Channel Estimation and Detection for Symbiotic Radio Systems Over High-Mobility Channels

Qin Tao, *Member, IEEE*, Kai Jin, Jiabao Gao, Weijie Yuan, *Senior Member, IEEE*, Ben Wang, and Chau Yuen, *Fellow, IEEE*

***Abstract*—In symbiotic radio (SR), the secondary system not only shares the spectrum and power of the primary system but also enhances its performance by providing multipath gains, fostering a cooperative mutualism between the two systems. However, in high-mobility channels, time-frequency selective fading presents significant challenges for reliable SR communications. The recently introduced orthogonal time-frequency space (OTFS) technique, which processes signals in the delay-Doppler (DD) domain, is expected to improve SR communication performance in high-speed mobile scenarios. In this paper, we propose embedding primary information symbols in the DD domain using amplitude-phase modulation, while employing a combinatorial frequency (CF) modulation strategy for secondary information transmission. To obtain channel state information (CSI) and detect secondary symbols, for some special scenarios, we propose an off-grid sparse Bayesian learning (SBL)-based method. This method first estimates the equivalent CSI and then detects the symbols by leveraging the highly structured Doppler shifts. For more general scenarios, we introduce a model-driven equivalent CSI estimation-net (ECSIEst-Net) and a data-driven secondary symbol detection-Net (SSymDet-Net). Numerical results are provided to guide parameter selection and demonstrate the effectiveness of the proposed methods.**

***Index Terms*—Symbiotic radio, orthogonal time frequency space, channel estimation, sparse Bayesian learning, deep learning, symbol detection.**

I. Introduction

A. Background

THE Internet of things (IoT) network has fundamentally transformed our interaction with the world and is experiencing rapid advancement. It is predicted that the number of IoT devices will surpass 30 billion by 2030. The rapid proliferation of connected devices presents significant challenges in spectrum management, battery maintenance, and energy efficiency [1]. Addressing these challenges, symbiotic radio (SR) emerges as a novel solution in IoT communications, leveraging a cooperative relationship between primary and secondary users to enhance energy and spectral efficiency [2], [3], [4], [5], [6], [7]. Specifically, in an SR system, IoT devices typically serve as secondary emitters (SE), transmitting data by backscattering or reflecting signals from a primary emitter (PE) without generating radio frequency (RF) signals themselves. Note that the SE can be implemented using either a tag or an intelligent reflecting surface (IRS). This transmission strategy not only conserves energy but also enhances the primary system's performance through multipath gains, fostering cooperative mutualism between the primary and secondary systems. Owing to these attributes, SR has attracted substantial interest from both industry and academia in recent years [8], [9], [10], [11], [12].

B. Related Works

1) SR Systems: Despite the high energy and spectrum efficiency offered by SR communications, several challenges remain, including the modulation of secondary information, and the separation and detection of combined primary and secondary signals at the receiver. To address these issues, the authors in [13] proposed using the on-off keying (OOK) modulation for secondary information and derived an energy-based detector. To enhance robustness against channel estimation errors, [14] suggested using Manchester coding for binary symbols at the SE, converting the energy hard decision detector in [13] into a relative comparison of energy between the first and second halves of each symbol duration. Furthermore, to increase the SE symbol rate, the authors in [15] and [16] proposed using higher-order modulation schemes such as M-frequency shift keying (MFSK) [15] and M-phase shift keying (MPSK) [16] at SE, and developed corresponding energy detectors and non-coherent detectors. To enhance the flexibility of SE information modulation and further improve system

---

Received 16 July 2024; revised 20 January 2025; accepted 27 April 2025. Date of publication 8 May 2025; date of current version 14 October 2025. The work of Qin Tao was supported by Zhejiang Provincial Natural Science Foundation of China under Grant LQ24F010005 and in part by the National Natural Science Foundation of China under Grant 62301206. The work of Weijie Yuan was supported by the National Natural Science Foundation of China under Grant 62471208. The work of Chau Yuen was supported by the National Research Foundation and Infocomm Media Development Authority under its Future Communications Research and Development Program under Grant FCP-NTU-RG-2024-025. The associate editor coordinating the review of this article and approving it for publication was K. Tourki. *(Corresponding author: Ben Wang.)*

Qin Tao, Kai Jin, and Ben Wang are with the School of Information Science and Technology, Hangzhou Normal University, Hangzhou 311121, China (e-mail: taoqin@hznu.edu.cn; jinkai@stu.hznu.edu.cn; 20170056@hznu.edu.cn).

Jiabao Gao is with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: gao_jiabao@zju.edu.cn).

Weijie Yuan is with the Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: yuanwj@sustech.edu.cn).

Chau Yuen is with the School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore 639798 (e-mail: chau.yuen@ntu.edu.sg).

Digital Object Identifier 10.1109/TWC.2025.3566173

performance, the authors in [17] and [18] proposed using intelligent reflecting surfaces (IRS) as the SE and employing spatial OOK reflection modulation for secondary transmission. Moreover, leveraging the signal sparsity, algorithms such as bilinear generalized approximate message passing (BiG-AMP) and generalized approximate message passing (GAMP) were introduced for detecting secondary symbols.

2) OTFS Technique: Most current research on SR focuses on static or slow-moving scenarios, where the channels remain quasi-static during the coherence time. However, in high-speed mobile receiver scenarios, the channel coherence time is significantly reduced, and the Doppler effect causes the channels to exhibit doubly selective characteristics in both time and frequency domains. This creates substantial challenges for channel characterization in the time-frequency (TF) domain, rendering traditional TF signal processing techniques, such as OFDM or enhanced OFDM [19], [20], unsuitable for these environments. To achieve reliable communication in such challenging environments, orthogonal time frequency space (OTFS) modulation has gained considerable attention in recent years [21], [22], [23], [24], [25]. Specifically, the OTFS systems process signals in the delay-Doppler (DD) domain, where channels exhibit sparse and quasi-static characteristics, allowing them to be effectively described by a limited number of Doppler and delay parameters. Moreover, the inherent DD domain signal processing ensures that each DD symbol experiences the entire TF domain channel response within one OTFS frame, allowing OTFS to fully exploit the degrees of freedom in high-mobility channels [26], [27]. Furthermore, by utilizing DD domain signal processing, OTFS effectively distributes signal energy more uniformly and reduces peaks in the time-domain signal, leading to a lower peak-to-average power ratio (PAPR) [28], [29], [30].

A substantial portion of studies on OTFS modulation have focused on primary systems. For example, the authors in [31], [32], [33] characterized the relationship between channels and transmitted symbols in the DD domain, addressing the input-output relationship with ideal pulse or rectangular waveforms, while considering integer/fractional delay and Doppler shifts. Building on these foundations and considering fractional Doppler shifts, [34] introduced an embedded pilot-aided strategy that organizes pilot and data symbols within each OTFS symbol frame, employing full or partial guards to mitigate interference. Additionally, the authors in [35] proposes transforming the estimation of fractional delay and Doppler shifts into a two-dimensional sparse signal recovery problem. By decoupling these parameters, an off-grid sparse Bayesian learning (SBL) method was also introduced for precise channel estimation [35]. For reliable symbol detection in OTFS systems, linear detectors like zero-forcing [36] and linear minimum mean square error (LMMSE) detectors [37] have been proposed. To further enhance detection performance, non-linear detectors, such as the variational Bayes approach [38] and the cross-domain iterative detection algorithm [39], have been developed, achieving significant performance improvements compared with linear detectors.

3) IRS-Assisted OTFS Systems: Parallel to channel estimation and symbol detection design, employing an IRS can enhance OTFS system spectrum efficiency by providing multipath gain with limited resource consumption [40]. Specifically, the authors in [41] and [42] analyzed the input-output relationship of IRS-assisted OTFS systems, assuming that the IRS reflection coefficients remain time-invariant over the duration of each OTFS frame. To fully leverage IRS's passive beamforming capabilities, [43] designed a phase adjustment strategy to maximize the strongest delay-Doppler channel response (DDCR), thereby optimizing the achievable rate. Additionally, [44] proposed an angle domain-based phase shift strategy for IRS in MIMO scenarios to maximize channel capacity. To ensure timely and effective IRS passive beamforming, the authors in [45] suggested utilizing the estimated delay and Doppler shifts to predict the IRS passive beamforming. Furthermore, [46] developed an optimization framework to design the IRS coefficients to minimize the bit error rate (BER) of a low-complexity zero-forcing (ZF) equalizer.

C. Motivation and Challenges

Most of the existing research on SR systems typically focuses on static or slow-moving scenarios, whereas this work addresses SR in high-mobility environments, necessitating the use of OTFS modulation. However, most OTFS research has primarily concentrated on primary transmission. Even in IRS-assisted OTFS systems, the IRS typically supports the primary system using time-invariant reflecting coefficients, without transmitting any information. In the considered SR-OTFS system, the coexistence of both primary and secondary transmissions within the OTFS framework gives rise to new issues and challenges:

1) Since the secondary information is modulated onto the primary signals, significant interference between the primary and secondary information is inevitable. It is crucial to jointly design modulation methods for both systems to facilitate the independent transmission and separation of the two types of information. Unlike time or frequency domain signal processing in slow-moving scenarios, DD domain signal processing in high-mobility scenarios leads to distinct modulation designs for both primary and secondary transmissions. Moreover, although our previous work [47] explored the SR-OTFS system by proposing modulating secondary information into the frequencies of a periodic rectangular wave, this simple secondary modulation strategy introduces harmonic interference, which significantly reduces sparsity and would lead to low data rate.
2) The reflection coefficients of the informative IRS vary over time during each OTFS frame, and the complex interaction between the time-varying IRS modulation patterns and the DD-domain channels creates an input-output relationship that differs significantly from the time-invariant IRS case. Additionally, the input-output relationship varies with different time-varying IRS modulation designs. Accurately capturing this input-output relationship under the influence of time-varying secondary information is essential for designing effective channel estimation and symbol detection methods.

3) The time-varying IRS reflecting pattern would reconfigure and interfere with the DD domain channels. Building on the proposed modulation strategy, it is crucial to develop methods for extracting secondary information from unknown channel state information (CSI) and accurately estimating the CSI.

D. Contributions

In response to the above challenges and difficulties, this paper investigates the transmission of SR systems using OTFS modulation, including information modulation schemes, input-output relationships, channel estimation, and symbol detection methods. The contributions of this work are summarized as follows:

- **Modulation Strategy Design:** To facilitate the aliasing and differentiation of independent primary and secondary information, this paper proposes embedding the primary informative symbols in the DD domain using amplitude-phase modulation, while employing a combinatorial frequency (CF) modulation strategy for secondary information transmission, as shown in Section II-B.
- **Input-Output Relationship Characterization:** Based on the modulation strategy outlined above, we then analyze the input-output relationships, highlighting the relationships between the secondary information and the channels, as shown in Section II-C. Our analysis shows that CF modulation in the secondary system shifts and expands the Doppler shifts of the original channels in a highly structured manner, thereby aiding in the detection of symbols from unknown CSI.
- **Equivalent CSI Estimation and Symbol Detection Design:** To acquire the equivalent CSI and secondary symbols, considering the *ultra-wideband or/and short package* scenarios, we propose an off-grid SBL-based method in Section III, which first estimates the equivalent CSI and subsequently detects symbols by leveraging the obtained highly structured Doppler shifts; For more *general scenarios*, the above method is no longer applicable, we then introduce a model-driven Equivalent CSI Estimation-Net (ECSIEst-Net) and a data-driven Secondary Symbol Detection-Net (SSymDet-Net) in Section IV.
- **Numerical Results:** In Section V, numerical results are provided to guide parameter selection and demonstrate the superiority of the proposed methods.

E. Organization and Notations

The rest of the paper is organized as follows. Section II presents the system model. Section III introduces the off-grid SBL-based joint estimation and detection approach. Section IV proposes data-driven and model-driven methodologies for channel estimation and symbol detection, respectively. Section V presents our simulation results, and Section VI concludes the paper.

Notations: Scalars are represented using non-bold letters, while vectors and matrices are indicated by lowercase and uppercase boldfaced letters, respectively. The function $\delta(\cdot)$ represents the unit pulse function. $\mathcal{CN}(\mu, \sigma^2)$ denotes the complex Gaussian distribution with mean μ and variance σ^2. The Hermitian transpose and transpose of a matrix $\boldsymbol{A}$ are denoted by $\boldsymbol{A}^H$ and $\boldsymbol{A}^T$, respectively. The operator $\bullet$ represents the element-wise product. $\mathbb{C}^{N\times M}$ represents an $N \times M$ complex matrix set. $\boldsymbol{a}(i)$ denotes the i-th entry of vector $\boldsymbol{a}$. $\Re\{a\}$ gets the real part of the complex value a. $\boldsymbol{I}_K$ denotes the identity matrix of size K. $\mathrm{diag}(\boldsymbol{x})$ returns a square diagonal matrix with the elements of vector $\boldsymbol{x}$ on the main diagonal, while $\mathrm{diag}(\boldsymbol{X})$ returns a column vector of the main diagonal elements of $\boldsymbol{X}$. $|\boldsymbol{x}|.^2$ represents the element-wise square.

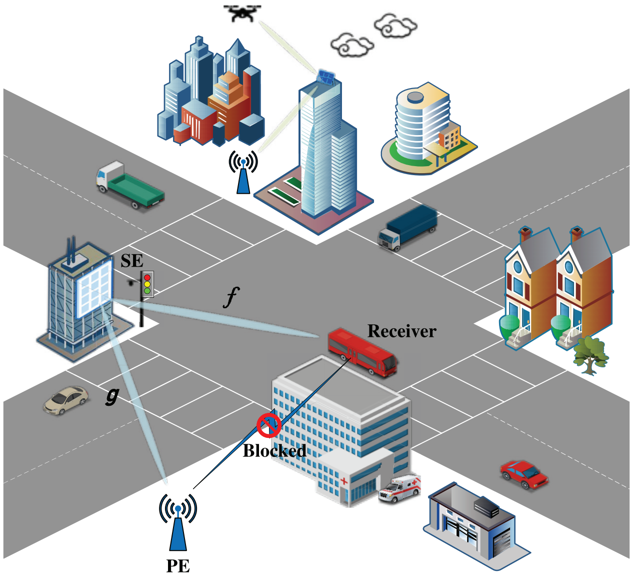


Fig. 1. Illustration of the SR-OTFS system model, where a primary emitter (PE) and secondary emitters (SEs) simultaneously transmit independent information to a high-mobility receiver. The SEs assists communication under blockage conditions between the PE and receiver.

II. System Model

This paper investigates the downlink transmission of SR systems over high-mobility environments using OTFS technology, as depicted in Fig. 1. In this configuration, both the PE and the SE simultaneously transmit independent information to a high-speed, single-antenna receiver. Specifically, the PE can be a single-antenna base station (BS) actively transmitting informative RF signals, such as high-data-rate or critical traffic information. Meanwhile, the SE could be an IRS equipped with a $K_x \times K_y$ uniform planar array, mounted on infrastructure surfaces, and passively modulating its information—such as data from road signs, traffic lights, or sensors—onto the impinging RF signal.[1] Since the PE typically covers a large-scale range, roads, buildings, and other structures often cause obstructions, making a continuous direct channel link between the PE and high-mobility receiver relatively uncommon. Therefore, in this work, we primarily focus on a scenario

---

[1] In the case with multiple SEs, the design of the handover protocol between SEs is intricately tied to system requirements. This is a complex task and falls outside the scope of the current work that deals with channel estimation and symbol detection, which remains a focus of our future research.

where there is blockage between the SE and the receiver.[2] Moreover, the SE will provide additional communication links and assist the primary transmission. OTFS modulation technology processes signal in the DD domain with a duration of *NT* and a bandwidth of $M\Delta f$, which is discretized into the grid $\mathit{\Gamma} = \left\{\left(\frac{k}{NT}, \frac{l}{M\Delta f}\right), k = 0, \ldots, N-1, l = 0, \ldots, M-1\right\}$ [48]. Here, $\Delta f = 1/T$ represents the subcarrier spacing.

A. Channel Representation

The high-mobility receiver induces significant time and frequency double-selective fading in the channels. To effectively characterize these rapidly changing channels, the OTFS modulation technology propose to handle the channels in the DD domain. In the context of mm-wave scenarios, given the proximity of the SE to the PE, we focus on the line of sight (LoS) path without Doppler shifts for the SE-PE channel, expressed as

$$\boldsymbol{g} = \sigma_1^{(g)} \delta(\tau - \tau_1^{(g)}) \boldsymbol{a}_{K_x}\left(\vartheta_{1,x}^{(g)}\right) \otimes \boldsymbol{a}_{K_y}\left(\vartheta_{1,y}^{(g)}\right). \tag{1}$$

Here, $\sigma_1^{(g)}$ and $\tau_1^{(g)}$ denote the complex channel gains and delays for the channel $\boldsymbol{g} \in \mathbb{C}^{K_x K_y \times 1}$, respectively; $\vartheta_{1,x/y}^{(g)}$ represents the effective direction of arrival (DOA) along the x/y axis; $\boldsymbol{a}_{K_{x/y}}(\vartheta) = \left[1, e^{j2\pi \frac{d}{\lambda}\vartheta}, \cdots, e^{j2\pi\frac{d}{\lambda}\left(K_{x/y}-1\right)\vartheta}\right]^T$ denotes the steering vector, where d represents the distance between two adjacent reflecting elements, typically set as $d = \lambda/2$ with λ being the wavelength [49].

The multi-path channel for the SE-receiver can be expressed as

$$\boldsymbol{f} = \sum_{\tau_i^{(f)} \in \mathcal{T}^{(f)}} \sum_{n=1}^{P_i^{(f)}} \left\{ \sigma_{i,n}^{(f)} \delta(\tau - \tau_i^{(f)}) \delta(\nu - \nu_{i,n}^{(f)}) \right.$$
$$\left. \times \boldsymbol{a}_{K_x}\left(\vartheta_{i,n,x}^{(f)}\right) \otimes \boldsymbol{a}_{K_y}\left(\vartheta_{i,n,y}^{(f)}\right) \right\}, \tag{2}$$

where $\mathcal{T}^{(f)}$ represents the set of delays for channel $\boldsymbol{f} \in \mathbb{C}^{K_x K_y \times 1}$, and $P_i^{(f)}$ is the number of paths on delay $\tau_i^{(f)}$; $\sigma_{i,n}^{(f)}$, $\tau_i^{(f)}$ and $\nu_{i,n}^{(f)}$ represent the complex channel gains, delays and Doppler shifts corresponding to *n*-th path on delay $\tau_i^{(f)}$, respectively; $\vartheta_{i,n,x/y}^{(f)}$ denotes the effective angle of departure (AOD) along the x/y axis.

It is important to note that the delays and Doppler shifts undergo relatively gradual changes. Consequently, the rapidly changing TF domain channel parameters for SR systems can be accurately characterized using limited and quasi-static parameters in the DD domain, enabling more efficient channel estimation.

B. Modulation Strategies

1) Primary Information Modulation: For the OTFS systems, the symbols $x[k,l]$ for $k = 0, \ldots, N-1$ and $l = 0, \ldots, M-1$ are multiplexed across the entire DD plane $\mathit{\Gamma}$, forming the corresponding symbol frame $\boldsymbol{X} \in \mathbb{C}^{N \times M}$. This symbol frame $\boldsymbol{X}$ comprises both pilot symbols utilized for channel estimation and data symbols conveying information. Specifically, we employ a single pilot $x[k,l] = x_p$ at the grid $l = k = 0$ and full guards $x[k,l] = 0$ at the grids ($l \in [0, l_{\max}] \bigcup [M - l_{\max}, M - 1], k \in [0, N-1]$) to mitigate the interference from Doppler spreading [34], [50], where $l_{\max}$ denotes the maximum delay tap. The remaining grids carry informative symbols $x[k,l] = x_d[k,l]$ for primary transmission, adopting arbitrary *amplitude-phase modulation*. Consequently, the structure of each OTFS frame can be described as follows

$$x[k,l] = \begin{cases} x_p, l = 0, k = 0, \\ 0, l \in [0, l_{\max}] \bigcup [M - l_{\max}, M-1], k \in [0, N-1], \\ x_d[k,l], \text{otherwise}, \end{cases}$$

where $x_d[k,l]$ and x_p represent the data symbols and the pilot, respectively. To transmit DD domain symbols through wireless channels, the PE initially employs the inverse Symplectic finite Fourier transform (ISFFT), yielding the time-frequency domain symbols $\boldsymbol{X}_{\mathrm{TF}}$. Subsequently, $\boldsymbol{X}_{\mathrm{TF}}$ are proceeded with the Heisenberg transform and pulse shaping to obtain the time-domain signals $x(t)$, which are then transmitted via antennas.

2) Secondary Information Modulation: In this section, we outline the strategy for secondary information modulation within the framework of SR. As previously mentioned, the SE not only conveys its own information but also facilitates primary transmissions. Thus, the SE's or IRS's reflecting vector can be expressed as $x_{\mathrm{SE}}(\mathfrak{f}, t)\boldsymbol{\theta}$. Specifically, the time-varying component $x_{\mathrm{SE}}(\mathfrak{f}, t) = \sum_{s=1}^{S} e^{j2\pi f_s t}$ contains the secondary information, with this information embedded in the selection of the combination $\mathfrak{f} = \{f_1, \ldots, f_S\}$. Here, each f_s for $s = 1, \ldots, S$ belongs to a different frequency set $\mathcal{F}_s$ with a size of $|\mathcal{F}_s| = \mathcal{M}$. By properly designing all $\mathcal{F}_s$ jointly, $\mathfrak{f}$ can have a maximum of $\mathcal{M}^S$ possible selections. Since CF modulation enables different secondary frequency sets to reshape the Doppler structure of the equivalent channels in distinct ways, it allows the extraction of secondary information from tightly coupled and unknown CSI by analyzing its structure in the Doppler domain. Additionally, the frequency modulation of the secondary information is orthogonal to the amplitude-phase modulation of the primary information, ensuring independent transmission of both signals without interference. As for $\boldsymbol{\theta}$, given the time-frequency double-selective channels, the optimal IRS phase $\boldsymbol{\theta}$ may vary over time. However, the intricate coupling between time-varying $\boldsymbol{\theta}$ and DD domain channels renders the input-output relationship intractable, posing significant challenges for both analysis and optimal $\boldsymbol{\theta}$ design. As this paper primarily focuses on channel estimation and symbol detection for SR systems, the explicit design of IRS phase shifts $\boldsymbol{\theta}$ is beyond its scope and remains an endeavor for future work. For analytical simplicity in this work, we treat $\boldsymbol{\theta}$ as time-invariant and set $\mathrm{diag}(\boldsymbol{\theta}) = \boldsymbol{I}_{K_x K_y}$, providing a path gain of $K_x K_y$.

C. Received Signals

The time-domain signals $x(t)$ emitted by the PE initially reach the SE. The received signals at the SE can be expressed

---

[2] The proposed transmission methods can still be applied in a non-blockage scenario by first estimating the PE-receiver channel with the SE turned off, and then subtracting it from the equivalent CSI.

as

$$\boldsymbol{y}_{\mathrm{SE}}^{\mathrm{rec}}(t) = \sigma_1^{(g)} x(t-\tau_1^{(g)}) \boldsymbol{a}_{K_x}\left(\vartheta_{1,x}^{(g)}\right) \otimes \boldsymbol{a}_{K_y}\left(\vartheta_{1,y}^{(g)}\right). \tag{3}$$

Upon receiving $\boldsymbol{y}_{\mathrm{SE}}^{\mathrm{rec}}(t)$, the SE instantaneously modulates $x_{\mathrm{SE}}(\mathfrak{f}, t)\boldsymbol{\theta}$ onto it and reflects it without introducing any delay.

Therefore, the reflected signals at the SE can be expressed as[3]

$$\boldsymbol{y}_{\mathrm{SE}}^{\mathrm{ref}}(t) = \sigma_1^{(g)} x(t-\tau_1^{(g)}) x_{\mathrm{SE}}(\mathfrak{f}, t-\tau_1^{(g)}) \mathrm{diag}(\boldsymbol{\theta}) \times \boldsymbol{a}_{K_x}\left(\vartheta_{1,x}^{(g)}\right) \otimes \boldsymbol{a}_{K_y}\left(\vartheta_{1,y}^{(g)}\right). \tag{4}$$

Following transmission through the channel $\boldsymbol{f}$, the received signals at the receiver's end is given by

$$y_{\mathrm{R}}(t) = \sum_{\tau_i^{(f)} \in \mathcal{T}^{(f)}} \sum_{n=1}^{P_i^{(f)}} \Bigg\{ \sigma_{i,n}^{(f)} e^{j2\pi\nu_{i,n}^{(f)}(t-\tau_i^{(f)})} \times \left[\boldsymbol{a}_{K_x}^H\left(\vartheta_{i,n,x}^{(f)}\right) \otimes \boldsymbol{a}_{K_y}^H\left(\vartheta_{i,n,y}^{(f)}\right)\right] \boldsymbol{y}_{\mathrm{SE}}^{\mathrm{ref}}(t-\tau_i^{(f)}) \Bigg\}. \tag{5}$$

Combining Eq. (4) and Eq. (5), and after some straightforward manipulations, $y_{\mathrm{R}}(t)$ can be rewritten as

$$y_{\mathrm{R}}(t) = \sum_{\tau_i^{(f)} \in \mathcal{T}^{(f)}} \sum_{n=1}^{P_i^{(f)}} \sum_{s=1}^{S} \gamma_{i,n} e^{j2\pi(f_s+\nu_{i,n}^{(f)})(t-\tau_i)} x(t-\tau_i), \tag{6}$$

where

$$\tau_i = \tau_1^{(g)} + \tau_i^{(f)}.$$

$$\gamma_{i,n} = \sigma_1^{(g)} \sigma_{i,n}^{(f)} e^{j2\pi\nu_{i,n}^{(f)}\tau_1^{(g)}} r(\boldsymbol{\theta}, i, n),$$

$$r(\boldsymbol{\theta}, i, n) = \left[\boldsymbol{a}_{K_x}^H\left(\vartheta_{i,n,x}^{(f)}\right) \otimes \boldsymbol{a}_{K_y}^H\left(\vartheta_{i,n,y}^{(f)}\right)\right] \mathrm{diag}(\boldsymbol{\theta}) \times \left[\boldsymbol{a}_{K_x}\left(\vartheta_{1,x}^{(g)}\right) \otimes \boldsymbol{a}_{K_y}\left(\vartheta_{1,y}^{(g)}\right)\right],$$

Observing Eq. (6), it is evident that the received signals $y_{\mathrm{R}}(t)$ are equivalent to the scenario where the transmitted signals $x(t)$ traverses through an equivalent channel with a total of $\sum_{\tau_i^{(f)} \in \mathcal{T}^{(f)}} P_i^{(f)} S$ propagation paths, with each path contributes a complex gains, delays, and Doppler shifts characterized by $\gamma_{i,n}$, τ_i, and $\kappa_{i,n,s} = \left(f_s + \nu_{i,n}^{(f)}\right) NT$, respectively.

To facilitate signal processing, the received time-domain signals $y_{\mathrm{R}}(t)$ are subsequently converted to the DD domain $\boldsymbol{Y} \in \mathbb{C}^{N \times M}$ by applying the Wigner transform followed by the Symplectic finite Fourier transform (SFFT) [34]. Mathematically, the resulting DD domain symbols $y[k, l]$ can be represented as

$$y[k, l] = \sum_{\tau_i \in \mathcal{T}} \sum_{n=1}^{\mathcal{P}_i} \sum_{z=0}^{N-1} \sum_{s=1}^{S} \frac{\gamma_{i,n}}{N} e^{j2\pi \frac{(l-l_{\tau_i})\kappa_{i,n,s}}{MN}} \alpha(\kappa_{i,n,s} - z) \times x\left[[k-z]_N, [l-l_{\tau_i}]_M\right] + w[k, l], \tag{7}$$

with

$$\alpha(\kappa_{i,n,s} - z) = \sum_{q=0}^{N-1} e^{\frac{j2\pi\kappa_{i,n,s}q}{N}} e^{-\frac{j2\pi zq}{N}}, \tag{8}$$

where the $\mathcal{T} = \{\tau_i^{(f)} + \tau_1^{(g)} | \tau_i^{(f)} \in \mathcal{T}^{(f)}\}$ denotes the delay collections of the equivalent channels; $\mathcal{P}_i$ denotes the number of paths at delay τ_i; $w[k, l] \sim \mathcal{CN}(0, \sigma^2)$. To this end, combining Eq. (7) and Eq. (8) yields the formulation presented in Eq. (9), as shown at the bottom of the page.

Lemma 1: For clarity and convenience in analysis, we transform Eq. (9) into vector form as

$$\boldsymbol{y}_l = \frac{1}{N} \boldsymbol{A}^H \sum_{\tau_i \in \mathcal{T}} \sum_{n=1}^{\mathcal{P}_i} \mathrm{diag}(\boldsymbol{\chi}_{l-l_{\tau_i},n}) \boldsymbol{A} \boldsymbol{x}_{l-l_{\tau_i}} + \boldsymbol{w}_l, \tag{10}$$

where

$$\boldsymbol{\chi}_{l-l_{\tau_i},n} = \boldsymbol{x}_{\mathrm{SE}}\left(\mathfrak{f}, \frac{l-l_{\tau_i}}{M\Delta f}\right) \bullet \boldsymbol{h}_{l-l_{\tau_i},n} \tag{11}$$

with

$$\boldsymbol{x}_{\mathrm{SE}}\left(\mathfrak{f}, \frac{l-l_{\tau_i}}{M\Delta f}\right) = \left[x_{\mathrm{SE}}\left(\mathfrak{f}, \frac{l-l_{\tau_i}}{M\Delta f}\right), \cdots, x_{\mathrm{SE}}\left(\mathfrak{f}, \frac{l-l_{\tau_i}}{M\Delta f} + (N-1)T\right)\right],$$

and

$$\boldsymbol{h}_{l-l_{\tau_i},n} = \gamma_{i,n}\left[e^{j2\pi\nu_{i,n}^{(f)}\left(\frac{l-l_{\tau_i}}{M\Delta f}\right)}, \ldots, e^{j2\pi\nu_{i,n}^{(f)}\left(\frac{l-l_{\tau_i}}{M\Delta f}+(N-1)T\right)}\right];$$

$\boldsymbol{y}_l$ and $\boldsymbol{x}_l$ represent the l-th column of $\boldsymbol{Y}$ and $\boldsymbol{X}$; the (p, q) element of $\boldsymbol{A}$ is given by $A(p, q) = e^{\frac{j2\pi(p-1)(q-1)}{N}}$; $\boldsymbol{w}_l = [w[0, l], w[1, l], \ldots, w[N-1, l]]^T$;

Remark 1: Given the IRS's capability of reconfiguring the CSI, the term $\boldsymbol{\chi}_{l-l_{\tau_i},n}$ in Eq. (10) can be regarded as a decisive part of the equivalent channel, which is directly related to the SE informative signals $\boldsymbol{x}_{\mathrm{SE}}\left(\mathfrak{f}, \frac{l-l_{\tau_i}}{M\Delta f}\right)$ and cascaded channels $\boldsymbol{h}_{l-l_{\tau_i},n}$. Specifically, the secondary signal $\boldsymbol{x}_{\mathrm{SE}}\left(\mathfrak{f}, \frac{l-l_{\tau_i}}{M\Delta f}\right)$ can shift and expand the Doppler shifts of the original channels, resulting a strong coupling between the SE signal and the original CSI in the equivalent CSI $\boldsymbol{\chi}_{i,n,l_{\tau_i}}$. This coupling poses significant challenges in extracting secondary information from the unknown CSI.

Given the properly arranged OTFS symbol frame, the received signals corresponding solely to the pilots and guards can be utilized for equivalent-channel estimation, given by $\boldsymbol{y}_l$ for $0 \le l \le l_{\max}$. To determine the value of τ_i, we can compare the energy of those $\boldsymbol{y}_l$ with a predetermined threshold, as described in [34]. Specifically, those $\boldsymbol{y}_l$ corresponding to l_{τ_i} can be derived as

$$\boldsymbol{y}_{l_{\tau_i}} = \sum_{n=1}^{\mathcal{P}_i} \frac{1}{N} \boldsymbol{A}^H \boldsymbol{\chi}_{0,n} x_p + \boldsymbol{w}_{l_{\tau_i}}. \tag{12}$$

$$y[k, l] = \sum_{\tau_i \in \mathcal{T}} \sum_{n=1}^{\mathcal{P}_i} \sum_{z=0}^{N-1} \sum_{q=0}^{N-1} \frac{\gamma_{i,n}}{N} e^{j2\pi\nu_{i,n}^{(f)}\left(\frac{l-l_{\tau_i}}{M\Delta f}+qT\right)} x_{\mathrm{SE}}\left(\mathfrak{f}, \frac{l-l_{\tau_i}}{M\Delta f} + qT\right) e^{-\frac{j2\pi zq}{N}} x\left[[k-z]_N, [l-l_{\tau_i}]_M\right] + w[k, l]. \tag{9}$$

---

[3] The reflection delays introduced by the IRS are negligible compared to the propagation delays.

Corollary 1: Due to the fractional Doppler shifts, the DD-domain received signals $\boldsymbol{y}_{l_{\tau_i}}$ experiences significant interference from Doppler spreading effects. Therefore, taking the inverse discrete Fourier transform (IDFT) of Eq. (12) yields the time-delay (TD) domain signal as

$$z_{l_{\tau_i}} = \boldsymbol{A}\boldsymbol{y}_{l_{\tau_i}} = \sum_{n=1}^{\mathcal{P}_i} x_p \boldsymbol{\chi}_{0,n} + \boldsymbol{w}'_{l_{\tau_i}}, \tag{13}$$

where $\boldsymbol{w}'_{l_{\tau_i}} = \boldsymbol{A}\boldsymbol{w}_{l_{\tau_i}}$ represents the resulting noise, following $\boldsymbol{w}'_{l_{\tau_i}} \sim \mathcal{CN}(\boldsymbol{0}, \tilde{\sigma}^2\boldsymbol{I}_N)$ with $\tilde{\sigma}^2 = N\sigma^2$.

Drawing upon Corollary 1, the estimation of the equivalent channel $\boldsymbol{h}^{\mathrm{Eq}} = \sum_{n=1}^{\mathcal{P}_i} \boldsymbol{\chi}_{0,n}$ can be viewed as a sparse signal recovery problem with a single snapshot and $\mathcal{P}_i S$ non-zero elements. Nonetheless, detecting SE information from the unknown original channels poses challenges. To tackle this, we first propose an off-grid SBL method to acquire the equivalent CSI, and then exploit the structured nature of the equivalent Doppler shifts to extract the SE information. Acknowledging the limitations of traditional off-grid SBL-based methods, we subsequently introduce a deep learning (DL)-based approach to improve communication performance.[4]

III. OFF-GRID SBL-BASED CSI ESTIMATION AND DETECTION

In this section, we address the aforementioned problem as a frequency estimation problem and introduce an off-grid SBL-based algorithm. We begin by transforming Eq. (13) into

$$z_{l_{\tau_i}} = \boldsymbol{E}_i\boldsymbol{\gamma}_i + \boldsymbol{w}'_{l_{\tau_i}}, \tag{14}$$

where the columns of $\boldsymbol{E}_i$ are defined by $\boldsymbol{e}(\kappa_{i,n,s}) = \left[1, e^{j2\pi\kappa_{i,n,s}/N}, \ldots e^{j2\pi\kappa_{i,n,s}(N-1)/N}\right]^T$ for all $n = 1, \ldots, N$ and $s = 1, \ldots, S$. Here, $\boldsymbol{\gamma}_i$ contains the amplitudes for all $\kappa_{i,n,s}$, with accordingly amplitude given by $x_p\gamma_{i,n}$.

A. *Off-Grid SBL Method*

1) Off-Grid Model Formulation: Given the fact that $\kappa_{i,n,s} \in [0, N)$, adhering to the principle of SBL, we partition the interval $[0, N)$ into G discrete grid points with $\Phi_g = \frac{g}{G}N$, $g = 0, \ldots, G-1$, which forms a dictionary matrix $\boldsymbol{\Xi} \in \mathbb{C}^{N\times G}$ with the g-th column defined by $\boldsymbol{e}(\Phi_g)$. Based on the dictionary matrix $\boldsymbol{\Xi}$, we obtain the following sparse-approximate received signals

$$z_{l_{\tau_i}} \approx \boldsymbol{\Xi}\boldsymbol{\Upsilon}_i + \boldsymbol{w}'_{l_{\tau_i}}, \tag{15}$$

where $\boldsymbol{\Upsilon}_i \in \mathbb{C}^{G\times 1}$ denotes the corresponding sparse vector. However, in a finite-resolution grid system, the actual angles may inevitably not perfectly align with the discrete grid points, resulting in significant errors. To mitigate these quantization errors, we employ a more accurate expression for the received signal based on the first-order Taylor expansion as

$$z_{l_{\tau_i}} \approx [\boldsymbol{\Xi} + \boldsymbol{\xi}\mathrm{diag}(\boldsymbol{\Delta})]\boldsymbol{\Upsilon}_i + \boldsymbol{w}'_{l_{\tau_i}}, \tag{16}$$

where $\boldsymbol{\xi} \in \mathbb{C}^{N\times G}$ denotes the first derivative of $\boldsymbol{\Xi}$, with its g-th column given by $[0, \ldots j2\pi(n-1)/N, \cdots j2\pi(N-1)/N]^T \bullet \boldsymbol{e}(\Phi_g)$. Here, $\boldsymbol{\Delta} \in \mathbb{C}^{G\times 1}$ represents the bias between the true value of $\kappa_{i,n,s}$ and the grids.

2) Hierarchical Hyperprior Model: Given that the noise follows a circularly symmetric complex Gaussian process, the conditional distribution of the received signal $\boldsymbol{z}_{l_{\tau_i}}$ is expressed as

$$\Pr(z_{l_{\tau_i}}|\gamma_i, \boldsymbol{\Delta}) = \mathcal{CN}\left(z_{l_{\tau_i}}|\boldsymbol{\Psi}\boldsymbol{\Upsilon}_i, \tilde{\sigma}^2\boldsymbol{I}_N\right), \tag{17}$$

where $\boldsymbol{\Psi} = \boldsymbol{\Xi} + \boldsymbol{\xi}\mathrm{diag}(\boldsymbol{\Delta})$.

To capture the sparsity of $\boldsymbol{\Upsilon}_i$, we formulate its probability density function (PDF) as

$$\Pr(\boldsymbol{\Upsilon}_i|\boldsymbol{\zeta}_i) = \mathcal{CN}\left(\boldsymbol{\Upsilon}_i|0, \mathrm{diag}(\boldsymbol{\zeta}_i)\right), \tag{18}$$

where $\boldsymbol{\zeta}_i = [\zeta_i(1), \zeta_i(2), \cdots, \zeta_i(G)]^T$ denotes the hyper-parameter describing the sparsity of $\boldsymbol{\Upsilon}_i$. To mitigate the impact of random initialization of $\boldsymbol{\zeta}_i$ on the estimation performance, we further assume the following hyper distribution for each element

$$\Pr(\boldsymbol{\zeta}_i; \rho) = \prod_{g=1}^{G} \Gamma\left(\zeta_i(g)|1, \rho\right), \tag{19}$$

where ρ is a fixed value. Consequently, this two-stage hierarchical prior encourages the sparsity of $\boldsymbol{\Upsilon}_i$, thereby facilitating efficient representation and inference processes.

The vector $\boldsymbol{\Delta}$ lies within the quantization bounds and is assumed to follow a uniform distribution given by

$$\boldsymbol{\Delta} \sim \mathcal{U}\left[-\frac{N}{2G}, \frac{N}{2G}\right]^G. \tag{20}$$

3) Bayesian Inference: Obtaining the exact distribution of $\Pr(\boldsymbol{\Upsilon}_i, \boldsymbol{\zeta}_i, \boldsymbol{\Delta}|\boldsymbol{z}_{l_{\tau_i}})$ is challenging. Leveraging the analytical framework outlined in [51], we instantiate a recursive evidentiary scheme within our Bayesian inferential process. Initially, we establish that the conditional posterior distribution of $\boldsymbol{\Upsilon}_i$ follows a complex Gaussian distribution, as given by

$$\Pr(\boldsymbol{\Upsilon}_i|z_{l_{\tau_i}}, \boldsymbol{\zeta}_i, \boldsymbol{\Delta}) = \mathcal{CN}\left(\boldsymbol{\Upsilon}_i|\boldsymbol{\vartheta}_i, \boldsymbol{\Omega}_i\right), \tag{21}$$

where the mean and variance are determined by

$$\boldsymbol{\vartheta}_i = \tilde{\sigma}^{-2}\boldsymbol{\Omega}_i\boldsymbol{\Xi}^H z_{l_{\tau_i}}, \tag{22}$$

and

$$\boldsymbol{\Omega}_i = \left(\tilde{\sigma}^{-2}\boldsymbol{\Xi}^H\boldsymbol{\Xi} + \mathrm{diag}(\boldsymbol{\zeta}_i)^{-1}\right)^{-1}, \tag{23}$$

respectively, following the relationship

$$\boldsymbol{\zeta}_i = |\boldsymbol{\vartheta}_i|^2 + \mathrm{diag}(\boldsymbol{\Omega}_i). \tag{24}$$

Moreover, the hyper-parameters $\boldsymbol{\zeta}_i$ can be obtained via maximum a posterior (MAP) principle by maximizing the posterior distribution $\Pr(\boldsymbol{\zeta}_i, \boldsymbol{\Delta} \mid z_{l_{\tau_i}})$. It is noteworthy that maximizing $\Pr(\boldsymbol{\zeta_i}, \boldsymbol{\Delta} \mid \boldsymbol{z}_{l_{\tau_i}})$ is equivalent to maximize $\Pr(\boldsymbol{z}_{l_{\tau_i}}, \boldsymbol{\zeta_i}, \boldsymbol{\Delta})$ since $\Pr(\boldsymbol{z}_{l_{\tau_i}})$ is independent of the hyper-parameters. Hence, we first treat $\boldsymbol{\Upsilon}_i$ as a latent variable, and $\boldsymbol{\zeta}_i$ can be obtained via

$$\boldsymbol{\zeta_i}^{\mathrm{new}} = \arg\max_{\boldsymbol{\zeta_i}} \mathbb{E}_{\boldsymbol{\Upsilon}_i|z_{l_{\tau_i}};\boldsymbol{\zeta_i},\boldsymbol{\Delta}}\{\log\Pr(z_{l_{\tau_i}}, \boldsymbol{\zeta_i}, \boldsymbol{\Delta}, \boldsymbol{\Upsilon}_i)\} \tag{25}$$

---

[4] It is important to note that based on the obtained $\boldsymbol{\chi}_{l-l_{\tau_i},n}$, the detection of PE information becomes a classical detection problem, which can be tackled using various existing detectors [37], [38], [39].

by utilizing the expectation-maximization (EM) algorithm, then the closed-form solution for the g-th element of $\boldsymbol{\zeta}_i$ can be obtained as

$$\zeta_i(g)=\frac{\sqrt{1+4\rho\mathbb{E}\{\|\Upsilon_i(g)\|_2^2\}}-1}{2\rho}, g=0,\dots G-1, \tag{26}$$

where $\Upsilon_i(g)$ represents the g-th element of $\boldsymbol{\Upsilon}_i$ [52].

Regarding the estimation of $\boldsymbol{\Delta}$, it is equivalent to solve the following problem

$$\text{(P1)}\quad \boldsymbol{\Delta}^{\text{new}}=\arg\min_{\boldsymbol{\Delta}\in\left[-\frac{1}{2G},\frac{1}{2G}\right]^G}\boldsymbol{\Delta}^T\mathbb{P}\boldsymbol{\Delta}+2\mathbb{V}^T\boldsymbol{\Delta}, \tag{27}$$

where $\mathbb{P}=\mathcal{R}\Big\{\boldsymbol{\xi}^H\boldsymbol{\xi}\odot(\boldsymbol{\vartheta}_i\boldsymbol{\vartheta}_i^H+\boldsymbol{\Omega}_i)\Big\}$ and $\mathbb{V}=\mathcal{R}\left\{\text{diag}(\boldsymbol{\xi}^H\boldsymbol{\Psi}\boldsymbol{\Omega}_i)\right\}-\mathcal{R}\Big\{\text{diag}(\boldsymbol{\vartheta}_i\boldsymbol{\xi}^H\left[\boldsymbol{z}_{l_{\tau_i}}-\boldsymbol{\Xi}\boldsymbol{\vartheta}_i\right]\Big\}$. If the matrix $\mathbb{P}$ is invertible and $\widetilde{\boldsymbol{\Delta}}=\mathbb{P}^{-1}\mathbb{V}\in\left[-\frac{1}{2G},\frac{1}{2G}\right]$, the optimal solution for (P1) is $\boldsymbol{\Delta}^{\text{new}}=\widetilde{\boldsymbol{\Delta}}$. Otherwise, $\boldsymbol{\Delta}^{\text{new}}$ should be obtained element-wise. Specifically, when update the g-th element $\Delta^{\text{new}}(g)$, we fix the other entries of $\boldsymbol{\Delta}^{\text{new}}$ and compute $\widetilde{\Delta}(g)=\frac{\vartheta(g)-\mathbb{P}^T_{-g}(\boldsymbol{\Delta}^{\text{new}})_{-g}}{\mathbb{P}(g,g)}$ for $g=1,\dots,G$, where $\boldsymbol{x}_{-g}$ represents the vector $\boldsymbol{x}$ without the g-th element. Then, combining the constraint $\widetilde{\Delta}(g)\in\left[-\frac{1}{2G},\frac{1}{2G}\right]$, the solution for $\Delta^{\text{new}}(g)$ can be obtained as

$$\Delta^{\text{new}}(g)=\begin{cases}\widetilde{\Delta}(g), & \text{if}\quad \widetilde{\Delta}(g)\in\left[-\frac{1}{2G},\frac{1}{2G}\right],\\ -\dfrac{1}{2G}, & \text{if}\quad \widetilde{\Delta}(g)<-\frac{1}{2G},\\ \dfrac{1}{2G}, & \text{if}\quad \widetilde{\Delta}(g)>\frac{1}{2G}.\end{cases} \tag{28}$$

To mitigate the error stemming from the first-order approximation on uniformly distributed grids, we propose updating the grids using the estimated values, which can definitely enhance the estimation performance [53].

B. Equivalent CSI Estimation

The above process proceeds iteratively until convergence is achieved. Based on the obtained estimate $\widehat{\boldsymbol{\delta}}=\boldsymbol{\delta}^{\text{new}}$ and $\widehat{\boldsymbol{\Delta}}=\boldsymbol{\Delta}^{\text{new}}$, the MAP estimate of $\boldsymbol{\Upsilon}_i$ is determined by the mean of its posterior distribution

$$\boldsymbol{\Upsilon}_i^{\text{MAP}}=\boldsymbol{\vartheta}_i. \tag{29}$$

Based on relationships in Eq. (14) and Eq.(16), by combining the obtained $\boldsymbol{\Upsilon}_i^{\text{MAP}}$ and $\widehat{\Delta}$, the equivalent CSI based on the traditional off-grid SBL algorithm can be obtained as

$$\widehat{\boldsymbol{h}}_{\text{TR}}^{\text{Eq}}\approx\frac{[\boldsymbol{\Xi}+\boldsymbol{\xi}\text{diag}(\widehat{\boldsymbol{\Delta}})]\boldsymbol{\Upsilon}_i^{\text{MAP}}}{x_p}. \tag{30}$$

According to Eq. (30), the accuracy of the estimated equivalent CSI is influenced by the quantization precision of the grid points G and the signal sparsity. Specifically, as the grid becomes finer (i.e., larger G), the error in the first-order approximation $\boldsymbol{\Xi}+\boldsymbol{\xi}\text{diag}(\widehat{\boldsymbol{\Delta}})$ decreases, leading to improved estimation accuracy. On the other hand, an increase in the number of paths at a given delay tap (i.e., $\mathcal{P}_i$) reduces the signal sparsity, making it more difficult to distinguish between channels from different paths (i.e., the non-zero elements in Υ_i^{MAP}), which in turn decreases the accuracy of the off-grid SBL CSI estimator. Additionally, increasing the number of observations N (i.e., the number of samples for each frequency/direction) also improves the estimation accuracy.

C. Secondary Symbol Detection

It has been elucidated that the secondary modulated signals induce frequency expansion of the original Doppler shifts due to their channel reconfiguring properties. However, the challenge lies in distinguishing secondary information $\mathfrak{f}$ from the unknown original CSI. Firstly, the estimated parameters related to the Doppler shifts in the equivalent CSI are given by

$$\widehat{\kappa_{i,n,s}}=\Phi_{\mathcal{I}}+\widehat{\Delta}(\mathcal{I}), \tag{31}$$

where $\mathcal{I}$ is the index of the largest $\mathcal{P}_iS$ energy values in the vector $\boldsymbol{\zeta}_i$. Notably, all $\widehat{\kappa_{i,n,s}}$ typically fall within the range $[0,\kappa_{\max}]$, where $\kappa_{\max}=\frac{f_c v_{\max}N}{c\Delta f}$, with f_c being the carrier frequency, $v_{\max}$ representing the maximum effective speed in m/s, and c denoting the speed of light [54]. In scenarios with ultra-wideband or/and short package, where Δf is large or N is relatively low, the range of $\widehat{\kappa_{i,n,s}}$ is very limited, indicating that the Doppler shifts of all paths are within a relatively small range. If we can properly design $\mathcal{F}_s$ for all $s=1,\dots,S$ to ensure that the gap between any two elements in $\mathfrak{f}$ is different and larger than the feasible range $\nu_{\max}=\frac{\kappa_{\max}}{NT}$, then we can guarantee that all $\nu_{i,n}^{(f)}$ fall within S separable clusters. In this case, we can obtain $\mathfrak{f}$ without the information of $\nu_{i,n}^{(f)}$ by distinguishing the gap between clusters. For example, when $\nu_{i,n}^{(f)}\in[0,0.2]$, $S=2$, $f_1\in\{0.25,0.35\}$, and $f_2\in\{0.55,0.75\}$. Denote the sorted $\widehat{\kappa_{i,n,s}}$ as $\overline{\boldsymbol{\phi}}$ and equally divide $\overline{\boldsymbol{\phi}}$ into S parts, denoted as $\overline{\boldsymbol{\phi}_s}$ for $s=1,\dots S$. We then have

$$|f_1-f_2|=\frac{\|\overline{\boldsymbol{\phi}}_1-\overline{\boldsymbol{\phi}}_2\|_1}{NT\mathcal{P}_i}. \tag{32}$$

By employing carefully designed $\mathcal{F}_s$ for all $s=1,\dots,S$, we can map the obtained $|f_1-f_2|$ with $\mathfrak{f}$ via minimum Euclidean distance matching. In addition, since the secondary information modulation has the same effect on $\boldsymbol{z}_{l_{\tau_i}}$ for all $\tau_i\in\mathcal{T}$, using multiple sets of $\boldsymbol{z}_{l_{\tau_i}}$ may help improve the detection accuracy.

The method described above is limited to certain scenarios. In cases where the number of paths $\mathcal{P}_i$ is large, effectively separating dense frequencies from narrow range $[0,\kappa_{\max}]$ using the off-grid SBL method becomes challenging [55], leading to poor estimation performance. Additionally, the accuracy of CSI estimation significantly influences the detection of SE symbols, further compromising detection performance. Moreover, if the value of $\kappa_{\max}$ is relatively large, Eq. (32) cannot be reliably applied for SE symbol detection due to spectrum overlapping. To address these challenges, we propose the following DL-based estimation and detection methods.[4]

IV. DL-Based CSI Estimation and Detection

In the preceding analysis, it becomes evident that traditional methods encounter challenges in obtaining equivalent CSI and secondary information in general scenarios. Leveraging the learning capacity of deep neural networks, we expect

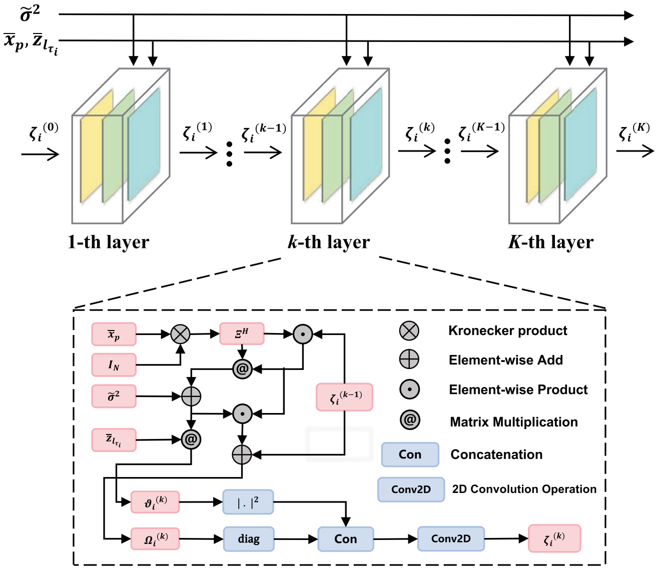


Fig. 2. The structure of the ECSIEst-Net.

to capture the spectrum characteristics of the equivalent CSI more accurately, even with dense or overlapping frequencies. Moreover, this approach might allow for the extraction of SE information without relying on prior CSI, facilitating accurate and versatile detection performance. In this section, our aim is to obtain CSI in Eq. (14) and disentangle the intertwined secondary information $\mathfrak{f}$ by proposing the ECSIEst-Net and SSymDet-Net, respectively.

A. Equivalent CSI Estimation: ECSIEst-Net

Note that the first-order sparse approximation of the received signals in Eq. (16) can be inaccurate due to the limited number of grids. This inaccuracy causes power spreading along adjacent grid points, leading to error propagation in the subsequent analysis. Moreover, as aforementioned, computing the distribution of $\Pr(\boldsymbol{\Upsilon}_i, \boldsymbol{\delta}, \boldsymbol{\Delta} | \boldsymbol{z}_{l_{\tau_i}})$ presents significant challenges. Consequently, we analyze the distribution of $\Pr(\boldsymbol{\Upsilon}_i | \boldsymbol{z}_{l_{\tau_i}}, \boldsymbol{\delta}, \boldsymbol{\Delta})$ as a compromise and iteratively update $\boldsymbol{\zeta}_i$. This approximation may lead to suboptimal updates. To address these issues, we propose the ECSIEst-Net by unfolding the SBL algorithm with numerous iterations into a deep neural network (DNN) with a moderate number of layers. This approach may help mitigate the errors arising from the aforementioned derivation and analysis. Note that, since the Doppler shifts of the original channels with different delays vary, we need to estimate the equivalent CSI separately for each delay using corresponding $\boldsymbol{z}_{l_{\tau_i}}$.

1) The Structure of ECSIEst-Net: For the proposed ECSIEst-Net consisting of K layers, the value of $\boldsymbol{\zeta}_i$ in each layer is obtained through a well-designed DNN, rather than using Eq. (26) in SBL framework. Specifically, as shown in Fig. 2, in the k-th layer, the inputs consist of the re-parameterized pilot symbol $\bar{\boldsymbol{x}}_p = \begin{bmatrix} \Re(x_p) \\ \Im(x_p) \end{bmatrix}$, the re-parameterized received signals $\bar{\boldsymbol{z}}_{l_{\tau_i}} = \begin{bmatrix} \Re(\boldsymbol{z}_{l_{\tau_i}}) \\ \Im(\boldsymbol{z}_{l_{\tau_i}}) \end{bmatrix}$, the variance of the noise $\tilde{\sigma}^2$, and the outputs of the previous layer $\boldsymbol{\zeta}_i^{(k-1)}$. The outputs of the k-th layer is the updated $\boldsymbol{\zeta}_i^{(k)}$. To establish a mapping function from inputs to outputs, namely, $\boldsymbol{\zeta}_i^{(k)} = f\left(\bar{\boldsymbol{x}}_p, \bar{\boldsymbol{z}}_{l_{\tau_i}}, \tilde{\sigma}^2, \boldsymbol{\zeta}_i^{(k-1)}\right)$, we propose using a DNN to learn the parameters that fit the function $f(\cdot)$. Note that the initial inputs for $\boldsymbol{\zeta}_i^{(0)}$ is an all-one vector. It is expected that the DNN can better capture the spectral features of the received signals, thereby achieving a more accurate mapping between $\boldsymbol{\zeta}_i$ and the input parameters, despite the first-order sparse approximation. Based on the output of the ECSIEst-Net, the equivalent CSI using DL-based algorithm can be obtained as $\widehat{\boldsymbol{h}}_{\mathrm{DL}}^{\mathrm{Eq}} \approx \boldsymbol{\Xi}\boldsymbol{\vartheta}_i^{(K)}/x_p$, where the hyperparameter $\boldsymbol{\vartheta}_i^{(K)}$ is obtained using Eq. (22) $\sim$ (24).

2) The Architecture of Each Layer: For each layer, a consistent structure shown in Fig. 2 is maintained to ensure a stable learning process throughout the DNN. Since each frequency in the secondary information $\mathfrak{f}$ shifts the original Doppler cluster with a common offset, resulting in a highly structured and correlated spectrum in the equivalent CSI, a two-dimensional convolution layer is employed to effectively capture the spectrum features through convolution operations. As the convolution kernel traverses the dimensions of the spectrum, it enables the CNN to learn and represent intricate patterns inherent in the input data, aiding in the precise estimation of channel characteristics. To preserve the dimensions of the tensor after the convolution operation, zero-padding (ZP) and circular-padding (CP) are applied to the feature map before convolution.

Algorithm 1 The Training Process of ECSIEst-Net

```
Input: The pilot symbol $\bar{\boldsymbol{x}}_p$, received signal $\bar{\boldsymbol{z}}_{l_{\tau_i}}$, noise
       variance $\tilde{\sigma}^2$, variance parameter matrix $\boldsymbol{\zeta}_i^{(0)}$
Output: Updated $\boldsymbol{\zeta}^{(K)}$
Initialization: Set the number of network layers K,
maximum number of epochs Epoch, batchsize;
for epoch from 1 to Epoch do
    for layer $k$ from 1 to $K$ do
        Update $\boldsymbol{\vartheta}_i^{(k)}, \boldsymbol{\Omega}_i^{(k)}$ according to $\boldsymbol{\zeta}_i^{(k-1)}$;
        Concatenate new feature matrix;
        With forward propagation through the network to
        obtain the updated $\boldsymbol{\zeta}_i^{(k)}$;
    end
    Calculate the validation loss;
    Computing gradients through backpropagation;
    Updating network weights;
end
```

3) Training Policy: Unlike data-driven methods, which require statistical analysis to learn the distribution and characteristics of data from a substantial dataset, a model-driven neural network training typically does not heavily rely on large amounts of labeled data, since the analysis process can be well exploited to assist the training process. Details on training the ECSIEst-Net are illustrated in Algorithm 1. The network parameters are updated using backpropagation to minimize the loss function, where we use mean squared error (MSE) of the equivalent channels as the loss function, given by

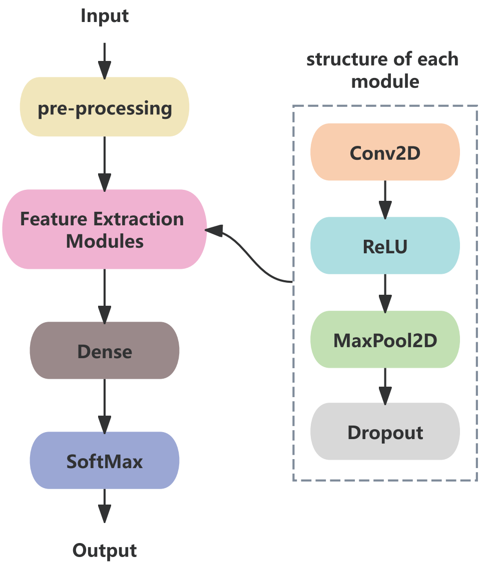


Fig. 3. The structure of the SSymDet-Net.

$\mathcal{L}_{\text{Est}} = |\widehat{\boldsymbol{h}}_{\text{DL}}^{\text{Eq}} - \boldsymbol{h}^{\text{Eq}}|^2$. To prevent overfitting during training, we implement an early stopping strategy that monitors the model's performance on the validation set. Additionally, learning rate scheduling is employed to avoid falling into local optima and facilitate finer parameter tuning in the later stages of training.

B. Secondary Symbol Detection: SSymDet-Net

Note that using the obtained equivalent CSI to detect the secondary symbol can result in significant error propagation and performance loss in certain scenarios. To enhance the robustness of the detector, we propose a data-driven SSymDet-Net to extract the secondary symbol without any CSI.

1) The Structure of SSymDet-Net: Given that all possible sets of $\mathfrak{f}$ are predefined, the detection of SE symbols becomes a multi-label classification problem within $\mathcal{M}^S$ classes. In this paper, we tackle this problem by employing a CNN-based detection network, termed SSymDet-Net, as illustrated in Fig. 3, which aims to extract the spectrum features by alternately connecting convolutional layers and pooling layers. Specifically, during the pre-processing stage, combining all $\boldsymbol{z}_{l_{\tau_i}}$ with $\tau_i \in \mathcal{T}$ holds potential for enhancing the accuracy of secondary symbol detection. This is because the IRS frequencies for all delay taps are identical, implying that we can exploit all $\boldsymbol{z}_{l_{\tau_i}}$ to determine the IRS common frequency set $\mathfrak{f}$. With this in mind, we rearrange all sets of $\boldsymbol{z}_{l_{\tau_i}}$ as $\boldsymbol{Z}_{l_{\tau_i}} \in \mathbb{C}^{N \times |\mathcal{T}|}$, where $|\mathcal{T}|$ denotes the size of set $\mathcal{T}$, presenting the number of delay taps in the channels. Since the complex values are difficult to process in CNNs, we present the complex values as a two-channel tensor, with one channel containing the real part and the other channel containing the imaginary part, given by

$$\bar{\boldsymbol{Z}}_{l_{\tau_i}} = \begin{bmatrix} \Re(\boldsymbol{Z}_{l_{\tau_i}}) \\ \Im(\boldsymbol{Z}_{l_{\tau_i}}) \end{bmatrix}. \tag{33}$$

Following the pre-processing stage, the inputs $\bar{\boldsymbol{Z}}_{l_{\tau_i}}$ and $\bar{\boldsymbol{x}}_p$ are then processed through multiple feature extraction modules. Within each module, a 2D convolution layer is firstly employed to extract local attributes and abstract representations from the input sequences. To introduce non-linearities into the model, a rectified linear unit (ReLU) activation function is then applied, given by $f(x) = \max(0, x)$. Following this, a 2D max pooling is utilized for down-sampling by selecting the maximum value within each local region, thereby retaining the most significant features of the signals. Finally, dropout layers are carefully included to alleviate overfitting. The successive feature extraction modules yield feature maps of the signals, which are then flattened and fed into the fully connected layer (or dense layer) for integration. To obtain a probability distribution across result categories, the raw output of the dense layer undergoes a transformation using the softmax activation function, defined as

$$\hat{p_i} = \frac{e^{x_i}}{\sum_{j=1}^{N} e^{x_j}}, \tag{34}$$

where x_i represents the *i*-th element in the output tensor, while $\hat{p_i}$ signifies the predicted probability of the result falling into the *i*-th category, enabling the prediction of the most probable outcome.

2) Training Policy: The SSymDet-Net is trained offline, which improves computation efficiency for online detection and enhances the stability of the trained model. Throughout the training process, trainable parameters in $\boldsymbol{\theta}$ are iteratively adjusted to minimize the discrepancy between true and predicted values over numerous epochs. Specifically, in the *t*-th training epoch, the updating rule for $\boldsymbol{\theta}$ can be expressed as

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \eta \cdot \frac{\partial \mathcal{L}_{\text{Det}}(\boldsymbol{\theta}_t)}{\partial \boldsymbol{\theta}_t}, \tag{35}$$

where η represents the learning rate along the gradient direction $\frac{\partial \mathcal{L}_{\text{Det}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_t}$. Building on this, we introduce the Adam optimizer for its adaptive rate update feature, which enables dynamic adjustment of the learning rate for each parameter, thereby facilitating accelerated convergence during training.

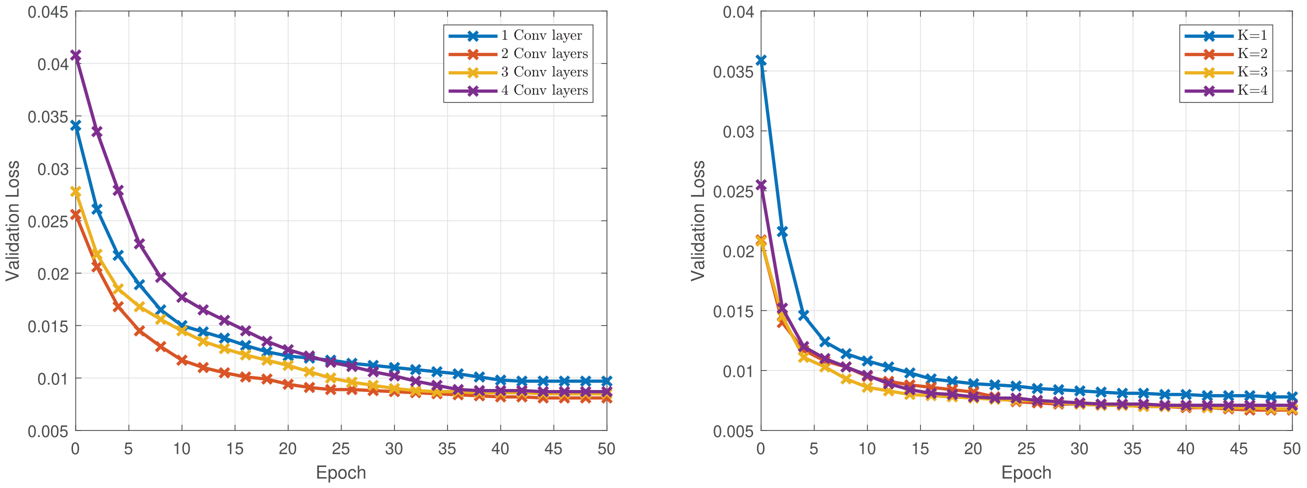
To evaluate the fitting performance, we employ a cross-entropy loss function tailored for multi-label classification tasks as the evaluation metric, given by

$$\mathcal{L}_{\text{Det}} = -\sum_{i} p_i \log\left(\hat{p_i}\right), \tag{36}$$

where p_i represents the probability of the *i*-th category in true labels. In each epoch, the loss value of $\mathcal{L}_{\text{Det}}$ is computed and propagated backward through the network.

V. NUMERICAL RESULTS

In this section, we present numerical results to validate the performance of CSI estimators and secondary symbol detectors. Unless stated otherwise, we use the following default settings: The complex gains of the channel $\boldsymbol{g}$ follows the distribution $\sigma_1^{(g)} \sim \mathcal{CN}(0, \boldsymbol{I}_{K_x K_y})$ and those of $\boldsymbol{f}$ follows the distribution $\sigma_{i,n}^{(f)} \sim \mathcal{CN}(0, 0.01 \times \boldsymbol{I}_{K_x K_y})$. The SNR is defined as SNR=$|x_p|^2/\widetilde{\sigma}^2$. For IRS frequency modulation, we assume $S = 2$ with frequency sets $f_1 \in \{0.5, 0.3\}$ and



(a) Impact of the convolution layers on the validation loss.

(b) Impact of K on the validation loss.

Fig. 4. Impact of convolution layers and K on the validation loss.

TABLE I
DEFAULT SIMULATION PARAMETERS

Parameter	Value
Carrier frequency	28 GHz
Bandwidth	20 MHz
Subcarrier spacing	60 kHz
Delay gridsize	333
Dimension of Doppler grid N	10
Size of IRS ($K_x \times K_y$)	10 × 10
Maximum number of delay taps $l_{\max}$	10
Maximum paths per delay tap $\mathcal{P}_i$	3
Epoch	50
Batch size	512
Learning rate	$1e^{-4}$
Weight decay	$5e^{-4}$

$f_2 \in \{0.2, 0.16\}$. The other default parameter settings are listed in Table I. The total dataset is divided into training, validation, and testing sets in a 7:2:1 ratio. First, the model is trained offline to determine the parameters/weights. Then, the validation set is used to assess the performance of the trained models. Finally, the selected optimal model is applied to the testing set to evaluate its performance.

A. Estimation Performance of the Equivalent CSI

Fig. 4 depicts the impact of the number of convolution layers within each ECSIEst-Net layer, as well as the number of layers in ECSIEst-Net denoted as K, on the validation loss. Since a single-layer convolution operation or network cannot fully capture data patterns, it results in slow convergence and suboptimal performance. Conversely, indiscriminate layer stacking may lead to performance saturation, merely escalating the computational complexity of the model without bringing significant performance improvements. Considering the tradeoff between performance and computational complexity of the network, we find that employing 2 convolution layers within each ECSIEst-Net layer and setting $K = 3$ would be a good choice, according to the results shown in Fig. 4(a) and Fig. 4(b).

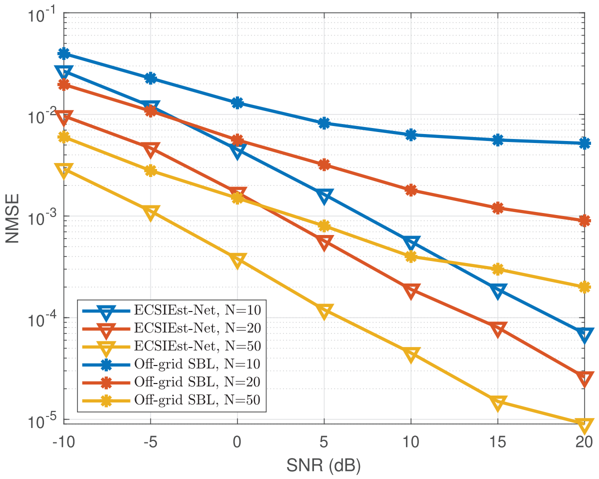


Fig. 5. Impact of N on NMSE performance.

Fig. 5 illustrates the impact of N on the CSI estimation performance. In this paper, we utilize the normalized mean squared error (NMSE) as the evaluation matric, which is given by

$$\mathrm{NMSE} = \mathbb{E}\left\{ \frac{\|\widehat{\boldsymbol{h}}^{\mathrm{Eq}}_{\mathrm{TR/DL}} - \boldsymbol{h}^{\mathrm{Eq}}\|^2}{N\mathcal{P}_i S \|\boldsymbol{h}^{\mathrm{Eq}}\|^2} \right\}.$$

To avoid the impact of the number of non-zero elements, the NMSE is normalized by $N\mathcal{P}_i S$. The off-grid SBL algorithm has historically provided a high-accuracy solution in low-mobility scenarios or sparse channels, however, its performance can degrade in more complex environments. As can be readily observed, the NMSE performance of the ECSIEst-Net surpasses that of the off-grid SBL method under identical parameters, suggesting the ECSIEst-Net's superior capability in distinguishing dense and correlated spectrum clusters in the Doppler domain. Moreover, comparing each approach with

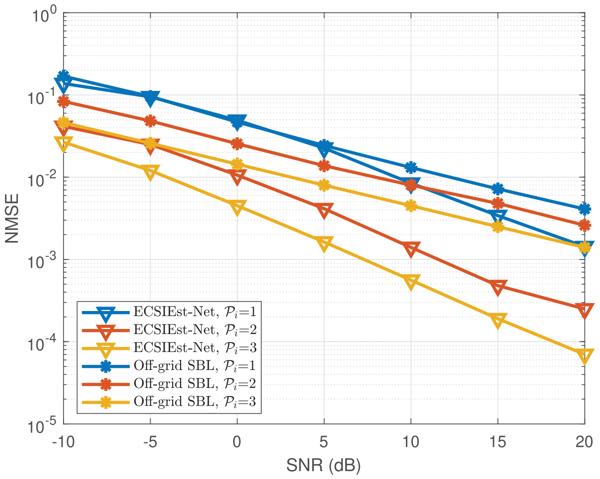


Fig. 6. Impact of $\mathcal{P}_i$ on NMSE performance.

varying values of N, we observe that increasing N consistently enhances CSI estimation. This improvement stems from the larger signal samples, which effectively characterize signal features and result in superior estimation performance.

Fig. 6 demonstrates the impact of $\mathcal{P}_i$ on the NMSE performance. We observe that the NMSE decreases as $\mathcal{P}_i$ increases for the given setups. This trend can be attributed to the fact that increasing $\mathcal{P}_i$ reduces sparsity while simultaneously enhancing the strength of the received signals, resulting in a tradeoff between sparsity and signal strength. Under the considered circumstances, the impact of signal energy outweighs that of sparsity. Furthermore, increasing the SNR consistently enhances the NMSE performance for both ECSIEst-Net and off-grid SBL method.

Fig. 7 compares the performance of the proposed off-grid SBL-based estimator and ECSIEst-Net with several benchmarks:

- Least square (LS) method: The LS method aims to minimize the mean square error (MSE) between the true and estimated outputs.
- Pattern-coupled sparse Bayesian learning (PC-SBL) method in [56]: The PC-SBL exploits block sparsity in sparse vectors to enhance the performance of the on-grid SBL algorithm.
- Data-driven method in [57]: Data-driven deep learning methods leverage large amounts of wireless communication data to train neural networks. Specifically, CNNs are employed for spectral feature extraction, and the network parameters are optimized using the backpropagation algorithm to minimize the error between the predicted and true channels.

For traditional detectors, the NMSE performance of the off-grid SBL-based detector surpasses that of the LS and PC-SBL detectors by fully capturing the sparse features of the received signals and leveraging accurate first-order approximations. Additionally, the proposed model-driven ECSIEst-Net outperforms data-driven methods by effectively incorporating prior system knowledge. Furthermore, ECSIEst-Net demonstrates superior performance compared to off-grid SBL methods, especially in high SNR environments, highlighting its robust capabilities and potentials in high-mobility scenarios.

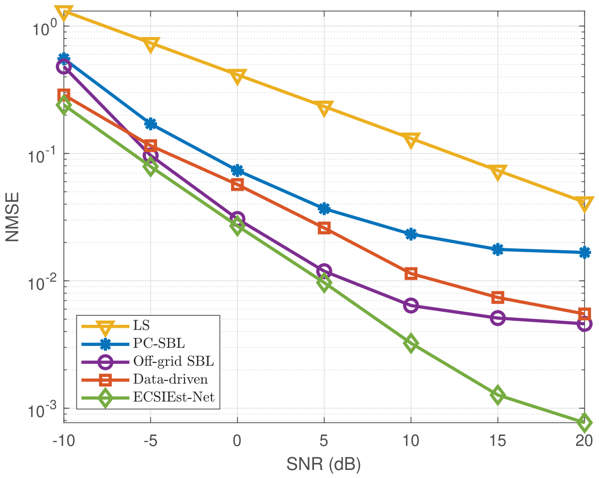


Fig. 7. Comparison among different CSI estimators.

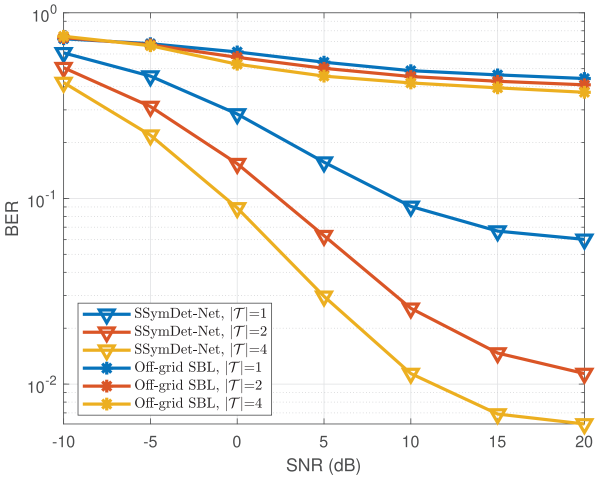


Fig. 8. Impact of $|\mathcal{T}|$ on BER performance.

B. Detection Performance of the Secondary Transmission

Fig. 8 illustrates the impact of $|\mathcal{T}|$ on the bit error rate (BER) performance for secondary transmission using both off-grid SBL-based and SSymDet-Net approaches. As can be readily observed, for both approaches, the detection accuracy improves as $|\mathcal{T}|$ increases since more samples would provide more useful information in extracting the secondary symbols, indicating that channels with more delay taps would facilitate the secondary transmission. Additionally, we see that the BER performance of the off-grid SBL method are inferior to that of the SSymDet-Net. This is primarily because the off-grid SBL algorithm struggles with separating dense frequencies from the narrow range $[0, \kappa_{\max}]$ and handling spectrum overlapping. In contrast, the data-driven SSymDet-Net method helps learn and extract features of dense, correlated spectrum clusters based on large number of training data, effectively recovering the secondary symbol even without any CSI.

Fig. 9 illustrates the impact of the IRS reflecting elements (i.e., $K_x \times K_y$) on the BER performance for secondary

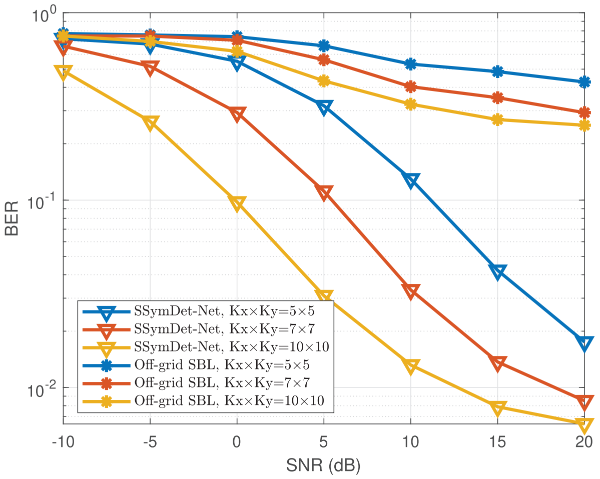


Fig. 9. Impact of IRS size on BER performance.

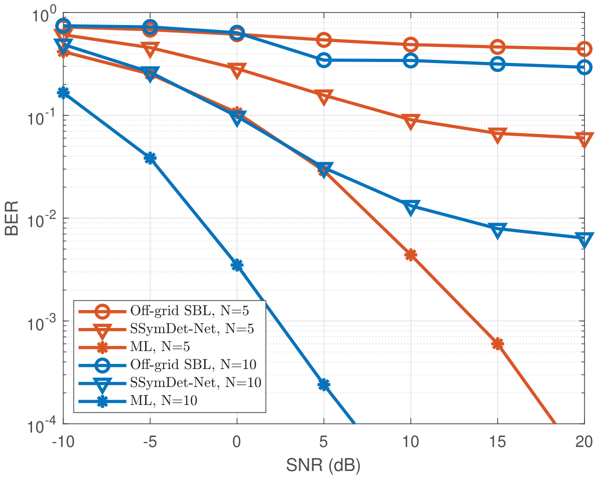


Fig. 10. Comparison of BER performance with varying *N*.

transmission, using either off-grid SBL-based approach or SSymDet-Net approach. As depicted, the BER performance improves for both methods as the IRS size increases. This improvement is attributed to the fact that an IRS with a larger number of reflecting elements is more effective at capturing and reflecting electromagnetic wave energy, resulting in a stronger reflected signal and enhanced detection performance. When comparing the BER performance of the two proposed approaches, the off-grid SBL-based approach shows only limited improvement due to difficulties in addressing spectrum overlap. In contrast, the SSymDet-Net approach demonstrates a significant BER gain as $K_x \times K_y$ increases. This highlights the advantage of SSymDet-Net in leveraging the IRS elements to enhance secondary transmission performance.

Fig. 10 compares the BER performance of the proposed off-grid SBL-based detector, SSymDet-Net and maximum likelihood (ML) detector with different *N*. As observed, the BER performance of the off-grid SBL-based detector is inferior to that of the SSymDet-Net, and the off-grid SBL-based detector reaches an error floor in high SNR regime due to spectrum overlapping. Additionally, the SSymDet-Net performs worse than the ML detector. As widely known, the ML detector is an optimal detector that examines all possible symbol sets and selects the one that maximizes the likelihood function. Although it offers optimal BER performance, its complexity is rather high. Furthermore, the detection performance of all three detectors improves with increasing *N*, indicating that longer symbol frame durations can enhance the spectrum identification in the Doppler domain.

C. Delay and Complexity Analysis

The end-to-end delay is primarily influenced by the transmission delay and the signal processing delay at the transceiver. The transmission delay is related to the channel distance, while the signal processing delay depends on the algorithm's complexity. In this work, we use the number of real floating-point operations (FLOPs) as the metric to evaluate the computational complexity of the proposed CSI estimator and detectors. Specifically, under the default parameters, the off-grid SBL-based CSI estimator and detector require a total of 319.5 GFLOPs (Giga FLOPs), while SSymDet-Net and ECSIEst-Net require 84.74 MFLOPs (Mega FLOPs) and 59.83 MFLOPs, respectively.[5] It is important to note that the computational complexity of the DL-based methods is primarily concentrated during the online stage, as the offline training process occurs less frequently. As observed, the DL-based methods exhibit lower complexity compared to the off-grid SBL-based methods. This is because the off-grid SBL-based algorithms require frequent iterative computations, while the DL-based algorithms mainly involve basic matrix operations that can be parallelized. Despite the lower complexity of the DL-based methods, they also achieve better performance compared with traditional SBL methods due to the powerful learning capabilities of deep learning, making them a promising solution for practical SR-OTFS systems.

VI. CONCLUSION

SR is a promising technology for future large-scale IoT networks. To tackle the time-frequency selective high-mobility channels, we propose to adopt OTFS modulation in SR system. For secondary information modulation, we proposed a CF modulation scheme by embedding the information within the selection of the frequency combinations. Since the time-varying secondary signals would induce complex coupling between the secondary signals and the channels, posing significant challenges in channel estimation and secondary symbol detection. Considering the scenario with ultra-wideband or/and short package, we propose an off-grid SBL-based approach, which first obtained the equivalent CSI and then detected secondary symbols based on the gap between shifted Doppler clusters. Then considering more general scenarios, we introduce a ECSIEst-Net for channel estimation and a SSymDet-Net for symbol detection even independent of CSI. Numerical results show that the proposed ECSIEst-Net and SSymDet-Net not only adapt better to various scenarios but also outperform traditional off-grid

---

[5] Noted that the complexity of the proposed DL methods slightly increases with non-stationary channels and the number of delay taps, but remains much lower than that of the off-grid SBL-based methods.

SBL algorithms in both estimation and detection performance. Additionally, we find that channels with more delay taps enhance secondary transmission.

REFERENCES

[1] L. Chettri and R. Bera, "A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 16–32, Jan. 2020.

[2] H. Guo, Y. C. Liang, R. Long, and Q. Zhang, "Cooperative ambient backscatter system: A symbiotic radio paradigm for passive IoT," *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1191–1194, Aug. 2019.

[3] Y.-C. Liang, Q. Zhang, E. G. Larsson, and G. Y. Li, "Symbiotic radio: Cognitive backscattering communications for future wireless networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 4, pp. 1242–1255, Dec. 2020.

[4] Y. Xu et al., "Robust resource allocation for wireless-powered backscatter communication systems with NOMA," *IEEE Trans. Veh. Technol.*, vol. 72, no. 9, pp. 12288–12299, Sep. 2023.

[5] H. Zhou, B. Cai, Q. Zhang, R. Long, Y. Pei, and Y.-C. Liang, "Modulation design and optimization for RIS-assisted symbiotic radios," *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 13022–13038, Oct. 2024, doi: 10.1109/TWC.2024.3398118.

[6] H. Zhou, X. Kang, Y.-C. Liang, S. Sun, and X. Shen, "Cooperative beamforming for reconfigurable intelligent surface-assisted symbiotic radios," *IEEE Trans. Veh. Technol.*, vol. 71, no. 11, pp. 11677–11692, Nov. 2022.

[7] H. Zhou, Q. Zhang, Y.-C. Liang, and Y. Pei, "Assistance-transmission tradeoff for RIS-assisted symbiotic radios," *IEEE Trans. Wireless Commun.*, vol. 23, no. 7, pp. 6838–6855, Jul. 2024.

[8] C. Zhou et al., "Energy-efficient maximization for RIS-aided MISO symbiotic radio systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 10, pp. 13689–13694, Oct. 2023.

[9] N. Van Huynh, D. T. Hoang, X. Lu, D. Niyato, P. Wang, and D. I. Kim, "Ambient backscatter communications: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2889–2922, 4th Quart., 2018.

[10] X. Xu, Y.-C. Liang, G. Yang, and L. Zhao, "Reconfigurable intelligent surface empowered symbiotic radio over broadcasting signals," *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 7003–7016, Oct. 2021.

[11] R. Long, Y.-C. Liang, H. Guo, G. Yang, and R. Zhang, "Symbiotic radio: A new communication paradigm for passive Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1350–1363, Feb. 2020.

[12] Q. Tao, X. Hu, S. Zhang, and C. Zhong, "Integrated sensing and communication for symbiotic radio systems in mobile scenarios," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 11213–11225, Sep. 2024.

[13] J. Qian, F. Gao, G. Wang, S. Jin, and H. Zhu, "Semi-coherent detection and performance analysis for ambient backscatter system," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5266–5279, Dec. 2017.

[14] Q. Tao, C. Zhong, H. Lin, and Z. Zhang, "Symbol detection of ambient backscatter systems with Manchester coding," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 4028–4038, Jun. 2018.

[15] Q. Tao, C. Zhong, K. Huang, X. Chen, and Z. Zhang, "Ambient backscatter communication systems with MFSK modulation," *IEEE Trans. Wireless Commun.*, vol. 18, no. 5, pp. 2553–2564, May 2019.

[16] J. Qian, A. N. Parks, J. R. Smith, F. Gao, and S. Jin, "IoT communications with M-PSK modulated ambient backscatter: Algorithm, analysis, and implementation," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 844–855, Feb. 2019.

[17] W. Yan, X. Yuan, Z.-Q. He, and X. Kuai, "Passive beamforming and information transfer design for reconfigurable intelligent surfaces aided multiuser MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1793–1808, Aug. 2020.

[18] W. Yan, X. Yuan, and X. Kuai, "Passive beamforming and information transfer via large intelligent surface," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 533–537, Apr. 2020.

[19] G. D. Ntouni, V. M. Kapinas, and G. K. Karagiannidis, "On the optimal tone spacing for interference mitigation in OFDM-IM systems," *IEEE Commun. Lett.*, vol. 21, no. 5, pp. 1019–1022, May 2017.

[20] M. Wen, B. Ye, E. Basar, Q. Li, and F. Ji, "Enhanced orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4786–4801, Jul. 2017.

[21] Z. Wei et al., "Orthogonal time-frequency space modulation: A promising next-generation waveform," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 136–144, Aug. 2021.

[22] Z. Wei, S. Li, W. Yuan, R. Schober, and G. Caire, "Orthogonal time frequency space modulation—Part I: Fundamentals and challenges ahead," *IEEE Commun. Lett.*, vol. 27, no. 1, pp. 4–8, Jan. 2023.

[23] W. Yuan, Z. Wei, S. Li, R. Schober, and G. Caire, "Orthogonal time frequency space modulation—Part III: ISAC and potential applications," *IEEE Commun. Lett.*, vol. 27, no. 1, pp. 14–18, Jan. 2023.

[24] Q. Tao, T. Xie, X. Hu, S. Zhang, and D. Ding, "Channel estimation and detection for intelligent reflecting surface-assisted orthogonal time frequency space systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 8419–8431, Aug. 2024.

[25] S. Li, W. Yuan, Z. Wei, R. Schober, and G. Caire, "Orthogonal time frequency space modulation—Part II: Transceiver designs," *IEEE Commun. Lett.*, vol. 27, no. 1, pp. 9–13, Jan. 2023.

[26] G. D. Surabhi, R. M. Augustine, and A. Chockalingam, "On the diversity of uncoded OTFS modulation in doubly-dispersive channels," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3049–3063, Jun. 2019.

[27] S. Li, J. Yuan, W. Yuan, Z. Wei, B. Bai, and D. W. K. Ng, "Performance analysis of coded OTFS systems over high-mobility channels," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 6033–6048, Sep. 2021.

[28] G. D. Surabhi, R. M. Augustine, and A. Chockalingam, "Peak-to-average power ratio of OTFS modulation," *IEEE Commun. Lett.*, vol. 23, no. 6, pp. 999–1002, Jun. 2019.

[29] Z. Wei, W. Yuan, S. Li, J. Yuan, and D. W. K. Ng, "Transmitter and receiver window designs for orthogonal time-frequency space modulation," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2207–2223, Apr. 2021.

[30] Y. Ge, Q. Deng, P. C. Ching, and Z. Ding, "Receiver design for OTFS with a fractionally spaced sampling approach," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4072–4086, Jul. 2021.

[31] R. Hadani et al., "Orthogonal time frequency space modulation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017, pp. 1–6.

[32] P. Raviteja, K. T. Phan, Y. Hong, and E. Viterbo, "Interference cancellation and iterative detection for orthogonal time frequency space modulation," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6501–6515, Oct. 2018.

[33] T. Thaj and E. Viterbo, "Low complexity iterative rake decision feedback equalizer for zero-padded OTFS systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 15606–15622, Dec. 2020.

[34] P. Raviteja et al., "Embedded pilot-aided channel estimation for OTFS in delay-Doppler channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4906–4917, May 2019.

[35] Z. Wei et al., "Off-grid channel estimation with sparse Bayesian learning for OTFS systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7407–7426, Sep. 2022.

[36] Z. Ding, R. Schober, P. Fan, and H. V. Poor, "OTFS-NOMA: An efficient approach for exploiting heterogenous user mobility profiles," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7950–7965, Nov. 2019.

[37] S. Tiwari, S. S. Das, and V. Rangamgari, "Low complexity LMMSE receiver for OTFS," *IEEE Commun. Lett.*, vol. 23, no. 12, pp. 2205–2209, Dec. 2019.

[38] W. Yuan, Z. Wei, J. Yuan, and D. W. K. Ng, "A simple variational Bayes detector for orthogonal time frequency space (OTFS) modulation," *IEEE Trans. Veh. Technol.*, vol. 69, no. 7, pp. 7976–7980, Jul. 2020.

[39] S. Li, W. Yuan, Z. Wei, and J. Yuan, "Cross domain iterative detection for orthogonal time frequency space modulation," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2227–2242, Apr. 2022.

[40] E. Basar et al., "Reconfigurable intelligent surfaces for 6G: Emerging hardware architectures, applications, and open challenges," *IEEE Veh. Technol. Mag.*, vol. 19, no. 3, pp. 27–47, Sep. 2024.

[41] V. S. Bhat, G. Harshavardhan, and A. Chockalingam, "Input–output relation and performance of RIS-aided OTFS with fractional delay-Doppler," *IEEE Commun. Lett.*, vol. 27, no. 1, pp. 337–341, Jan. 2023.

[42] A. S. Bora, K. T. Phan, and Y. Hong, "IRS-assisted high mobility communications using OTFS modulation," *IEEE Wireless Commun. Lett.*, vol. 12, no. 2, pp. 376–380, Feb. 2023.

[43] A. Thomas, K. Deka, S. Sharma, and N. Rajamohan, "IRS-assisted OTFS system: Design and analysis," *IEEE Trans. Veh. Technol.*, vol. 72, no. 3, pp. 3345–3358, Mar. 2023.

[44] M. Li, S. Zhang, Y. Ge, F. Gao, and P. Fan, "Joint channel estimation and data detection for hybrid RIS aided millimeter wave OTFS systems," *IEEE Trans. Commun.*, vol. 70, no. 10, pp. 6832–6848, Oct. 2022.

[45] Z. Li, W. Yuan, B. Li, J. Wu, C. You, and F. Meng, "Reconfigurable intelligent surface aided OTFS: Transmission scheme and channel estimation," *IEEE Internet Things J.*, vol. 10, no. 22, pp. 19518–19532, Nov. 2023.

[46] R. K. Yadav, H. B. Mishra, S. Mukhopadhyay, and R. Mishra, "IRS-OTFS systems: Design of reflection coefficients for low-complexity ZF equalizer," *IEEE Trans. Veh. Technol.*, vol. 73, no. 10, pp. 15721–15726, Oct. 2024, doi: 10.1109/TVT.2024.3400529.

[47] Q. Tao, T. Xie, Z. Yang, and Z. Zhang, "Symbiotic radio with orthogonal time frequency space modulation over high-mobility channels," *IEEE Trans. Wireless Commun.*, vol. 23, no. 12, pp. 18070–18082, Dec. 2024.

[48] W. Yuan, J. Zou, Y. Cui, X. Li, J. Mu, and K. Han, "Orthogonal time frequency space and predictive beamforming-enabled URLLC in vehicular networks," *IEEE Wireless Commun.*, vol. 30, no. 2, pp. 56–62, Apr. 2023.

[49] X. Hu, C. Zhong, Y. Zhang, X. Chen, and Z. Zhang, "Location information aided multiple intelligent reflecting surface systems," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7948–7962, Dec. 2020.

[50] J. Francis and V. P. Reddy, "Delay-Doppler channel estimation in OTFS systems using DoA estimation techniques," in *Proc. IEEE 95th Veh. Technol. Conf.*, Jun. 2022, pp. 1–5.

[51] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, Jan. 2013.

[52] F. Zhu, A. Liu, and V. K. N. Lau, "Channel estimation and localization for mmWave systems: A sparse Bayesian learning approach," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[53] J. Gao, C. Zhong, G. Y. Li, J. B. Soriaga, and A. Behboodi, "Deep learning-based channel estimation for wideband hybrid mmWave massive MIMO," *IEEE Trans. Commun.*, vol. 71, no. 6, pp. 3679–3693, Jun. 2023.

[54] L. Xiao, S. Li, Y. Qian, D. Chen, and T. Jiang, "An overview of OTFS for Internet of Things: Concepts, benefits, and challenges," *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7596–7618, May 2022.

[55] A. Jagannath, J. Jagannath, and T. Melodia, "Redefining wireless communication for 6G: Signal processing meets deep learning with deep unfolding," *IEEE Trans. Artif. Intell.*, vol. 2, no. 6, pp. 528–536, Dec. 2021.

[56] J. Fang, Y. Shen, H. Li, and P. Wang, "Pattern-coupled sparse Bayesian learning for recovery of block-sparse signals," *IEEE Trans. Signal Process.*, vol. 63, no. 2, pp. 360–372, Jan. 2015.

[57] W. Yu, F. Sohrabi, and T. Jiang, "Role of deep learning in wireless communications," *IEEE BITS Inf. Theory Mag.*, vol. 2, no. 2, pp. 56–72, Nov. 2022.

Qin Tao (Member, IEEE) received the B.S. degree from the Electronic Information School, Wuhan University, Wuhan, China, in 2017, and the Ph.D. degree in communication and information systems from Zhejiang University, Hangzhou, China, in 2022. She is currently an Associate Professor with the School of Information Science and Technology, Hangzhou Normal University, Hangzhou. Her research interests include backscatter communication, intelligent reflecting surface, symbiotic radio, and orthogonal time frequency space modulation. She and her co-authors have been awarded the Best Paper Award at the ICCC 2018, WCSP 2020, and WCSP 2023.

Kai Jin received the B.S. degree from the School of Computer Science and Engineering, Ningbo University of Technology, Ningbo, China, in 2023. He is currently pursuing the M.S. degree with the School of Information Science and Technology, Hangzhou Normal University, Hangzhou, China. His research interests include intelligent reflecting surface, deep learning-based intelligent communication, and semantic communication.

Jiabao Gao received the B.S. degree in information engineering from Zhejiang University, Hangzhou, China, in 2019, where he is currently pursuing the Ph.D. degree with Zhejiang Provincial Key Laboratory of Information Processing, Communication and Networking. From September 2021 to March 2023, he was a Visiting Student with the Department of Electrical and Electronic Engineering, Imperial College London, U.K. His current research interests include millimeter wave and THz massive MIMO and deep learning for wireless communications.

Weijie Yuan (Senior Member, IEEE) received the joint Ph.D. degree from the University of Technology Sydney, Australia, and Beijing Institute of Technology, China, in 2019. In 2016, he was a Visiting Ph.D. Student with the Institute of Telecommunications, Vienna University of Technology, Austria. He was a Research Assistant with The University of Sydney, a Visiting Associate Fellow with the University of Wollongong, and a Visiting Fellow with the University of Southampton from 2017 to 2019. From 2019 to 2021, he was a Research Associate with The University of New South Wales. He is currently the Series Lead Editor of *IEEE Communications Magazine* and an Associate Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, IEEE COMMUNICATIONS LETTERS, IEEE OPEN JOURNAL OF COMMUNICATIONS SOCIETY, and *EURASIP Journal on Advances in Signal Processing*. He is the Lead Editor of two feature topics in *IEEE Communications Magazine*.

Ben Wang received the M.S. degree from Zhejiang University of Technology, China, in 2002, and the Ph.D. degree from University of Essex in 2007. He is currently a Professor with the School of Information Science and Technology (SIST), Hangzhou Normal University, Hangzhou, China. His research interests include digital forensics, the Internet of Things (IoT), and intelligent information processing.

Chau Yuen (Fellow, IEEE) received the B.Eng. and Ph.D. degrees from Nanyang Technological University, Singapore, in 2000 and 2004, respectively. He was a Post-Doctoral Fellow with Lucent Technologies Bell Laboratories, Murray Hill, in 2005. From 2006 to 2010, he was with the Institute for Infocomm Research, Singapore. From 2010 to 2023, he was with the Engineering Product Development Pillar, Singapore University of Technology and Design. Since 2023, he has been with the School of Electrical and Electronic Engineering, Nanyang Technological University, where he is currently the Provosts Chair of wireless communications and the Assistant Dean of the Graduate College. He received the IEEE Marconi Prize Paper Award in Wireless Communications in 2021, the EURASIP Best Paper Award for *Journal on Wireless Communications and Networking* in 2021, the IEEE Communications Society Fred W. Ellersick Prize in 2023, the IEEE APB Outstanding Paper Award in 2023, the IEEE Communications Society Leonard G. Abraham Prize in 2024, and the IEEE Communications Society Best Tutorial Paper Award in 2024. He is a Distinguished Lecturer of the IEEE Vehicular Technology Society, Top 2% Scientists by Stanford University, and also a Highly Cited Researcher by Clarivate Web of Science.