For Security and Higher Spectrum Efficiency: A Variable Packing Ratio Transmission System Based on Faster-Than-Nyquist and Deep Learning

Peiyang Song, *Student Member, IEEE*, Nan Zhang, Lin Cai, *Fellow, IEEE*, Guo Li, *Member, IEEE*, Tong Wu, and Feng-Kui Gong, *Member, IEEE*

***Abstract*— With the rapid development of various services in wireless communications, spectrum resource has become increasingly valuable. Faster than Nyquist (FTN) signaling, proposed in the 1970s, is a promising paradigm for improving spectrum utilization. This paper proposes the variable-packing-ratio (VPR)-based transmissions for high spectrum efficiency (SE) and security, respectively. Aided by deep learning (DL)-based estimation, the proposed scheme for high SE can achieve a higher capacity than the conventional Nyquist-criterion transmission with negligible modification to existing communication paradigms (e.g., spectrum allocation or frame structure). More importantly, for VPR-based secure transmission, a dynamic generation scheme is proposed to produce randomly distributed positions to switch the packing ratio, which can effectively avoid detections and attacks. In addition, we propose a simplified DL-based packing ratio estimation for both of these two scenarios so that the receiver can estimate the packing ratio without any in-band or out-band control messages. Simulation results show that the proposed simplified estimation achieves nearly the same accuracy and convergence speed as the original multi-branch fully-connected structure with a complexity reduction of 20 folds. Finally, we derive the SE of the proposed VPR transmission under different channels. The numerical results validate the correctness of the derivation and demonstrate the SE gains of the VPR scheme beyond conventional Nyquist transmission.**

***Index Terms*— Faster than Nyquist signaling, spectrum efficiency, variable packing ratio, deep learning.**

I. Introduction

THE last several decades have witnessed the rapid development of terrestrial wireless communications, including the widely concerned fifth-generation mobile communications (5G) and the increasing demands for data traffic by various communication services.

Faster than Nyquist (FTN) signaling was firstly proposed in the 1970s by *Bell Laboratories* and has been investigated and studied since the 2000s. It is promising to provide a higher symbol rate and spectrum efficiency (SE) in future terrestrial and satellite communications.

In conventional Nyquist-criterion communications, the symbol duration must be set as $T > T_N = 1/(2W)$ to guarantee the performance of the transmission system, where W is the transmission bandwidth. In such scenarios, the receiver can effectively recover the transmitted symbols from received ones benefiting from the strict orthogonality between different symbols. FTN signaling, in contrast, destroys the orthogonality and introduces unavoidable inter-symbol interference (ISI) by applying a smaller symbol duration $T < T_N$. It can improve the transmission rate, at the cost of higher complexity in the receiver to recover the transmitted symbols.

Mazo [1] has proved that the FTN signaling can improve nearly 25% rate than the conventional Nyquist-criterion communications in the additive white Gaussian noise (AWGN) channel without loss of bit error rate (BER) performance, which is known as *the Mazo limit*.

Many pieces of research have been conducted on the signal detection for FTN signaling, including time-domain [2], [3], [4], [5], [6], [7], [8] and frequency-domain [9], [10] algorithms. Also, for sake of available high SE, some researchers attempt to merge FTN signaling with various conventional technologies such as frequency division multiplexing (FDM) [11], [12], [13], [14], [15], multiple input multiple output (MIMO) [16], [17], [18], [19], multi-path fading channel [20], [21], [22], etc. The comprehensive review of the latest study on FTN signaling can be found in [23], [24], and [25]. Especially, [26] firstly derives the analytical-form capacity of FTN signaling, which inspires our work to extends it to more scenarios.

The packing ratio is a key parameter that can directly affect the symbol rate and the strength of ISI. Conventional FTN signaling considers a fixed packing ratio which may not always achieve the maximum capacity during variable transmission conditions. A variable packing ratio (VPR)-based FTN signaling is a promising solution to this issue. Also, although

Manuscript received 26 May 2022; revised 5 October 2022 and 21 November 2022; accepted 12 January 2023. Date of publication 26 January 2023; date of current version 12 September 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 62001354 and in part by the Natural Sciences and Engineering Research Council of Canada (NSERC). The associate editor coordinating the review of this article and approving it for publication was G. Geraci. *(Corresponding author: Feng-Kui Gong.)*

Peiyang Song, Nan Zhang, Guo Li, and Feng-Kui Gong are with the State Key Laboratory of ISN, School of Telecommunications Engineering, Xidian University, Xi'an, Shaanxi 710071, China (e-mail: pysong@stu.xidian.edu.cn; nzhang@xidian.edu.cn; gli@xidian.edu.cn; fkgong@xidian.edu.cn).

Lin Cai is with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8W3P6, Canada (e-mail: cai@ece.uvic.ca).

Tong Wu is with the China Academy of Space Technology (CAST), Xi'an, Shaanxi 710071, China (e-mail: wut40@cast504.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2023.3238174.

Digital Object Identifier 10.1109/TWC.2023.3238174

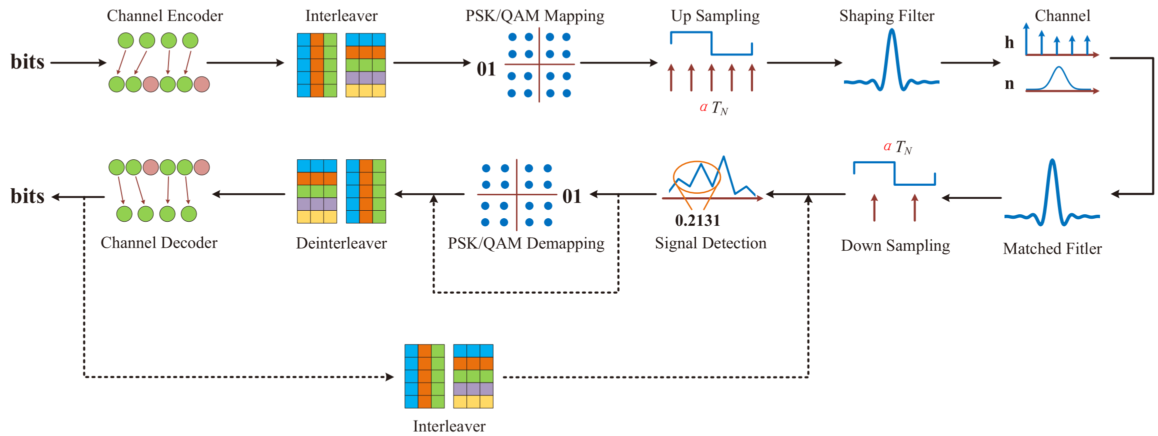


Fig. 1. System model of conventional FTN signaling.

the hopping roll-off factors [27] have been successfully employed to improve the security of communications, the VPR-based secure transmission has not been studied yet. Last but not least, the success of deep learning (DL) in packing ratio estimation inspires us to develop DL and VPR-based transmissions for the high SE and security.

The contribution of this paper can be summarized as follows.

- We present a high SE VPR-based transmission based on FTN and DL. The transmitter can change the packing ratio based on specific conditions (e.g., channel state information (CSI) or cooperative strategy). No in-band or out-band control messages are required to notify the receiver of the packing ratio values, which means the scheme doesn't need to conduct a complex modification for existing communication paradigms (e.g., spectrum allocation and frame structure).
- We propose a VPR-based secure transmission, where the positions to change the packing ratio are secret and known only by the transmitter and the receiver. Also, no control messages are required since the receiver can infer the packing ratio with the DL-based simplified estimation and the information of positions.
- We propose a dynamic generation scheme for positions to change the packing ratio. With the measured CSI between the transmitter and the receiver, randomly distributed positions can be generated, which are secret to any other eavesdroppers.
- We propose a simplified DL-based packing ratio estimation, which achieves nearly the same performance as the original architecture while reducing the computing cost by 20 times.
- We derive the expression of the capacity for the proposed VPR scheme in different channels and validate the theoretical results by Monte Carlo simulations. The derived capacities are also applicable to conventional FTN signaling.
- We conduct comprehensive evaluations and verify the SE gain between the proposed VPR-based and conventional Nyquist-criterion transmissions under different channels. Also, with the same SE, the BER degradations of the proposed VPR-based transmission over FTN signaling are demonstrated to be small enough.

Herein, we give the definition of notations throughout the rest of the paper. Bold-face lower case letters (e.g. $\boldsymbol{x}$) are applied to denote column vectors. Light-face italic letters (e.g. x) denote scalers. x_i is the i-th element of vector $\boldsymbol{x}$. $x(t)*y(t)$ denotes the convolution operation between $x(t)$ and $y(t)$. And $\|\boldsymbol{W}\|_0$ represents the number of non-zero items in matrix $\boldsymbol{W}$.

The rest of the paper is organized as follows. In Section II, we present the system model of FTN signaling. In Section III, the structure of the proposed VPR system is introduced. Section IV presents the proposed dynamic generation for positions of segments. And the simplified DL-based packing ratio estimation is presented in Section V. The capacity of the proposed VPR system under different channels is derived in Section VI. In Section VII, comprehensive simulations are conducted to evaluate the performance and the complexity of the proposed VPR system and the DL-based estimation. Also, the derived capacity for the proposed VPR system is verified. Section VIII concludes this paper.

II. System Model of Conventional FTN Signaling

Fig. 1 illustrates the conventional architecture of FTN signaling. In the transmitter, the signal that has passed through the shaping filter can be written as

$$s(t) = \sqrt{P_s} \sum_{k=-\infty}^{+\infty} x_k h(t - k\alpha T_N), \tag{1}$$

where P_s is the average power of the bandwidth signals, x_k $(k = 0, \pm 1, \pm 2, \cdots)$ is the k-th symbol and α $(0 < \alpha \leq 1)$ is the symbol packing ratio. $h(t)$ is the function of shaping filters. Since the value of the filter function is 0 at every multiple of T_N, when $\alpha < 1$, the filtered symbols are no longer orthogonal and become the weighted sum of several successive symbols.

Corresponding to the shaping filter, a filter with a conjugate structure named matched filter is employed in the receiver to maximize the received symbols' signal-to-noise ratio (SNR).

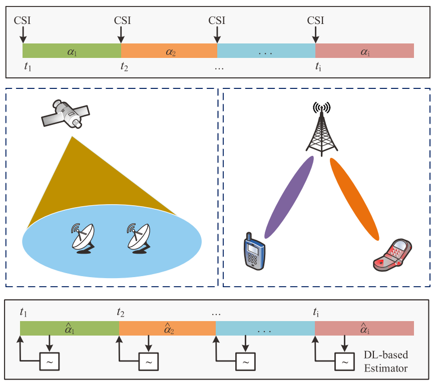


Fig. 2. Architecture of the proposed VPR transmission scheme with CSI as the indicator to adjust the packing ratio.

The filtered symbols can be written as

$$\begin{aligned} y(t) &= (s(t) + n(t)) * h(t) \\ &= \sqrt{E_s} \sum_{k=-\infty}^{+\infty} x_k g(t - k\alpha T_N) + \widetilde{n}(t), \end{aligned} \tag{2}$$

where $g(t) = \int h(x)h(t-x)dx$, $\widetilde{n}(t) = \int n(x)h(t-x)dx$, and $n(t)$ is the Gaussian white noise.

Finally, the samples of the received symbols can be formulated as

$$y_n = \sqrt{E_s} \sum_{k=-\infty}^{n-1} x_k g((n-k)\alpha T_N) + \sqrt{E_s} x_n g(0) \tag{3}$$

$$+ \sqrt{E_s} \sum_{k=n+1}^{+\infty} x_k g((n-k)\alpha T_N) + \widetilde{n}(n\alpha T_N)). \tag{4}$$

Different from the conventional Nyquist-criterion transmission system, each sampled symbol in FTN signaling contains both the expected symbol and the adjacent ones. Meanwhile, due to the non-orthogonality between different samples in the matched filter, the noise becomes colored noise. All these new features make it difficult to recover the original symbols in the FTN receiver.

III. The Proposed Variable Packing Ratio Transmission System

A. System Architecture

As shown by Fig. 2, in the proposed VPR transmission system, the transmitter changes the symbol packing ratio at every specific moment, which divides the transmitted symbols into different segments and results in individual transmission rates within each part. The determination of each packing ratio is based on CSI, cooperative target or other possible strategies. Different from the conventional variable coding and modulation (VCM) schemes [28], the receiver in the proposed VPR system does not need to know precisely the current symbol packing ratio. The only necessary knowledge is when the parameter changes, which can be appointed in advance. Then, the DL-based estimation will help the receiver infer the packing ratio within a short time.

There are two advantages to employ the DL-based estimation for α in the receiver instead of directly sending it by the transmitter. Considering high SE, since no control message is required, the proposed VPR conducts negligible modification to the existing communication paradigms (e.g., spectrum allocation or frame structure). Also, when considering security, if α is put into the frame head, the repeated specific modulation type and UW word will make it easy for the eavesdropper to locate and decode the information.

B. VPR-Based Transmission for High SE

In this scenario, the packing ratio should be determined to balance the SE and the performance constraint (e.g., BER or ISI strength). For example, the SNR where $BER = 10^{-3}$ can be employed as the threshold to select the packing ratio to achieve the maximum SE with acceptable BER performance, as demonstrated in Section VII. Also, the signal-to-interference-plus-noise ratio (SINR) can be applied for the base station and satellite to adjust the packing ratio in non-orthogonal multiple access (NOMA) which divides users into pairs and the multi-beam satellite serving users within a certain area.

The positions when the packing ratio changes can be generated by the following approaches.

1) **Fixed interval.** After the synchronization, the transmitter checks the transmission status after every fixed interval and decides whether to change the packing ratio. The receiver should carry out the estimation at the same positions.
2) **Static storage.** A preset vector of starting positions is determined with the practical characteristics of the transmission environment and the requirement.
3) **Pilot or dedicated channel.** Without considering security, a public object (e.g., dedicated channel or pilot) can directly carry the information of packing ratio, at the expense of extra resources consumed and the modification of existing resources allocation.

C. VPR-Based Secure Transmission

The VPR-based transmission is a promising paradigm to improve the security of communications. For one thing, the change of the symbol packing ratio only affects the baseband symbols and can not be caught by analysis of the frequency spectrum. For another, the blind estimation cannot indicate the accurate starting position. Once the eavesdropper employs a wrong symbol packing ratio, the sampled points will severely deviate from their correct positions, making it meaningless to detect the signals and further estimate the following symbol packing ratio.

In this scenario, the packing ratio should be employed randomly with the same probability to avoid possible detection and attack, as assumed for the roll-off factors in [27]. The positions when the packing ratio changes can be generated by the following approaches.

1) **Static storage.** A preset vector of starting positions should be stored in advance. Although such assumptions have been widely employed [27], it suffers from the

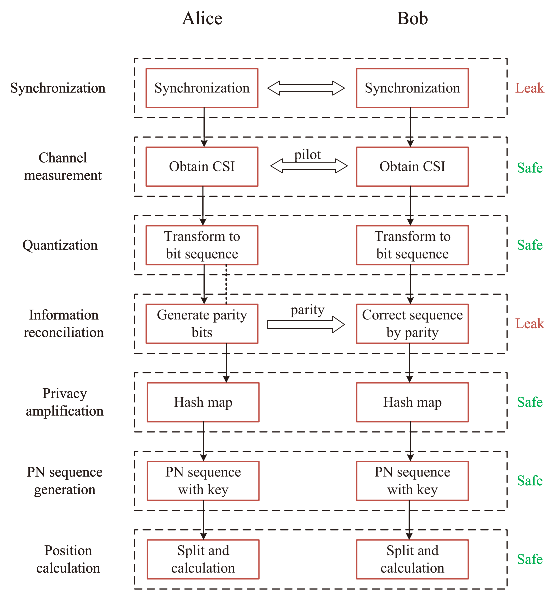


Fig. 3. The proposed dynamic generation scheme for the information of starting positions.

risk that the expected security will disappear once the information is stolen by the eavesdropper.

2) **Dynamic generation.** A dynamic generation can effectively avoid the risk resulting from information leakage. In this paper, we utilize the fact that the CSI is known only by the two sides of communications and propose a dynamic scheme to generate a secret sequence of starting positions. The following section will give a detailed introduction on it.

IV. The Proposed Dynamic Generation Scheme for Positions of Segments

A. The Architecture of the Dynamic Generation Scheme

This section presents the proposed dynamic generation scheme for the starting positions of each segment where a new packing ratio is employed. The architecture of the scheme is shown in Fig. 3. Alice, Bob, and Eve represent the transmitter, the receiver and the eavesdropper, respectively. The detailed steps are presented as follows.

- **Channel measurement.** In this stage, the two sides of communications send pilots to each other to measure the channel characteristics (e.g., signal intensity or channel response).
- **Quantization.** In this step, the transmitter and receiver quantize the measured information into a bit sequence by single [29], double [30] or multiple-threshold [31] quantization.
- **Information reconciliation.** This step is conducted to correct the errors between the quantized bit sequence of the transceivers. For example, in Cascade-based information reconciliation [32], Alice splits the bits into segments and sends parity check information to Bob. Bob checks the parity states with corresponding pieces. Once the parity bits mismatch, a binary search is conducted by changing as few bits as possible to satisfy the parity requirement.
- **Privacy amplification.** Generally speaking, there always exists some information that leaks to Eve in the information reconciliation stage. By mapping the quantized information into a new bit sequence (i.e., the secret key) with the hash function (e.g., the message-digest algorithm 5 (MD5)), the risk resulting from the partly leaked information can be eliminated.
- **Pseudo-noise (PN) sequence generation.** In this step, the secret key is employed as the seed for the PN generator, which can produce random and unrelated bits within its period.
- **Position calculation.** The offsets can be easily obtained by splitting the PN sequence into segments with the same length and transforming them into signed integers. Then, the starting positions can be calculated by adding them to the original positions with the fixed step.

B. Performance Analysis for the Dynamic Generation Scheme

1) **Security.** The CSI between Alice and Bob can only be measured by both of them, Eve cannot obtain it even if the pilot is eavesdropped. Except the information reconciliation, the other stages are safe since they are executed internally without any information exchanged. Although limited information may be leaked by the parity sent to Bob, the hash mapping operation enhances the system security by mapping the information bits to a new bit sequence (i.e., the secret key) which cannot be inferred by partial original information.
So, throughout this paper, the private key, as well as the generated positions, are considered to be secret and cannot be obtained by Eve.
2) **Randomness.** In fact, this issue has been studied in the physical layer security field. The national institute of standards and technology (NIST) test [33] is widely employed to measure the randomness of the generated secret key. There have been many existing CSI-based key generation schemes that pass the NIST test.
3) **Robustness.** Another import metric is the robustness, which means the proposed scheme can work well under various scenarios and guarantee enough randomness.
The work about this issue can also be found in the existing literature. For example, the secret key generation under different transmission, e.g., MIMO [34] and multi-carrier communications [35], have been widely studied. The solution in different channels (e.g., multi-path channel [36] and even the static channel [31], [35]) have also been presented.
4) **Period.** The PN generator has a period of 2^{N_r}, where N_r is the number of the registers (i.e., the number of bits in the PN generator's state). During the period, the generated bits has an excellent autocorrelation feature which achieves nearly an impulse function. When the generated bits are divided by step N_s, the period of the calculated offset is $\mathrm{lcm}(2^{\mathrm{N}}_{\mathrm{r}}, \mathrm{N_s})$, where $\mathrm{lcm}(a, b)$ means the least common multiple of a and b.

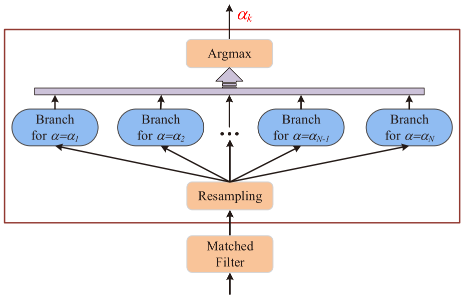


Fig. 4. Structure of the symbol packing ratio estimation employed in the proposed system.

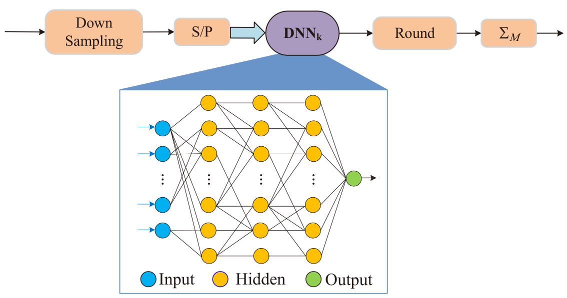


Fig. 5. Structure of the analysis for α_k in the proposed simplified symbol packing ratio estimation.

V. A Simplified Symbol Packing Ratio Estimation for FTN Signaling

In this part, we present a simplified symbol packing ratio estimation for FTN signaling. Fig. 4 illustrates the complete architecture of the proposed estimation. The symbols that have passed through the matched filter and then been sampled are applied as the input of several analysis models. The main task of the analysis for α_k is to decide whether $\alpha = \alpha_k$, where α is the correct symbol packing ratio employed by the transmitter.

Fig. 5 shows the detailed structure of the branch for analyzing whether $\alpha = \alpha_k$. Firstly, the input symbols are down-sampled by the shared knowledge of starting position and interval $\alpha_i T_N$. Then, through serial-parallel conversion (S/P), the sampled serial symbols are reformed and fed into the deep neural network (DNN) [37]. The output of DNN can be regarded as the probability of $\alpha_A = \alpha_k$ and will be transformed into integer 0 (false) or 1 (true). And finally, the number of true decisions during a specific time will be counted.

The DNN we employed in Fig. 5 contains an input layer, three hidden layers and an output layer. Each hidden layer is a sparsely connected layer with ReLU as its activation function. The system function of the DNN can be written as

$$\mathbf{y} = g_4\left(f\left(g_3\left(f\left(g_2\left(f\left(g_1(\boldsymbol{x})\right)\right)\right)\right)\right)\right), \tag{5}$$

where $f(\boldsymbol{x})_i = \max(x_i, 0)$ is the item-wise ReLU function to vector $\boldsymbol{x}$. $g_i(\boldsymbol{x}) = \boldsymbol{W}_i\boldsymbol{x} + \boldsymbol{b}_i$, where $\boldsymbol{W}_i$ and $\boldsymbol{b}_i$ are the weight matrix and bias vector in the i-th layer of the DNN.

Benefiting from that the information of starting position for each transmission segment with a new α is known by both the transceiver, the receiver does not need to divide the signal into several streams [38] to avoid the sampling offset. So, the multiplexer (MUX), the demultiplexer (DEMUX) and the decision model in the original structure can be removed.

Meanwhile, we focus on the simplification of DNN. The main idea is to reduce the amounts of items in the weight matrices. Here, we employ an iterative strategy. After the model is well trained, we remove the items in $\boldsymbol{W}_i$ that are small enough and then train the remaining network. The process will be iteratively carried out until the target sparsity ratio is reached.

VI. Spectrum Efficiency of Proposed VPR System in Different Channels

Generally speaking, the proposed VPR-based system can work well on various channels as long as FTN signaling are applicable. Here, we consider the AWGN, Rayleigh [39] and Nakagami-m [40] channels as the examples. The Rayleigh and Nakagami-m channels are typical fading channels that were first studied in 1940s and 1960s, repectively.

A. Theoretical SE of VPR System in AWGN Channel

In the receiver, α can be easily obtained with the help of blind estimation and the exact starting position. So, α can be regarded as the shared information between the transmitter and the receiver. And the transmission can be considered to be a conventional FTN signaling. With power $\sigma_s^2 = P_s \alpha T_N$ for the transmitted signal, the analytical-form capacity of FTN signaling can be been formulated by Rusek as [26]

$$R_A(\alpha) = \frac{1}{2\pi\alpha T_N}\int_0^{\pi} \log_2\left(1 + \frac{2\sigma_s^2}{N_B} H(\alpha,\omega)\right) \mathrm{d}\omega, \tag{6}$$

where $N_B/2$ is the power spectrum density of the Gaussian noise in the AWGN channel. And $H(\alpha,\omega)$ is defined by

$$H(\alpha,\omega) = \frac{1}{\alpha T_N}\sum_{k=-\infty}^{\infty}\left|G\left(\frac{\omega}{2\pi\alpha T_N} + \frac{k}{\alpha T_N}\right)\right|^2, \tag{7}$$

where $G(f)$ is the Fourier transform of $h(t)$. $|G(f)|^2$ can be expressed as [41] (8), shown at the bottom of the page.

Here, we define three bound functions $b_1(\alpha) = \alpha\pi(1-\beta)$, $b_2(\alpha) = 2\pi - \alpha\pi(1+\beta)$ and $b_3(\alpha) = \alpha\pi(1+\beta)$. And the following conclusion can be derived.

Lemma 1: When $\omega \in [0,\pi]$, for any $k \geq 1$, it always holds that

$$G^2\left(\frac{\omega + 2k\pi}{2\pi\alpha T_N}\right) = 0. \tag{9}$$

Proof: Here, we firstly assume that for any $k \geq 1$, it holds that

$$\frac{\omega + 2\pi k}{2\pi\alpha T_N} \geq \frac{1+\beta}{2T_N}. \tag{10}$$

Since that $\omega \in [0,\pi]$, (10) can be proved by

$$\frac{\omega + 2\pi k}{2\pi\alpha T_N} \geq \frac{1+\beta}{2T_N} \Leftarrow \frac{2\pi k}{2\pi\alpha T_N} \geq \frac{1+\beta}{2T_N} \Leftrightarrow \frac{k}{\alpha} \geq \frac{1+\beta}{2} \Leftrightarrow 2k \geq \alpha(1+\beta). \tag{11}$$

Considering that $k > 1$ and $0 \leq \alpha, \beta \leq 1$, (11) can be further obtained by

$$2k \geq \alpha(1+\beta) \Leftarrow 2 \geq \alpha\left(1+\beta\right) \Leftarrow 2 \geq 2. \tag{12}$$

It's obviously that $2 > 2$ always holds. So, the assumption (10) is proved. And finally, *Lemma* 1 can be proved by combining (8) with (10). ■

Lemma 2: When $\omega \in [0, \pi]$, for any $k \leq -2$, it always holds that

$$G^2\left(\frac{\omega + 2k\pi}{2\pi\alpha T_N}\right) = 0. \tag{13}$$

Proof: Firstly, we assume that for any $k \leq -2$, it holds that

$$\frac{\omega + 2\pi k}{2\pi\alpha T_N} \leq -\frac{1+\beta}{2T_N}. \tag{14}$$

Since that $\omega \in [0, \pi]$, (14) can be proved by

$$\begin{aligned}\frac{\omega + 2\pi k}{2\pi\alpha T_N} &\leq -\frac{1+\beta}{2T_N} \Leftarrow \frac{\pi + 2\pi k}{2\pi\alpha T_N} \leq -\frac{1+\beta}{2T_N}\\ &\Leftrightarrow 2k \leq -\alpha\left(1+\beta\right) - 1.\end{aligned} \tag{15}$$

Considering that $k \leq -2$ and $0 \leq \alpha, \beta \leq 1$, (15) can be obtained by

$$\begin{aligned}2k &\leq -\alpha\left(1+\beta\right) - 1 \Leftarrow -4 \leq -\alpha\left(1+\beta\right) - 1\\ &\Leftarrow -4 \leq -3.\end{aligned} \tag{16}$$

It's obviously that $-4 \leq -3$ always holds. So, (14) is proved. And finally, *Lemma* 2 can be proved by combining (8) with (14). ■

Theorem 1: For $\omega \in [0, \pi]$, $H(\alpha, \omega)$ can be expressed as

$$H(\alpha,\omega) = \frac{1}{\alpha T_N}\left(G^2\left(\frac{\omega}{2\pi\alpha \mathrm{T}_N}\right) + G^2\left(\frac{\omega - 2\pi}{2\pi\alpha \mathrm{T}_N}\right)\right). \tag{17}$$

Proof: The theorem can be proved by combining *Lemma* 1, *Lemma* 2 and (7). ■

Lemma 3: For $\omega \in [0, b_1(\alpha))$, it always holds that

$$H(\alpha,\omega) = \frac{1}{\alpha}. \tag{18}$$

Proof: Since that $\omega \in [0, b_1(\alpha))$, it can be obtained that

$$0 \leq \frac{\omega}{2\pi\alpha T_N} \leq \frac{b_1(\alpha)}{2\pi\alpha T_N} = \frac{1-\beta}{2T_N}. \tag{19}$$

Considering (8) and (19), for $\omega \in [0, b_1(\alpha))$, it can be derived that

$$G^2\left(\frac{\omega}{2\pi\alpha T_N}\right) = \frac{1}{\alpha}. \tag{20}$$

Also, since $\omega \in [0, b_1(\alpha)$, it can be obtained that

$$\frac{\omega - 2\pi}{2\pi\alpha T_N} \leq \frac{b_1(\alpha) - 2\pi}{2\pi\alpha T_N} = \frac{\alpha\pi\left(1-\beta\right) - 2\pi}{2\pi\alpha T_N}. \tag{21}$$

Considering that $0 \leq \alpha, \beta \leq 1$, (21) can be further written as

$$\begin{aligned}\frac{\omega - 2\pi}{2\pi\alpha T_N} &\leq \frac{\alpha\pi\left(1-\beta\right) - 2\pi}{2\pi\alpha T_N}\\ &\leq \frac{\alpha\pi\left(1-\beta\right) - 2\alpha\pi}{2\pi\alpha T_N} = -\frac{1+\beta}{2T_N}.\end{aligned} \tag{22}$$

Considering (8) and (22), it can be derived that

$$G^2\left(\frac{\omega - 2\pi}{2\pi\alpha T_N}\right) = 0. \tag{23}$$

Finally, *Lemma* 3 can be proved by combining *Theorem* 1, (20) and (23). ■

Lemma 4: For $\omega \in [b_3(\alpha), \pi]$ and $b_2(\alpha) \geq \pi$, it always holds that

$$G^2\left(\frac{\omega}{2\pi\alpha T_N}\right) = 0. \tag{24}$$

Proof: Firstly, it can be obtained by $b_2(\alpha) \geq \pi$ that

$$\begin{aligned}b_2\left(\alpha\right) &= 2\pi - \alpha\pi\left(1+\beta\right) \geq \pi\\ &\Leftrightarrow \alpha \leq \frac{1}{1+\beta}.\end{aligned} \tag{25}$$

Then, since that $\omega \in [b_3(\alpha), \pi]$, it can be obtained that

$$\frac{\omega - 2\pi}{2\pi\alpha T_N} \leq \frac{\pi - 2\pi}{2\pi\alpha T_N} = -\frac{1}{2\alpha T_N} \leq -\frac{1+\beta}{2T_N}. \tag{26}$$

Considering (8) and (26), it can be derived that

$$G^2\left(\frac{\omega - 2\pi}{2\pi\alpha T_N}\right) = 0. \tag{27}$$

Also, since that $\omega \in [b_3(\alpha), \pi]$, it can be obtained that

$$\frac{\omega}{2\pi\alpha T_N} \geq \frac{\alpha\pi\left(1+\beta\right)}{2\pi\alpha T_N} = \frac{1+\beta}{2T_N}. \tag{28}$$

Considering (8) and (28), it can be derived that

$$G^2\left(\frac{\omega}{2\pi\alpha T_N}\right) = 0. \tag{29}$$

Finally, *Lemma* 4 can be proved by combining *Theorem* 1, (27) and (29). ■

By combining *Theorem* 1, *Lemma* 3 and *Lemma* 4, $H(\alpha, \omega)$ can be expressed as

$$H\left(\alpha,\omega\right) = \begin{cases} H_1\left(\alpha,\omega\right), & b_2(\alpha) < \pi \\ H_2\left(\alpha,\omega\right), & b_2\left(\alpha\right) \geq \pi \end{cases}, \tag{30}$$

where (31) and (32), shown at the bottom of the next page.

$$|G\left(f\right)|^2 = \begin{cases} T_N, & |f| \in \left[0, \frac{1-\beta}{2T_N}\right] \\ \frac{T_N}{2}\left\{1 + \cos\left[\frac{\pi T_N}{\beta}\left(|f| - \frac{1-\beta}{2T_N}\right)\right]\right\}, & |f| \in \left[\frac{1-\beta}{2T_N}, \frac{1+\beta}{2T_N}\right] \\ 0, & |f| \in \left[\frac{1+\beta}{2T_N}, +\infty\right] \end{cases}. \tag{8}$$

With the system bandwidth that can be calculated by $W = 1/(2T_N) = W_T/(1+\beta)$, where W_T is the total bandwidth of the channel, SE of FTN signaling can be written as

$$C_A(\alpha) = \frac{1}{\pi\alpha(1+\beta)} \underbrace{\int_0^{\pi} \log_2\left(1+\frac{2\sigma_s^2}{N_B}H(\alpha,\omega)\right) \mathrm{d}\omega}_{C_B(\alpha)}. \tag{33}$$

Then, we split $C_B(\alpha)$ into several subsection integral and calculate them respectively. For $\omega \in [0, b_1(\alpha))$, the integral can be expressed as

$$\begin{aligned} C_1(\alpha) &= \int_0^{b_1(\alpha)} \log_2\left(1+\frac{2\sigma_s^2}{N_B}H(\alpha,\omega)\right)\mathrm{d}\omega \\ &= \alpha\pi(1-\beta)\log_2\left(1+\frac{2\sigma_s^2}{\alpha N_B}\right). \end{aligned} \tag{34}$$

Theorem 2 (Chebyshev-Gauss Quadrature Rule): For a given function $f(x)$, its integration between -1 and 1 can be approximated as [42]

$$\int_{-1}^{1} \frac{f(x)}{\sqrt{1-x^2}}dx \approx \sum_{i=1}^{n} w_i f(\xi_i), \tag{35}$$

where $\xi_i = \cos\left(\frac{2i-1}{2n}\pi\right)$ and $w_i = \frac{\pi}{n}$.

Theorem 3: For given function $f(x)$, its integration between a and b can be approximated as

$$\int_a^b f(x)\,dx \approx \frac{b-a}{2}\sum_{i=1}^{n} w_i\sqrt{1-\xi_i^2}\, f\left(\frac{b-a}{2}\xi+\frac{b+a}{2}\right), \tag{36}$$

where the values of ξ_i and w_i are the same as those in *Theorem* 2.

Proof: Here, we set

$$x = \frac{b-a}{2}\xi+\frac{b+a}{2}.$$

Then, the integration of $f(x)$ can be rewritten as

$$\begin{aligned} &\int_a^b f(x)\,dx \\ &= \int_{arg_\xi(a)}^{arg_\xi(b)} f\left(\frac{b-a}{2}\xi+\frac{b+a}{2}\right) d\left[\frac{b-a}{2}\xi+\frac{b+a}{2}\right] \\ &= \frac{b-a}{2}\int_{-1}^{1} f\left(\frac{b-a}{2}\xi+\frac{b+a}{2}\right)d\xi \\ &= \frac{b-a}{2}\int_{-1}^{1}\frac{\sqrt{1-\xi^2}\psi(\xi)}{\sqrt{1-\xi^2}}d\xi, \end{aligned} \tag{37}$$

where $\psi(\xi)$ is defined as

$$\psi(\xi) = f\left(\frac{b-a}{2}\xi+\frac{b+a}{2}\right).$$

Considering *Theorem* 2, (37) can be written as

$$\begin{aligned} \int_a^b f(x)\,dx &\approx \frac{b-a}{2}\sum_{i=1}^{n} w_i\sqrt{1-{\xi_i}^2}\psi(\xi_i) \\ &= \frac{b-a}{2}\sum_{i=1}^{n} w_i\sqrt{1-{\xi_i}^2} f\left(\frac{b-a}{2}\xi_i+\frac{b+a}{2}\right). \end{aligned} \tag{38}$$

■

According to the *Theorem* 3, for $\omega \in [b_1(\alpha), \pi]$, the integral can be written as

$$\begin{aligned} C_2(\alpha) &= \int_{b_1(\alpha)}^{\pi} \log_2\left(1+\frac{2\sigma_s^2}{N_B}H(\alpha,\omega)\right)d\omega \\ &\approx A_1\sum_{i=1}^{N} m_i\sqrt{1-\omega_{1i}^2}\log_2\left(1+\frac{2\sigma_s^2}{N_B}H(\alpha,\omega_{1i})\right), \end{aligned} \tag{39}$$

where

$$A_1 = \frac{\pi\left[1+\alpha(\beta-1)\right]}{2}, \quad m_i = \frac{\pi\left|\sin\left(\frac{\pi(2i-1)}{2N}\right)\right|}{N}$$

and

$$\omega_{1i} = \frac{\pi}{2}\left\{\left[1+\alpha(\beta-1)\right]\cos\left(\frac{\pi(2i-1)}{2n}\right) + 1 - \alpha(\beta-1)\right\}.$$

Similarly, for $\omega \in [b_1(\alpha), b_3(\alpha))$, the integral can be written as

$$\begin{aligned} C_3(\alpha) &= \int_{b_1(\alpha)}^{b_3(\alpha)} \mathrm{l}_2\left(1+\frac{2\sigma_s^2}{N_B}H(\alpha,\omega)\right)d\omega \\ &\approx A_2\sum_{i=1}^{N} m_i\sqrt{1-\omega_{2i}^2}\log_2\left(1+\frac{2\sigma_s^2}{N_B}H(\alpha,\omega_{2i})\right), \end{aligned} \tag{40}$$

where

$$A_2 = \pi\alpha\beta$$

and

$$\omega_{2i} = \pi\alpha\left(1+\beta\cos\left(\frac{\pi(2i-1)}{2n}\right)\right).$$

$$H_1(\alpha,\omega) = \begin{cases} \frac{1}{\alpha}, & \omega \in [0, b_1(\alpha)) \\ \frac{1}{\alpha T_N}\left(G^2\left(\frac{\omega}{2\pi\alpha \mathrm{T}_N}\right)+G^2\left(\frac{\omega-2\pi}{2\pi\alpha \mathrm{T}_N}\right)\right), & \omega \in [b_1(\alpha), \pi] \end{cases}, \tag{31}$$

$$H_2(\alpha,\omega) = \begin{cases} \frac{1}{\alpha}, & \omega \in [0, b_1(\alpha)) \\ \frac{1}{\alpha T_N}\left(G^2\left(\frac{\omega}{2\pi\alpha \mathrm{T}_N}\right)\right), & \omega \in [b_1(\alpha), b_3(\alpha)) \\ 0, & \omega \in [b_3(\alpha), \pi] \end{cases}. \tag{32}$$

For the convenience of implementation, the set of available α values is usually finite. Finally, for a specific α value, SE of the proposed VPR system in the AWGN channel can be written as

$$C_A(\alpha) = \begin{cases} \dfrac{1}{\alpha\pi(1+\beta)}\left(C_1(\alpha)+C_2(\alpha)\right), & b_2(\alpha) < \pi \\ \dfrac{1}{\alpha\pi(1+\beta)}\left(C_1(\alpha)+C_3(\alpha)\right), & b_2(\alpha) \geq \pi. \end{cases} \tag{41}$$

To avoid the possible detection and attack when the VPR system is employed to improve the security, every α is preferred to be applied with the same probability, just as the roll-off factor in [27]. So, for the proposed VPR system, the average SE in such a scenario can be written as

$$C'_A = \frac{1}{N_\alpha}\sum_{i=1}^{N_\alpha} C_A(\alpha_i), \tag{42}$$

where α_i $(i = 1, 2 \cdots, N_\alpha)$ is the i-th symbol packing ratio that is employed in the transmission.

B. Theoretical SE of VPR System in Rayleigh Channel

For the Rayleigh and Nakagami-m channel, the channel gain is considered and can be regarded as a constant during every data block in this paper. So, the power of the signal in the receiver with channel gain h can be written as

$$\sigma^2_{s'}(h) = h^2 P_s \alpha T_N. \tag{43}$$

The capacity of FTN signaling with specific h can be obtained as

$$R'(\alpha) = \frac{1}{2\pi\alpha T_N}\int_0^\pi \log_2\left(1 + \frac{2\sigma^2_{s'}(h)}{N_B}H(\alpha,\omega)\right) \mathrm{d}\omega. \tag{44}$$

Considering that h is a random variable, the mean SE of FTN signaling with packing ratio α in Rayleigh channel can be formulated as (45), shown at the bottom of the page, where $f_R(h)$ is the probability density function (PDF) of h, which can be written as [39]

$$f_R(h) = \frac{h}{\sigma^2}e^{-\frac{h^2}{2\sigma^2}}, \tag{46}$$

where σ^2 is the power parameter. Then, by applying $C_{o1}(\alpha,\omega) = 2P_s\alpha T_N H(\alpha,\omega)/N_B$, $C_{i1}(\alpha,\omega)$, which has been defined in (45), can be written as

$$C_{i1}(\alpha,\omega) = \int_0^{+\infty} -\log_2\left(1 + C_{o1}(\alpha,\omega)h^2\right)\left(-\frac{h}{\sigma^2}e^{-\frac{h^2}{2\sigma^2}}\right)\mathrm{d}h. \tag{47}$$

By extracting the integral items as $F_1(\alpha,\omega) = -\log_2\left(1 + C_{o1}(\alpha,\omega)h^2\right)$ and $F_2(h) = e^{\frac{-h^2}{2\sigma^2}}$, $C_{i1}(\alpha,\omega)$ can be expressed as

$$C_{i1}(\alpha,\omega) = \int_0^{+\infty} F_1(h,\alpha,\omega)F_2{}'(h)\ \mathrm{d}h. \tag{48}$$

According to the principle of integral by parts [43], $C_i(h,\alpha,\omega)$ can be further written as

$$C_{i1}(\alpha,\omega) = F_1(h,\alpha,\omega)F_2(h)\Big|_0^{+\infty} - \int_0^{+\infty} F_1{}'(h,\alpha,\omega)F_2(h)\,\mathrm{d}h. \tag{49}$$

Due to the fact that

$$F_1(0,\alpha,\omega)F_2(0) = -\log2(1)\cdot e^0 = 0, \tag{50}$$

$$\begin{aligned}&\lim_{h\to+\infty} F_1(h,\alpha,\omega)F_2(h) \\ &= \lim_{h\to+\infty}\left(-\log_2\left(1 + C_{o1}(\alpha,\omega)h^2\right) \times e^{\frac{-h^2}{2\sigma^2}}\right) = 0,\end{aligned} \tag{51}$$

(49) can be expressed as

$$\begin{aligned} C_{i1}(\alpha,\omega) &= -\int_0^{+\infty}\frac{2C_{o1}(\alpha,\omega)h}{\ln 2\cdot(1 + C_{o1}(\alpha,\omega)h^2)}e^{-\frac{h^2}{2\sigma^2}}\mathrm{d}h \\ &= -\frac{e^{\frac{1}{2\sigma^2 C_{o1}(\alpha,\omega)}}}{\ln 2}\int_{\frac{1}{2\sigma^2 C_{o1}(h,\alpha,\omega))}}^{+\infty}\frac{2\sigma^2 e^{-\frac{\frac{1}{C_{o1}(\alpha,\omega)}+h^2}{2\sigma^2}}}{\left(\frac{1}{C_{o1}(\alpha,\omega)}+h^2\right)} \times d\left(\frac{\frac{1}{C_{o1}(\alpha,\omega)}+h^2}{2\sigma^2}\right) \\ &= -\frac{e^{\frac{N_B}{4\sigma^2 P_s\alpha T_N H(\alpha,\omega)}}}{\ln 2}\mathrm{Ei}\left(-\frac{N_B}{4\sigma^2 P_s\alpha T_N H(\alpha,\omega)}\right), \end{aligned} \tag{52}$$

where $\mathrm{E_i}(x)$ is the exponential integral function which is defined as $\mathrm{E_i}(x) = \int_{-x}^{+\infty}\frac{e^{-t}}{t}\mathrm{d}t$.

Now, by applying $C_{o2}(\alpha) = -4\sigma^2 P_s\alpha T_N/N_B$, (45) can be written as

$$C_R(\alpha) = -\frac{1}{\pi\alpha(1+\beta)\ln 2}\underbrace{\int_0^\pi e^{-\frac{C_{o2}(\alpha)}{H(\alpha,\omega)}}\mathrm{Ei}\left(\frac{C_{o2}}{H(\alpha,\omega)}\right)d\omega}_{C_{i2}(\alpha,\omega)}. \tag{53}$$

Then, we split $C_{i2}(\alpha,\omega)$ into several subsection integral and calculate them respectively. For $\omega \in [0, b_1(\alpha))$, the

$$C_R(\alpha) = \frac{1}{\pi\alpha(1+\beta)}\cdot\int_0^\pi \underbrace{\int_0^{+\infty} f_R(h)\cdot\log_2\left(1 + \frac{2h^2 P_s\alpha T_N}{N_B}H(\alpha,\omega)\right)\mathrm{d}h\mathrm{d}\omega}_{C_{i1}(\alpha,\omega)}, \tag{45}$$

integral can be calculated as

$$\begin{aligned} C_4(\alpha) &= \int_0^{b_1(\alpha)} e^{-\frac{C_{o2}(\alpha)}{H(\alpha,\omega)}} \mathrm{Ei}\left(\frac{C_{o2}}{H(\alpha,\omega)}\right) d\omega \\ &= \int_0^{b_1(\alpha)} e^{-\alpha C_{o2}(\alpha)} \mathrm{Ei}\left(\alpha C_{o2}(\alpha)\right) d\omega \\ &= \alpha\pi(1-\beta) e^{-\alpha C_{o2}} \mathrm{Ei}\left(\alpha C_{o2}(\alpha)\right). \end{aligned} \tag{54}$$

According to the *Theorem* 3, for $\omega \in [b_1(\alpha), \pi]$, the integral can be written as

$$\begin{aligned} C_5(\alpha) &= \int_{b_1(\alpha)}^{\pi} e^{-\frac{C_{o2}(\alpha)}{H(\alpha,\omega)}} \mathrm{Ei}\left(\frac{C_{o2}}{H(\alpha,\omega)}\right) d\omega \\ &\approx A_1 \sum_{i=1}^{N} m_i \sqrt{1-\omega_{1i}^2} e^{-\frac{C_{o2}(\alpha)}{H(\alpha,\omega_{1i})}} \mathrm{Ei}\left(\frac{C_{o2}}{H(\alpha,\omega_{1i})}\right). \end{aligned} \tag{55}$$

Similarly, for $\omega \in [b_1(\alpha), b_3(\alpha))$, the integral can be written as

$$\begin{aligned} C_6(\alpha) &= \int_{b_1(\alpha)}^{b_3(\alpha)} e^{-\frac{C_O(\alpha)}{H(\alpha,\omega)}} \mathrm{Ei}\left(\frac{C_{o2}}{H(\alpha,\omega)}\right) d\omega \\ &\approx A_2 \sum_{i=1}^{N} m_i \sqrt{1-\omega_{2i}^2} e^{-\frac{C_{a2}(\alpha)}{H(\alpha,\omega_{2i})}} \mathrm{Ei}\left(\frac{C_{o2}}{H(\alpha,\omega_{2i})}\right). \end{aligned} \tag{56}$$

Finally, the SE of the proposed VPR scheme in Rayleigh channel can be written as

$$C_R(\alpha) = \begin{cases} \dfrac{1}{\alpha\pi(1+\beta)}\left(C_4(\alpha)+C_5(\alpha)\right), & b_2(\alpha) < \pi \\ \dfrac{1}{\alpha\pi(1+\beta)}\left(C_4(\alpha)+C_6(\alpha)\right), & b_2(\alpha) \geq \pi. \end{cases} \tag{57}$$

The average SE for VPR-based secure transmission in Rayleigh channel can be obtained as

$$C_R' = \frac{1}{N_\alpha} \sum_{i=1}^{N_\alpha} C_{R'}(\alpha_i). \tag{58}$$

C. Theoretical SE of VPR System in Nakagami-m Channel

Similar to (45), the SE of FTN signaling in Nakagami-m channel can be formulated as (59), shown at the bottom of the page, where $f_N(h)$ is the PDF of h in Nakagami-m channel which can be written as [40]

$$f_N(h) = \frac{2m^m h^{2m-1}}{\Gamma(m) {P_r}^m} e^{-\frac{mh^2}{P_r}}, \tag{60}$$

where m ($m > 0$) is the fading parameter, P_r is the average power, $\Gamma(m)$ is the Gamma function which can be expressed as [44]

$$\Gamma(m) = \int_0^{+\infty} t^{m-1} e^{-t}\, \mathrm{d}t \quad (m > 0). \tag{61}$$

By applying $C_{o3} = 2m^m / (\Gamma(m) P_r^m)$, the integral of channel gain h can be written as

$$\begin{aligned} C_{i,2}(\alpha,\omega) &= \int_0^{+\infty} \mathrm{C}_{o3} \mathrm{h}^{2m-1} e^{-\frac{mh^2}{P_r}} \log_2 \\ &\quad \times \left(1 + C_{o1}(\alpha,\omega) h^2\right) \mathrm{d}h \\ &= C_{o3} \mathcal{M}\left[e^{-\frac{mh^2}{P_r}} \log_2 (1 \right. \\ &\quad \left. + C_{o1}(\alpha,\omega) h^2); 2m\right], \end{aligned} \tag{62}$$

where $\mathcal{M}[f(x); s]$ means the Mellin transform [45] of $f(x)$.

Theorem 4: (Mellin Convolution Theorem) For functions $f(x)$ and $g(x)$, it holds that

$$\begin{aligned} \mathcal{M}[f(x)g(x); s] &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \\ &\quad \times \mathcal{M}[f(x); s] \mathcal{M}[g(x); s-u]\, du. \end{aligned} \tag{63}$$

Considering *Theorem* 4, (62) can be further written as

$$\begin{aligned} C_{i,2}(\alpha,\omega) &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+\infty} \mathcal{M}\left[e^{-\frac{mh^2}{P_r}}; 2m-u\right] \\ &\quad \times \mathcal{M}\left[\log_2\left(1 + C_{o1}(\alpha,\omega) h^2\right); u\right] du. \end{aligned} \tag{64}$$

Property 1: The Mellin transform has the properties as follows [46].

$$\mathcal{M}[f(\nu x); s] = \nu^{-s} f^*(s), \ \nu > 0, \tag{65}$$

$$\mathcal{M}\left[e^{-x^2}; s\right] = \frac{1}{2}\Gamma\left(\frac{s}{2}\right), \ \mathrm{Re}(s) > 0, \tag{66}$$

$$\mathcal{M}[f(x^\nu); s] = \frac{1}{|\nu|} f^*\left(\frac{s}{\nu}\right), \ s/\nu \text{ is feasible}, \tag{67}$$

$$\mathcal{M}[\ln(1+x); s] = \frac{\pi}{s \sin(\pi s)}, -1 \leq \mathrm{Re}(s) \leq 0. \tag{68}$$

Considering (65) and (66), it can be obtained that

$$\begin{aligned} \mathcal{M}\left[e^{-\frac{mh^2}{P_r}}; 2m-u\right] &= \left(\frac{m}{P_r}\right)^{-\left(m-\frac{u}{2}\right)} \mathcal{M}\left[e^{-h^2}; 2m-u\right] \\ &= x\frac{1}{2}\left(\frac{m}{P_r}\right)^{-\left(m-\frac{u}{2}\right)} \Gamma\left(m-\frac{u}{2}\right). \end{aligned} \tag{69}$$

Considering (65), (67) and (68), it can be obtained that

$$\begin{aligned} &\mathcal{M}\left[\log_2\left(1 + C_{o1}(\alpha,\omega) h^2\right); u\right] \\ &= \frac{1}{2\ln 2} \left[C_{o1}(\alpha,\omega)\right]^{-\frac{u}{2}} \mathcal{M}\left[\ln(1+h); \frac{u}{2}\right] \\ &= \frac{1}{\ln 2} \left[C_{o1}(\alpha,\omega)\right]^{-\frac{u}{2}} \frac{\pi}{u \sin\left(\frac{\pi u}{2}\right)}. \end{aligned} \tag{70}$$

$$C_N(\alpha) = \frac{1}{\pi\alpha(1+\beta)} \cdot \int_0^{\pi} \underbrace{\int_0^{+\infty} f_N(h) \log_2\left(1 + \frac{2h^2 P_s \alpha T_N}{N_B} H(\alpha,\omega)\right) \mathrm{d}h}_{C_{i2}(\alpha,\omega))} \mathrm{d}\omega, \tag{59}$$

Combining (69) and (70), (64) can be written as

$$\begin{aligned} &C_{i,2}\left(\alpha,\omega\right) \\ &= \frac{C_{o,3}}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{\pi\left(\frac{P_r}{m}\right)^{m-\frac{u}{2}}\Gamma\left(m-\frac{u}{2}\right)}{4\ln 2\cdot\left(-\frac{u}{2}\right)\left[C_{o,1}\left(\alpha,\omega\right)\right]^{\frac{u}{2}}\sin\left(-\frac{\pi u}{2}\right)}du. \end{aligned} \tag{71}$$

Property 2: Gamma function has the properties as follows.

$$\Gamma\left(1-x\right)\Gamma\left(x\right) = \frac{\pi}{\sin\left(\pi x\right)}, \tag{72}$$

$$\Gamma\left(x+1\right) = x\Gamma\left(x\right). \tag{73}$$

Considering *Property* 2, (71) can be further written as (74), shown at the bottom of the page, where $G_{p,\,q}^{m,\,n}\left(\begin{matrix} a_1, a_2\cdots a_p \\ b_1, b_2\cdots b_q\end{matrix}\,|z\right)$ represents the Meijer-G function [47].

By applying $C_{o4} = mN_B/\left(2P_s\alpha T_N P_r\right)$, for $\omega \in [0, b_1(\alpha))$, the integral can be written as

$$\begin{aligned} C_7(\alpha) &= \int_0^{b_1(\alpha)}\frac{1}{\Gamma\left(m\right)\ln(2)} \\ &\quad\cdot G_{1,\,0}^{3,\,1}\left(\begin{matrix}0,1\\0,0,m\end{matrix}\,\middle|\,\alpha C_{o4}\left(\alpha,\omega\right)\right)\mathrm{d}\omega \\ &= \frac{\alpha\pi\left(1-\beta\right)}{\Gamma\left(m\right)\ln(2)} \\ &\quad\cdot G_{1,\,0}^{3,\,1}\left(\begin{matrix}0,1\\0,0,m\end{matrix}\,\middle|\,\alpha C_{o4}\left(\alpha,\omega\right)\right). \end{aligned} \tag{75}$$

For $\omega \in [b_1(\alpha), \pi)$, the integral can be written as

$$\begin{aligned} C_8(\alpha) &= \int_{b_1(\alpha)}^{\pi}\frac{1}{\Gamma\left(m\right)\ln(2)} \\ &\quad\cdot G_{1,\,0}^{3,\,1}\left(\begin{matrix}0,1\\0,0,m\end{matrix}\,\middle|\,\frac{C_{o4}}{H\left(\alpha,\omega\right)}\right)d\omega \\ &\approx \frac{C_1}{\Gamma\left(m\right)\ln(2)}\sum_{i=1}^{N} m_i\sqrt{1-\omega_{1i}^2} \\ &\quad\cdot G_{1,\,0}^{3,\,1}\left(\begin{matrix}0,1\\0,0,m\end{matrix}\,\middle|\,\frac{C_{o4}}{H\left(\alpha,\omega_{1i}\right)}\right). \end{aligned} \tag{76}$$

And for $\omega \in [b_1(\alpha), b_3(\alpha))$, the integral can be written as

$$\begin{aligned} C_9(\alpha) &= \int_{b_1(\alpha)}^{b_3(\alpha)}\frac{1}{\Gamma\left(m\right)\ln(2)} \\ &\quad\cdot G_{1,\,0}^{3,\,1}\left(\begin{matrix}0,1\\0,0,m\end{matrix}\,\middle|\,\frac{C_{o4}}{H\left(\alpha,\omega\right)}\right)d\omega \\ &\approx \frac{C_2}{\Gamma\left(m\right)\ln(2)}\sum_{i=1}^{N} m_i\sqrt{1-\omega_{2i}^2} \\ &\quad\cdot G_{1,\,0}^{3,\,1}\left(\begin{matrix}0,1\\0,0,m\end{matrix}\,\middle|\,\frac{C_{o4}}{H\left(\alpha,\omega_{2i}\right)}\right). \end{aligned} \tag{77}$$

Finally, the SE of the proposed VPR system in Nakagami-m channel can be written as

$$C_N\left(\alpha\right) = \begin{cases}\frac{1}{\pi\alpha(1+\beta)}\left(C_7(\alpha)+C_8(\alpha)\right), & b_2(\alpha) < \pi \\ \frac{1}{\pi\alpha(1+\beta)}\left(C_7(\alpha)+C_9(\alpha)\right), & b_2(\alpha) \geq \pi\end{cases}. \tag{78}$$

The average SE of the VPR-based secure system in Nakagami-m channel can be written as

$$C_N' = \frac{1}{N_\alpha}\sum_{i=1}^{N_\alpha} C_N(\alpha_i). \tag{79}$$

VII. Numerical Results

This section carries out comprehensive analysis and evaluation for the proposed VPR transmission systems. The simulation employs the binary phase shift keying (BPSK) modulation and SRRC filter with roll-off factor β. And the training parameters for the DNN in the proposed simplified symbol packing ratio estimation are listed in Table I. Each group mentioned in the table consists of 20 received symbols.

A. SE of the Proposed VPR System in AWGN Channel

The average SEs of the proposed VPR system in AWGN channel are illustrated in Fig. 6(a) and Fig. 6(b) with roll-off factors $\beta = 0.5$ and $\beta = 0.3$ respectively. The curves labeled *Monte-Carlo* or without special label are obtained by numerical simulation. While the curve labeled *theoretical* is calculated by (42). To avoid the confusion resulting from too many curves and marks, only the curve for average theoretical capacity in Section VI is plotted. And the perfect match of the results by theoretical derivation and numerical simulation proves the correctness of the SE presented in Section VI.

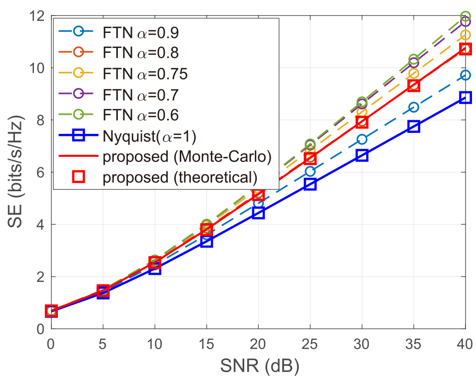
It should be noticed that, the SE of FTN signaling only increases when $\alpha > 1/(1+\beta)$, which has been proved by [26].

$$\begin{aligned} C_{i,2}\left(\alpha,\omega\right) &= \frac{C_{o,3}}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{\left(\frac{P_r}{m}\right)^{m-\frac{u}{2}}\Gamma\left(m-\frac{u}{2}\right)\Gamma\left(1+\frac{u}{2}\right)\Gamma\left(-\frac{u}{2}\right)\Gamma\left(-\frac{u}{2}\right)}{4\ln 2\cdot\Gamma\left(1-\frac{u}{2}\right)\cdot\left[C_{o,1}\left(\alpha,\omega\right)\right]^{\frac{u}{2}}}du \\ &= \frac{C_{o,3}\left(\frac{P_r}{m}\right)^m}{2\ln 2}\cdot\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{\Gamma\left(m-\frac{u}{2}\right)\Gamma\left(1+\frac{u}{2}\right)\Gamma\left(-\frac{u}{2}\right)\Gamma\left(-\frac{u}{2}\right)}{\Gamma\left(1-\frac{u}{2}\right)} \\ &\quad\times\left(\frac{m}{P_r C_{o1}\left(\alpha,\omega\right)}\right)^{\frac{u}{2}}d\left(\frac{u}{2}\right) \\ &= \frac{C_{o,3}\left(\frac{P_r}{m}\right)^m}{2\ln 2}\cdot G_{1,0}^{3,1}\left(\begin{matrix}0,1\\0,0,m\end{matrix}\,\middle|\,\frac{m}{P_r C_{o1}\left(\alpha,\omega\right)}\right), \end{aligned} \tag{74}$$

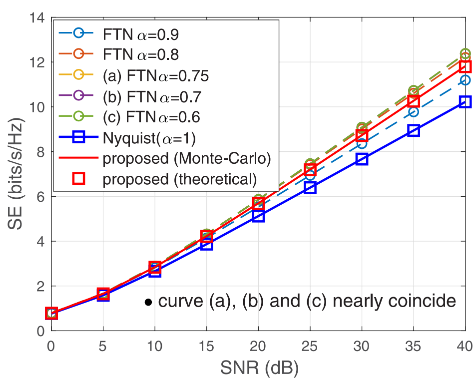
TABLE I

TRAINING AND TESTING PARAMETERS OF THE DNN IN THE PROPOSED SIMPLIFIED SYMBOL PACKING RATIO ESTIMATION

item	value
number of neurons	(20, 1000, 500, 250, 1)
training data size	3×10^6 groups
training E_b/N_0	4dB
training epoch	50
optimizer	Adam
loss function	mean square error (MSE)
learning rate	0.001
testing data size	3×10^6 groups
start / end sparsity	0 / 0.5



(a) $\beta = 0.5,\ \alpha_{Th} \approx 0.667$.



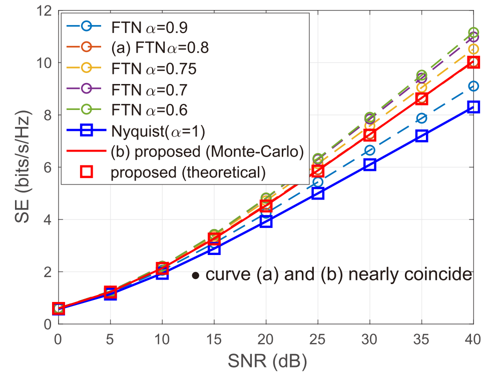
(b) $\beta = 0.3,\ \alpha_{Th} \approx 0.769$.

Fig. 6. SE of the proposed VPR system versus Nyquist-criterion transmission in AWGN channels.

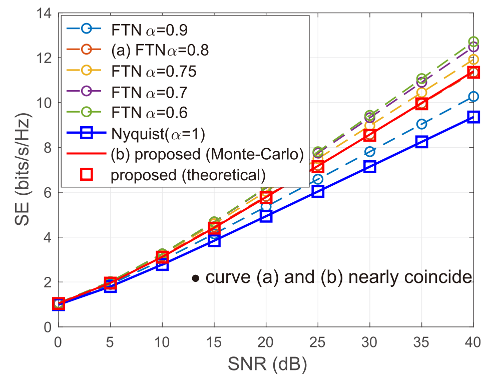
So, the curves with $\alpha \leq 1/(1+\beta)$ coincide and show the same SE, as demonstrated in the figures. To make it more clearly, we add the threshold $\alpha_{Th} = 1/(1+\beta)$ in the subtitles of each figure.

B. SE of the Proposed VPR System in Rayleigh and Nakagami-m Channels

Fig. 7(a) and Fig. 7(b) illustrate the SE of the proposed scheme in Rayleigh and Nakagami-m ($m = 3$ and $P_r = 2$) channels. The curve labeled *Monte-Carlo* is obtained by independent repeated trials with randomly generated channel gain h values. And the curve labeled *theoretical* is calculated by (58) and (79). 7 points are considered for the Chebyshev-Gauss quadrature. The curves labeled *proposed* demonstrate the average SE of the VPR-based secure transmissions.



(a) Rayleigh channel, $\beta = 0.5,\ \alpha_{Th} \approx 0.667$.



(b) Nakagami-m channel, $\beta = 0.5$, $m = 3$, $P_r = 2$, and $\alpha_{Th} \approx 0.667$.

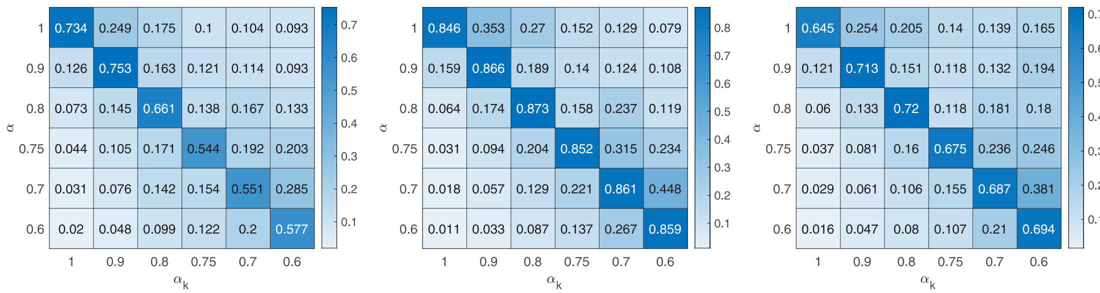
Fig. 7. SE of the proposed VPR system versus Nyquist-criterion transmission in Rayleigh and Nakagami-m channels.

As can be seen, the Monte-Carlo simulation fits the curve with theoretical result provided in Section VI. It shows that (58) and (79) accurately describe the capacities of VPR scheme in Rayleigh and Nakagami-m channels. Also, as shown, with security performance, the SE of VPR-based scheme is still higher than conventional Nyquist transmissions.

C. Performance of the Proposed Simplified Estimation for FTN Signaling in Different Channels

For the proposed scheme, an effective blind estimation for the packing ratio is required to make the communications available. Fig. 8 illustrates the accuracy of the proposed packing ratio estimation in different channels. α is the real packing ratio of the input data. Every grid represents the probability of outputting 1 in the estimation branch for whether $\alpha = \alpha_k$. It should be noticed that the estimations for all α values are carried out independently and the α_k with the most 1 output is considered the correct packing ratio of the data. Hence, the sum value of any row or column in Fig. 8 does not have to be 1.

As seen, the correct α value always corresponds to the highest probability to output 1. After a specific time to count the number of 1 in each branch, the right α_k will be chosen as the estimated α value. Hence, the simplified estimation for α is proved to be effective.



(a) AWGN, Eb/No=4dB. (b) Rayleigh, Eb/No=25dB. (c) Nakagami-m, Eb/No=18dB.

Fig. 8. The accuracy of the proposed simplified estimation for packing ratio of FTN signaling.

D. SE Gain of the Proposed VPR-Based Scheme Over Conventional Nyquist Transmissions

In this part, we provide an example of implementation for the proposed VPR-based high SE transmission, where the maximum a priori probability (MAP) [48] is employed as the detection algorithm, as shown in Table II. The parameters for the Nakagami-m channel here are set as $\mu = 3$ and $\omega = 2$. Under a certain SNR, we will choose the smallest one of the optional α values with which the BER is lower than 10^{-3} to achieve the highest SE. And to better compare the SE gain in different channels, the simulated SNR range is set as $[0, 60]$ (dB) for all scenarios.

Fig. 9 detailed illustrates the SE comparison between the proposed scheme and the conventional Nyquist system. Obvious SE gain, as seen, can be achieved by the proposed VPR system under all simulated channels and roll-off factors. A flexible switching strategy can help the system take advantage of high SNR to achieve a higher SE up to 47% without any extra spectrum consumed.

E. BER Performance for Alice-Bob and Alice-Eve

Fig 10 demonstrates the BER performance of Alice-Bob and Alice-Eve links. As can be seen, the Alice-Bob link can achieve nearly the same BER performance as that in the ISI-free AWGN channel. For the Alice-Eve link, when $\alpha_E \neq \alpha_A$, it will not be able to sample the received signals by the expected interval. Despite the assumption that when $\alpha_E = \alpha_A$, sampling offset is not taken into consideration, the average BER of the Alice-Eve link is still poor enough.

F. The Power of Random Segment Starting Positions to Avoid Attack and Detection

Eve's estimations on the exampled frame with sample-based and range-based sliding windows (presented in Fig. 11) are demonstrated in Fig. 12(a) and Fig. 12(b), respectively. And the frame is constructed with $\alpha = 0.9, 0.8, 0.7, 0.6$, where α for each segment has been marked in the figures.

The estimation in Fig. 12(a) is based on η_1 continuous decisions with fixed interval αT_N, where η_1 is the length of the sliding window. As shown, the result is messy and it is difficult to find a pattern to map the estimation to the original packing ratio for each segment. In fact, the information of starting positions helps the receiver carry out the estimation at the perfect times to eliminate the interference of other erroneous results.

Another way for Eve's estimation is to employ the continuous η_2 decisions with fixed interval $\alpha T_N / U_s$, where U_s is the up-sampling times. As shown in Fig. 12(b), the estimation is still confusing. And especially, the starting position cannot be inferred by the estimation results.

G. BER Degradation of VPR-Based Secure Transmission

According to the simulation results, the VPR-based secure transmission performs nearly the same SE with FTN signaling where $\alpha = 0.8$, $\beta = 0.5$ and $\alpha = 0.9$, $\beta = 0.3$. So, we compare the BER of them under such two cases, where the following channel codings are considered.

- **Low density parity check (LDPC) code.** We employ the (1296, 648) LDPC code with a rate of 1/2. The parity matrix is defined in [49]. And the back propagation (BP) is employed as the decoding algorithm.
- **Turbo code.** We employ the (6298, 1256) Turbo code with a rate of 628/3149 and the constraint length of 4. The parity bits are obtained by $y_1 = x^3 + x^1 + x^0$ and $y_2 = x^3 + x^2 + x^1 + x^0$, where x^κ represents the κ-th bits in the state of shift registers. And the feedback bit is calculated by $r_b = x^3 + x^2 + x^0$. The MAP is applied as the decoding algorithm.
- **Convolutional code (CC).** We employ the (3768, 1256) CC with a rate of 1/3. The constraint length and the structure of the shift registers are the same as that of the turbo code presented in the previous item. No tail bits are required in this case. And Viterbi decoding with hard decisions is employed as the decoding algorithm.

The simulation results are demonstrated by Fig. 13(a) and Fig. 13(b).

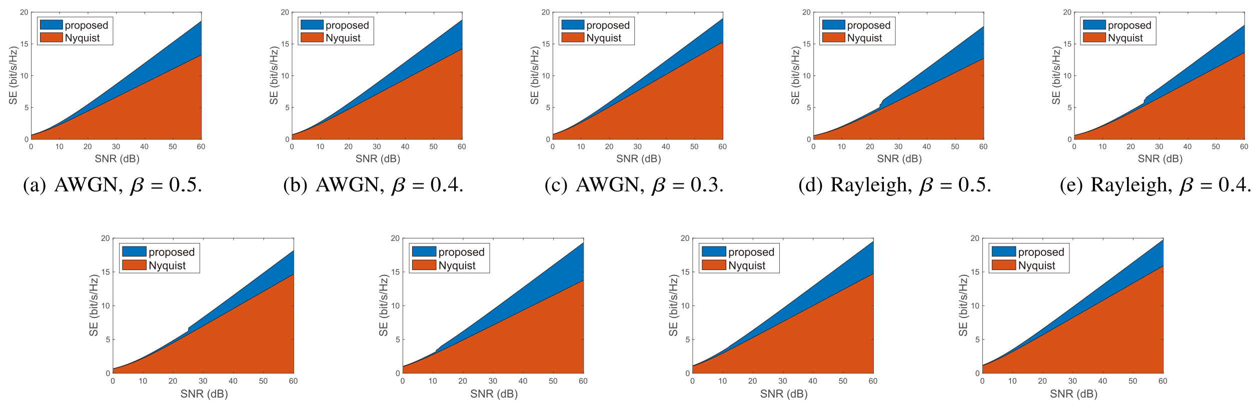
As seen, with the same SE, the proposed VPR-based secure transmission can achieve nearly the same BER performance as the conventional FTN signaling. It means that the proposed scheme can achieve security at the expense of negligible BER performance degradation.

H. Comparison Between the Simplified Packing Ratio Estimation and the Original Architecture

In this part, we compare the simplified packing ratio estimation and its original architecture [38] by the complexity and the accuracy. For the convenience of representation,

TABLE II
THE SIMULATED PACKING RATIO FOR DIFFERENT CHANNELS AND SNRS

Eb/No range (dB)	$\alpha = 1.0$	$\alpha = 0.9$	$\alpha = 0.8$	$\alpha = 0.75$	$\alpha = 0.7$	$\alpha = 0.6$
AWGN ($\beta = 0.5$)	–	–	$(-\infty, 7]$	$(7, 7.1]$	$(7.1, 8.6]$	$(8.6, +\infty)$
AWGN ($\beta = 0.4$)	–	–	$(-\infty, 7.1]$	$(7.1, 7.3]$	$(7.3, 9.9]$	$(9.9, +\infty)$
AWGN ($\beta = 0.3$)	–	–	$(-\infty, 7.6]$	$(7.6, 8.6]$	$(8.6, 12.1]$	$(12.1, +\infty)$
Rayleigh ($\beta = 0.5$)	–	$(-\infty, 23.5]$	$(23.5, 24.1]$	$(24.1, 24.4]$	$(24.4, 25.1]$	$(25.1, +\infty)$
Rayleigh ($\beta = 0.4$)	–	$(-\infty, 24.6]$	$(24.6, 25.3]$	$(25.3, 25.6]$	$(25.6, 26.9]$	$(26.8, +\infty)$
Rayleigh ($\beta = 0.3$)	–	$(-\infty, 25.2]$	–	$(25.2, 26.2]$	$(26.2, 27]$	$(27, +\infty)$
Nakagami-m ($\beta = 0.5$)	–	$(-\infty, 11.2]$	$(11.2, 12]$	$(12, 12.6]$	$(12.6, 13.2]$	$(13.2, +\infty)$
Nakagami-m ($\beta = 0.4$)	–	–	$(-\infty, 12]$	$(12, 12.6]$	$(12.6, 14]$	$(14, +\infty)$
Nakagami-m ($\beta = 0.3$)	–	–	$(-\infty, 12]$	$(12, 13]$	$(13, 15]$	$(15, +\infty)$



(a) AWGN, $\beta = 0.5$. (b) AWGN, $\beta = 0.4$. (c) AWGN, $\beta = 0.3$. (d) Rayleigh, $\beta = 0.5$. (e) Rayleigh, $\beta = 0.4$.

(f) Rayleigh, $\beta = 0.3$. (g) Nakagami, $\beta = 0.5$. (h) Nakagami, $\beta = 0.4$. (i) Nakagami, $\beta = 0.3$.

Fig. 9. Comparison for SE of the proposed scheme and Nyquist scheme.

TABLE III
THE COMPLEXITY COMPARISON BETWEEN THE PROPOSED SIMPLIFIED ESTIMATION AND ITS ORIGINAL STRUCTURE

Algorithm	MUX	DEMUX	sum	max	S/P	$\|\mathbf{W}_1\|_0$	$\|\mathbf{W}_2\|_0$	$\|\mathbf{W}_3\|_0$	$\|\mathbf{W}_4\|_0$	multi-add
Original Structure	1	2	10	1	10	20k	500k	125k	0.25k	645.25k
Proposed Structure	0	0	1	0	1	10k	250k	62.5K	0.125k	about 32.263k

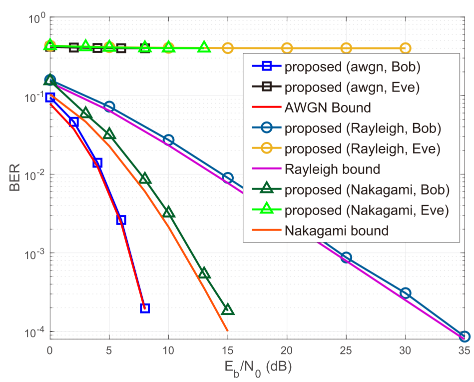


Fig. 10. BER performance of the proposed system for Alice-Bob and Alice-Eve links.

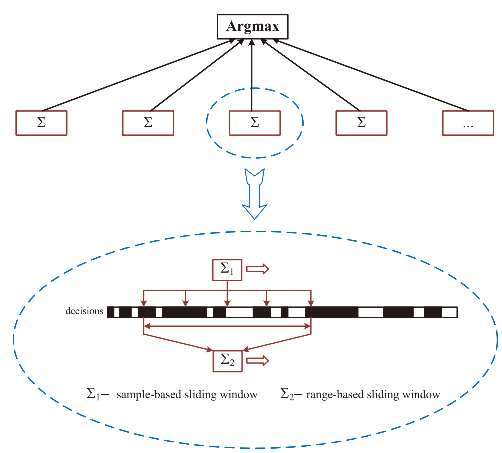
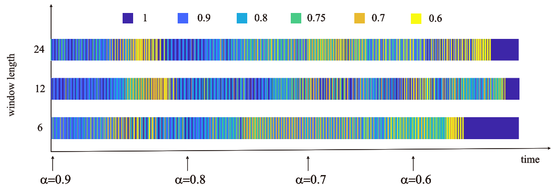


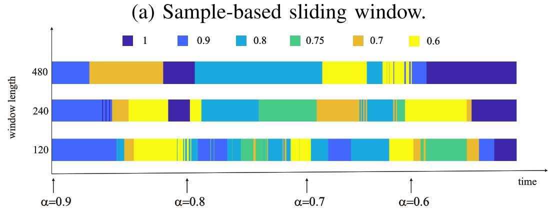
Fig. 11. The sample-based and range-based sliding windows for the simulation of the proposed estimation.

we only provide the complexity of the branch for analysis on $\alpha = 0.7$, while the total complexity is approximately proportional to it. An up-sampling times of 10 is considered. And for the convenience of comparison, the number of multi-add operations of the proposed structure is based on every 1/10 decision, since it requires only one decision when the original structure needs 10. Table III provides the complexity comparison between these two schemes.

The proposed structure nearly removes all the MUX, DEMUX, sum, maximum and S/P operations in the original design. Also, in the sparse DNN employed in our proposed simplified estimation, the number of non-zero weights in each layer has been reduced to half of that in the original network. Significantly, benefiting from the sparse DNN and the single branch structure, the number of multiply-add operations required for each estimation has been reduced to 5% of that in the original architecture. This allows more flexibility for researchers to balance the resource of time and space in practical implementation.

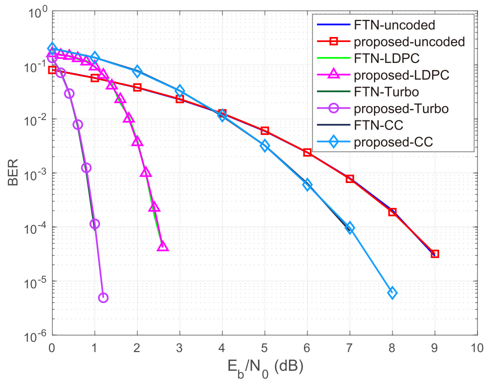


(a) Sample-based sliding window.

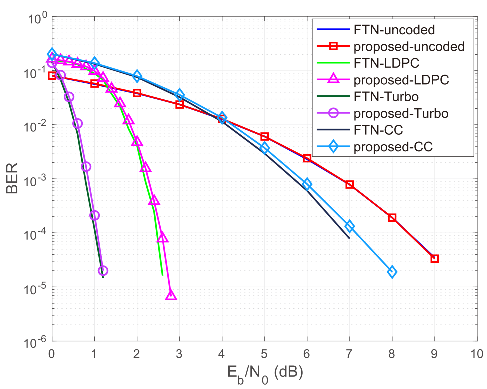


(b) Range-based sliding window.

Fig. 12. Estimation by Eve at the exampled frame.



(a) $\beta = 0.5$ and $\alpha = 0.8$.



(b) $\beta = 0.3$ and $\alpha = 0.9$.

Fig. 13. BER performance of the proposed scheme and the FTN signaling.

To more visually demonstrate the performance, we employ the accuracy as [38]

$$P_{\text{acc}} = \sum_{m=1}^{M} \sum_{n=0}^{m-1} \left(C_M^m C_M^n (p_1)^m (1-p_1)^{(M-m)} \times (p_2)^n (1-p_2)^{(M-n)} \right), \tag{80}$$

where M is the number of decisions applied to determine the final estimated value of α. p_1 is the probability that the analysis branch for $\alpha_k = \alpha$ outputs integer 1 (i.e., the diagonal items in Fig. 8). And p_2 is the maximum probabilities that the analysis branches for $\alpha_k \neq \alpha$ produce integer 1 (i.e. the maximum one of non-diagonal items within each row in Fig. 8).

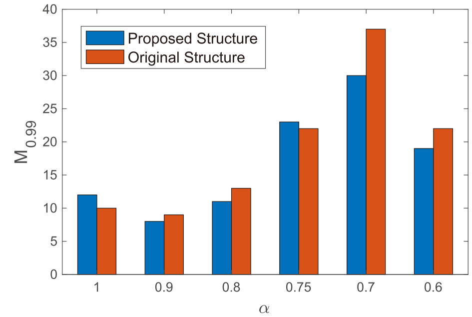


Fig. 14. The comparison of the proposed simplified estimation and its original structure in the minimum times of decisions required to achieve a 99% accuracy.

TABLE IV

COMPLEXITY COMPARISON BETWEEN THE PROPOSED SCHEME AND SOME COMMON NETWORKS

Network	Parameters	FLOPs
Proposed	0.645M	6.5×10^4
Transformer (base)	65M	3.3×10^{18}
Transformer (big)	213M	2.3×10^{19}
Inception-v4	48M	1.3×10^{10}

TABLE V

PERFORMANCE OF THE PROPOSED ESTIMATION WHICH IS TRAINED AT SNR = 4dB AND TESTED FOR DIFFERENT SNRs

SNR	4dB	3dB	2dB	1dB
p_1	0.7534	0.5084	0.4372	0.2914
p_2	0.1632	0.1596	0.1463	0.1485
$M_{0.99}$	8	21	28	95
$M_{0.999}$	13	35	47	163

Fig. 14 shows the minimum number of decisions required to achieve a 99% accuracy ($P_{acc} > 0.99$). As seen, the proposed simplified estimation can converge nearly as fast as the original structure within 35 decisions, while the complexity has been greatly reduced.

I. Complexity of the Simplified Estimation and Other Common DL Networks

Here, two common deep learning networks named Transformer [50] and Inception-v4 [51] are considered for the comparison of complexity. They are both proposed by Google and have been widely employed in natural language processing (NLP) and computer vision (CV) research fields. Their effectiveness and complexity have been verified by mass researchers and applications. Without loss of generality, the 0-value items of the proposed network are also taken into consideration in parameter counts.

The complexity comparison is shown in Table IV. As seen, the proposed scheme has an obviously lower complexity than the selected widely employed networks.

J. The Robustness of the Simplified Estimation to SNR Values

Here, the performance of the proposed simplified estimation for $\alpha = 0.9$ in AWGN channels under different SNR values is listed in Table V. As shown, although the model is trained at SNR=4dB, it can work well for other SNR values. It can effectively reduce the resource required for the proposed estimation during both the training and the implementation stages.

VIII. Conclusion

This paper proposed VPR transmissions for high SE and security, respectively, based on FTN and DL. The VPR-based system achieved a higher SE than the conventional Nyquist transmission without consuming extra spectrum resources and modifying the existing communication paradigms (e.g., spectrum allocation or frame structure). More importantly, considering security, a dynamic generation scheme was proposed to produce secret and randomly distributed positions for the segments of the VPR system. The scheme was demonstrated to be effective in avoiding detection and attack. In addition, we derived the expression for the capacity of the proposed VPR system in different channels, which were also effective for conventional FTN signaling. Finally, a simplified symbol packing ratio, which had been employed in the proposed system, was developed in this paper. Simulation results proved that it achieved nearly the same performance as the original structure with only 5% of the complexity in the original design.

In fact, there are still many open issues with the proposed VPR system beckoning further research. For example, how to design an effective switching strategy for the VPR system considering practical factors (e.g., interference, relay, energy harvesting, etc.), especially the nondeterministic polynomial (NP)-hard scenario is considered? How to derive the SE of the proposed scheme in other channels? How to analyze the security performance of the proposed VPR system more comprehensively? These issues will be studied in our future works.

References

[1] J. E. Mazo, "Faster-than-Nyquist signaling," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1451–1462, 1975.

[2] A. D. Liveris and C. N. Georghiades, "Exploiting faster-than-Nyquist signaling," *IEEE Trans. Commun.*, vol. 51, no. 9, pp. 1502–1511, Sep. 2003.

[3] J. B. Anderson, A. Prlja, and F. Rusek, "New reduced state space BCJR algorithms for the ISI channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, South Korea, Jun. 2009, pp. 889–893.

[4] E. Bedeer, M. H. Ahmed, and H. Yanikomeroglu, "A very low complexity successive symbol-by-symbol sequence estimator for faster-than-Nyquist signaling," *IEEE Access*, vol. 5, pp. 7414–7422, 2017.

[5] P. Song, F. Gong, Q. Li, G. Li, and H. Ding, "Receiver design for faster-than-Nyquist signaling: Deep-learning-based architectures," *IEEE Access*, vol. 8, pp. 68866–68873, 2020.

[6] B. Liu, S. Li, Y. Xie, and J. Yuan, "A novel sum-product detection algorithm for faster-than-Nyquist signaling: A deep learning approach," *IEEE Trans. Commun.*, vol. 69, no. 9, pp. 5975–5987, Sep. 2021.

[7] T. Petitpied, R. Tajan, P. Chevalier, S. Traverso, and G. Ferre, "Circular faster-than-Nyquist signaling for high spectral efficiencies: Optimized EP-based receivers," *IEEE Trans. Commun.*, vol. 69, no. 8, pp. 5487–5501, Aug. 2021.

[8] A. Ibrahim, E. Bedeer, and H. Yanikomeroglu, "A novel low complexity faster-than-Nyquist signaling detector based on the primal-dual predictor-corrector interior point method," *IEEE Commun. Lett.*, vol. 25, no. 7, pp. 2370–2374, Jul. 2021.

[9] S. Sugiura, "Frequency-domain equalization of faster-than-Nyquist signaling," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 555–558, Oct. 2013.

[10] T. Ishihara and S. Sugiura, "Frequency-domain equalization aided iterative detection of faster-than-Nyquist signaling with noise whitening," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[11] F. Rusek and J. B. Anderson, "Multistream faster than Nyquist signaling," *IEEE Trans. Commun.*, vol. 57, no. 5, pp. 1329–1340, May 2009.

[12] H. Che, K. Zhu, and Y. Bai, "Multicarrier faster-than-Nyquist based on efficient implementation and probabilistic shaping," *IEEE Access*, vol. 9, pp. 63943–63951, 2021.

[13] T. Ishihara and S. Sugiura, "Reduced-complexity FFT-spread multicarrier faster-than-Nyquist signaling in frequency-selective fading channel," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 530–542, 2022.

[14] Y. Ma, N. Wu, J. A. Zhang, B. Li, and L. Hanzo, "Generalized approximate message passing equalization for multi-carrier faster-than-Nyquist signaling," *IEEE Trans. Veh. Technol.*, vol. 71, no. 3, pp. 3309–3314, Mar. 2021.

[15] J. B. Anderson and F. Rusek, "Improving OFDM: Multistream faster-than-Nyquist signaling," in *Proc. 4th Int. Symp. Turbo Codes Rel. Topics, 6th Int. ITG-Conf. Source Channel Coding*, 2006, pp. 1–5.

[16] F. Rusek, "On the existence of the Mazo-limit on MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1118–1121, Mar. 2009.

[17] A. T. Abebe and C. G. Kang, "FTN-based MIMO transmission as a NOMA scheme for efficient coexistence of broadband and sporadic traffics," in *Proc. IEEE 87th Veh. Technol. Conf. (VTC Spring)*, Jun. 2018, pp. 1–5.

[18] M. Yuhas, Y. Feng, and J. Bajcsy, "On the capacity of faster-than-Nyquist MIMO transmission with CSI at the receiver," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2015, pp. 1–6.

[19] M. McGuire, A. Dimopoulos, and M. Sima, "Faster-than-Nyquist single-carrier MIMO signaling," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2016, pp. 1–7.

[20] S. Wen, G. Liu, C. Liu, H. Qu, L. Zhang, and M. A. Imran, "Joint precoding and pre-equalization for faster-than-Nyquist transmission over multipath fading channels," *IEEE Trans. Veh. Technol.*, vol. 71, no. 4, pp. 3948–3963, Apr. 2022.

[21] T. Ishihara and S. Sugiura, "Iterative frequency-domain joint channel estimation and data detection of faster-than-Nyquist signaling," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6221–6231, Sep. 2017.

[22] Q. Li, F.-K. Gong, P.-Y. Song, G. Li, and S.-H. Zhai, "Joint channel estimation and precoding for faster-than-Nyquist signaling," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13139–13147, Nov. 2020.

[23] T. Ishihara, S. Sugiura, and L. Hanzo, "The evolution of faster-than-Nyquist signaling," *IEEE Access*, vol. 9, pp. 86535–86564, 2021.

[24] J. Zhou et al., "Digital signal processing for faster-than-Nyquist non-orthogonal systems: An overview," in *Proc. 26th Int. Conf. Telecommun. (ICT)*, Apr. 2019, pp. 295–299.

[25] J. Fan, S. Guo, X. Zhou, Y. Ren, G. Ye Li, and X. Chen, "Faster-than-Nyquist signaling: An overview," *IEEE Access*, vol. 5, pp. 1925–1940, 2017.

[26] F. Rusek and J. B. Anderson, "Constrained capacities for faster-than-Nyquist signaling," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 764–775, Feb. 2009.

[27] J. Wang, W. Tang, X. Li, and S. Li, "Filter hopping based faster-than-Nyquist signaling for physical layer security," *IEEE Wireless Commun. Lett.*, vol. 64, no. 5, pp. 2122–2128, May 2018.

[28] ETSI. (2015). *Digital Video Broadcasting (DVB); Implementation Guidelines for the Second Generation System for Broadcasting, Interactive Services, News Gathering and Other Broadband Satellite Applications; Part 1: DVB-S2*," [Online]. Available: https://www.etsi.org/deliver/etsi_tr/102300_102399/10237601/()01.02.01_60/tr_10237601v010201p.pdf

[29] T. Aono, K. Higuchi, T. Ohira, B. Komiyama, and H. Sasaoka, "Wireless secret key generation exploiting reactance-domain scalar response of multipath fading channels," *IEEE Trans. Antennas Propag.*, vol. 53, no. 11, pp. 3776–3784, Nov. 2005.

[30] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, "Radio-telepathy: Extracting a secret key from an unauthenticated wireless channel," in *Proc. 14th ACM Int. Conf. Mobile Comput. Netw.*, Sep. 2008, pp. 128–139.

[31] N. Patwari, J. Croft, S. Jana, and S. K. Kasera, "High-rate uncorrelated bit extraction for shared secret key generation from channel measurements," *IEEE Trans. Mobile Comput.*, vol. 9, no. 1, pp. 17–30, Jan. 2010.

[32] G. Brassard and L. Salvail, "Secret-key reconciliation by public discussion," in *Proc. Workshop Theory Appl. Cryptograph. Techn.* Cham, Switzerland: Springer, 1993, pp. 410–423.

[33] A. L. Rukhin et al., *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*. Gaithersburg, MD, USA: National Institute of Standards and Technology, Apr. 2010.

[34] M. G. Madiseh, S. W. Neville, and M. L. McGuire, "Applying beamforming to address temporal correlation in wireless channel characterization-based secret key generation," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 4, pp. 1278–1287, Aug. 2012.

[35] N. Aldaghri and H. Mahdavifar, "Physical layer secret key generation in static environments," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 2692–2705, 2020.

[36] A. Sayeed and A. Perrig, "Secure wireless communications: Secret keys through multipath," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2008, pp. 3013–3016.

[37] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[38] P. Song, F. Gong, and Q. Li, "Blind symbol packing ratio estimation for faster-than-Nyquist signalling based on deep learning," *Electron. Lett.*, vol. 55, no. 21, pp. 1155–1157, Oct. 2019.

[39] S. O. Rice, "Mathematical analysis of random noise," *Bell Syst. Tech. J.*, vol. 23, no. 3, pp. 282–332, 1944.

[40] M. Nakagami, "The m-distribution: A general formula of intensity distribution of the rapid fading," in *Statistical Methods in Radio Wave Propagation*. Oxford, U.K.:Pergamon, 1960.

[41] E. Cubukcu, "Root raised cosine (RRC) filters and pulse shaping in communication systems," in *Proc. AIAA Conf.*, 2012, pp. 1–30.

[42] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, vol. 55. Washington, DC, USA: U.S. Government printing office, 1964.

[43] G. B. Thomas and R. L. Finney, *Calculus*. Reading, MA, USA: Addison-Wesley, 1961.

[44] C. C. Ross, *Differential Equations: An Introduction With Mathematica*. New York, NY, USA: Springer, 2004.

[45] P. Flajolet, X. Gourdon, and P. Dumas, "Mellin transforms and asymptotics: Harmonic sums," *Theor. Comput. Sci.*, vol. 144, nos. 1–2, pp. 3–58, Jun. 1995.

[46] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals Series and Products*. New York, NY, USA: Academic, 2000.

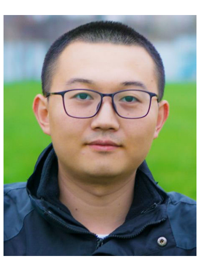
[47] H. Bateman, *Higher Transcendental Functions [Volumes I-III]*, vol. 1. New York, NY, USA: McGraw-Hill, 1953.

[48] S. Li, B. Bai, J. Zhou, P. Chen, and Z. Yu, "Reduced-complexity equalization for faster-than-Nyquist signaling: New methods based on Ungerboeck observation model," *IEEE Trans. Commun.*, vol. 66, no. 3, pp. 1190–1204, Mar. 2018.

[49] *IEEE Standard for Information Technology–Telecommunications and Information Exchange Between Systems—Local and Metropolitan Area Networks–Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications—Redline*, IEEE Standard 802.11-2020 (Revision IEEE Standard 802.11-2016), 2021.

[50] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.

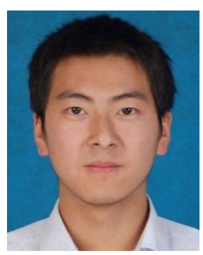
[51] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 1–12.



Peiyang Song (Student Member, IEEE) was born in Henan, China, in 1994. He received the B.Sc. degree from the School of Communications Engineering, Xidian University, in 2017, where he is currently pursuing the Ph.D. degree with the State Key Laboratory of Integrated Services Networks (ISN). His research interests include AI-aided wireless communications, faster than Nyquist signaling, and multibeam satellite.

Nan Zhang received the B.Sc. and M.Sc. degrees in telecommunication engineering in 2003 and 2006, respectively, and the Ph.D. degree in telecommunication and information system from Xidian University, Shaanxi, China, in 2012. Her current research interests include study and implementation of key technologies of communication systems under complex environments.

Lin Cai (Fellow, IEEE) received the M.A.Sc. and Ph.D. degrees in electrical and computer engineering from the University of Waterloo, Waterloo, Canada, in 2002 and 2005, respectively. Since 2005, she has been with the Department of Electrical and Computer Engineering, University of Victoria, and she is currently a Professor. She is also an NSERC E. W. R. Steacie Memorial Fellow and an Engineering Institute of Canada (EIC) Fellow. In 2020, she was elected as a member of the Royal Society of Canada's College of New Scholars, Artists and Scientists, and the 2020 "Star in Computer Networking and Communications" by N2Women. Her research interests span several areas in communications and networking, with a focus on network protocol and architecture design supporting emerging multimedia traffic and the Internet of Things. She has co-founded and chaired the IEEE Victoria Section Vehicular Technology and Communications Joint Societies Chapter. She has been elected to serve the IEEE Vehicular Technology Society Board of Governors (2019–2024) and served its VP Mobile Radio in 2023. She has been a Voting Board Member of IEEE Women in Engineering (2022–2023). She has served as the Associate Editor-in-Chief for IEEE Transactions on Vehicular Technology, a member of the Steering Committee of the IEEE Transactions on Mobile Computing, IEEE Transactions on Big Data, and IEEE Transactions on Cloud Computing, an Associate Editor of the IEEE Internet of Things Journal, IEEE/ACM Transactions on Networking, IEEE Transactions on Wireless Communications, IEEE Transactions on Vehicular Technology, and IEEE Transactions on Communications, and as the Distinguished Lecturer of the IEEE VTS Society and the IEEE Communications Society. She received the Award Outstanding Achievement in Graduate Studies for M.A.Sc. and Ph.D. degrees.



Guo Li (Member, IEEE) received the B.S., master's, and Ph.D. degrees from Xidian University, Xi'an, China, in 2011, 2014, and 2017, respectively. He is currently a Lecturer with the School of Telecommunication Engineering, Xidian University. His research interests include 6G communications, terrestrial-satellite communication systems, beamforming and its applications, and optimization theory.

Tong Wu was born in Weinan, Shaanxi. He received the Ph.D. degree from the University of Chinese Academy of Sciences in 2017. He is currently a Senior Engineer with the China Academy of Space Technology (CAST), Xi'an. His main research directions are satellite communications.

Feng-Kui Gong (Member, IEEE) was born in Shandong, China, in 1979. He received the M.S. and Ph.D. degrees from Xidian University, Xi'an, China, in 2004 and 2007, respectively. From 2011 to 2012, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. He is currently a Professor with the State Key Laboratory of Integrated Services Networks, Department of Communication Engineering, Xidian University. His research interests include cooperative communication, distributed space-time coding, digital video broadcasting systems, satellite communication, and 4G/5G techniques.