Improved Label Design for Timing Synchronization in OFDM Systems against Multi-path Uncertainty

Chaojin Qing, Member, IEEE, Shuhai Tang, Na Yang, Chuangui Rao, and Jiafan Wang

Abstract—Timing synchronization (TS) is vital for orthogonal frequency division multiplexing (OFDM) systems, which makes the discrete Fourier transform (DFT) window start at the intersymbol-interference (ISI)-free region. However, the multi-path uncertainty in wireless communication scenarios degrades the TS correctness. To alleviate this degradation, we propose a learning-based TS method enhanced by improving the design of training label. In the proposed method, the classic crosscorrelator extracts the initial TS feature for benefiting the following machine learning. Wherein, the network architecture unfolds one classic cross-correlation process. Against the multi-path uncertainty, a novel training label is designed by representing the ISI-free region and especially highlighting its approximate midpoint. Therein, the closer to the region boundary of ISI-free the smaller label values are set, expecting to locate the maximum network output in ISI-free region with a high probability. Then, to guarantee the correctness of labeling, we exploit the priori information of line-of-sight (LOS) to form a LOS-aided labeling. Numerical results confirm that, the proposed training label effectively enhances the correctness of the proposed TS learner against the multi-path uncertainty.

Index Terms—Timing synchronization, OFDM, label designing, multi-path uncertainty, machine learning.

I. INTRODUCTION

ORTHOGONAL frequency division multiplexing (OFDM) technology has been widely applied in the modern wireless and mobile communication systems, e.g., the fifth generation (5G) system [1]. At an OFDM receiver, the correct timing synchronization (TS) locates the starting of discrete Fourier transform (DFT) window per OFDM symbol within its inter-symbol-interference (ISI)-free region [2]. However, this task is hardly complete in multi-path propagation scenarios. In fact, wireless propagation scenarios emerge a lot of multi-path uncertainty, e.g., uncertain multipath delays and complex path gains, etc [3]. Under such uncertainty, the error probability of TS is increased.

To suppress the multi-path uncertainty, the joint TS and channel estimation method via the iterative interference cancellation is proposed in [4]. While this iterative processing requires high computational complexity and large processing delay. Machine learning, due to its powerful ability in tackling nonlinear problems [5], can be an alternative way to improve the TS correctness against the impact of multi-path interference. The authors in [6] investigate a neural network (NN)-based signal detection with impacts of TS error and multi-path uncertainty, yet this study does not focus on the TS task. In a recent study [7], a convolutional NN (CNN)based TS is investigated, which features a specially designed network architecture to improve the TS correctness. However, this method takes a long time for training and is not conducive to practical application. In [8], the residual timing offset estimation is accomplished by assuming the achievement of coarse TS and channel estimation, and therefore it omits the impacts of multi-path uncertainty. Since training labels to be learned are usually helpful to improve the model training without fine-tuning [9], the authors in [10] improve the TS correctness against multi-path interference by specially designing training labels. While the uncertainty of multi-path delay is hardly considered in [10], which greatly affects the correctness of label designing. Thus, the incorrect labeling limits the improvement of TS correctness provided by [10]. To our best knowledge, there is limited literature addressing this issue by designing the training label against uncertain multi-path delay.

1

In this paper, a learning-based TS method aided by the improved label designing is proposed, which aims to improve its adaptability against the multi-path uncertainty. Specifically, we design novel training label by assigning nonzero values to the label values indexed in the ISI-free region while setting other label values to zeros. In the designed training label, the values of these labels are set smaller when they are closer to the ISI-free region boundary, which highlights the ISI-free middle region to reduce the risk of timing error. Nevertheless, the uncertain multi-path delay may result in a variation of ISI-free region with environmental changes, leading to the incorrect labeling and affecting the TS correctness. To overcome this issue, we further relax the labeling restrictions to line-of-sight (LOS) cases to reserve an enough region for accommodating the uncertain multi-path delay for non-lineof-sight (NLOS) cases. This avoids the highlighted midpoint being outside the ISI-free region and thus forms the LOSbased priori information. Without excessively increasing the computational complexity, we combine a single hidden backpropagation NN (BPNN) with the classic TS processing to form the learning-based TS (so called TS learner), in which the employed BPNN only unfolds one cross-correlation process. Numerical results show that the proposed training label can effectively enhance the adaptability and correctness of TS learner against the multi-path uncertainty.

Notations: $[\cdot]^T$, $E\{\cdot\}$, $|\cdot|$, $\lceil\cdot\rceil$, and $(\cdot)^*$ denote the operations of transpose, statistical expectation, absolute, ceiling, and complex conjugate, respectively.

This work is supported in part by the Sichuan Science and Technology Program (Grant No. 2021JDRC0003, 23ZDYF0243, 2021YFG0064), the Demonstration Project of Chengdu Major Science and Technology Application (Grant No. 2020-YF09- 00048-SN), the Special Funds of Industry Development of Sichuan Province (Grant No. zyf-2018-056), and the Industry-University Research Innovation Fund of China University (Grant No. 2021ITA10016).

C. Qing, S. Tang, N. Yang, C. Rao, and J. Wang are with the School of Electrical Engineering and Electronic Information, Xihua University, Chengdu 610039, China (e-mails: qingchj@mail.xhu.edu.cn, tang-shh@stu.xhu.edu.cn, yangna6717@163.com, raochuangui5621@163.com, and jifanw@gmail.com).



Fig. 1. System model.

II. SYSTEM MODEL

We consider an OFDM system with N subcarriers, as shown in Fig. 1. At the transmitter, a transmitted OFDM symbol is expressed as

$$s(n) = \begin{cases} \sum_{k=0}^{N-1} d(k) e^{j\frac{2\pi k}{N}n}, & 0 \le n \le N\\ s(n+N), & -N_g \le n < 0 \end{cases}$$
(1)

where d(k) represents the data symbol or the element of training sequence modulated on the kth subcarrier. N_g is the length of cyclic prefix (CP). In (1), $E\{|s(n)|^2\} = P_t$ with P_t being the transmitted signal power. After transmitting s(n) over multi-path channel [11], the received sample is given by

$$y(n) = e^{j\frac{2\pi n}{N}\varepsilon} \sum_{l=1}^{L} h_l s(n-\theta-\tau_l) + w(n), \qquad (2)$$

where θ denotes the timing offset to be estimated, ε is the normalized carrier frequency offset, and w(n) stands for the additive white Gaussian noise with zero-mean and variance σ^2 . In (2), h_l and τ_l are the complex gain and normalized multipath delay of the *l*-th resolvable path, respectively. Besides, the multi-path delays are shorten than the CP length to prevent ISI, i.e., $N_g > \tau_l$, and $\tau_l = l - 1$ is considered in this paper.

By considering a N_w -length observed interval, N_w -samples of the received samples are buffered as the observed vector $\mathbf{y} \in \mathbb{C}^{N_w \times 1}$, i.e.,

$$\mathbf{y} = [y(0), y(1), \cdots, y(n), \cdots, y(N_w - 1)]^T$$
, (3)

where $N_w = 2N + N_g$ aims to observe a complete training sequence, and correspondingly, the length of searching range for unknown θ is $N_s = N_w - N = N + N_g$. In Section III, the proposed TS method for estimating timing offset is elaborated, in which the estimation value of θ is denoted as $\hat{\theta}$.

III. LEARNING-BASED TS AIDED BY LABEL DESIGNING

A. Improvement of Label Designing

By paying special attention to the uncertain multi-path delay, a novel training label $\mathbf{t} \in \mathbb{R}^{N_s \times 1}$ is designed to enhance the TS learner, i.e.,

$$\mathbf{t} = \begin{bmatrix} \underbrace{0, \cdots, 0}_{\theta + \tau_L}, \underbrace{\zeta(1), \cdots, \zeta(D)}_{\text{ISI-free region}}, \underbrace{0, \cdots, 0}_{N_s - \theta - N_g - 1} \end{bmatrix}^T, \quad (4)$$

where $D = N_g - \tau_L + 1$, denoting the length of ISI-free region. For convenience, a discrete interval for searching θ is denoted as Ω and defined as that $\Omega : \{m \mid 0 \le m \le N_s - 1, \forall m \in \mathbb{Z}\}$. Therein, the ISI-free region (denoted as Ω_{free}) is defined as that $\Omega_{\text{free}} : \{m \mid \tau_L \le m - \theta \le N_g\}$, and the ISI region, denoted as Ω_{ISI} , corresponds to a complementary set of Ω_{free} in Ω , i.e., $\Omega_{\text{ISI}} \cap \Omega_{\text{free}} = \emptyset$ and $\Omega_{\text{ISI}} \cup \Omega_{\text{free}} = \Omega$. For $\forall m \in \Omega_{\text{free}}$, the *m*-th entry in t is denoted as $\zeta(d)$, $d = 1, 2, \dots, D$, i.e.,

$$\zeta(d) = \begin{cases} d, & 1 \le d < \lceil D + 1/2 \rceil \\ D - d + 1, & \lceil D + 1/2 \rceil \le d \le D \end{cases}, \quad (5)$$

where each value of $\zeta(d)$ satisfies that the closer to the leftmost or rightmost boundary of Ω_{free} , the smaller values are set. Different from [10], the designed t in (4) not only represents Ω_{free} , but also highlights its middle region. Since the minimal training loss can be achieved after BPNN training [12], the maximum value of network output will be concentrated nearby the midpoint of Ω_{free} with a high probability. That is, the probability of correct TS is increased.

In (4), an approximate midpoint of Ω_{free} , denoted as $\mu = \theta + \lceil (\tau_L + N_g)/2 \rceil$, is considered as the ideal case of correct labeling for (4), i.e., $\mu \in \Omega_{\text{free}}$. However, due to the uncertain multi-path delay in NLOS cases, $\mu \notin \Omega_{\text{free}}$ may appear, which makes error labeling. If $\mu \notin \Omega_{\text{free}}$ has appeared in (4), the trained TS learner will learn the incorrect label, and thus degrades its TS correctness significantly. To tackle this issue, the labeling restriction is relaxed to LOS cases to reserve the enough region for accommodating the uncertain multi-path delays of NLOS cases.

B. Labeling by LOS-based Priori Information

By separately denoting ξ , G_r , G_t , and λ as the propagation distance, received antenna gain, transmitted antenna gain, and wavelength, the received signal power in a LOS scenario, denoted as $P_r(\xi)$, is given by $P_r(\xi) = \frac{\lambda^2}{(4\pi\xi)^2} P_t G_t G_r$ [13]. With the given transmitted power P_t and the resolvable received power (i.e., a constant value denoted by P_{res}), the maximum propagation distance in a LOS scenario (denoted as ξ_{LOS}) is $\xi_{\text{LOS}} = \max_{\xi} \{P_r(\xi) \ge P_{\text{res}}\}$. Therein, the resolvable received power is defined as the received power of the minimal resolvable path during the synchronization phase [14]. For $\forall \xi$ with a given P_t , relative to LOS scenarios, $P_r(\xi)$ of a NLOS scenario is inevitably reduced due to obstacles [13]. Accordingly, with a given P_t , the maximum propagation distance in a NLOS scenario, denoted as ξ_{NLOS} , satisfies $\xi_{\text{NLOS}} \leq \xi_{\text{LOS}}$. Correspondingly, $\tau_L \leq \frac{\xi_{LOS}}{c_r T}$ is satisfied, where T and c denote the sampling interval and light speed, respectively. Usually, τ_L of NLOS scenario is difficult to obtain due to the uncertain multi-path delay, while ξ_{LOS} seems to be much easier to obtain, so that $\frac{\xi_{\text{LOS}}}{c \cdot T}$ can be easily captured. Without loss of generality, $\frac{\xi_{\text{LOS}}}{c.T} < N_g$ is assumed. From [15], the CP length is much greater than the maximum propagation delay. Thus, we explore the prior information of $\frac{\xi_{\text{LOS}}}{c \cdot T} < N_g$ to improve the TS correctness. In fact, this prior information is a loose upper bound, which is only utilized to reflect the heuristic idea of this letter. With the development of integrated sensing and communication (ISAC), the sensing technology (e.g., [16]) can be employed to obtain a tight bound to further improve the TS correctness. Then, in a real propagation environment, $\tau_L \leq \frac{\xi_{\text{LOS}}}{c \cdot T} < N_g$ can be satisfied, yielding

$$\theta + \tau_L \le \theta + \frac{\xi_{\text{LOS}}}{c \cdot T} < \theta + N_g.$$
 (6)

TABLE I NETWORK ARCHITECTURE

Layer Name	Neuron Nodes	Activation Function	Normalization
Input Layer	N_s	None	ℓ_2 -norm
Hidden Layer	N	Sigmoid	None
Output Layer	N_s	None	None

The above (6) can be viewed as the LOS-based priori information, obtaining a narrowed interval of ISI-free region, i.e.,

$$\Omega_{\text{nfree}} : \left\{ m | \frac{\xi_{\text{LOS}}}{c \cdot T} \le m - \theta < N_g \right\} s.t. \ \Omega_{\text{nfree}} \subseteq \Omega_{\text{free}}.$$
(7)

Although the uncertain multi-path delay makes $\mu \in \Omega_{\text{free}}$ hardly, it can be easily achieved by $\mu \in \Omega_{\text{nfree}}$. Then, m in (7) is substituted by μ , in which $\frac{\xi_{\text{LOS}}}{c \cdot T} \leq \mu - \theta < N_g$. By using (6), we have

$$\theta + \tau_L \le \mu < \theta + N_g,\tag{8}$$

and thus $\mu \in \Omega_{\text{free}}$ is satisfied. By considering $\mu = \theta + \lceil (\tau_L + N_g)/2 \rceil$, the τ_L for training can be obtained via $\frac{\xi_{\text{LOS}}}{c \cdot T}$. In (6), if $\lceil N_g/2 \rceil \leq \frac{\xi_{\text{LOS}}}{c \cdot T} < N_g$, the τ_L for training can be set as $2\frac{\xi_{\text{LOS}}}{c \cdot T} - N_g \leq \tau_L \leq \frac{\xi_{\text{LOS}}}{c \cdot T}$. Otherwise, $1 \leq \tau_L \leq \frac{\xi_{\text{LOS}}}{c \cdot T}$ can be considered. For the both cases, a training option of τ_L aided by the LOS priori information of (6) is expressed as

$$\begin{cases} \tau_L \sim U\left[1, \left\lceil \frac{\xi_{\text{LOS}}}{c \cdot T} \right\rceil\right], & \frac{\xi_{\text{LOS}}}{c \cdot T} \in \left(0, \left\lceil \frac{N_g}{2} \right\rceil\right) \\ \tau_L \sim U\left[2 \left\lceil \frac{\xi_{\text{LOS}}}{c \cdot T} \right\rceil - N_g, \left\lceil \frac{\xi_{\text{LOS}}}{c \cdot T} \right\rceil\right], & \frac{\xi_{\text{LOS}}}{c \cdot T} \in \left[\left\lceil \frac{N_g}{2} \right\rceil, N_g\right) \end{cases}$$
(9)

According to (6)–(9), the τ_L for training is derived to promise the correct labeling given in (4). Since the ISI-free's approximate midpoint μ is highlighted with $\mu = \theta + \lceil (\tau_L + N_g)/2 \rceil$ and $\tau_L = 2\lceil \mu - \theta \rceil - N_g$, the setting of τ in (9) can guarantee the condition that $\mu \in \Omega_{\text{nfree}} \subseteq \Omega_{\text{free}}$. On the basis of this, the NN trained by the label designed in (4) can achieve $\hat{\theta} \in \Omega_{\text{free}}$ with a high probability. Furthermore, the more cases of τ_L that are learned, the greater robustness against the varied τ_L that can be achieved. Therefore, it is reasonable to relax τ_L to a random value instead of a constant value, as done in (9).

C. Learning-based TS Method

1) Network Architecture: The architecture of the designed TS learner is illustrated in TABLE I, which has an input layer, a hidden layer and an output layer. To avoid excessively increasing complexity, the neurons of input layer, hidden layer, and output layer are set as N_s , N, and N_s , respectively. Therein, the N-neuron hidden layer is derived from considering unfolding one cross-correlation process. Since the input scale of NN may differ from those of other layers, it is practical to employ the ℓ_2 -norm in the input layer. Besides, the hidden layer employs the sigmoid activation function, i.e., $f(x) = 1/(1 + e^{-x})$, for the reason that it is easy to calculate and suitable for shallow NNs [17].

2) Initial Feature Extraction: By employing the classic TS method, its timing metric is utilized as the initial N_s features of TS to facilitate the model learning [18], i.e.,

$$F(m) = \left|\sum_{n=0}^{N-1} x^*(n) y(m+n)\right|^2, 0 \le m \le N_s - 1, \quad (10)$$

where x(n) is a local training sequence, e.g., the Zadoff-Chu sequence in [19]. By using the ℓ_2 -norm of F(m), 3

the network input, denoted as Q(m), is normalized to the interval [0,1), i.e., $0 \leq Q(m) < 1$, which is expressed as $Q(m) = F(m)/\sqrt{\sum_{m=0}^{N_s-1} |F(m)|^2}$. The vector form of Q(m), denoted as $\mathbf{Q} \in \mathbb{R}^{N_s \times 1}$, is expressed as $\mathbf{Q} = [Q(0), \dots, Q(N_s - 1)]^T$.

3) Offline Training: First, by using (1)–(10), the training data set $\{\mathbf{Q}_i, \mathbf{t}_i\}_{i=1}^{N_t=10^4}$ are obtained. For the *i*-th training sample, we employ the exponentially decayed channel model with its decayed factor η_i being $\eta_i \sim U[0.01, 0.5]$ for combating the uncertainty of complex path gain. Meanwhile, $\frac{\xi_{\text{LOS}}}{c \cdot T} = \lceil 7N_g/8 \rceil$ is given, which derives $\tau_{L,i} \sim U[\lceil 3N_g/4 \rceil, \lceil 7N_g/8 \rceil]$ for the label designing in (4) to combat the uncertain multi-path delay. Besides, the SNR for generating the *i*-th training sample is randomly selected from $\{-2\text{dB}, \text{odB}, 2\text{dB}, ..., 10\text{dB}\}$ with a probability of 1/7.

In the developed TS learner, the optimizer employs the stochastic gradient descent (SGD) algorithm and sets the initial learning rate as $\alpha = 0.001$ [20]. In this paper, the loss function is defined as

$$L_{\boldsymbol{\Theta}} = \frac{1}{N_t} \sum_{i=1}^{N_t} \|G_{\boldsymbol{\Theta}}\left(\mathbf{Q}_i\right) - \mathbf{t}_i\|_2^2, \tag{11}$$

where Θ is a set of network parameters (i.e., weights and biases) to be optimized, and $G_{\Theta}(\cdot)$ is a mapping function parameterized by Θ . By setting J as the total iterative steeps, the network optimization is defined as [21]

$$\boldsymbol{\Theta}_{j+1} \leftarrow \boldsymbol{\Theta}_j - \alpha \nabla L_{\boldsymbol{\Theta}_j}, \tag{12}$$

where Θ_j , $j = 1, 2, \dots, J$, denotes the network parameters after the *j*th optimization, and ∇L_{Θ_j} is the gradient of L_{Θ_j} .

4) Online Deployment: According to (1)–(2) and (10), the initial TS features are first extracted, and then normalized to form the network input **Q**. With the trained G_{Θ} , the network output $\mathbf{O} \in \mathbb{R}^{N_s \times 1}$ is obtained by $\mathbf{O} = G_{\Theta}(\mathbf{Q})$. Finally, the output **O** is expressed as $[O(0), \dots, O(m), \dots, O(N_s - 1)]^T$ for timing offset estimation, i.e., $\hat{\theta} = \arg \max_{m \in \Omega} |O(m)|$.

IV. NUMERICAL RESULTS

A. Parameter Setting

In the simulations, we consider that N = 128, $N_g = 32$, $N_w = 288$, and $N_s = 160$. The error probability of TS is utilized to evaluate the TS correctness, which is defined as the probability of estimating the timing offset outside of the ISI-free region. The OFDM technology is primarily employed to combat the ISI caused by the multi-path propagation [22], and thereby the frequency selective fading channels are mainly considered in the simulations. We leverage the delay spread profile to quantify the uncertainty of multi-path fading [23]. To simulate the multi-path uncertainty, Fig. 3 to Fig. 5 depict the TS correctness of the proposed method for the cases where the training delay spread profile differs from the testing ones. The case of the maximum LOS propagation delay (i.e., $\frac{\xi_{\text{LOS}}}{c.T}$) is set only to provide the LOS priori information (i.e., (6)) for assisting the improvement of label designing in (4). Consequently, we consider a relatively large value of $\frac{\xi_{\text{LOS}}}{c \cdot T}$ to guarantee the correctness of LOS priori information, i.e., $\frac{\xi_{\text{LOS}}}{c \cdot T}$ is set as $[7N_q/8]$. In the simulations, the exponentially decayed Rayleigh fading channel in [11] and different 5G tapped-delayline (TDL) channel models (e.g., TDL-B and TDL-C given in

TABLE II COMPUTATIONAL COMPLEXITY AND PROCESSING DELAY

Method	Computational Complexity	Example (CM)	Processing Delay (sec)
Ref [4]	$LNN_s + \sum_{l=1}^{L} (3lN_s + l^3 + l^2N_s)$	2167396	81.41
Ref [10]	$1.5N + 4(N_s - 1) + 16N_s^2$	410428	1.614
Ref [7]	$0.5N_s^2 + 2NN_s + 1.5N_gN_s + N_s$	70240	0.349
Prop	$NN_s + 0.5NN_s$	30720	0.141



Fig. 2. Computational complexity analysis.

3GPP TR 38.901 [24]) are employed. Besides, the signal-tonoise ratio (SNR) is defined as SNR = $10 \log_{10}(P_t/\sigma_n^2)$ [25], and correspondingly, $\sum_{l=1}^{L} |h_l|^2 = 1$ is considered [11].

For the ease of description, the TS learner with the training labels proposed in this paper and [10] are referred as to "Prop" and "Ref [10]", respectively. "Ref [7]" is the TS method given in [7]. "Ref [4]" stands for the iterative-based TS method in [4]. Besides, the classic TS method in [26], denoted as "Ref [26]", serves as the baseline.

B. Computational Complexity and Processing Delay

TABLE II presents the expression of computational complexity, in which the complex multiplication (CM) is used for evaluating the computational complexity. By considering the impact of searching length N_s , Fig. 2 plots the CM of each given TS method in terms of N_s , where N_s increases from 160 to 1024 with the interval being 16. All the evaluations are carried out on an Intel Core i5-11300H, 3.10GHz CPU, and L = 28 is considered for 10^4 experiments in both TABLE II and Fig. 2. In Fig. 2, the CM of "Prop" is smaller than those of "Ref [4]", "Ref [7]", and "Ref [10]". Similarly, TABLE II reflects that "Prop" obtains a lower computational complexity and processing delay, and the case where $N_s = 160$ is given. The reason is that "Prop" unfolds one cross-correlation process, while others unfold at least two iterations of the crosscorrelation process. From TABLE I, the CM of the designed NN of "Prop" is $0.5NN_s$, which does not exceed the CM of one cross-correlation process, i.e., NNs. Naturally, the CM of "Prop" will not exceed the CM of two cross-correlation process. Thus, the relatively lightweight NN is employed by "Prop" compared with the given TS methods.

C. Effectiveness Analysis

Fig. 3 plots the error probability of TS to evaluate the TS performance of different methods against multi-path interference and its uncertain multi-path delay. Meanwhile, the



Fig. 3. Effectiveness analysis.



Fig. 4. Robustness analysis.

case of $\{\tau_L = 27, \eta = \ln 10/(L-1)\}$ is employed to simulate the uncertainty of multi-path delay, while the case of $\{\tau_L \sim U[24, 28], \eta \sim U[0.01, 0.5]\}$ is utilized for the offline training presented in Section III-C. For each given SNR in Fig. 3, "Prop" obtains significantly smaller error probability of TS compared with other given TS methods. Compared with "Prop", "Ref [26]" presents a weaker robustness against the increased numbers of multi-path due to the lack of the consideration of rich multi-path scenarios (only 3 or 6 taps are considered in the simulations in [26]). Relative to "Ref [10]", the superiority of "Prop" is that the proposed training label is not only specially designed for decreasing the risk of estimating the timing offset outside the ISI-free region, but also improved by LOS-based labeling for increasing the correctness of labeling. Although "Prop" and "Ref [7]" both exploit the LOS-based priori information for correct labeling, the training label designed by "Prop" is superior due to the highlighting of ISI-free region and its approximate middle point. This reveals the feasibility of enhancing the label designing in improving the TS correctness. Last but not least, compared with "Ref [4]", "Prop" achieves a lower error probability with the reduced computational complexity, due to the powerful ability of machine learning in coping with the deficiencies such as noise and multi-path interference. This reflects that the learning-based TS method with an appropriate training label can improve the TS correctness with the reduced complexity, and thereby the proposed method provides an alternative solution for alleviating the demand of complexly modeling in a practical system. To sum up, relative to "Ref [26]", "Ref [4]", "Ref [7]", and "Ref [10]", "Prop" can effectively improves the TS correctness against uncertain multi-path delay.



Fig. 5. Generalization analysis.

D. Robustness Analysis

To analyze the robustness against the training sequence length (i.e., N), Fig. 4 plots the error probability of TS, where N = 96, N = 128, and N = 160 are considered. Except for N, other parameters remain the same as those mentioned in Section III-C. For each given value of N in Fig. 4, "Prop" reaches the smallest error probability compared with other given TS methods. This reflects the robustness of "Prop" against the change of N. Meanwhile, it is worth noting that the error probability of "Prop" reduces with the increases of N. This is due to the fact that the ability of anti-noise in cross-correlator is enhanced (deteriorated) with the increase (decrease) of N. Therefore, the proposed learning-based TS method assisted by the improved label design can robustly improve the TS correctness.

E. Generalization Analysis

Fig. 5 presents the comparison of error probability in different channel models to analyze the generalization performance of "Prop". Except for the testing wireless channel models, other parameters are the same as those mentioned in Section IV-B. Importantly, the NN adopted in this letter dose not need to be retrained when the testing channel model differs from the training one. For each given channel model in Fig. 5, "Prop" reaches the lowest error probability among the given TS methods. Furthermore, it is not obvious that, for all given SNRs, the fluctuations of the error probabilities of "Prop" caused by different channel models. Therefore, the TS correctness of "Prop" is superior to those of other TS methods. This shows that the proposed TS methods possesses a good generalization performance against different 5G TDL channel models.

V. CONCLUSION

Against the influence of uncertain multi-path delay, the proposed learning-based TS method aided by the improved label designing is investigated in OFDM systems. By highlighting the ISI-free region and its the approximate midpoint against uncertain multi-path delay, the designed training label effectively reduces the risk that the DFT window starts at the ISI region. Meanwhile, with the LOS-based priori information, the incorrect labeling affected by the uncertain multi-path delay is further rectified. Simulation results validates the effectiveness and generalization of designed training label in improving the TS correctness of the learning-based TS method against multi-path uncertainty. In our future works, we will investigate the sensing-aided TS from the perspective of ISAC.

REFERENCES

- I. B. F. de Almeida, L. L. Mendes, J. J. P. C. Rodrigues, and M. A. A. da Cruz, "5G waveforms for IoT applications," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2554–2567, Apr. 2019.
- [2] B. Yang, K. Letaief, R. Cheng, and Z. Cao, "Timing recovery for OFDM transmission," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2278– 2291, Nov. 2000.
- [3] A. Goldsmith, Wireless communications. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [4] B. Lopes, S. Catarino, N. M. B. Souto, R. Dinis, and F. Cercas, "Robust joint synchronization and channel estimation approach for frequencyselective environments," *IEEE Access*, vol. 6, pp. 53180–53190, Sep. 2018.
- [5] J. Wang, C. Jiang, H. Zhang, Y. Ren, K.-C. Chen, and L. Hanzo, "Thirty years of machine learning: The road to pareto-optimal wireless networks," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1472–1514, Jan. 2020.
- [6] Z. He and X. Huang, "Improved deep learning in OFDM systems with imperfect timing synchronization," in *Proc. IEEE 91th Veh. Tech. Conf.*, May 2020, pp. 1–5.
- [7] C. Qing, S. Tang, X. Cai, and J. Wang, "Lightweight 1-D CNN-based timing synchronization for OFDM systems with CIR uncertainty," *IEEE Wireless Commun. Lett.*, vol. 11, no. 11, pp. 2375–2379, Nov. 2022.
- [8] J. Liu, K. Mei, X. Zhang, D. McLernon, D. Ma, J. Wei, and S. A. R. Zaidi, "Fine timing and frequency synchronization for MIMO-OFDM: An extreme learning approach," *IEEE Trans. Cogn. Commun. Netw.*, vol. 8, no. 2, pp. 720–732, June 2022.
- [9] Y. Liu, C. An, and X. Qiu, "*Y*-tuning: An efficient tuning paradigm for large-scale pre-trained models via label representation learning," *arXiv* preprint arXiv:2202.09817, 2022.
- [10] C. Qing, S. Tang, C. Rao, Q. Ye, J. Wang, and C. Huang, "Label designbased ELM network for timing synchronization in OFDM systems with nonlinear distortion," in *Proc. IEEE 94th Veh. Tech. Conf.*, Sep. 2021, pp. 01–05.
- [11] Z. Xiao, D. Jin, and N. Ge, "Matched filter-autocorrelation (MF-AC) for packet detection in multipath channels," *IEEE Commun. Lett.*, vol. 17, no. 8, pp. 1608–1611, Aug. 2013.
- [12] R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Neural Netw. Percept.* Elsevier, June 1992, pp. 65–93.
- [13] T. S. Rappaport et al., Wireless communications: principles and practice. 2nd ed. Upper Saddle River, NJ: Prentice Hall PTR, 2001.
- [14] O.-S. Shin and K. B. Lee, "Differentially coherent combining for doubledwell code acquisition in DS-CDMA systems," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1046–1050, Jul. 2003.
- [15] C. Zhang, X. Chen, and Y. Luo, "Threshold optimization for isifree region detection in high-mobility fading channels," in *Proc. Int. Workshop High Mobil. Wirel. Commun.*, Oct. 2015, pp. 46–50.
- [16] J. A. Zhang, K. Wu, X. Huang, Y. J. Guo, D. Zhang, and R. W. Heath, "Integration of radar sensing into communications with asynchronous transceivers," *IEEE Commun. Mag.*, vol. 60, no. 11, pp. 106–112, Nov. 2022.
- [17] B. Ding, H. Qian, and J. Zhou, "Activation functions and their characteristics in deep neural networks," in *Proc. Chin. Control Decis. Conf.*, Jul. 2018, pp. 1836–1841.
- [18] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern. Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Mar. 2013.
- [19] D. Chu, "Polyphase codes with good periodic correlation properties (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 531–532, Jul. 1972.
- [20] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint, arXiv:1412.6980, 2014.
- [21] S. Ruder, "An overview of gradient descent optimization algorithms," arXiv preprint, arXiv:1609.04747, 2016.
- [22] C. Y. Wong, R. Cheng, K. Lataief, and R. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.
- [23] I. E. Telatar and D. N. C. Tse, "Capacity and mutual information of wideband multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1384–1400, Jul. 2000.
- [24] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," Tech. Rep. 3GPP TR 38.901 (V17.0.0), Mar. 2022.
- [25] X. Yao, Y. Liu, and G. Lin, "Evolutionary programming made faster," *IEEE Trans. Evol. Comput.*, vol. 3, no. 2, pp. 82–102, Jul. 1999.
- [26] M. Ruan, M. C. Reed, and Z. Shi, "Training symbol based coarse timing synchronization in OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2558–2569, May 2009.