Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 4181626, 10 pages
https://doi.org/10.1155/2018/4181626

*Research Article*

# Superposition Coded Modulation Based Faster-Than-Nyquist Signaling

**Shuangyang Li**,[1,2] **Baoming Bai**,[1] **Jing Zhou**,[3] **Qingli He**,[1] **and Qian Li**[1]

[1]*State Key Laboratory of ISN, Xidian University, Xi'an, China*
[2]*Science and Technology on Communication Networks Laboratory, Shijiazhuang, China*
[3]*Department of EEIS, University of Science and Technology of China, Hefei, China*

Correspondence should be addressed to Baoming Bai; bmbai@mail.xidian.edu.cn

Received 24 November 2017; Accepted 31 March 2018; Published 16 May 2018

Academic Editor: Luca Reggiani



A structure of faster-than-Nyquist (FTN) signaling combined with superposition coded modulation (SCM) is considered. The so-called FTN-SCM structure is able to achieve the constrained capacity of FTN signaling and only requires a low detection complexity. By deriving a new observation model suitable for FTN-SCM, we offer the power allocation based on a proper detection method. Simulation results show that, at any given spectral efficiency, the bit error rate (BER) curve of FTN-SCM lies clearly outside the minimum signal-to-noise ratio (SNR) boundary of orthogonal signaling with a larger alphabet. The achieved data rates are also close to the maximum data rates of the certain shaping pulse.

## 1. Introduction

With the demand and the growth of advanced signal processing capabilities at base stations, the need of efficient backhauling solutions to transmit a large amount of data increases significantly. Thus, as one of the most important parts of deploying the fifth-generation (5G) cellular network, more efficient backhauling techniques need to be applied [2]. Conventionally, the capacity of networks is enlarged by consuming more time/bandwidth/spatial resources. However, this solution may not always be possible, due to the practical reasons. Hence, as an alternative method to gain more capacity, FTN signaling has recently received a lot of attention. An overview of FTN signaling for 5G communication systems was provided in [3].

FTN signaling is an extension of traditional linear modulation and a classical way of nonorthogonal signal transmission, which was first proposed by Mazo in 1975 [4]. Mazo discovered that, with *sinc* pulse as the shaping pulse, the minimum squared Euclidean distance of binary phase shift keying (BPSK) modulated pulses remains the same even when the symbol rate is, to some extent, higher than the Nyquist criterion. His work indicates that there are roughly 25% more bits that could be transmitted in the same bandwidth compared to that of Nyquist signaling, with almost the same error performance over additive white Gaussian noise (AWGN) channels. Recently, Rusek et al. proved that FTN signaling is able to bring more degrees of freedom (DoF) over the AWGN channel [5, 6] compared to orthogonal signaling. As a result, a higher spectral efficiency is expected for FTN signaling and, indeed, fascinating simulation results have already been reported. In [7], a precoded FTN system with quadrature phase shift keying (QPSK) modulation was presented, which, as simulation results imply, requires lower SNR to reach the BER $< 10^{-5}$ compared to that of the constrained capacity of 8-PSK for orthogonal signaling with the same spectral efficiency. However, there is still no such signaling method existing in the literature that is able to outperform orthogonal signaling constrained by a larger alphabet at any preferred spectral efficiency. The reason for this problem lies in the complexity of maximum-likelihood (ML) detection for FTN signals growing exponentially with the size of the alphabet and with the number of taps of intersymbol interference (ISI), respectively. When the system requires high spectral efficiency, conventional FTN signaling systems need either an alphabet with a larger size or a compression factor of a smaller

value to meet the requirements. Consequently, the required ML detection complexity becomes prohibitively high and a suboptimal detection method has to be utilized, which in return somehow damages the performance. Hence, in this paper, we attempt to solve such an issue by considering SCM [8–13].

SCM is a special case of multilevel coding (MLC) [8], which offers an excellent solution to transmissions with severe interference. With the use of the fast Fourier transform- (FFT-) based technique proposed in [9], the detection complexity of SCM system is $O(\log N_{\text{frame}})$, where $N_{\text{frame}}$ is denoted as the frame length [10]. Moreover, with proper Gaussian assumption, the optimization for SCM systems is easier than that of conventional bit interleaved coded modulation (BICM) systems [10]. SCM has also been proven to have promising performance over a variety of channels [11, 12]. More advantages of SCM can be found in [10] and the references therein.

The idea of combining SCM with FTN signaling first appeared in [14], where FTN signals are treated as the sum of several orthogonal signals with different time delays; thus it allows the successive interference cancellation (SIC) detection at the receiver. However, in [1], it has been proven that the aforementioned structure cannot bring any gain in terms of DoF. Hence, a so-called "full-FTN" structure has been proposed in [1] along with its proof of achieving the constrained capacity of FTN signaling. In this structure, the signals are viewed as the sum of several FTN signals of the same compression factor and the SIC is also utilized to reduce the detection complexity. Different from the traditional FTN signaling, to gain a higher spectral efficiency, this structure utilizes more layers rather than a small compression factor. Since, with SIC detection, the detection complexity grows linearly with the number of layers and exponentially with the number of ISI taps, the overall detection complexity of this structure is normally very low. On the other hand, since at each layer, the symbol rate still exceeds the signal bandwidth, the DoF gain of FTN signaling is maintained. However, this structure still lacks a well-designed equalizer to perform SIC, because the common equalizers for FTN signaling, such as the one in [15], require the utilization of the whitening filter in the receiver. This is rather difficult and even impossible when the FTN signal, at each layer, is corrupted by both the colored noise and the signals from other layers. Hence, it is needed to derive a new observation model, which allows the SIC and the detection for each individual layer at the same time. It should also be noted that the combination of FTN signaling and SCM bypasses the obstacle of designing the channel code in terms of different compression factors. Similar to the traditional SCM, an identical code can be utilized for all layers of FTN-SCM, which makes the design and implementation of FTN-SCM system very easy. By simply adjusting the number of superposition layers and the power allocation, FTN-SCM is able to provide a wide range of spectral efficiencies with excellent performance.

The main contributions of this paper are summarized in the following:

(1) We adapt the idea from [1] and provide a more generalized FTN-SCM scheme.

(2) A new channel observation model suitable for FTN-SCM is introduced.

(3) The detection method and the corresponding power allocation for FTN-SCM are also discussed.

(4) Simulation results show that, for BER $< 10^{-5}$, FTN-SCM requires lower SNR than that of the orthogonal signaling with a larger alphabet at any given spectral efficiency.

The rest of this paper is organized as follows. The diagram of FTN-SCM is provided in Section 2. Then, the new channel observation is derived in Section 3. In Section 4, the detection method is described, along with the power allocation derivation. Our numerical results are presented in Section 5, and finally a brief conclusion is provided in Section 6.

## 2. System Model

The transmitter structure of FTN-SCM is illustrated in Figure 1. Assume that the sequence $\mathbf{u}$ carrying information bits is separated into $K$ substreams, namely, $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{K-1}$. Each subsequence of $\mathbf{u}$, say $\mathbf{u}_k$, is then encoded by its corresponding encoder generating the codeword $\mathbf{c}_k$ of length $N$. $\mathbf{c}'_k$ is the permuted version of $\mathbf{c}_k$, which is afterward modulated in the form of BPSK with an average symbol energy $E_k = P_k \tau T$, where $P_k$ is a pregiven power, $\tau$ is the compression factor, and $\tau T$ is the symbol time. $\mathbf{x}_k = \{x_k[1], x_k[2], \ldots, x_k[n], \ldots, x_k[N]\}^{\mathrm{T}}$ represents the modulated symbols at the $k$th layer, which are then superposed directly with the modulated symbols from other layers. The transmitted symbol sequence $\mathbf{x}$ is obtained as the superposition is finished, where the $n$th symbol of $\mathbf{x}$ is given as $x[n] = \sum_{k=0}^{K-1} x_k[n]$. The FTN modulator is able to shape the transmitted signal $s(t)$ for the given input $\mathbf{x}$ based on a certain $T$-orthogonal pulse $h(t)$. Without loss of generality, an FTN-SCM signal can be expressed as

$$s(t) = \sum_{n=1}^{N} x[n]\, h(t - n\tau T) = \sum_{k=0}^{K-1} \sum_{n=1}^{N} x_k[n]\, h(t - n\tau T). \quad (1)$$

A brief diagram of FTN-SCM signal is given in Figure 2, where $K = 2$ and $\tau = 0.5$. As shown in the figure, the pulse of each individual symbol is interfered by the pulses from both the current layer and the other layers. It should be mentioned that, in this case, a symbol rate that is higher than the Nyquist criterion is maintained at each layer. Thus, the capacity gain of FTN signaling is preserved. Note that the different power assignment for each layer is not the only way of performing SIC in the receiver, similar to that of the orthogonal signaling; choosing codes of different rate for each layer may also do the work.

Figure 3 illustrates the receiver structure of FTN-SCM. In this paper, as we only focus on AWGN channels, the received signal $r(t)$ is presented as $r(t) = s(t) + w(t)$, where $w(t)$ has one side power spectral density (PSD) $N_0$. Let $g_n$

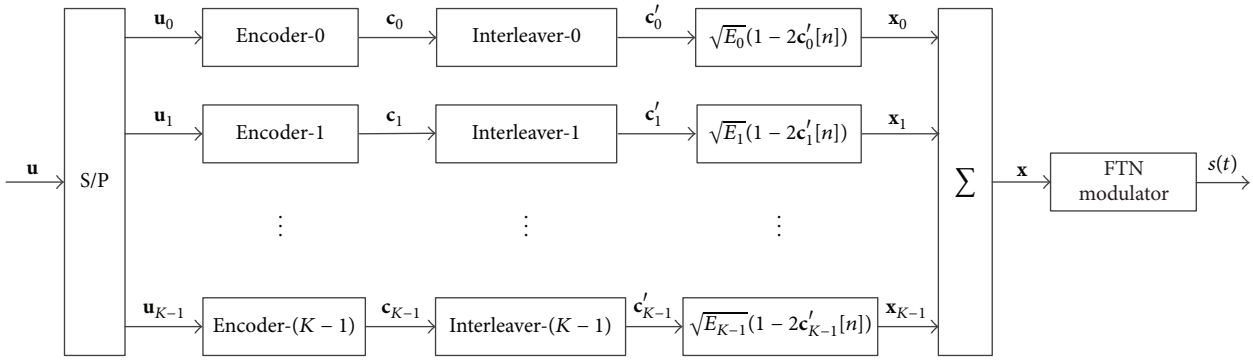


Figure 1: The transmitter structure of FTN-SCM.

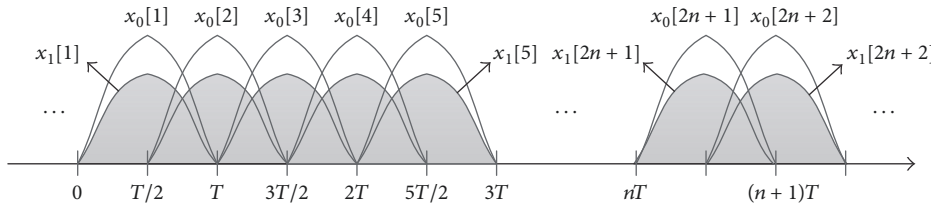


Figure 2: A brief diagram of an FTN-SCM signal with $K = 2$ and $\tau = 0.5$.

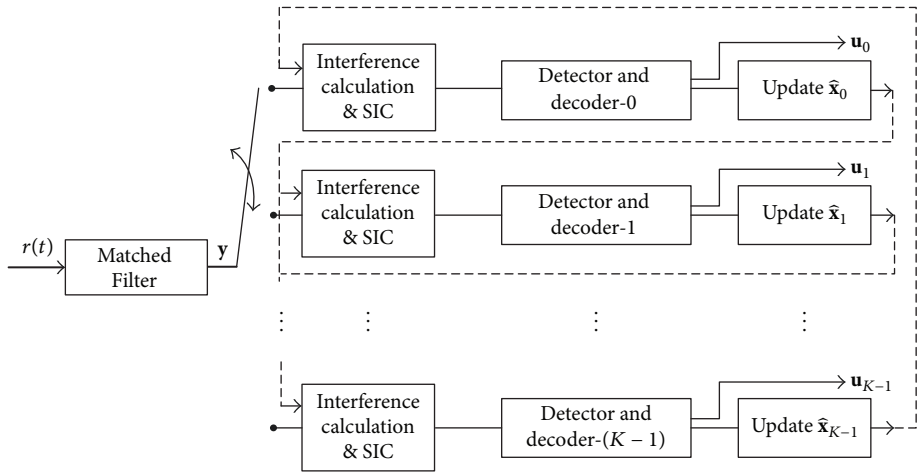


Figure 3: The receiver structure of FTN-SCM.

represent the autocorrelation function samples of $h(t)$. We have

$$g_n = \int_{-\infty}^{\infty} h(t)\, h^*(t - n\tau T)\, \mathrm{d}t, \quad -L_\mathrm{I} \le n \le L_\mathrm{I}, \tag{2}$$

where $(\cdot)^*$ denotes the complex conjugation and $L_\mathrm{I}$ is the length of finite ISI tap. The output of the matched filter is denoted as $\mathbf{y}$. We thus have

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \boldsymbol{\eta}, \tag{3}$$

where $\mathbf{G}$ is a Toeplitz matrix given as

$$\mathbf{G} = \begin{pmatrix} 1 & g_{-1} & \cdots & g_{-L_\mathrm{I}} & 0 & 0 & 0 & 0 & \cdots \\ g_1 & 1 & g_{-1} & \cdots & g_{-L_\mathrm{I}} & 0 & 0 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & & \ddots & & \ddots & \\ g_{L_\mathrm{I}} & \cdots & g_1 & 1 & g_{-1} & \cdots & g_{-L_\mathrm{I}} & 0 & \cdots \\ 0 & g_{L_\mathrm{I}} & \cdots & g_1 & 1 & g_{-1} & \cdots & g_{-L_\mathrm{I}} & \cdots \\ \vdots & & \ddots & & \ddots & \ddots & \ddots & & \ddots \end{pmatrix} \tag{4}$$

and $\boldsymbol{\eta}$ is the colored-noise vector with zero mean and the covariance matrix $\mathrm{E}[\boldsymbol{\eta}\boldsymbol{\eta}^\mathrm{H}] = (N_0/2)\mathbf{G}$. Here, $\mathrm{E}[\cdot]$ is the expectation operator and $(\cdot)^\mathrm{H}$ is the Hermitian (conjugate) transpose.

Without loss of generality, the detection can start at the first layer, and after the detection of each layer, the estimation of current layer inputs, say $\widehat{\mathbf{x}}_k$, is stored for the interference calculation of the following layers. Note that, after one complete sweep over all layers, the updated estimation can be reused to perform other sweeps, which is able to further improve the interference calculation. Normally, three complete sweeps would be enough for FTN-SCM systems.

## 3. Channel Observation Model

We consider the minimum distance detector in this paper. Based on the receiver structure, for detecting the $k$th layer, we have

$$\widehat{\mathbf{x}}_k = \arg\min_{\mathbf{x}_k} \int_{-\infty}^{\infty} \left| r(t) - \sum_{\substack{i=0 \\ i\neq k}}^{K-1} \widehat{s}_i(t) - s_k(t) \right|^2 \mathrm{d}t, \tag{5}$$

where

$$\widehat{s}_i(t) = \sum_{n=1}^{N} \widehat{x}_i[n]\, h(t - n\tau T) \tag{6}$$

represents the estimation of the signal of the $i$th layer. By expanding the equation and dropping the irrelevance, (5) yields

$$\widehat{\mathbf{x}}_k = \arg\max_{\mathbf{x}_k} \int_{-\infty}^{\infty} \left\{ \mathrm{Re}\left\{ \left( r(t) - \sum_{\substack{i=0 \\ i\neq k}}^{K-1} \widehat{s}_i(t) \right) s_k^*(t) \right\} - \frac{1}{2} s_k(t)\, s_k^*(t) \right\} \mathrm{d}t, \tag{7}$$

where $\mathrm{Re}(\cdot)$ represents the real part of a complex number. By switching the integral sequence, (7) can be further derived as

$$\widehat{\mathbf{x}}_k = \arg\max_{\mathbf{x}_k} \mathrm{Re}\left\{ \sum_{n=1}^{N} x_k^*[n] \cdot \int_{-\infty}^{\infty} \left( r(t) - \sum_{\substack{i=0 \\ i\neq k}}^{K-1} \widehat{s}_i(t) \right) h^*(t - n\tau T)\, \mathrm{d}t \right\} - \frac{1}{2} \int_{-\infty}^{\infty} s_k(t)\, s_k^*(t)\, \mathrm{d}t. \tag{8}$$

Thus, with respect to the matched filter outputs, we get

$$\begin{aligned} \widehat{\mathbf{x}}_k = \arg\max_{\mathbf{x}_k} & \sum_{n=1}^{N} \mathrm{Re}\left\{ \left( y[n] - a_k[n] \right) x_k^*[n] \right\} \\ & - \frac{1}{2} \sum_{m=1}^{N} \sum_{n=1}^{N} x_k^*[m]\, x_k[n]\, g_{m-n}, \end{aligned} \tag{9}$$

where

$$a_k[n] = \sum_{\substack{i=0 \\ i\neq k}}^{K-1} \sum_{j=-L_1}^{L_1} \widehat{x}_i[n+j]\, g_j. \tag{10}$$

It is obvious that (9) enables the implementation of the Viterbi algorithm [16]. Similarly, as for soft-in soft-out (SISO) algorithms, for example, the BCJR algorithm [17], (9) implies the recursive probabilistic factorization of the form

$$\begin{aligned} P(\mathbf{y} \mid \mathbf{x}_k, \mathbf{a}_k) \propto \prod_{n=1}^{N} \exp & \left\{ \frac{1}{N_0/2 + \sigma_a^2[n]} \right. \\ & \left. \cdot \mathrm{Re}\left\{ x_k^*[n] \cdot \left( y_n - a_k[n] - \frac{1}{2} g_0 x_k[n] - \sum_{l=1}^{L_1} g_l x_k[n-l] \right) \right\} \right\}, \end{aligned} \tag{11}$$

in which $a_k[n]$ is assumed to be Gaussian, and its variance is denoted as $\sigma_a^2[n]$. Hence, the derivation of the channel observation model is complete.

## 4. Detection Method and Power Allocation

To detect FTN-SCM signal, a well-designed equalizer that accepts nonwhite noise is necessary. Thus, we choose the original method from [18] as the method detecting each layer. The method has been proven to offer promising performance based on the Ungerboeck observation model [19]. As an extension of the traditional $M$-algorithm BCJR ($M$-BCJR) algorithm, the detection method selects the best $M$ states not only based on the current symbol but also considering the influence of some "future" symbols. At each trellis section, for each possible state $S_n$, the method calculates the metrics of the path $\mathbf{v} = x_1^n$ that induces $S_n$ and all possible paths from the section $n+1$ till $n+L$ that are extended from $S_n$. However, new concerns arise due to the presence of other layers. In the following, we aim to offer a performance analysis for the detection of each layer and we further give the power allocation of each layer.

We believe slightly abusing the notation is acceptable. We henceforth use $x_n$ and $a_n$ in place of $x_k[n]$ and $a_k[n]$, respectively, and then the sequence of $\{x_1, x_2, \ldots, x_N\}$ can be represented as $x_1^N$. Without loss of generality, we assume

$x_n$ equiprobably taking values in the alphabet. Hence, in our case, for detecting the $k$th layer, based on the description in [18] and the aforementioned observation model, the metric of path $v_1^{n+L} = x_1^{n+L} + e_1^{n+L}$ with a random error pattern $e_1^{n+L}$ is represented as

$$J\left(v_1^{n+L}\right) = \mathrm{Re}\left\{\left(v_1^{n+L}\right)^{\mathrm{H}}\left(y_1^{n+L} - a_1^{n+L}\right) - \frac{1}{2}\left\|v_1^{n+L}\right\|^2 - \left(v_1^{n+L}\right)^{\mathrm{H}}\mathbf{G}_{\mathrm{L}_{(n+L)\times(n+L)}}\left(v_1^{n+L}\right) + \left(v_1^{n+L}\right)^{\mathrm{H}}\eta_1^{n+L}\right\}, \tag{12}$$

where $\|\cdot\|^2$ is the norm operator and $\mathbf{G}_{\mathrm{L}_{(n+L)\times(n+L)}}$ is the lower triangular matrix with zero main diagonal of the size $(n+L)\times(n+L)$ that is denoted as

$$\mathbf{G}_{\mathrm{L}_{(n+L)\times(n+L)}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ g_1 & 1 & 0 & 0 & 0 & \cdots \\ \vdots & \ddots & 1 & 0 & 0 & \cdots \\ g_{L_{\mathrm{I}}} & \cdots & g_1 & 1 & 0 & \cdots \\ 0 & g_{L_{\mathrm{I}}} & \cdots & g_1 & 1 & \ddots \\ \vdots & & \ddots & & \ddots & \ddots \end{pmatrix}. \tag{13}$$

As we define size A $\triangleq (n+L)\times(n+L+L_{\mathrm{I}})$ and size B $\triangleq (n+L)\times(n+L)$, by substituting (3) into (12), we have

$$\begin{aligned} J\left(v_1^{n+L}\right) = \mathrm{Re}\Big\{&\left(x_1^{n+L} + e_1^{n+L}\right)^{\mathrm{H}}\mathbf{G}_{\mathrm{A}}x_1^{n+L+L_{\mathrm{I}}} \\ &- \frac{1}{2}\left\|x_1^{n+L} + e_1^{n+L}\right\|^2 \\ &- \left(x_1^{n+L} + e_1^{n+L}\right)^{\mathrm{H}}\mathbf{G}_{\mathrm{L}_{\mathrm{B}}}\left(x_1^{n+L} + e_1^{n+L}\right) \\ &+ \left(x_1^{n+L} + e_1^{n+L}\right)^{\mathrm{H}}\mathbf{G}_{\mathrm{B}}b_1^{n+L} + \left(x_1^{n+L} + e_1^{n+L}\right)^{\mathrm{H}}\eta_1^{n+L}\Big\}, \end{aligned} \tag{14}$$

in which $b_n$ represents the accuracy of the estimation and is given as

$$b_n = \sum_{\substack{i=0 \\ i\neq k}}^{K-1} x_i[n] - \widehat{x}_i[n], \tag{15}$$

with the variance $\sigma_b^2[n]$.

Furthermore, we consider the difference of the metrics of two individual erroneous paths. Define the two paths as $v_1^{n+L} = x_1^{n+L} + e_1^{n+L}$ and ${v'}_1^{n+L} = x_1^{n+L} + {e'}_1^{n+L}$ and further define $m_1^{n+L} = {e'}_1^{n+L} - e_1^{n+L}$. After several manipulations [18], we obtain

$$\begin{aligned} J\left({v'}_1^{n+L}\right) - J\left(v_1^{n+L}\right) = \mathrm{Re}\Big\{&\left(m_1^{n+L}\right)^{\mathrm{H}}\mathbf{T}x_{n+L+1}^{n+L+L_{\mathrm{I}}} \\ &- \left(m_1^{n+L}\right)^{\mathrm{H}}\mathbf{G}_{\mathrm{B}}\left(e_1^{n+L} + \frac{1}{2}m_1^{n+L}\right)\Big\} \\ &+ \mathrm{Re}\left\{\left(m_1^{n+L}\right)^{\mathrm{H}}\mathbf{G}_{\mathrm{B}}b_1^{n+L} + \left(m_1^{n+L}\right)^{\mathrm{H}}\eta_1^{n+L}\right\}, \end{aligned} \tag{16}$$

where

$$\mathbf{T} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \\ g_{-L_{\mathrm{I}}} & 0 & 0 & \cdots & 0 \\ g_{-(L_{\mathrm{I}}-1)} & g_{-L_{\mathrm{I}}} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ g_{-2} & g_{-3} & \cdots & g_{-L_{\mathrm{I}}} & 0 \\ g_{-1} & g_{-2} & \cdots & g_{-(L_{\mathrm{I}}-1)} & g_{-L_{\mathrm{I}}} \end{pmatrix}. \tag{17}$$

In the following, we focus on the detection performance of each stage. Since the detection at each stage is interfered by the signals from the other stages, it is necessary to make sure that the algorithm is still able to offer a correct detection, for which we offer the following theorem.

**Theorem 1** (correct detection criterion). *In FTN-SCM systems, the kth stage can be successfully detected without the presence of noise, if the maximum variance of $b[n]$, say $\sigma_{max}^2 = \max\{\sigma_b^2[n],\ 1 \le n \le N\}$, satisfies*

$$\sqrt{2\sigma_{max}^2}\sum_{l=-L_{\mathrm{I}}}^{L_{\mathrm{I}}}|g_l| < \frac{1}{2}\sqrt{P_k\tau T}d_{min}^2 - 2\sqrt{P_k\tau T}\sum_{l=1}^{L_{\mathrm{I}}-L} l\left|g_{-(l+L)}\right|, \tag{18}$$

*where $(d_{min}^2)$ represents the minimum squared Euclidean distance of BPSK modulated FTN signals with normalized signal energy, i.e. BPSK modulated FTN signal $s'(t)$ satisfies $\int_{-\infty}^{\infty}|s'(t)|^2 = 1$.*

*Proof.* The proof is given in Appendix A. □

Clearly, Theorem 1 is a sufficient condition for the $k$th stage being successfully detected. Thus, we have proved that the algorithm is able to provide a correct detection, as long as (18) is satisfied. It is straightforward to offer the power allocation based on (18). However, this may not be a good choice for three reasons. Firstly, the derivation for (18) involves the scaling of inequalities; thus the power allocation based on (18) is not the best. Secondly, the algorithm operates on a reduced ISI trellis, where the certain error patterns that have a larger metric may not be included during the detection. Thirdly, practically speaking, Theorem 1 is not the necessary condition for the successful detection. Since error-correcting codes are normally implemented in FTN-SCM systems which helps in the detection in a certain level, thus the power allocation should also take the influence of the corresponding codes into account.

All these three reasons suggest that the power allocation may not necessarily be derived in such a strict way. Hence, we slightly adjust the detection criterion by calculating

Table 1: Power allocation for the simulations in Figure 4.

| $K$ | $P_1/P$ | $P_2/P$ | $P_3/P$ | $P_4/P$ | $P_5/P$ | $P_6/P$ | $P_7/P$ |
|---|---|---|---|---|---|---|---|
| 2 | 0.6714 | 0.3286 | - | - | - | - | - |
| 3 | 0.5774 | 0.2837 | 0.1389 | - | - | - | - |
| 4 | 0.5403 | 0.2655 | 0.1304 | 0.0638 | - | - | - |
| 5 | 0.5237 | 0.2573 | 0.1264 | 0.0621 | 0.0304 | - | - |
| 6 | 0.5159 | 0.2535 | 0.1246 | 0.0612 | 0.0301 | 0.0147 | - |
| 7 | 0.5122 | 0.2517 | 0.1237 | 0.0608 | 0.0299 | 0.0147 | 0.0072 |

the probability of $P(x_n \mid y_1^N, b_1^N, S_{n-1})$ instead of $P(x_n \mid y_1^N, b_1^N)$, where the detection algorithm is assumed with $M = 1$ and $S_{n-1}$ is the correct state that is preserved at the $(n-1)$th section. Thus, at $n$th section, the log likelihood ratio (LLR) of the input $L(x_n)$ can be obtained by the following theorem.

**Theorem 2** (the correct tail path). *We claim that* $\mathbf{v}$ *is the correct tail path (CTP) if and only if the last L elements are correct, which is* $e_{n+1}^{n+L} = \mathbf{0}^{\mathrm{T}}$. *Then at nth section, for the CTPs* $\mathbf{v}$ *and* $\mathbf{v}'$ *of the states* $s_+$ *and* $s_-$ *that are induced by the correct state* $S_{n-1}$, *we have*

$$\begin{aligned} L(x_n) &\triangleq \ln \frac{P\left(x_n = +\sqrt{P_k \tau T} \mid y_1^N, b_1^N, S_{n-1}\right)}{P\left(x_n = -\sqrt{P_k \tau T} \mid y_1^N, b_1^N, S_{n-1}\right)} \\ &= \ln \frac{P\left(s_+ \mid y_1^N, b_1^N, S_{n-1}\right)}{P\left(s_- \mid y_1^N, b_1^N, S_{n-1}\right)} \propto J(\mathbf{v}) - J\left(\mathbf{v}'\right). \end{aligned} \tag{19}$$

*Proof.* The proof is given in Appendix B. □

With Theorem 2, it is possible to evaluate the error event rate (EER) of each layer. We assume the two CTPs are $v_1^{n+L} = x_1^{n+L}$ and $v'^{n+L}_1 = x_1^{n+L} + e_1^{n+L}$ with $e_1^{n+L} = [0, \ldots, 0, e_n, 0, \ldots, 0]^{\mathrm{T}}$, respectively. Thus, for the error event $\varepsilon \triangleq e_1^{n+L}$, we have

$$\begin{aligned} P(\varepsilon) &= P\left(J\left(v'^{n+L}_1\right) - J\left(v_1^{n+L}\right) > 0\right) \\ &= P\left(e_n \left(\sum_{l=L+1}^{L_1} g_{-l} x_{n+l}\right) - |e_n|^2 + e_n \eta_n \right. \\ &\quad \left. + e_n \left(\sum_{l=-L_1}^{L} g_{-l} b_{n+l}\right) > 0\right). \end{aligned} \tag{20}$$

By considering the Gaussian assumption, and the steep decrease of $Q$ function, (20) can be simplified as

$$P(\varepsilon) \simeq Q\left(\sqrt{\frac{P_k \tau T}{\sigma^2}}\right), \tag{21}$$

where

$$\sigma^2 = \frac{N_0}{2} + \sum_{l=L+1}^{L_1} |g_{-l}|^2 P_k \tau T + \sum_{l=-L_1}^{L} |g_{-l}|^2 \sigma_b^2[n]. \tag{22}$$

In the following, we offer a power allocation with respect to individual error-correcting codes. Without loss of generality, we assume that the code at the $k$th layer successfully recovers the information sequence at SNR $= \rho_k$ over the FTN channel. Meanwhile, according to (21), the signal-to-interference plus noise ratio (SINR) for the $k$th layer is defined as

$$\begin{aligned} \mathrm{SINR}_k &\triangleq \frac{E\left(x_k^2[n]\right)}{E\left(\sigma_b^2[n]\right) + N_0/2} \\ &= \frac{P_k \tau T}{N_0/2 + \sum_{l=L+1}^{L_1} |g_{-l}|^2 P_k \tau T + \sum_{l=-L_1}^{L} |g_{-l}|^2 \left(\sum_{i=k+1}^{K-1} P_i \tau T\right)}. \end{aligned} \tag{23}$$

Therefore, to successfully decode $u_k$, $\mathrm{SINR}_k \geq \rho_k$ must hold. Thus, we have

$$P_k \geq \frac{\rho_k \left(N_0/2 + \sum_{l=-L_1}^{L} |g_{-l}|^2 \left(\sum_{i=k+1}^{K-1} P_i \tau T\right)\right)}{\tau T \left(1 - \rho_k \sum_{l=L+1}^{L_1} |g_{-l}|^2\right)}. \tag{24}$$

Hence, by noticing the natural power assignment constraint that $P_0 + P_1 + \cdots + P_{K-1} = 1$, the required $P_k$ for all $k$ can be obtained recursively starting from $k = K - 1$. The numerical results based on the above power allocation are demonstrated in the next section.

## 5. Numerical Results

We choose the root raised cosine (RRC) (with roll-off factor $\beta = 0.3$ and a time-truncation to $\pm 15T$ around $t = 0$) (without loss of generality, we assume $T = 1$) as the shaping pulse $h(t)$. Meanwhile, the outer code is chosen as the code rate $R = 1/3$ asymmetric Turbo code in [18], with the generator polynomial $g_1(D) \triangleq [1 \quad (1 + D + D^2)/(1 + D^2)]$ and $g_2(D) \triangleq [1 \quad (1 + D + D^3)/(1 + D^2 + D^3)]$. As we transmit 20000 information bits per layer, we have $N = 60010$ as the codeword length, wherein 10 redundant bits are included to terminate the trellis. The two-dimensional normalized spectral efficiency is defined as $\eta = 2RK/\tau(1 + \beta)$.

The simulation results of FTN-SCM with $K = 2$ to 7 and $\tau = 2/3$ are plotted in Figure 4, wherein the power allocation is shown in Table 1. The parameters for the detection method are chosen as $M = 4$ and $L = 2$. There are 50 Turbo iterations between the FTN and Turbo code at each layer and 3 complete sweeps in total. Figure 5 shows the corresponding achieved data rate (for details on the data rate, please refer to [5]) at BER $< 10^{-5}$. As figures imply, the BER results of FTN-SCM

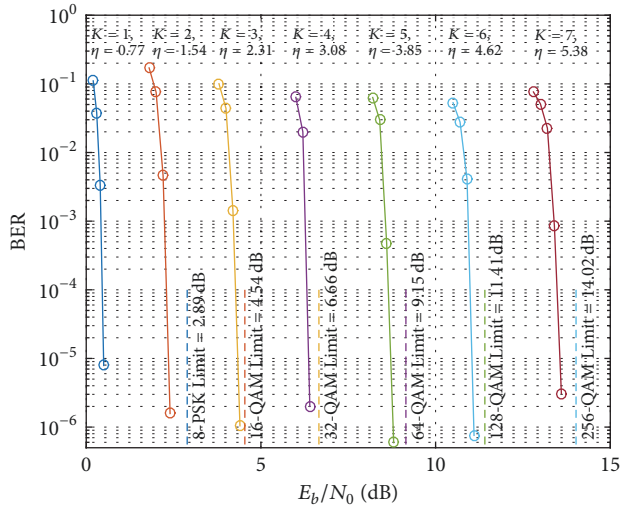


FIGURE 4: BER results of FTN-SCM with $K$ layers and $\tau = 2/3$ compared to constrained capacities of orthogonal signaling, where $\eta$ (bits/s/Hz) represents the two-dimensional normalized spectral efficiency.

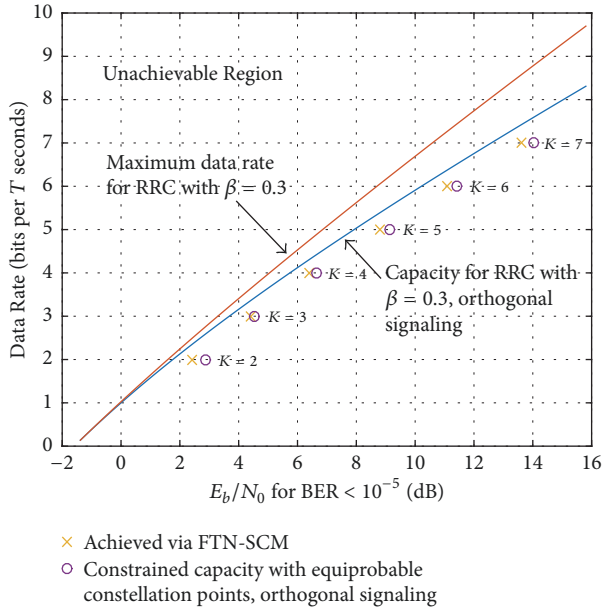


FIGURE 5: Achieved data rate (BER < $10^{-5}$) compared to both the constrained and unconstrained capacity of orthogonal signaling as well as the ultimate capacity of RRC with $\beta = 0.3$.

lie clearly outside the constrained capacity boundary with a larger alphabet of orthogonal signaling. We also observe that FTN-SCM is able to achieve a wide range of spectral efficiencies with a simple binary modulation format at each layer. In order to better illustrate the performance of FTN-SCM, we make a comparison between our method and the method in [1]. The BER results of two methods are given in Figure 6, where the simulation parameters are the same as the case of $K = 2$ in Figure 4, except the power allocation for method in [1] is given as $P_1/P = 0.6705$ and $P_2/P = 0.3295$. Note that, in [1], the authors utilize an *optimal* FTN equalizer in order to gain a very good performance. However, such equalizers are normally impractical, especially when there are a lot of layers; for example, $K = 7$. On the other hand, for BER < $10^{-5}$, our method only needs no more than 0.1 dB to achieve the same performance as the *optimal* result, but with much less complexity, which proves that our method exhibits a better trade-off between performance and complexity than the method in [1]. For a detailed complexity comparison, please refer to [18]. It should be mentioned that the BER results can be improved by choosing a better outer code or a better detection method, which is a future topic for us.

## 6. Conclusion

In this paper, we considered the FTN-SCM structure. Based on the transceiver structure, we derived a new observation model and further offered the power allocation with respect

to the detection method of each layer. Simulation results show that, in a wide range of spectral efficiencies, FTN-SCM requires lower SNR than that for orthogonal signaling with a larger alphabet. It should be noted that the proposed scheme is easy to be extended to nonbinary modulation cases and other types of channels.

## Appendix

## A. Proof of Theorem 1

According to [18], the probabilities of the correct state $S_n = s$ and the wrong state $S_n = s'$ with the error sequence $e_1^n$ satisfy the following equation:

$$\ln \frac{P\left(S_n = s', y_1^N, b_1^N\right)}{P\left(S_n = s, y_1^N, b_1^N\right)} \propto \sum_{i=1}^{2^L} J\left(\mathbf{v}'_i\right) - J\left(\mathbf{v}_i\right), \tag{A.1}$$

where $\mathbf{v}'_i$ and $\mathbf{v}_i$ represent the $i$th path of the probability calculation for $S_n = s'$ and $S_n = s$, respectively. Without loss of generality, we assume that $\mathbf{v}'_i$ and $\mathbf{v}_i$ have the same error pattern $e_{n+1}^{n+L}$. Thus, by considering (16), (A.1) can be further simplified as

$$\begin{aligned} \ln \frac{P\left(S_n = s', y_1^N, b_1^N\right)}{P\left(S_n = s, y_1^N, b_1^N\right)} &\propto 2^L \\ &\times \left\{ \mathrm{Re}\left\{ \left(e_1^n\right)^{\mathrm{H}} \mathbf{T}' x_{n+L+1}^{n+L_{\mathrm{I}}} - \frac{1}{2} d^2\left(e_1^n\right) \right\} \right. \\ &\left. + \mathrm{Re}\left\{ \left(e_1^n\right)^{\mathrm{H}} \mathbf{G}' b_1^{n+L} + \left(e_1^n\right)^{\mathrm{H}} \eta_1^n \right\} \right\}, \end{aligned} \tag{A.2}$$

where

$$\mathbf{T}' = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & & & \\ 0 & & & \vdots \\ g_{-L_{\mathrm{I}}} & 0 & & \\ g_{-(L_{\mathrm{I}}-1)} & g_{-L_{\mathrm{I}}} & \ddots & \\ \vdots & & \ddots & 0 \\ g_{-(L+1)} & g_{-(L+2)} & \cdots & g_{-L_{\mathrm{I}}} \end{pmatrix},$$

$$\mathbf{G}' = \begin{pmatrix} 1 & g_{-1} & \cdots & g_{-L} & g_{-L-1} & \cdots & g_{-L_{\mathrm{I}}} & 0 & \cdots & 0 \\ g_1 & 1 & g_{-1} & \cdots & g_{-L} & g_{-L-1} & \cdots & g_{-L_{\mathrm{I}}} & 0 & \cdots \\ g_2 & g_1 & 1 & g_{-1} & \cdots & g_{-L} & g_{-L-1} & \cdots & g_{-L_{\mathrm{I}}} & 0 \\ & & \ddots & \ddots & \ddots & & \ddots & \ddots & & \\ \ddots & & & \ddots & \ddots & \ddots & & \ddots & \ddots & \\ \cdots & 0 & g_{L_{\mathrm{I}}} & \cdots & g_1 & 1 & g_{-1} & \cdots & g_{-L} & g_{-L-1} \\ 0 & \cdots & 0 & g_{L_{\mathrm{I}}} & \cdots & g_1 & 1 & g_{-1} & \cdots & g_{-L} \end{pmatrix}, \tag{A.3}$$

and $d^2(e_1^n) = (e_1^n)^{\mathrm{H}} \mathbf{G}_{n\times n} e_1^n$, representing the squared Euclidean distance between the erroneous path and the correct path at the current stage. As $b[n]$ is assumed to be Gaussian with variance $\sigma_b^2[n]$, the right-hand side of (A.2) can be upper-bounded in the noiseless regime by

$$\begin{aligned} 2^L &\times \left\{ \mathrm{Re}\left\{ \left(e_1^n\right)^{\mathrm{H}} \mathbf{T}' x_{n+L+1}^{n+L_{\mathrm{I}}} - \frac{1}{2} d^2\left(e_1^n\right) \right\} \right. \\ &\left. + \mathrm{Re}\left\{ \left(e_1^n\right)^{\mathrm{H}} \mathbf{G}' b_1^{n+L} \right\} \right\} < 2^L \\ &\times \left\{ \mathrm{Re}\left\{ \max_{e_1^n, x_{n+L+1}^{n+L_{\mathrm{I}}}} \left\{ \left(e_1^n\right)^{\mathrm{H}} \mathbf{T}' x_{n+L+1}^{n+L_{\mathrm{I}}} \right\} \right\} \right. \\ &\quad - \frac{1}{2} \mathrm{Re}\left\{ \min_{e_1^n} \left\{ d^2\left(e_1^n\right) \right\} \right\} \\ &\left. \quad + \mathrm{Re}\left\{ \max_{e_1^n} \left\{ \left(e_1^n\right)^{\mathrm{H}} \mathbf{G}' b_1^{n+L} \right\} \right\} \right\}. \end{aligned} \tag{A.4}$$

By noticing the fact that $e_n \in \{-\sqrt{2P_k\tau T}, 0, 2\sqrt{P_k\tau T}\}$, the right-hand side of (A.4) can further be extended as

$$\begin{aligned} 2^L &\times \left\{ \mathrm{Re}\left\{ \max_{e_1^n, x_{n+L+1}^{n+L_{\mathrm{I}}}} \left\{ \left(e_1^n\right)^{\mathrm{H}} \mathbf{T}' x_{n+L+1}^{n+L_{\mathrm{I}}} \right\} \right\} \right. \\ &\quad \left. - \frac{1}{2} \mathrm{Re}\left\{ \min_{e_1^n} \left\{ d^2\left(e_1^n\right) \right\} \right\} \right. \end{aligned}$$

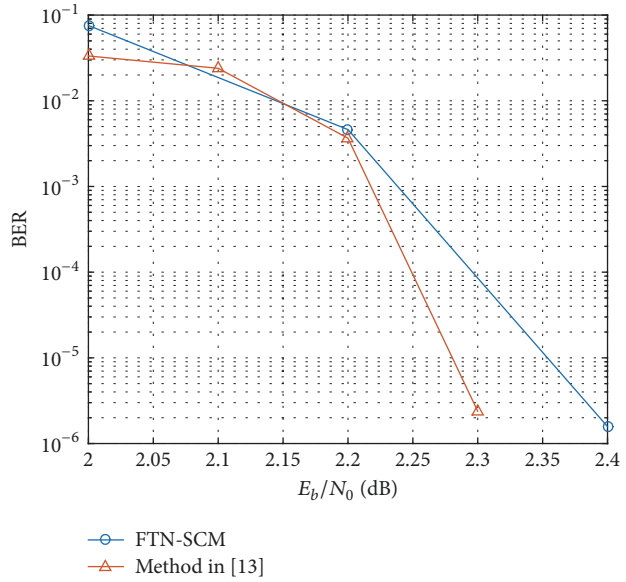


FIGURE 6: BER result of FTN-SCM with $K = 2$ compared to the result of the method in [1].

$$+ \operatorname{Re}\left\{\max_{e_1^n}\left\{\left(e_1^n\right)^{\mathrm{H}}\mathbf{G}'b_1^{n+L}\right\}\right\} < 2^L \times \left\{P_k\tau T\left(2\sum_{l=1}^{L_1-L} l\left|g_{-(L+l)}\right| - \frac{1}{2}d_{\min}^2\right) + \sqrt{2\sigma_{\max}^2 P_k\tau T}\sum_{l=-L_1}^{L_1}\left|g_l\right|\right\}. \tag{A.5}$$

This completes the proof of Theorem 1.

## B. Proof of Theorem 2

Clearly, since $S_{n-1}$ is the correct state at section $(n-1)$ and the states in the trellis are Markovian, the LLR of the input $x_n$ is determined by the probabilities of the states $s_+$ and $s_-$. According to the description in [18], in our case, the probability of $S_n = s$ follows

$$P\left(S_n = s \mid y_1^N, b_1^N, S_{n-1}\right) \propto \sum_{i=1}^{2^L}\exp\left[J\left(\mathbf{v}_i\right)\right], \tag{B.1}$$

where $\mathbf{v}_i = x_1^{n+L} + e_1^{n+L}$, with $e_1^{n+L} = [0,\ldots,0,e_n,e_{n+1},\ldots,e_{n+L}]^{\mathrm{T}}$, representing the $i$th possible path that is extended from $S_n = s$. The calculation implies a process of generating the marginal probability from all joint probabilities, as $2^L$ combinations are all taken into account.

For derivation brevity, we use $\mathbf{v}$ and $\mathbf{v}'$ representing the paths extended from $s_+$ and $s_-$, respectively. We further require that the same subscript $i$ represents the same error pattern. Thus, it is fair to assume that $\mathbf{v}_k$ and $\mathbf{v}'_k$ are the CTPs from states $s_+$ and $s_-$, respectively. Hence, we obtain

$$\frac{P\left(s_+ \mid y_1^N, b_1^N, S_{n-1}\right)}{P\left(s_- \mid y_1^N, b_1^N, S_{n-1}\right)} \propto \exp\left[J\left(\mathbf{v}_k\right) - J\left(\mathbf{v}'_k\right)\right] \times \frac{\exp\left[J\left(\mathbf{v}_1\right) - J\left(\mathbf{v}_k\right)\right] + \exp\left[J\left(\mathbf{v}_2\right) - J\left(\mathbf{v}_k\right)\right]\cdots}{\exp\left[J\left(\mathbf{v}'_1\right) - J\left(\mathbf{v}'_k\right)\right] + \exp\left[J\left(\mathbf{v}'_2\right) - J\left(\mathbf{v}'_k\right)\right]\cdots}. \tag{B.2}$$

Without loss of generality, we consider $J(\mathbf{v}_i) - J(\mathbf{v}_k)$, where $\mathbf{v}_i$ is not the CTP. Recall (16); we have

$$J\left(\mathbf{v}_i\right) - J\left(\mathbf{v}_k\right) = \operatorname{Re}\left\{\left(m_{n+1}^{n+L}\right)^{\mathrm{H}}\mathbf{T}''x_{n+L+1}^{n+L+L_1} - \left(m_{n+1}^{n+L}\right)^{\mathrm{H}}\mathbf{G}_{\mathrm{B}}\left(\frac{1}{2}m_{n+1}^{n+L}\right)\right\} + \operatorname{Re}\left\{\left(m_{n+1}^{n+L}\right)^{\mathrm{H}}\mathbf{G}_{L\times L}b_{n+1}^{n+L} + \left(m_{n+1}^{n+L}\right)^{\mathrm{H}}\eta_{n+1}^{n+L}\right\}, \tag{B.3}$$

where

$$\mathbf{T}'' = \begin{pmatrix} g_{-L} & g_{-(L+1)} & \cdots & g_{-L_1} & 0 & \cdots & 0 \\ \vdots & \ddots & & & \ddots & & \vdots \\ g_{-2} & \cdots & g_{-L} & g_{-(L+1)} & \cdots & g_{-L_1} & 0 \\ g_{-1} & g_{-2} & \cdots & g_{-L} & g_{-(L+1)} & \cdots & g_{-L_1} \end{pmatrix}. \tag{B.4}$$

Since $e_1^{n+L}$ is no longer part of the equation, the only variable in the equation is $m_{n+1}^{n+L}$. Thus, as all the combinations of $m_{n+1}^{n+L}$ are included in (B.2), we can safely draw the conclusion that

$$\frac{\exp\left[J\left(\mathbf{v}_1\right) - J\left(\mathbf{v}_k\right)\right] + \exp\left[J\left(\mathbf{v}_2\right) - J\left(\mathbf{v}_k\right)\right]\cdots}{\exp\left[J\left(\mathbf{v}'_1\right) - J\left(\mathbf{v}'_k\right)\right] + \exp\left[J\left(\mathbf{v}'_2\right) - J\left(\mathbf{v}'_k\right)\right]\cdots} = 1. \tag{B.5}$$

This completes the proof of Theorem 2.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grants 61771364, 91438101, and 61601346, by the Science and Technology on Communication Networks Laboratory under Grant 614210403070717, by the Fundamental Research Funds for the Central Universities under Grant WK2100060020, by the Key Research Program of Frontier Sciences of CAS under Grant QYZDY-SSW-JSC003, by the China Postdoctoral Science Foundation under Grant 2015M580819, and by the Shaanxi Province Postdoctoral Science Foundation.

## References

[1] Y. J. D. Kim, J. Bajcsy, and D. Vargas, "Faster-Than-Nyquist Broadcasting in Gaussian Channels: Achievable Rate Regions and Coding," *IEEE Transactions on Communications*, vol. 64, no. 3, pp. 1016–1030, 2016.

[2] CISCO, Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015-2020, http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.pdf, 2016.

[3] J. B. Anderson, "Faster-than-Nyquist signaling for 5G communication," in *Signal Processing for 5G: Algorithms and Implementations*, pp. 24–46, Wiley-IEEE Press, 2016.

[4] J. E. Mazo, "Faster-than-Nyquist signaling," *Bell Labs Technical Journal*, vol. 54, no. 8, pp. 1451–1462, 1975.

[5] F. Rusek and J. B. Anderson, "Constrained capacities for faster-than-Nyquist signaling," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 764–775, 2009.

[6] J. B. Anderson, F. Rusek, and V. Öwall, "Faster-than-Nyquist signaling," *Proceedings of the IEEE*, vol. 101, no. 8, pp. 1817–1830, 2013.

[7] F. Rusek and J. B. Anderson, "Serial and parallel concatenations based on faster than nyquist signaling," in *Proceedings of the IEEE International Symposium on Information Theory*, pp. 1993–1997, Seattle, WA, USA, 2006.

[8] T. M. Cover, "Broadcast Channels," *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 2–14, 1972.

[9] X. Yuan, Q. Guo, and L. Ping, "Low-complexity iterative detection in multi-user MIMO ISI channels," *IEEE Signal Processing Letters*, vol. 15, pp. 25–28, 2008.

[10] L. Ping, J. Tong, X. Yuan, and Q. Guo, "Superposition coded modulation and iterative linear MMSE detection," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 6, pp. 995–1004, 2009.

[11] K. Wu and L. Ping, "A quasi-random approach to space-time codes," *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 1073–1085, 2008.

[12] X. Chen, X. Wang, and X. Ma, "Superposition convolutional coded modulation for the Rayleigh fading channel," in *Proceedings of the International Conference on Communications, Circuits and Systems, ICCCAS 2007*, pp. 37–41, IEEE, Kokura, Japan, 2007.

[13] X. Ma and L. Ping, "Coded modulation using superimposed binary codes," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3331–3343, 2004.

[14] Y. J. Kim and J. Bajcsy, "Faster than Nyquist broadcast signaling," in *Proceedings of the 2012 26th Biennial Symposium on Communications (QBSC)*, pp. 186–189, Kingston, Canada, May 2012.

[15] G. Colavolpe, G. Ferrari, and R. Raheli, "Reduced-state BCJR-type algorithms," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 5, pp. 848–859, 2001.

[16] G. D. Forney, "The Viterbi Algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, 1973.

[17] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, 1974.

[18] 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 Transactions on Communications*, vol. 66, no. 3, pp. 1190–1204, 2018.

[19] G. Ungerboeck, "Adaptive Maximum-Likelihood Receiver for Carrier-Modulated Data-Transmission Systems," *IEEE Transactions on Communications*, vol. 22, no. 5, pp. 624–636, 1974.