

Received March 15, 2020, accepted April 4, 2020, date of publication April 8, 2020, date of current version April 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2986679

Receiver Design for Faster-Than-Nyquist Signaling: Deep-Learning-Based Architectures

PEIYANG SONG¹, (Student Member, IEEE), FENGKUI GONG¹, (Member, IEEE),
QIANG LI¹, (Student Member, IEEE), GUO LI¹, (Student Member, IEEE),
AND HAIYANG DING², (Member, IEEE)

¹State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710000, China

²School of Information and Communications, National University of Defense Technology, Xi'an 710000, China

Corresponding author: Fengkui Gong (fkgong@xidian.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFE0100500, in part by the National Natural Science Foundation of Shaanxi Province under Grant 2019CGXNG-010, in part by the Youth Program of National Natural Science Foundation of China under Grant 61901325, in part by the China's Post-Doctoral under Grant 2018M640958, in part by the Fundamental Research Funds for the Central Universities under Grant 8002/20101196698, in part by the National Natural Science Foundation of Shaanxi Province under Grant 2019JQ-658, and in part by the Innovation Fund of Xidian University 5001-20109195456.

ABSTRACT Faster-than-Nyquist (FTN) is a promising paradigm to improve bandwidth utilization at the expense of additional intersymbol interference (ISI). In this paper, we apply state-of-the-art deep learning (DL) technology into receiver design for FTN signaling and propose two DL-based new architectures. Firstly, we propose an FTN signal detection based on DL and connect it with the successive interference cancellation (SIC) to replace traditional detection algorithms. Simulation results show that this architecture can achieve near-optimal performance in both uncoded and coded scenarios. Additionally, we propose a DL-based joint signal detection and decoding for FTN signaling to replace the complete baseband part in traditional FTN receivers. The performance of this new architecture has also been illustrated by simulation results. Finally, both the proposed DL-based receiver architecture has the robustness to signal to noise ratio (SNR). In a nutshell, DL has been proved to be a powerful tool for the FTN receiver design.

INDEX TERMS Faster-than-Nyquist, receiver design, signal detection, deep learning, intersymbol interference, channel coding.

I. INTRODUCTION

The last couple of decades have seen the exponential growth of wireless devices and data traffic. Nowadays, spectral efficiency has become extremely valuable. With increasingly demanding requirements for spectral resources, a promising technology named FTN is rediscovered and has attracted a lot of attention in both industrial and academic communities [1]–[15].

As known, in conventional Nyquist-criterion transmission, when available bandwidth is W Hz, the symbol interval T is always set as $T \geq T_N = 1/(2W)$. The strict orthogonality between transmitted symbols guarantees the signal recovery in the receiver. In contrast, the symbol interval reduces to $T < T_N$ in FTN signaling to achieve a higher transmission rate, which, at the same time, destroys the orthogonality and introduces unavoidable ISI. Although the additional

interference increases the complexity to recovery original signals in the receiver, the Mazo limit [1] proves that without the expansion of bandwidth and loss of BER performance, the FTN signaling can achieve an up to 25% higher transmission rate than conventional Nyquist-criterion design in the additive white Gaussian noise (AWGN) channel.

Traditional receiver design focuses on the detection algorithms to eliminate the ISI caused by the smaller symbol interval. Among time-domain equalizations, [2] and its simplified version [3] formulate the FTN signal as convolutionally encoded symbols and applies the Viterbi algorithm for detection. Reference [4] employs a symbol-by-symbol signal detection, which achieves the near-optimal performance with very low complexity. Channel shortening is employed for maximum-likelihood sequence estimation (MLSE) in [5]. Meanwhile, a few effective frequency-domain detection algorithms for FTN can also be found in the literature. Reference [6] designs a minimum mean square error (MMSE) frequency-domain equalization (FDE) for FTN detection.

The associate editor coordinating the review of this manuscript and approving it for publication was Zhanyu Ma¹.

An H-ARQ technique for FTN signaling is proposed in [7]. And [8] focuses on iterative FDE architectures to eliminate the introduced ISI. Recently, an interesting work that focuses on the blind estimation for the packing ratio in FTN transmission has been addressed [9] and is promising to make the receiver design more adaptive.

In recent years, a new trend has appeared to merge the two technologies of communications and DL [16]–[18]. Nowadays, DL has been widely employed in conventional communication scenarios, such as orthogonal frequency-division multiplexing (OFDM) systems [19]–[21], multi-antenna systems [22], channel estimation and prediction [23]–[25], channel coding [26]–[28], modulation classification [29], etc. However, DL-based receiver design for FTN signaling, as far as we know, has not been studied yet in the literature.

In fact, signal detection for FTN based on sequence estimation (as well as the channel decoding) can be regarded as a classification problem which aims to divide a multiple dimension space into several parts. For example, when we try to recover M transmitted symbols from M received symbols in BPSK modulation, we practically divide an M -dimension space into 2^M parts. The power of DL in solving such classification problems has been proved by its successful application in image and voice recognition. Also, with the development of artificial intelligence (AI) chips [30], the DL-based algorithms may show their advance in future communication systems. These facts have inspired us to employ DL into FTN receiver designs.

The contribution of this paper can be summarized as follows.

- We propose joint DL-based detection and SIC to replace traditional FTN detection algorithms, where SIC is used to eliminate the interference and obtain more accurate log-likelihood ratio (LLR) values.
- We develop DL-based joint detection and decoding for FTN signaling, which can replace the whole baseband part of conventional FTN receivers with the DL-based architecture.
- We investigate the robustness of both the proposed receiver designs to SNR values. And results show that after training by the FTN data set under a specific SNR value, the proposed designs can fit the scenarios with different SNRs and achieve near-optimal performance in the offline recovery.
- We have carried out comprehensive evaluations to verify and analyze the proposed DL-based FTN receiver architectures.

Herein, we give the definition of notations which we will encounter throughout the rest of the paper. Bold-face lower case letters (e.g. \mathbf{x}) are applied to denote column vectors. Light-face italic letters (e.g. x) denote scalars. x_i is the i th element of vector \mathbf{x} . $x(t) * y(t)$ denotes the convolution operation between $x(t)$ and $y(t)$. $\lfloor x \rfloor$ is the maximum integer less than or equal to x . And $P(\cdot)$ means the probability.

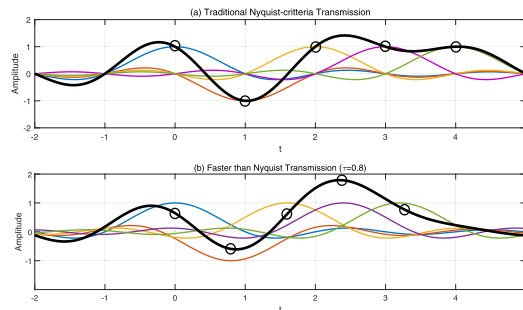


FIGURE 1. Waveform and the sampled symbols in Nyquist and FTN transmission.

II. SYSTEM MODEL

We consider the communication system with the complex-valued quadrature amplitude modulation (QAM) scheme and AWGN channel. As known, after constellation mapping, the baseband signal should pass through a shaping filter $h(t)$. Hence, the transmitted signal $s(t)$ can be written as

$$s(t) = \sqrt{E_s} \sum_{k=-\infty}^{+\infty} x_k h(t - k\tau T_N), \quad (1)$$

where E_s is the average energy of constellation symbols, x_k ($k = 0, \pm 1, \pm 2, \dots$) is the k th transmitted symbol and τ is the time acceleration factor which satisfies $0 < \tau \leq 1$. Due to the existence of τ , practical symbol interval T is smaller than Nyquist limit T_N , which helps the system achieve a higher transmission rate.

Actually, when $\tau = 1$, benefiting from the orthogonality between any two symbols, each sample of the symbols will not be influenced by the others. However, when $0 < \tau < 1$, each sample becomes a weighted sum of different symbols, which will make it difficult to recover the transmitted signals. The intersymbol effect of FTN has been illustrated in Fig. 1.

Fig. 2 shows the block diagram of the traditional communication system with channel coded FTN signaling, where τT_N is not only the practical symbol interval of the symbols passing through the shaping filter but also the sampling interval for the signal which has passed through the matched filter. The LLR values can be calculated by PSK/QAM demapping as

$$LLR_i = \ln \frac{P(x_i = 0 | \hat{y}_j)}{P(x_i = 1 | \hat{y}_j)}, \quad (2)$$

where $i = \{1, 2, 3, \dots\}$, $j = \lfloor i/r \rfloor + 1$ and r is the modulation order. x and \hat{y} represent the transmitted bits and detected symbols respectively.

The received signal after passing through the matched filter can be written as

$$\begin{aligned} y(t) &= (s(t) + n(t)) * h(t) \\ &= \sqrt{E_s} \sum_{k=-\infty}^{+\infty} x_k g(t - k\tau T_N) + \tilde{n}(t), \end{aligned} \quad (3)$$

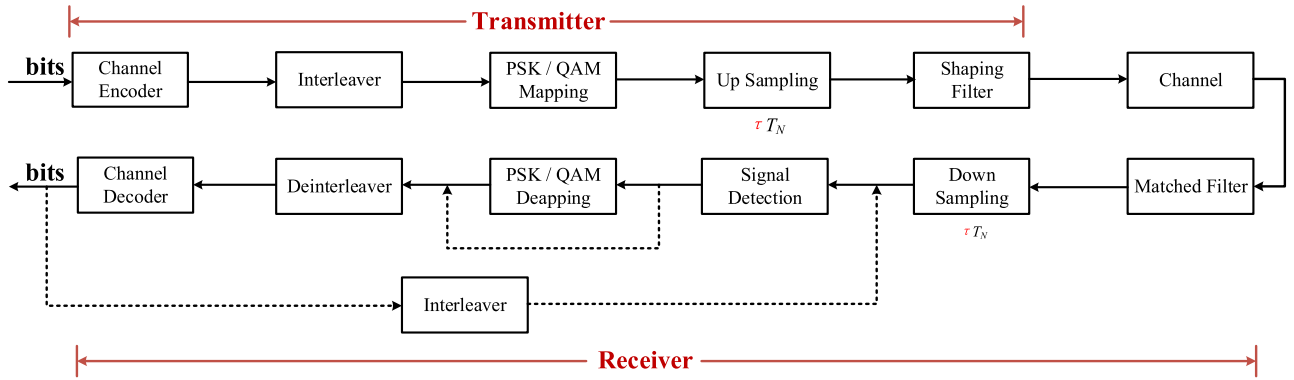


FIGURE 2. Block diagram of the traditional communication system with channel coded FTN signaling.

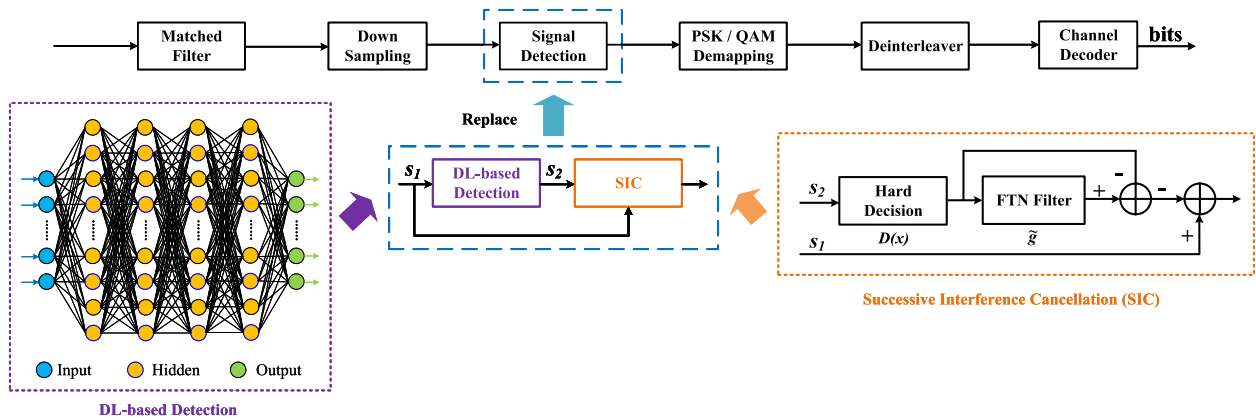


FIGURE 3. The architecture of the proposed FTN receiver with joint DL-based detection and SIC.

where $g(t) = \int h(x)h(t-x)dx$, $\tilde{n}(t) = \int n(x)h(t-x)dx$, and $n(t)$ is a zero mean complex-valued Gaussian random process with variance σ^2 . Throughout this letter, time synchronization error is not taken into consideration. Hence, the n th sample of received signal $y(t)$ can be obtained as (4).

As seen, each sample of the received waveform contains not only the expected symbol but also the weighted sum of both its previous and upcoming symbols. A key problem for receivers is eliminating the ISI from both directions and recover the transmitted sequence \mathbf{x} from the received symbols \mathbf{y} . It will certainly increase the complexity of signal detection, which can be regarded as the price of the higher transmission rate.

$$\begin{aligned}
 y_n &= \sqrt{E_s} \sum_{k=-\infty}^{+\infty} x_k g(n\alpha T_N - k\alpha T_N) + \tilde{n}(n\alpha T_N) \\
 &= \underbrace{\sqrt{E_s} \sum_{k=-\infty}^{n-1} x_k g((n-k)\alpha T_N)}_{\text{ISI from previous } L-1 \text{ symbols}} + \underbrace{\sqrt{E_s} x_n g(0)}_{\text{desired signal}} \\
 &\quad + \underbrace{\sqrt{E_s} \sum_{k=n+1}^{+\infty} x_k g((n-k)\alpha T_N) + \tilde{n}(n\alpha T_N)}_{\text{ISI from upcoming } L-1 \text{ symbols}}. \quad (4)
 \end{aligned}$$

III. THE PROPOSED FTN RECEIVER DESIGN WITH DL-BASED DETECTION AND SIC

The proposed FTN receiver architecture with joint DL-based detection and SIC has been illustrated in Fig. 3. The conventional signal detection is replaced by the proposed DL-based detection and SIC whose detailed structures have been shown in the part surrounded by the dotted line. In this section, we provide a detailed explanation of the proposed new architecture.

A. DL-BASED DETECTION

The proposed DL-based detection is essentially a deep neural network (DNN) which includes six layers (an input layer, an output layer and four hidden layers). As shown in Fig. 3, each hidden layer is composed of a fully connected (FC) sublayer and a rectifier linear unit (ReLU) function while the output layer is simply an FC sublayer. The numbers of neurons in each layer are L , 320, 160, 80, 40 and m , respectively. L and m are sliding window parameters which will be further introduced in the next subsection. It should be noted that these numbers of layers and neurons are employed for all the scenarios in the following part of this paper and are applicable for other general FTN signalings. The performance gain by simply enlarging the size of the network is disappointing.

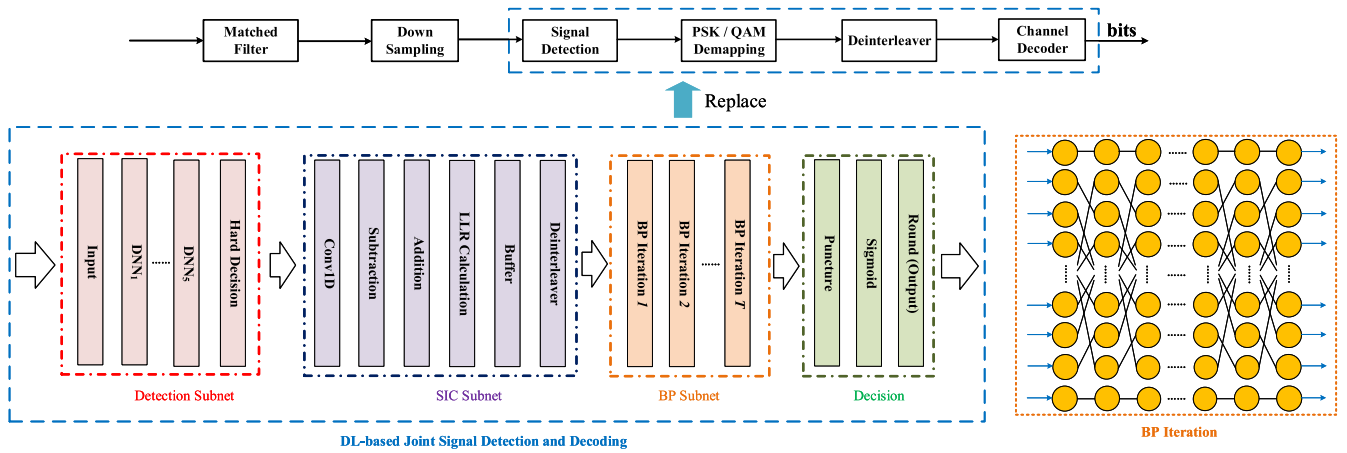


FIGURE 4. The architecture of the proposed FTN receiver with DL-based joint signal detection and decoding.

The input and output of the DL-based detection are real numbers. When BPSK is taken into consideration, one DNN should be employed on the real part of the received symbols. While in QPSK or higher QAM modulations, two identical structures (or multiplexing of one DNN) should be combined to detect both the real and imaginary parts of the received symbols.

B. DATASET

The training and testing data sets share the same structure. The input data set is composed of the real and imaginary parts of received symbols which have been downsampled and sliced by block length L and step size m . And the label set contains the real and imaginary parts of corresponding transmitted symbols sliced by m . All these symbols are generated by the software simulation. Also, the structure of the data set has been shown visually in Fig. 5.

C. WORKFLOW

A notable feature of our proposed DL-based detection is the *sliding window*, as shown in Fig. 5. The input in each recovery is a sliding window with length L on the real or imaginary part of the received symbols y . During each recovery, the DL-based detection tries to recover the middle m symbols of the input window. Then, the input window slides forward over m symbols to start the next recovery. This special architecture results from the uncertainty of edge symbols of the input window. They always suffer from severe ISI from either the previous or subsequent symbols while these symbols are unknown since they are not contained in the input.

Similar to most DL methods, our proposed detection includes two stages named *offline training* and *online recovery*.

- *Offline training*: During this stage, the model is trained by known transmitted symbols and the corresponding received symbols which are generated with the given β and τ . The process called back-propagation is required

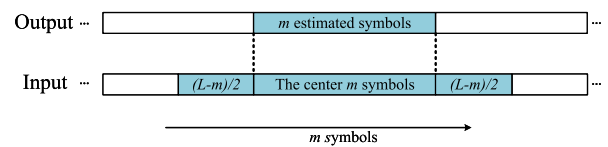


FIGURE 5. Workflow of the proposed DL-based signal detection for FTN signaling.

to calculate the loss function and its derivative to each neuron to update the network.

- *Online recovery*: During this stage, the DL-based detection can directly produce the estimated symbols by the received signals with the previously determined neural network. The back-propagation is not required, which can greatly reduce the delay and computational complexity of the detection.

It is worth noting that the two stages do not need to work iteratively. Once the offline training is completed, the neural network will be determined and does not need to be trained any more in the practical FTN detection.

D. SIGNAL RECONSTRUCTION BY SIC

Benefiting from the correcting and anti-interference ability, channel coding, nowadays, has become a necessary part of wireless communications. To gain the benefit of large rate enhancement from the FTN signaling with start-of-the-art channel coding technologies, we introduce SIC to obtain more accurate LLR values from the proposed DL-based detection. The main idea of the reconstruction is to calculate the interference and subtract it from the received symbols. The reconstructed symbols can be written as

$$\tilde{y} = s_1 - (D(s_2) * \tilde{g} - D(s_2)), \tag{5}$$

where s_1 and s_2 represent the received symbols and detected symbols respectively, as illustrated in Fig. 3. $D(\cdot)$ means the hard decision of a certain symbol sequence. The convolution operation is completed by the FTN filter, where the \tilde{g} is the ISI vector and $\tilde{g}_k = g((k - K)\tau T_N)$, $k = \{0, 1 \dots 2K\}$.

IV. THE PROPOSED FTN RECEIVER DESIGN WITH DL-BASED JOINT SIGNAL DETECTION AND DECODING

We propose a hybrid DL-based architecture, which has been illustrated in Fig. 4, to replace both the signal detection and channel decoding in conventional FTN receivers. In this section, the proposed new architecture is detailed presented.

A. DL-BASED JOINT SIGNAL DETECTION AND DECODING

In this section, we introduce the low-latency belief-propagation (BP) polar code decoder [31] into the FTN receiver design and proposed a DL-based complete baseband part in FTN receivers. The proposed DL-based joint signal detection and decoding is a cascade of a detection subnet, an SIC subnet, a BP subnet and finally a decision component. The detection subnet shares the same structure of the proposed DL-based detection in Section III. The SIC subnet eliminates the interference among different symbols by (5) with a convolution layer, a subtraction layer and an addition layer. Besides, a buffer is employed to output a complete polar code block so that the BP subnet can work normally.

The BP subnet, which can be considered as the unfolding of the conventional BP decoding structure, contains several left-to-right (LTO) and right-to-left (RTL) propagations. The value of each layer can be obtained by

$$\begin{cases} L_{i,j}^t = g\left(L_{i+1,2j-1}^t, L_{i+1,2j}^t + R_{i,j+N/2}^t\right) \\ L_{i,j+N/2}^t = g\left(R_{i,j}^t, L_{i+1,2j-1}^t\right) + L_{i+1,2j}^t \\ R_{i+1,2j-1}^t = g\left(R_{i,j}^t, L_{i+1,2j}^{t-1} + R_{i,j+N/2}^t\right) \\ R_{i+1,2j}^t = g\left(R_{i,j}^t, L_{i+1,2j-1}^{t-1}\right) + R_{i,j+N/2}^t, \end{cases} \quad (6)$$

where $L_{i,j}^t$ and $R_{i,j}^t$ represent the values of j -th node in i -th layer of RTL and LTR during the t -th iteration. $g(x, y) = 0.9375 * \text{sign}(x)\text{sign}(y)\min(|x|, |y|)$. Before running the iterations, the decoder should be initialized by

$$R_{1,j}^1 = \begin{cases} 0, & \text{if } j \in \mathcal{A}, \\ +\infty, & \text{if } j \in \mathcal{A}^c \end{cases} \quad (7)$$

and

$$L_{n+1,j}^1 = \frac{P(y_j|x_j = 1)}{P(y_j|x_j = 0)}, \quad (8)$$

where \mathcal{A} and \mathcal{A}^c are the information bits set and frozen bits set respectively.

Finally, a decision component is employed for the last layer to realize

$$f(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0, \end{cases} \quad (9)$$

where the sigmoid function, which is used to limit the output to $[0, 1]$, can be written as

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (10)$$

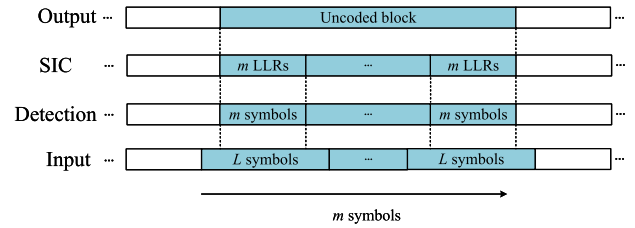


FIGURE 6. Workflow of the proposed FTN receiver design with the DL-based joint detection and decoding.

TABLE 1. Training and testing parameters of the proposed DL-based FTN detection.

Parameter	Value
Training Data Size	1.6×10^9 symbols
Training E_b/N_0	$E_b/N_0 @ \{\text{BER}=2 \times 10^{-4}\}$
Learning rate	0.001, 0.0002, 0.00004
Loss function	Mean square error (MSE)
Optimizer	Adam
Testing Data Size	3.2×10^7 symbols for each E_b/N_0

B. DATASET

As can be seen from the previous analysis, only the detection subnet contains trainable parameters. Hence, it is not the whole model but simply the detection subnet that needs to be trained. And the dataset of training and testing has been provided in Section III.

C. WORKFLOW

Fig. 6 illustrates the workflow of the proposed DL-based joint signal detection and decoding. Similar to Section III, the input in each recovery is a sliding window with length L on the received symbols while the output is the corresponding m estimated symbols. Aided by the buffer in the SIC subnet, every $L_b/(R_b r)$ LLRs are sent to the BP subnet, where L_b is the code length and R_b is the code rate. Finally, a complete uncoded bit block is obtained after BP and decision.

V. SIMULATION RESULTS

In this section, we assess the performance and robustness of our proposed two DL-based FTN detection and decoding architectures. All the simulation results are obtained on the test data set. The square root raised cosine (SRRC) filters with different roll-off factors are taken into consideration. A more detailed list of the parameters is provided in Table 1.

A. ROBUSTNESS TO SNR MISMATCHING

It is very important for the proposed DL-based detection to be robust to the SNR values, without which the proposed DL-based detection will be trained and employed for different SNR values independently and suffer from high complexity

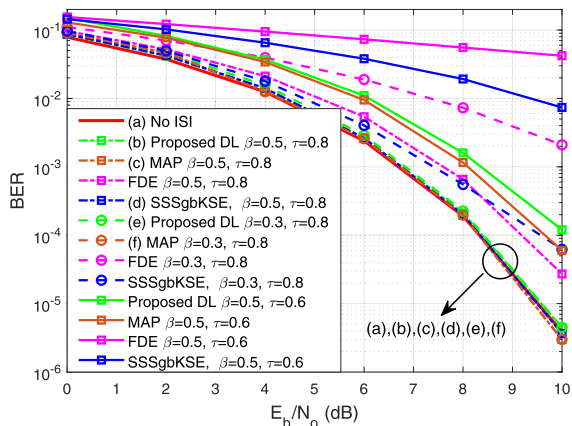


FIGURE 7. Performance of the proposed DL-based signal detection versus some baselines for uncoded FTN signaling.

resulting from SNR estimation and the store of massive DL network parameters corresponding to different SNR values.

Throughout a lot of simulations, fortunately, we find an interesting point that the DL network trained with $E_b/N_0@ \{BER = 2 \times 10^{-4}\}$ (e.g. 7.9dB in QPSK) always shows the near-optimal performance in the offline prediction stage. $E_b/N_0@ \{BER = x\}$ here means the E_b/N_0 value under which the ideal QAM modulation with Nyquist-criterion can achieve a BER performance of x in the AWGN channel.

B. PERFORMANCE OF THE PROPOSED DL-BASED DETECTION IN UNCODED SCENARIOS

To better evaluate the performance of our proposed DL-based detection, we choose three traditional algorithms maximum a posteriori (MAP) [15], MMSE FDE [6] and SSSgbKSE [4] as the baselines. And MAP algorithm is generally considered as the optimal detection scheme.

Fig. 7 compares the BER performance of our proposed DL-based detection with the baselines. As shown, all the detection schemes show similar performances when $\alpha = 0.5$ and $\tau = 0.8$. The SNR gain of our proposed DL algorithm over FDE is about 0.9dB, considering $BER = 10^{-4}$. And as ISI becomes more severe, the SNR gains of the proposed DL algorithm over SSSgbKSE and FDE get larger. When $\alpha = 0.5$ and $\tau = 0.6$, the gains increase to more than 10dB, while the performance of the proposed DL algorithm and MAP just appears to differ and the distance of their SNR at $BER = 10^{-4}$ is only 0.28dB. The simulation result confirms that the DL-based detection is applicable to FTN signal detection problems.

C. PERFORMANCE OF THE PROPOSED DL-BASED DETECTION IN HIGH ORDER MODULATIONS

Actually, signal detection for FTN will be difficult in high order modulations since the SNR gains are negligible. However, when parameters with relatively slight interference are employed, the FTN can still work well and contributes a rate enhancement. Fig. 8 illustrates the BER performance of the proposed DL-based detection in high order modulations with

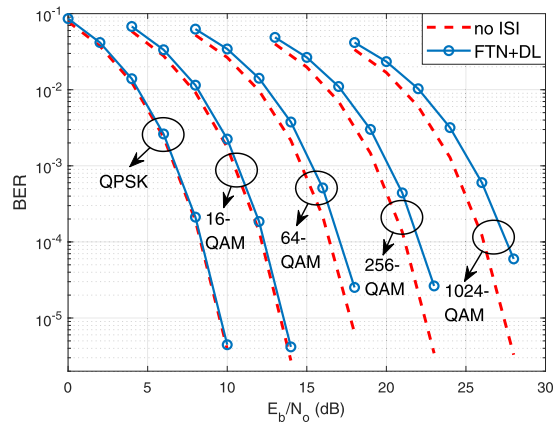


FIGURE 8. Performance of the proposed DL-based signal detection in high order modulations with $\beta = 0.5$ and $\tau = 0.8$.

$\beta = 0.5$ and $\tau = 0.8$, where the dashed lines represent the performance in conventional Nyquist-criterion systems. The result reveals the potential of DL-based detection for FTN signaling with high order modulations.

D. PERFORMANCE OF THE PROPOSED RECEIVER WITH JOINT DL-BASED DETECTION AND SIC

Fig. 9 illustrates the performance of our proposed DL-based receiver design with detection and SIC versus other baselines in FTN signaling with (64800, 32400) LDPC code [32]. The dashed lines show the performance in conventional coded and uncoded Nyquist-criterion systems. Since the SSSgbKSE method can not produce soft information independently (as stated in [4]), it is not taken into consideration here.

As shown, the proposed receiver design can achieve great performance gains with the help of the start-of-the-art channel coding scheme. Compared to other baselines, the proposed design achieves an SNR gain over FDE by 0.9dB and is only 0.1 dB worse than MAP detection, considering $\beta = 0.5$, $\tau = 0.7$ and $BER = 10^{-4}$. Moreover, the proposed DL-based detection and SIC can be directly cascaded with the existing channel decoders without changing their original structures or getting involved in their iterative decoding process, which is convenient for the practical implementation.

E. PERFORMANCE OF THE PROPOSED RECEIVER WITH DL-BASED JOINT DETECTION AND DECODING

Fig. 10 illustrates the performance of our proposed DL-based joint detection and decoding with (1024, 512) polar code [33]. The BP subnet with 50 iterations is taken into consideration. And similar to Fig. 9, the dashed lines are the performance of conventional coded and uncoded Nyquist-criterion systems.

As shown, compared to other baselines, the proposed DL-based joint detection and decoding can achieve an SNR gain over FDE by 1.16dB and is only 0.3dB worse than MAP detection, considering $\beta = 0.5$, $\tau = 0.7$ and $BER = 10^{-4}$. Additionally, the complete DL-based receiver design is applicable to the implementation by AI chips in the future.

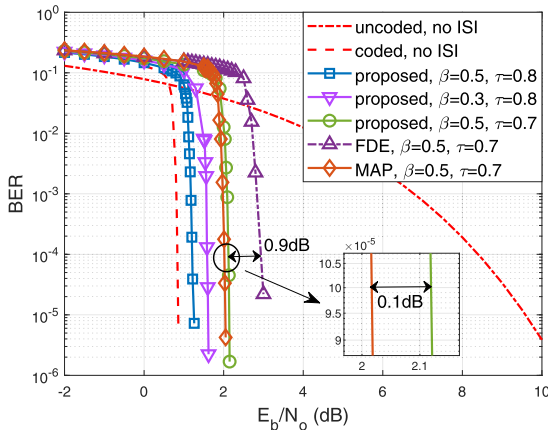


FIGURE 9. Performance of the proposed joint DL-based detection and SIC versus some baselines for (64800, 32400) LDPC-coded FTN signaling.

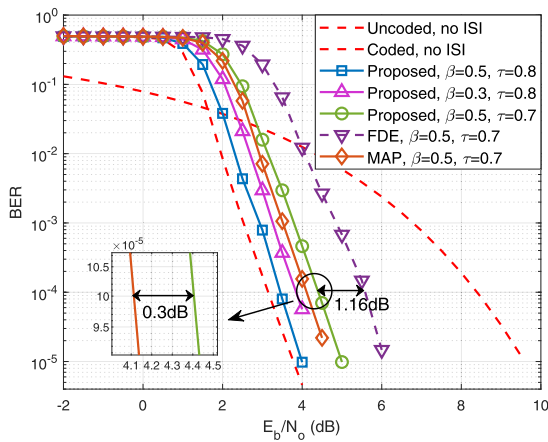


FIGURE 10. Performance of the proposed DL-based joint signal detection and decoding versus some baselines for (1024, 512) polar-coded FTN signaling.

TABLE 2. Computational complexity of different FTN detection schemes.

Scheme	Additions	Multiplications	Parallel
FDE [6]	8194	8196	Support
Proposed DL	9720	9720	Support
MAP [15]	12770	3980	No

F. THE COMPLEXITY ANALYSIS OF DIFFERENT FTN DETECTIONS

To compare the computational complexity of different FTN detections, for each scheme, we count the numbers of addition and multiplication operations consumed for every single estimated symbol and list them in Table 2. Also, the parallelizability of each scheme is taken into consideration to show whether the parallel implementation can be carried out to convert the computational complexity into space occupation to reduce the detection delay. In view of the fact that the offline training does not involve in the FTN signal detection in practical communication systems, as described

in Section III, we do not include the training process into the evaluation about computational complexity of the proposed DL algorithm.

As shown, among the schemes, the complexity of the proposed DL algorithms is very close to that of FDE detection. And both these two schemes can be implemented in parallel to reduce the detection delay. Although MAP algorithm requires fewer multiplications than FDE and the proposed DL algorithm, the lack of support for parallel implementation will lead to severe detection delay, which has been proved by its application in turbo decoders.

VI. CONCLUSION

FTN is a promising technology to improve spectrum efficiency. This paper, as far as we know, is the first attempt to apply DL into channel coded FTN detection and decoding. In this paper, we propose two DL-based FTN receiver designs, which have shown a near-optimal performance over some traditional algorithms. In particular, the proposed DL-based joint detection and decoding architecture can replace the whole baseband part of the FTN receiver, which can effectively improve the integration of the FTN receiver design. Moreover, with the rapid development of AI chips, the proposed DL-based FTN receiver designs are promised to show their advantages in the future implementation for AI-aided communications. Finally, in our future work, we will conduct more researches to extend the DL-based FTN detection to other practical channel environments (e.g. the multi-path fading channel).

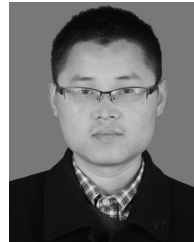
REFERENCES

- [1] J. E. Mazo, "Faster-Than-Nyquist signaling," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1451–1462, Oct. 1975.
- [2] A. D. Liveris and C. N. Georghiadis, "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. Yanikomeroğlu, "A very low complexity successive Symbol-by-Symbol sequence estimator for Faster-Than-Nyquist signaling," *IEEE Access*, vol. 5, pp. 7414–7422, 2017.
- [5] J. Fan, Y. Ren, Y. Zhang, and X. Luo, "MLSE equalizer with channel shortening for Faster-Than-Nyquist signaling," *IEEE Photon. Technol. Lett.*, vol. 30, no. 9, pp. 793–796, May 1, 2018.
- [6] S. Sugiura, "Frequency-domain equalization of faster-than-Nyquist signaling," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 555–558, Oct. 2013.
- [7] F. Ganhao, B. Cunha, L. Bernardo, R. Dinis, and R. Oliveira, "A high throughput H-ARQ technique with Faster-than-Nyquist signaling," in *Proc. Int. Conf. Telecommun. Multimedia (TEMU)*, Heraklion, Greece, Jul. 2014, pp. 122–126.
- [8] 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.
- [9] 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.
- [10] D. Chang, O. Omomukuyo, O. Dobre, R. Venkatesan, P. Gillard, and C. Rumbolt, "Tomlinson-harashima precoding with soft detection for faster than nyquist DP-16QAM coherent optical systems," in *Proc. Opt. Fiber Commun. Conf.*, 2015, pp. 1–3.
- [11] D. Chang, O. Omomukuyo, O. Dobre, R. Venkatesan, and P. Gillard, "A faster-than-Nyquist PDM-16QAM scheme enabled by tomlinson-harashima precoding," in *Proc. 17th Int. Conf. Transparent Opt. Netw. (ICTON)*, Jul. 2015, pp. 1–4.

- [12] D. Chang, O. Omomukuyo, X. Lin, S. Zhang, O. A. Dobre, and R. Venkatesan, "Robust Faster-Than-Nyquist PDM-mQAM systems with Tomlinson-Harashima precoding," *IEEE Photon. Technol. Lett.*, vol. 28, no. 19, pp. 2106–2109, Oct. 1, 2016.
- [13] M. Maso and S. Tomasin, "Pre-equalized faster than nyquist transmission for 5G cellular microwave backhaul," in *Proc. IEEE 17th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2016, pp. 1–6.
- [14] Q. Li, F.-K. Gong, P.-Y. Song, and S.-H. Zhai, "Symbol-based multi-layer iterative successive interference cancellation for faster-than-Nyquist signalling," *JET Commun.*, vol. 14, no. 2, pp. 243–250, Jan. 2020.
- [15] 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.
- [16] O. Simeone, "A very brief introduction to machine learning with applications to communication systems," *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 4, pp. 648–664, Dec. 2018.
- [17] T. Wang, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, "Deep learning for wireless physical layer: Opportunities and challenges," *China Commun.*, vol. 14, no. 11, pp. 92–111, Nov. 2017.
- [18] H. He, S. Jin, C.-K. Wen, F. Gao, G. Y. Li, and Z. Xu, "Model-driven deep learning for physical layer communications," *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 77–83, Oct. 2019.
- [19] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.
- [20] E. Balevi and J. G. Andrews, "One-bit OFDM receivers via deep learning," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4326–4336, Jun. 2019.
- [21] X. Gao, S. Jin, C.-K. Wen, and G. Y. Li, "ComNet: Combination of deep learning and expert knowledge in OFDM receivers," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2627–2630, Dec. 2018.
- [22] S. Xue, Y. Ma, N. Yi, and R. Tafazolli, "Unsupervised deep learning for MU-SIMO joint transmitter and noncoherent receiver design," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 177–180, Feb. 2019.
- [23] T. Wang, C.-K. Wen, S. Jin, and G. Y. Li, "Deep learning-based CSI feedback approach for time-varying massive MIMO channels," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 416–419, Apr. 2019.
- [24] Y. Yang, F. Gao, X. Ma, and S. Zhang, "Deep learning-based channel estimation for doubly selective fading channels," *IEEE Access*, vol. 7, pp. 36579–36589, 2019.
- [25] Y. Yang, F. Gao, G. Y. Li, and M. Jian, "Deep learning-based downlink channel prediction for FDD massive MIMO system," *IEEE Commun. Lett.*, vol. 23, no. 11, pp. 1994–1998, 2019.
- [26] E. Nachmani, E. Marciano, L. Lugosch, W. J. Gross, D. Burshtein, and Y. Be'ery, "Deep learning methods for improved decoding of linear codes," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 119–131, Feb. 2018.
- [27] A. Bennatan, Y. Choukroun, and P. Kisilev, "Deep learning for decoding of linear Codes—A syndrome-based approach," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 1595–1599.
- [28] C. Cao, D. Li, and I. Fair, "Deep learning-based decoding of constrained sequence codes," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2532–2543, Nov. 2019.
- [29] S. Huang, Y. Jiang, Y. Gao, Z. Feng, and P. Zhang, "Automatic modulation classification using contrastive fully convolutional network," *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1044–1047, Aug. 2019, doi: 10.1109/LWC.2019.2904956.
- [30] Basicmi. (2020). *AI Chip (ICs and IPs)*. [Online]. Available: <https://github.com/basicmi/AI-Chip>
- [31] B. Yuan and K. K. Parhi, "Early stopping criteria for energy-efficient low-latency belief-propagation polar code decoders," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6496–6506, Dec. 2014.
- [32] Digital Video Broadcasting (DVB). (2015). *Implementation Guidelines for the Second Generation System for Broadcasting, Interactive Services, News Gathering and Other Broadband Satellite Applications*. [Online]. Available: https://dvb.org/wp-content/uploads/2019/12/a171-1_s2_guide.pdf
- [33] Y. Zhang, A. Liu, K. Pan, C. Gong, and S. Yang, "A practical construction method for polar codes," *IEEE Commun. Lett.*, vol. 18, no. 11, pp. 1871–1874, Nov. 2014.



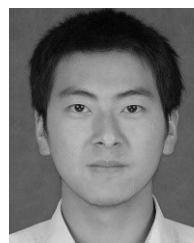
PEIYANG SONG (Student Member, IEEE) was born in Henan, China, in 1994. He received the B.Sc. degree from the School of Communications Engineering, Xidian University, in 2017. He is currently pursuing the Ph.D. degree with the State Key Laboratory of Integrated Services Networks (ISN), Xidian University. His research interests include AI-aided wireless communications, multi-beam satellite, and so on.



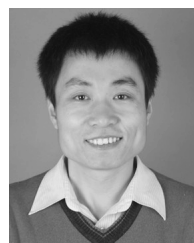
FENGKUI GONG (Member, IEEE) was born in Shandong, China, in 1979. He received the M.S. and Ph.D. degrees from Xidian University, Xi'an, China, in 2004 and 2007, respectively. From 2011 to 2012, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, McMaster University, 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.



QIANG LI (Student Member, IEEE) was born in Anhui, China, in 1994. He received the B.Sc. degree in communication engineering from Xidian University, Xi'an, China, in 2016, where he is currently pursuing the Ph.D. degree in communication and information system. His research interests include large capacity satellite communication technology, non-orthogonal transmission schemes, and pre-coding.



GUO LI (Student Member, IEEE) was born in Hanzhong, Shaanxi, China, in 1989. He received the B.S., M.S., and Ph.D. degrees in telecommunication engineering from Xidian University, Xi'an, in 2011, 2014, and 2017 respectively. From December 2015 to January 2017, he was a Visiting Ph.D. Student in electronic and computer engineering with McMaster University, Canada. Since 2018, he has been a Lecturer with the State Key Laboratory of Integrated Services Networks (ISN), Xidian University. His research interests include massive MIMO, cooperative relay transmission, space-time coding, array signal processing, and millimeter-wave communication.



HAIYANG DING (Member, IEEE) received the B.Sc. degree in communications engineering from the Xi'an Communications Institute, Xi'an, China, in 2003, the M.Sc. degree (Hons.) in electrical engineering from the Beijing University of Technology, Beijing, China, in 2006, and the Ph.D. degree (Hons.) in telecommunications engineering from Xidian University, Xi'an, in 2013. From 2016 to 2017, he was a Visiting Scholar with the Department of Electrical Engineering, Columbia University, New York. Since August 2017, he has been with the School of Information and Communications, National University of Defense Technology (NUDT), where he is currently an Associate Professor. His research interests include backscatter communications, energy harvesting, cooperative communications, and cognitive radio systems.

...