS-NMT/F. (b) Channel capacity of TS-NMT. (c) Channel capacity of TS-NMT with additional waves. (d) BER performance of TS-NMT.

3.4 Time Delay Multi-carrier Modulation Based on Time -Frequency -Pulse Shaping (TFP-TDMC)

TFP-TDMC is the evolution of TS-NMT, and it also belongs to the multi-carrier time-delay modulation. Compared with TS-NMT, the main difference lies in the different structure of base-subcarrier. A subcarrier of TFP-TDMC is divided into three parts in time: the front information wave, the main wave and the back information wave. Among them, the main wave does not

carry information, that is, once its amplitude is determined, it will not change any more, and only the amplitude of the information wave will change randomly according to the quantization amplitude levels. Figure 7 shows an example of a subcarrier.

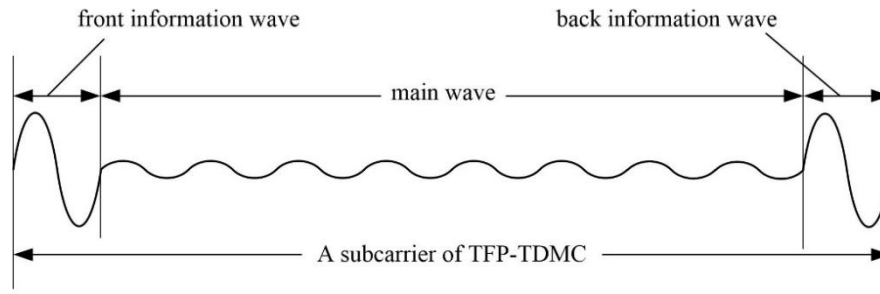


Fig. 7. Structure of the subcarrier of TFP-TDMC.

The modulation process of TFP-TDMC is similar to that of TS-NMT, and the structure of a TFP-TDMC symbol containing two groups of subcarriers is shown in Fig. 8. In the symbol of the TFP-TDMC, each group of information waves is equivalent to a TS-NMT symbol, and there can be many groups of TS-NMT in a single symbol of TFP-TDMC. In the demodulation process, each group of information waves is extracted from the TFP-TDMC symbol and demodulated to recover the transmitted message.

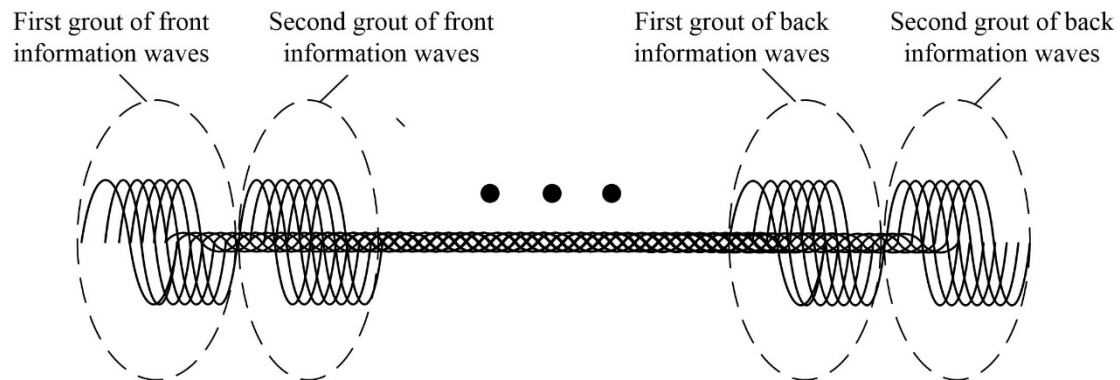


Fig. 8. Structure of the symbol of TFP-TDMC.

The performance of TFP-TDMC with various symbol parameters are simulated and shown in Fig. 9. Meanwhile, the peak data transmission rate and spectral efficiency corresponding to various parameters are shown in Table I. Firstly, thanks to the high spectral efficiency, TFP-TDMC can achieve a 10Gbps peak data transmission rate in 20MHz bandwidth, which is the peak data rate of uplink transmission in 5G. In other words, the TFP-TDMC can achieve the peak data transmission rate of 5G in a bandwidth of 4G, and the spectral efficiency is 500 bit/s/Hz, which is 100 times of the spectral efficiency of 4G. What's more, the rate in Table I is calculated in the single-input-single-output (SISO) system. When multiple-input-multiple-output (MIMO) is introduced to TFP-TDMC, the data transmission rate will be further increased. Meanwhile, according to the 5G standard 3GPP TS 38.101-1, the data transmission rate in a bandwidth of 100 MHz can be 1.57 Gbps (4×4MIMO) , and the TFP-TDMC can

achieve more than 50 Gbps in the same bandwidth with SISO, which is nearly 10 times of the 5G standard.

It is worth noting that, to achieve above peak data transmission rates, the TFP-TDMC requires an extremely high sampling rate², which is a remarkable characteristic of TFP-TDMC. For instance, to achieve a 10 Gbps data rate in 20 MHz bandwidth, the sampling rate requires to be nearly THz, which can hardly be achieved in actual engineering. However, symbol parameters of TFP-TDMC have a strong flexibility, and the sampling rate can be decreased through properly adjusting symbol parameters, meanwhile maintaining a relatively high spectral efficiency and data transmission rate. According to Table I, when the group is set to be 3, and each group contains 30 sub-carriers, TFP-TDMC can achieve a data transmission rate of more than 400 Mbps, or even 500 Mbps in 20 MHz bandwidth, and both the spectral efficiency and data rate is 4-5 times of that of 4G. At this time, the sampling rate is 8 GHz, and it can be achieved in actual engineering. Under same parameters, TFP-TDMC can achieve up to 2.9 Gbps in 100 MHz bandwidth with SISO, which is similar to the data rate of 5G using MIMO technology.

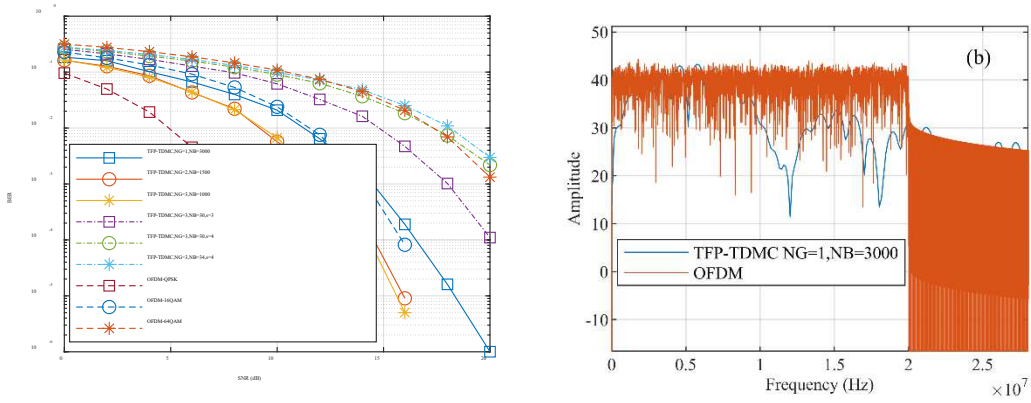


Fig. 9. Performance (a) and spectrum (b) of TFP-TDMC.

Table I. Parameters of TFP-TDMC

System Parameters			Bandwidth				Spectral Efficiency (bit/s/Hz)
			20 MHz		100 MHz		
Number of subcarriers per group (NB)	Number of groups (NG)	Number of bits per subcarrier	Duration of a symbol (μs)	Bit rate (Gbps)	Duration of a symbol (μs)	Bit rate (Gbps)	
30	3	3	1.37	0.4	0.27	1.9	20
30	3	4	1.37	0.4	0.27	2.6	20
34	3	4	1.37	0.5	0.27	2.9	25
3000	1	2	1.1667	10.28	0.2333	51.42	514.3
1500	2	2	1.2727	9.428	0.2545	47.14	471.4
1000	3	2	1.5556	7.714	0.3112	38.57	385.7

Finally, to demonstrate the superiority of TFP-TDMC, the physical layer communication technology of current 4G-LTE is chosen to be the benchmark. For fairness, the bandwidth and symbol power of 4G-LTE is set to be same with TFP-TDMC. According to 3GPP R8, the orthogonal frequency division multiplexing (OFDM) technology is used in physical layer, and the peak data rate can reach up to 100.8 Mbps with 64QAM for each sub-channel ignoring any channel encoding method. The spectral efficiency of 4G-LTE is 5 bit/s/Hz. The comparison between TFP-TDMC and OFDM with 16QAM and 64QAM is shown in Fig. 9. The performance of TFP-TDMC is similar to that of OFDM with 16QAM, and is better than OFDM with 64QAM. From Fig. 9(b), when achieving a 5G data rate, the bandwidth of TFP-TDMC is nearly 20MHz, and the out-of-band radiation is also similar to OFDM. This again proves that the TFP-TDMC has an extremely high spectral efficiency.

Next, some performance comparison and analysis are made between TFP-TDMC with those of TS-NMT:

(1) Firstly, both belong to non orthogonal multi carrier modulation, and both have much higher spectral efficiency than orthogonal frequency division multiplexing (OFDM). However, TFP-TDMC and TS-NMT have a common feature that they need high sampling rate and increase with the increase of subwave number.

(2) TFP-TDMC not only meets the requirements of 5G for transmission rate, but also meets the requirements of 5G for real-time performance. The period of one symbol of TS-NMT is twice that of OFDM, and it needs to solve the equation and demodulate, which more consumes on time. However, the time width of information wave of TFP-TDMC can be much smaller than that of one symbol of TS-NMT, which can greatly reduce the time consumption. For example, if the time width of information wave is less than 1ms, Therefore, it can meet the real-time requirements of 5G.

(3) Although the main wave of TFP-TDMC does not carry information, it can reduce the power of synthetic wave due to the delay dislocation of different waves

(4) Compared with TS-NMT, TFP-TDMC can reduce the sampling rate without reducing the transmission rate by reducing the number of waves of information wave and increasing the amplitude of subwaves and the number of groups of information waves while keeping the synthetic wave power unchanged.

Discussion

In the research of TS-NMT for more than 20 years, we have gradually realized that the delay overlapping method is an effective technique to improve the transmission efficiency. We have long believed that this technology can break through the Shannon Limit^[34]. However, for the sake of caution, for a long time we did not announce breaking Shannon limits.

In fact, we can explain the efficiency of TS-NMT more intuitively. A single physical channel is divided into several logical channels by using the method of time delay overlapping. Then, the transmission efficiency of a single physical channel can be improved by gathering all the modulated logical channels.

A hypothesis

Through the research on breaking the Shannon limit, especially after combining the uncertainty principle in quantum mechanics with our research, we find that the scientific research on information in modern science is not deep enough. Shannon only gave the measurement of information, but did not answer what is the physical essence of information. So far, people have always believed that the world is composed of matter, which can be infinitely divided into smaller elementary particles. In Jim's book 《The Quest to Understand Matter from Greek Atom to Quantum Fields》, he comprehensively summarizes human views on matter from ancient Greek atomic theory to today's physics. Finally, he concluded: “the greatest feature of the universe is the energy of quantum fields, not hard, impenetrable atoms.” Such an answer is not satisfactory. We can ask: in what form do the electric fields and energy exist?

We will answer this question with the assumption that the universe is a whole interconnected by infinite hierarchical information groups. The premise of this hypothesis is to admit that the relativistic theory of constant speed of light, the uncertainty principle of quantum mechanics, the big bang and cosmic expansion theory are correct.

In his book *the Tao of Physics —An Exploration of Parallels between Modern Physics and Eastern Mysticism*, American physicist f. kapura first discussed the cosmology, and believed that eastern mysticism emphasized the cosmology of universal unity at the beginning. The west views the universe with scientific thinking method. It can be roughly divided into two stages: classical physics represented by Newtonian physics and modern physics. Classical physics is a cosmology of separation and mechanism, which has not changed into a cosmology of universal unity until modern physical mechanics. We believe that although Oriental mysticism emphasizes the unity of the universe, compared with science, it lacks detailed analysis methods and cannot express the mathematical model given by scientific methods in ordinary language. This paper will explain our hypothesis with the method of scientific thinking.

In scientific methods, firstly should distinguish different objects and look for the relationship between objects, and then describe them through various rational thinking methods such as logical reasoning, finally explain them in general language. In order to further elaborate this theory, needing to simply sort out the process of human cognitive development. A baby starts by distinguishing light and light shadows, then identifies individual objects and their simple connections. From children to adults, through learning and practice, they have the ability to understand all kinds of things (including abstract ability). The whole humanity needs the intellectual accumulation of many generations to improve its cognitive ability. It can be seen that distinguishing objects is a basic ability. That is why the first step of scientific methods needs to separate objects, and then find the regularity of things by looking for the relationship between objects. It should be noted that distinguishing "objects" It is a basic ability for human beings to understand the objective world. But it does not mean that the object they know is the inevitable existence of the objective world³. However, if there aren't any objects the universe observed by human beings cannot be described more clearly, completely and consistently. For example, when people understand the cloud, they first see a large area cloud that seems to be continuous, then divide it into different parts according to different gray levels, and then analyze

the behavior of different parts of the cloud, such as whether it will rain? It was not until later that humans realized that clouds were mainly composed of unevenly distributed water molecules. With the development of science and technology, it is now possible to more accurately predict the meteorological conditions of the vast sky (so far, the weather forecast can only increase the probability of accuracy gradually, but it is impossible to be completely sure). The question is whether the object comprehended by human must be material. Today, although humans have built a huge electron Collider, they have only found some new quark particles at different levels. That is to say, since no smaller particles can be found, it is concluded that "the greatest feature of the universe is the energy of the quantum field". This is not a thinking in line with scientific methods. Energy and energy-field no longer contain other smaller objects, so we can't give clearer reasoning according to scientific methods. This is no different from kapura's thinking method of "Oriental mysticism".

We believe that the foundation of the universe is information! The information exchange between organisms has been recognized by the scientific community. In a broad sense, there is also information transmission between non creatures. For example, in sunlight, the thermal expansion and cold contraction of uneven parts of rock will lead to mechanical changes, and the infiltration of water into cracks will lead to chemical changes. Finally, the rock may turn into sand (biological effects are ignored). Among them, thermal expansion and contraction are mainly caused by infrared thermal radiation in sunlight. The thermal radiation formed by infrared belongs to electromagnetic radiation; Chemical changes also send signals. It can be assumed that all signals are presented in the form of electromagnetic waves. It is in a variety of signal flows that colorful changes are presented and colorful nature is presented in front of mankind! A rough model of the universe composed of hierarchical information clusters is given below. We see the moment before and after the big bang as the boundary at the beginning of the universe. Before this boundary is a black hole, followed by the beginning of the universe composed of countless smallest information groups. Each information group switches to each other at a very high speed in the two states, and the two states are recorded as 0 and 1 respectively. The information group is a flexible sphere. All information groups are seamlessly squeezed in the universe and vibrate in a local range. It can be imagined as the motion state of water molecules in the container. It is called layer 0 before the creation of the universe and layer 1 composed of the smallest information group. In the first layer, all information groups collide with each other in high-speed vibration, resulting in mutual attraction or repulsion. The mutual attraction is combined into some large information groups,² and many isolated small information groups are retained to form the second layer information group; Such repetition forms a layer of gradually increasing information. Information clusters can be distinguished by different codes. When the information group reaches the level that human beings can distinguish clearly, it is called matter, and then there is the concept of quality. Each layer will retain some small information groups from the lower layers (even layer 0).. These small information galaxy clusters may be the dark matter considered by the physical world today. Steve Hawking once gave a light cone model in *Brief History of Time*. Following Hawking's

² We believe that human beings cannot fully and truly understand the whole universe with scientific methods. There is a simple explanation as follows. Mathematical models given by scientific methods always have boundary conditions. If the universe is expanding, the boundary conditions should change, but human cognition has no ability to follow the change speed of boundary conditions.

light cone, we construct a rough model diagram of the universe composed of hierarchical information groups (Fig. 10).

The combination of information groups can be described by 2^n , where $n=1, \dots, N$.

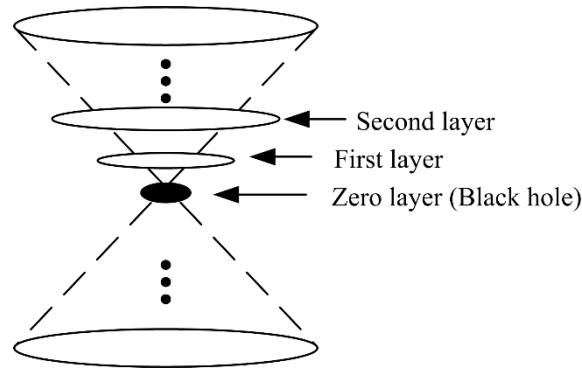


Fig. 10. Hierarchical information cone.

According to the above, we believe that the origin of the universe is information. If there is no information exchange, all particles can only be static, and the universe will not appear! The view that "the biggest unified feature of the universe is the energy of quantum field" is obviously flawed. Because it is called a "quantum field", it should have quantum motion and then produce energy. Quantum motion will produce information exchange between quantum. Information exchange will make information groups attract or repel each other, so the upper layer will produce larger information groups, while the lower layer will retain some unbound information groups. The effect of mutual attraction or repulsion is random, that is, it follows the uncertainty principle of quantum mechanics.

According to the theory of modern physics, hydrogen particles move at the speed of light in vacuum. In our hypothesis, the so-called "vacuum" is not a "vacuum" without anything. It is full of small information groups moving at the speed of light. The high-speed movement of these small information groups will produce signals. At present, the field of signal processing has been facing the problem of noise interference. In the face of the signal generated by the small information group in the vacuum, there is no noise problem. All noise is a useful signal for new particles that will be found. We used the non orthogonal frequency Fourier method to approximate the continuous function, and had prove that it has higher accuracy than the orthogonal Fourier method. We also assume that the upper layer of the information cone will always retain a large number of small information groups that originally belong to the lower layer. For example, particles now considered dark matter may be these small sets of information that remain from the bottom. If we use the method of signal processing, it is possible to perceive these small information groups and their relationships.

Scientific hypotheses need to be proved by experiments and described by mathematical methods. Unfortunately, the author is not a physicist and has no experimental conditions. Only two assumptions can be given. Firstly, we think it should look for evidence of chemical reaction information coding in the process of transformation from inorganic matter to organic matter. For example, looking for electrical signals from chemical reactions, and then converting them into quantitative coding by imitating the method of digital communication, different codes may be obtained in different chemical reactions. Is it possible to get some inspiration from it? If we

can observe and analyze the smallest particles from the perspective of signal, it will greatly reduce the cost of scientific research, rather than the current method of directly decomposing substances from top to bottom. Now, in order to find smaller particles, building a huge electron collider is an extremely huge investment! If we can find smaller particles from the signal, our main energy will be used to make more sensitive sensors. Noise in modern signal processing seems to be able to provide useful information for our discussion. We can also find some signal analysis method. This method will be used to decompose the signals sent by large information groups to obtain the signals of smaller information groups in the next layer. In this way, it is possible to gradually go deep into the signal of the lower information group

References

- [1] M. Shafi, A. F. Molisch, P. J. Smith, et al. 5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice. *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201-1221, June 2017.
- [2] I. Parvez, A. Rahmati, I. Guvenc, et al. A Survey on Low Latency Towards 5G: RAN, Core Network and Caching Solutions. *IEEE Communications Surveys & Tutorials (Early Access)*, 2018.
- [3] S. Buzzi, C. I. T. E. Klein, et al. A Survey of Energy-Efficient Techniques for 5G Networks and Challenges Ahead. *IEEE Journal on Selected Areas in Communications*, vol. 34, no.4, pp. 697-709, April 2016.
- [4] Z. Ding, Y. Liu, J. Choi, et al. Application of Non-Orthogonal Multiple Access in LTE and 5G Networks. *IEEE Communications Magazine*, vol. 55, no. 2, pp. 185-191, Feb. 2017.
- [5] Y. Li, P. Fan, A. Leukhin, et al. On the Spectral and Energy Efficiency of Full-Duplex Small-Cell Wireless Systems with Massive MIMO. *IEEE Transactions on Vehicular Technology*, vol. 66, no. 3, pp. 2339-2353, Mar. 2017.
- [6] B. Ai, K. Guan, R. He, et al. On Indoor Millimeter Wave Massive MIMO Channels: Measurement and Simulation. *IEEE Journal on Selected Areas in Communications*, vol. 35, no.7, pp. 1678-1690, July 2017.
- [7] S. Chen, B. Ren, Q. Gao, et al. Pattern Division Multiple Access—A Novel Nonorthogonal Multiple Access for Fifth-Generation Radio Networks. *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3185-3196, April 2017.
- [8] S. M. R. Islam, N. Avazov, O. A. Dobre, et al. Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges. *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 721-742, Second Quarter 2017.
- [9] KNRSV Prasad, E. Hossain, and VK Bhargava. Energy Efficiency in Massive MIMO-Based 5G Networks: Opportunities and Challenges. *IEEE Wireless Communications*, vol. 24, no. 3, pp. 86-94, Jun. 2017.

- [10] B.P. Rimal, D. P. Van, and M. Maier. Mobile Edge Computing Empowered Fiber-Wireless Access Networks in the 5G Era. *IEEE Communications Magazine*, vol. 55, no. 2, pp. 192-200, Feb. 2017.
- [11] E. Baser. Index Modulation Techniques for 5G Wireless Networks. *IEEE Communications Magazine*, vol. 54, no.7, pp. 168-175, Jul. 2016.
- [12] Z. Ding, P. Fan, and H.V. Poor. Random Beamforming in Millimeter-Wave NOMA Networks. *IEEE Access*, vol. 5, pp. 7667-7681, 2017.
- [13] H. Kim. Coding and Modulation Techniques for High Spectral Efficiency Transmission in 5G and Satcom. *23RD EUROPEAN SIGNAL PROCESSING CONFERENCE (EUSIPCO)*, 2015.
- [14] S. Chen, K. Peng, Y. Zhang, et al. Near Capacity LDPC coded MU-BICM-ID for 5G. *INTERNATIONAL WIRELESS COMMUNICATIONS & MOBILE COMPUTING CONFERENCE (IWCMC)*, 2015.
- [15] H. Gamage, N. Rajatheva, and M. Latva-Aho. Channel Coding for Enhanced Mobile Broadband Communication in 5G Systems. *EUROPEAN CONFERENCE ON NETWORKS AND COMMUNICATIONS (EUCNC)*, 2017.
- [16] L. Dong, H. Zhao, Y. Chen, et al. Introduction on IMT-2020 5G Trials in China. *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 8, pp. 1849-1866, Aug. 2017.
- [17] Telatar E. Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunication*, vol. 10, no. 6, pp. 585-596, 1999.
- [18] P. F. Driessen, G. J. Foschini. On the capacity formula for multiple input-multiple output wireless channels: a geometric interpretation. *IEEE Transactions on Communications*, vol. 47, no. 2, pp. 173-176, 1999.
- [19] G. Gampala, and C. J. Reddy. Massive MIMO - Beyond 4G and a Basis for 5G. *International Applied Computational Electromagnetics Society Symposium (ACES)*, 2018.
- [20] X. Gao, L. Dai, S. Han, et al. Energy-Efficient Hybrid Analog and Digital Precoding for MmWave MIMO Systems with Large Antenna Arrays. *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 998-1009, Apr. 2016.
- [21] N. Liang and W. Zhang. Mixed-ADC Massive MIMO. *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 983-97, Apr. 2016.
- [22] D. Yue and G. Wang. On the Approximation of MIMO Shannon Capacity Over Spatially Correlated Ricean Fading Channels, *First International Conference on Communications and Networking in China*, 2006.
- [23] C. E. Shannon. A mathematical Theory of Communication. *Bell System Technical Journal*, vol. 27, no. 3, pp. 379-423, Jul. 1948.

- [24] Cover, T.M. and Thomas, J.A. (1991) Elements of information theory. Wiley, New York.
- [25] Hartley R.V.L. Transmission of Information. Bell Systems Technical Journal, vol.7, pp. 535, 1928.
- [26] Peng Wang, and Jianping Li. Quantum Interpretation of Signal's Uncertainty Principle. Journal of University of Electronic Science and Technology of China, vol. 37, no. 1, pp. 14-16, 2008. (in Chinese)
- [27] L. Cohen. (1995) Time-frequency Analysis Theory and Application. Prentice Hall, New York.
- [28] M. Rosler, M. Voit. An Uncertainty Principle for Hankel Transforms. Proceedings of the American Mathematical Society, vol. 127, 1998.
- [29] C. E. Shannon. Communication in the Presence of Noise. Proceedings of the I.R.E. vol. 37, no. 1, pp. 10-21, Jan. 1949.
- [30] R. Clausius. (1865) The Mechanical Theory of Heat – with its Applications to the Steam Engine and to Physical Properties of Bodies. John van Voorst, London.
- [31] L. Boltzmann. On the Relationship between the Second Fundamental Theorem of the Mechanical Theory of Heat and Probability Calculations Regarding the Conditions for Thermal Equilibrium. Wien. Ber. Vol. 7, pp. 373-435, 1877.
- [32] E. Jaynes. Information Theory and Statistical Mechanics. Physical Review, vol. 106, no. 4, pp. 620-630, May 1957.
- [33] W. Heisenberg. The Actual Content of Quantum Theoretical Kinematics and Mechanics. Z. Phys., vol. 43, pp. 172, 1927.
- [34] Dequn Liang, Mingyan He, Qisheng Cao, Ganna Liu. A Multi-Carriers Modulation Technology with Higher Spectral Efficiency and Stronger Capability Tolerating the Frequency Shift, Intelligent Signal Processing and Communication Systems (ISPACS 2011), Chiangmai, Thailand, 7-9 December, 2011.
- [35] Qisheng Cao. Research on Modulation in Digital Communications Based on Non-orthogonal Ideas. Ph. D. Thesis (in Chinese), pp. 126, 2008.

Acknowledgement

First of all, we would like to express our deep gratitude to Mr. Sun Changnian. Since 1997, he has provided professor Liang with the opportunity to explore efficient digital modulation methods. He was also the first one to support the initial idea of TS-NMT. When the rationality of TS-NMT was proved by simulation, we have been cooperating for a long time since 2005 with the goal of using TS-NMT for xDSL. Mr. Sun Changnian was responsible for the development of hardware prototype and commercial financing. Unfortunately, due to the lack of funds and human resources, we have not been able to enter the business road! However, a prototype for experiment has been made, which helps us to complete the verification of short-

distance wired and wireless physical channels. We also thank three doctoral students and 33 master's students from Xi'an Jiaotong University and Dalian Maritime University who have done a lot of computer simulation and physical channel experiments from 1997 to 2010. Funding: This research was supported via the Natural Science Foundation of China (NSFC) (60272017, 60772160). Author contributions: Dequn Liang wrote the first draft and reviewed the final draft. Xinyu Dou completed the computer simulation and editing of the full text mentioned in the article. Competing interests: Dequn Liang is the author of two international patents related to this work (application. no. 00123342.4 and 200810119412.1). Xinyu Dou declares no competing interests. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper. Additional data related to this paper may be requested from the authors.