

Estimation of Dispersive High-Doppler Channels in the RIS-Aided mmWave Internet of Vehicles

Jiaxin Chen, Wenqian Shen, Shixun Luo, Siqi Ma, Chengwen Xing, and Lajos Hanzo, *Life Fellow, IEEE*

Abstract—Reconfigurable intelligent surfaces (RISs) have emerged as a promising candidate for improving the spectral- and energy- efficiency of millimeter-wave (mmWave) Internet of vehicles (IoV) communications, but the conception of their accurate channel estimation poses. Hence the existing estimation methods mainly focus on time-invariant channels, while ignoring the Doppler effect induced by the high-velocity vehicles, which will lead to significant performance degradation. In this paper, we investigate the problem of channel estimation in RIS-aided mmWave IoV systems considering the deleterious Doppler effect. Firstly, we derive the expression of the time-varying cascaded two-hop multiple-path channels, where each delay tap is subject to multiple paths instead of having a simple one-to-one correspondence. In order to decouple the paths, the problem is formulated in the delay-domain by a series of transformations and the cascaded two-hop channel can be estimated at each delay tap. Then we propose a pair of estimation strategies by considering different hardware constraints depending on the number of receiver antennas at the base station (BS). When a large receiver array is employed at the BS, we can exploit its high angular selectivity for distinguishing each resolvable path at a certain delay tap because they arrive from different directions. However, this cannot be achieved for small arrays, given their more limited angular resolution. Thus, the RIS reflection patterns are delicately designed for distinguishing multiple resolvable paths. After separating the paths, Doppler estimation can be performed by calculating the phase difference of the adjacent symbols. Our simulation results demonstrate the superior performance of the proposed methods within a wide range of Doppler shifts.

Index Terms—Channel estimation, reconfigurable intelligent surfaces (RISs), Doppler effect.

I. INTRODUCTION

Recently, Internet-of-vehicles (IoV) has attracted much attention from both academia and industry due to the increasing interest in both the vehicular Internet of things (IoT) and autonomous driving [1], [2]. However, the tremendous volume

of data traffic and high-mobility of users constitute a pair of issues limiting the performance of IoV. On the one hand, given its ample bandwidth reserves, millimeter-wave (mmWave) communication, can be integrated into the IoV for high-rate transmission [3]. However, the mmWave links can be blocked by obstacles, and their high hardware cost, energy consumption, complexity and severe interference issues still constitute challenges. On the other hand, the high mobility of users shortens the visibility range of base stations (BS). Hence, more BSs, active relays or roadside units (RSU) have to be deployed for providing reliable communications, which incurs high cost and power consumption [4]. Fortunately, the emerging reconfigurable intelligent surfaces (RIS) mitigate blockages at a low cost and easy deployment. Briefly, RISs are composed of a large number of passive software-controlled elements capable of adjusting the phase or even potentially the amplitude of the incident radio waves in near-real-time, which can turn the wholly uncontrollable wireless environment into a relatively controllable one [5], [6]. In other words, RISs provide an additional degree of freedom for the system design to some extent, hence much improved spectral efficiency can be achieved by appropriately designing the reflecting coefficients (RCs) of the RIS, which has been extensively studied in the literature [7]–[9].

Given these benefits, RISs are expected to be harvested also in vehicular networks for assisting IoV communication, which extends the coverage of BSs and enhances the transmission quality as a benefit of their nearly real-time-adjustable passive beamforming gain [10], [11]. Mensi *et al.* proposed RIS-based access points for vehicle-to-vehicle (V2V) networks and RIS-based relays for vehicular adhoc networks (VANET), where the passive RIS beamformer was optimized and hence quality-of-service (QoS) was improved. Moreover, it has been shown that RIS-assisted vehicular communication reduces the outage probability [12], while enhancing the robustness [13] and stability of links [14].

Most of the RIS-aided passive beamformer designs were conceived under the assumption that the channel state information (CSI) is perfectly known. However, acquiring CSI is a challenging task, because the passive elements of RISs operating without radio frequency (RF) chains cannot perform sophisticated signal processing. In order to estimate the channel of RIS-aided systems, sophisticated methods were proposed for narrowband systems over flat-fading channels, which include the simple ON/OFF-based reflection pattern [15] and numerous more refined designs. The specific characteristics of channels were exploited for reducing the training overhead using low-rank [16], two-timescale [17] or structured-sparsity [18] based designs, just to name a few. However, there is a paucity of channel estimation techniques for wideband sys-

This paper was supported in part by the the National Key Research and Development Program of China under No. 2019YFB1803200, and in part by the National Natural Science Foundation of China under Grant 62231004. L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/W016605/1 and EP/X01228X/1 as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028). (*Corresponding author: S. Luo.*)

J. Chen, W. Shen, S. Ma, C. Xing are with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (E-mails: jxchen.ee@gmail.com; shenwq@bit.edu.cn; sqma@bit.edu.cn; xingchengwen@gmail.com).

S. Luo is with the School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing 100081, China (E-mail: luoshixun@bit.edu.cn).

L. Hanzo is with the Department of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK (E-mail: lh@ecs.soton.ac.uk).

©Copyright (c) 20xx IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.”

tems communicating over frequency-selective fading channels, which is hence the focus of this paper. Having said that there are a few studies for this challenging scenario. Yang *et al.* [19] first proposed a RIS-element grouping strategy based on least square estimation for reducing the pilot overhead of RIS-assisted orthogonal frequency-division multiplexing (OFDM) systems. However, this method cannot attain a high passive RIS beamforming gain owing to the rudimentary nature of the ON/OFF control of the RIS groups. Zheng *et al.* [20] improved this method by pre-designing the RIS reflection pattern with all units being turned ON in the first subframe, which is equivalent to adding a new pilot pattern. Then, they took multi-user systems into consideration and proposed a pair of channel estimation schemes for different channel configurations, which simultaneously optimized the pilot tone allocations and RIS reflection pattern [21]. Despite employing the above grouping technique, both the pilot overhead and the computational complexity remain prohibitive for a large number of RIS elements. The authors of [22], [23] conceived compressed sensing (CS)-based algorithms by exploiting the inherent angular sparsity of mmWave channels for reducing the pilot overhead further, which, however, assumed that the channel parameters between the BS and RIS were known.

The above literature assumes quasi-static frequency-selective block fading channels, which cannot be directly applied in IoV communications due to the destructive Doppler effect experienced by high-mobility users. A high Doppler makes the channel time-varying, which may hence become outdated when relying on traditional methods and gravely degrades the performance. Hence a few authors aimed for mitigating the Doppler effect, for example, Wu [24] *et al.* performed Doppler estimation under the assumption that only a single cascaded path existed in the reflected link. In [25], the authors considered having multiple paths in the RIS-user channel, but a single line-of-sight (LoS) path between the BS and RIS. Besides Li *et al.* proposed a joint channel estimation and data detection method for high-mobility RIS-aided mmWave systems [26], which is based on the orthogonal time frequency space (OTFS) modulation framework. They assumed that only LoS propagation exists between the BS and the RIS. In these cases, the cascaded channel derived was in the form of one-to-one correspondence between a delayed tap and the propagation path (characterized by Doppler shift). Moreover, Xu *et al.* studied a RIS-aided point-to-point system operating in high-mobility Ricean channels and proposed two MMSE-based channel estimators respectively in [27], [28], which however cannot be extended to large BS arrays at a low pilot overhead. For the general scenario where there are multiple paths in both the BS-RIS link and the RIS-user link, the cascaded two-hop channel at an equivalent delay tap may actually be constituted by the summation of several combinations of these two-hop channel paths associated with coupled Doppler frequency shifts. This issue will be elaborated on in the next section.

Since the previously mentioned methods are unsuitable for this case, we formulate a more general system model directly considering the Doppler effect in our RIS-assisted mmWave IoV system and propose a pair of channel estimation methods.

In Table I, we explicitly contrast our contributions to the existing literature. The main contributions of this paper are summarized as follows:

- We deduce the input and output relationship of our OFDM mmWave IoV system by directly considering the Doppler effect, when there are multiple paths in both links of the two-hop channels, namely in the BS-RIS and RIS-user links. The cascaded channel of the reflected link is given by the convolution of the two separate channels, which can be characterized by their equivalent delays, angles, complex-valued gains and Doppler frequency shifts. Due to the various combinations of the two individual delays, a specific equivalent delay tap will correspond to several propagation paths, each of which possesses its own Doppler shift. In order to separate these mixed paths, we perform a series of transformations to derive the delay-domain relationship between the pilots transmitted from the user and the signals received at the BS. Thus, the cascaded channel coefficients can be estimated at each equivalent delay tap.
- When a large number of antennas are employed at the BS, the high angular resolution of the BS array can be exploited for distinguishing the paths at each delay tap. Therefore, we propose a two-stage strategy for estimating the equivalent parameters of the cascaded channel. During Stage I, we decouple the paths at a given delay tap by estimating the AoAs at the BS by a CS-based algorithm, where the RIS RCs are randomly generated. During Stage II, the Doppler shift of each individual path can be calculated according to the phase difference of the consecutive symbols, where the RIS RCs remain unchanged. Finally, the equivalent AoDs at the RIS and the equivalent gains are estimated after Doppler compensation by exploiting again the signals received in Stage I.
- When the angular resolution is insufficiently high for distinguishing the angles arriving at the BS having an insufficiently large array, the additional grade of freedom of the RIS inspires us to separate the paths by appropriately designing the RCs. Likewise, the strategy proposed in this scenario is similar to the last one, but with specifically designed RIS RCs during Stage II. We coarsely estimate the equivalent AoDs at the RIS by first ignoring the Doppler effect. Then the RIS RCs are set to certain values according to the rough estimate of the equivalent AoDs in an interleaved way, which can separate paths due to the asymptotic orthogonality among the steering vectors of the RIS. Eventually, we refine the estimate of the equivalent AoDs and gains by considering the estimated Doppler shifts. Our simulation results will show that the proposed algorithms outperform their traditional counterparts, despite their low pilot overhead.

The rest of this paper is organized as follows. In Section II, we present the system model and formulate the channel estimation problem of our RIS-aided mmWave IoV systems by directly considering the Doppler effect. Section III presents our cascaded channel estimation algorithm designed for large

TABLE I
CONTRASTING OUR CONTRIBUTION TO THE LITERATURE

	[15], [17]	[16], [18]	[19]–[21]	[22], [23]	[24]	[25]	[27], [28]	Proposed
Narrowband flat-fading channel	✓	✓			✓			
Wideband frequency-selective channel			✓	✓		✓	✓	✓
Quasi-static channel	✓	✓	✓	✓				
Time-varying channel					✓	✓	✓	✓
Single path between BS and RIS				✓	✓	✓		
Multiple paths between BS and RIS	✓	✓	✓				✓	✓
Non-parametric method	✓		✓		✓	✓	✓	✓
Parametric method		✓		✓				✓

BS arrays, which is then further refined for small BS arrays. Our simulation results are provided in Section V, followed by our conclusions.

Notations: We let a , \mathbf{a} , \mathbf{A} represents the scalar, vector, and matrix, respectively; $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^*$ and $(\cdot)^{-1}$ denote the transpose, conjugate transpose, conjugate and inverse of a matrix, respectively; $j = \sqrt{-1}$ is the imaginary unit; $[\mathbf{a}]_m$ denotes the m -th element of vector \mathbf{a} ; $[\mathbf{A}]_{m,n}$ denotes the (m,n) -th element of matrix \mathbf{A} ; $[\mathbf{A}]_{:,n}$ denotes the n -th column of matrix \mathbf{A} ; The $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes circularly symmetric complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. The M -dimensional vector having all zero entries is denoted by $\mathbf{0}_M$.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider the scenario of high-mobility mmWave uplink (UL) IoV transmission, as depicted in Fig.1. A RIS is employed to assist communication between a static BS having a uniform linear array (ULA) of M_b receiver antennas (RAs) and a fast moving single-antenna user, which results in severe Doppler effect that cannot be ignored. For ease of exposition, we assume that the M_r elements of the RIS are also arranged in the form of a ULA¹ and they are connected to a smart controller, which can adjust the RCs according to the commands of the BS. OFDM modulation is adopted to combat the frequency-selective channel effects caused by the multipath delay spread. The signal bandwidth is B and the number of subcarriers is N . To avoid inter symbol interference, we assume that the length of the cyclic prefix (CP) T_{cp} is longer than the maximum path delay. We neglect the estimation of the direct channel, since it can be estimated by various traditional methods [29], [30] via turning off the RIS [31]. Although the two-hop channels in the reflected link cannot be estimated separately due to the passivity of the RIS, the dominant cascaded BS-RIS-user channel has to be estimated for designing the joint active and RIS-based passive beamforming [7]. Hence this is what we will focus on in the following.

A. Time-Varying Cascaded Two-Hop Channel with Multiple Paths Coupled

In general, the BS-RIS channel is usually quasi-static, since the RIS is usually located at high building surfaces, whereas

¹The channel estimation method proposed in this paper can be readily extended to uniform planar arrays (UPA).

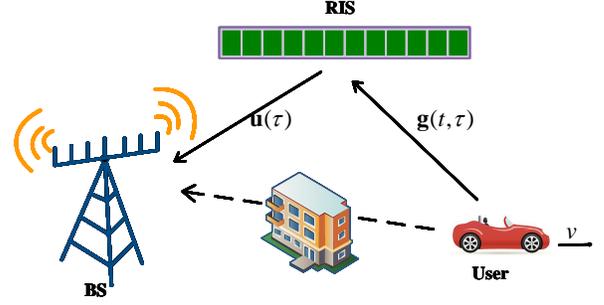


Fig. 1. RIS-aided mmWave uplink communication systems with fast moving user.

the RIS-user channel is time-varying due to the movement of the user. According to the widely used Saleh-Valenzuela channel model [32], we assume that there are L_1 paths between the BS and RIS. The channel impulse response (CIR) between the r -th element of the RIS and the b -th antenna of the BS is represented as

$$u_{b,r}(\tau) = \sum_{l_1=1}^{L_1} c_{l_1} e^{-j2\pi(b-1)\varphi_{l_1}^B} e^{j2\pi(r-1)\varphi_{l_1}^R} \delta(\tau - \tau_{l_1}), \quad (1)$$

where c_{l_1} and τ_{l_1} denote the complex gain and delay of the l_1 -th path, respectively. Furthermore, we separately define the normalized angle-of-arrival (AoA) at the BS and the angle-of-departure (AoD)² at the RIS of the l_1 -th path as $\varphi_{l_1}^B = d\sin(\phi_{l_1}^B)/\lambda$ and $\varphi_{l_1}^R = d\sin(\phi_{l_1}^R)/\lambda$, where λ is the wavelength of the carrier and $d = \lambda/2$ denotes the adjacent array spacing.

As for the rapidly time-varying RIS-user channel, we assume that L_2 paths exist. In contrast to the channel model between the BS and RIS, an extra phase shift varying over time caused by the Doppler frequency shift is added in each path [33], [34]. Thus, the CIR between the user and the r -th element of the RIS is expressed as

$$g_r(t, \tau) = \sum_{l_2=1}^{L_2} c_{l_2} e^{j2\pi f_{l_2} t} e^{-j2\pi(r-1)\varphi_{l_2}^R} \delta(\tau - \tau_{l_2}), \quad (2)$$

where c_{l_2} , τ_{l_2} and $\varphi_{l_2}^R$ have similar definitions as in (1). Here, the Doppler shift of the l_2 -th path is $f_{l_2} = f_m \cos(\gamma_{l_2})$, with $f_m = v/\lambda$ being the maximum Doppler shift, where v is the

²We omit the adjective *Normalized* throughout the paper for simplicity.

user's velocity and γ_{l_2} denotes the AoD of the wave relative to the moving direction of the user in the l_2 -th path. Therefore, the Doppler spread is $D_s = 2f_m$ and the channel's coherence time can be expressed as $T_c = 1/4D_s$ [35].

Moreover, since the user's location and speed change very little during the transmission frame considered, we assume that c_{l_2} , $\varphi_{l_2}^R$ and f_{l_2} remain approximately constant [24], [25], [36]. We term the transmission time of a single frame as the *relevance interval* for simplicity. Note that it is different from the channel's coherence time, which is less than one OFDM symbol duration owing to the Doppler shift. Hence, the channel variation is merely attributed to the Doppler-induced phase shift.

Then, the cascaded channel spanning from the user to the b -th BS RA passing through the r -th element of the RIS is given by the linear convolution of the two individual channels, which is

$$\begin{aligned} h_{b,r}(t, \tau) &= u_{b,r}(\tau) * g_r(t, \tau) \\ &= \sum_{l_1=1}^{L_1} \sum_{l_2=1}^{L_2} c_{l_1} c_{l_2} e^{j2\pi f_{l_2} t} e^{-j2\pi(b-1)\varphi_{l_1}^B} e^{j2\pi(r-1)(\varphi_{l_1}^B - \varphi_{l_2}^R)} \delta(\tau - \tau_{l_1} - \tau_{l_2}) \end{aligned} \quad (3)$$

$$= \sum_{l=1}^L \left(\sum_{s_l=1}^{S_l} c_{s_l} e^{j2\pi f_{s_l} t} e^{-j2\pi(b-1)\varphi_{s_l}^B} e^{j2\pi(r-1)\varphi_{s_l}^R} \right) \delta(\tau - \tau_l). \quad (4)$$

To elaborate, we define $\tau_l = \tau_{l_1} + \tau_{l_2}$ as the equivalent delay of the cascaded two-hop path. There will be a total of $L_1 L_2$ legitimate combinations of the delay sum, but, some of them may have duplicate values. More explicitly, the cascaded channel at a particular delay tap may be the amalgam of several cascaded paths, each of which has its corresponding Doppler shift. Let L be the total number of unique channel delay taps, among which the maximum delay is assumed to be τ_D associated with the D -th tap. We have $\tau_l \in (0, \tau_D]$, and the channel's coherence bandwidth can be formulated as $W_c = 1/2\tau_D$ [35]. In this way, the cascaded channel in (3) can be recast as (4), where S_l denotes the number of paths associated with the same delay τ_l . Furthermore, let us define the equivalent complex gain, the Doppler shift, AoA and the equivalent AoD of the s_l -th path at delay τ_l as

$$\begin{aligned} c_{s_l} &= c_{l_1} c_{l_2}, \\ f_{s_l} &= f_{l_2}, \\ \varphi_{s_l}^B &= \varphi_{l_1}^B \in \left[-\frac{1}{2}, \frac{1}{2}\right), \\ \varphi_{s_l}^R &= \varphi_{l_1}^B - \varphi_{l_2}^R \in [-1, 1), \end{aligned}$$

where $s_l \in \{1, 2, \dots, S_l\}^3$.

Let us now stack channels from all the RIS elements to all the BS RAs into a matrix form, which is expressed as

$$\mathbf{H}(t, \tau) = \sum_{l=1}^L \left(\sum_{s_l=1}^{S_l} c_{s_l} e^{j2\pi f_{s_l} t} \mathbf{a}_{M_b}(\varphi_{s_l}^B) \mathbf{a}_{M_r}^H(\varphi_{s_l}^R) \right) \delta(\tau - \tau_l), \quad (5)$$

³The period of $e^{j2\pi\varphi}$ is 1, so $\varphi_{l_1}^B - \varphi_{l_2}^R \in [-1/2, 1/2)$ actually.

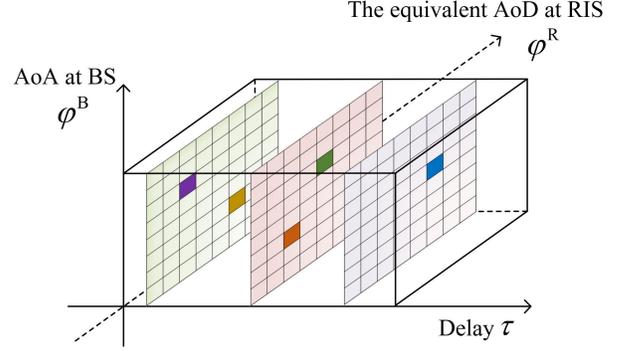


Fig. 2. Three dimensions to distinguish the coupled cascaded two-hop paths.

where $\mathbf{a}_{M_b}(\varphi_{s_l}^B) \in \mathbb{C}^{M_b \times 1}$ and $\mathbf{a}_{M_r}(\varphi_{s_l}^R) \in \mathbb{C}^{M_r \times 1}$ denote the steering vector at the BS and the RIS, respectively. When denoting the dimension by M and the normalized angle by φ , the steering vector $\mathbf{a}_M(\varphi)$ is formulated as

$$\mathbf{a}_M(\varphi) = [1, e^{-j2\pi\varphi}, \dots, e^{-j2\pi(M-1)\varphi}]^T. \quad (6)$$

Each cascaded two-hop path can be distinguished by its path delay, the AoA at BS or the equivalent AoD at the RIS. The support sets of the three dimensions are sparse, since the number of paths is limited in mmWave channels [37], [38], which is illustrated in Fig.2. A single delay tap may be associated with multiple paths, hence we will first separate the paths in the delay domain. Then the paths with the same delay can be separated by AoAs at the BS when large BS RAs are employed, or by the equivalent AoDs at the RIS when RAs of the BS are insufficiently large. We can then recover the cascaded channel according to (5), once the estimate of these parameters has been obtained by each path.

B. Problem Formulation in the Delay-Domain

For performing channel estimation, the user transmits his/her UL pilots and the BS will receive these signals characterizing the time-varying cascaded two-hop channel. In OFDM systems, the subcarrier spacing is denoted as Δf and the OFDM symbol duration is $T_s = 1/\Delta f$. We choose N_p out of N subcarriers in a comb-type structure to carry the pilots with a spacing of $\Delta = N/N_p$, where we assume that N_p is the divisor of N for simplicity and we have $N_p \geq D$ to combat the frequency selectivity of the channels [39]. Furthermore, we define the sampling interval as $T_s = 1/B$ and hence the duration of an OFDM symbol is $T = NT_s$. Traditional channel estimation methods assumed that the channel remains constant during the time of the entire transmission frame, which ignored the variation of the channel due to Doppler shift. By contrast, we explicitly take this variation among different OFDM blocks into consideration. For ease of exposition, we assume furthermore that the Doppler-induced variation within the duration of a single OFDM symbol is neglected, which is consistent with the authoritative literature [40], [41]. However, the Doppler-induced variation is explicitly considered when generating channels in Section V and our simulation results verify that the channel estimation performance is not sensitive to this

assumption. The beginning of the q -th symbol is denoted as $t^q = (T + T_{cp})(q - 1)$, thus the assumption above means that the values of the N sampling points in the q -th symbol duration are equal, i.e. $\mathbf{H}(t^q, \tau) = \mathbf{H}(t^q + T_s, \tau) = \dots = \mathbf{H}(t^q + (N - 1)T_s, \tau) = \mathbf{H}^q(\tau)$.

Upon denoting the transmitted OFDM symbol in the frequency domain (FD) by \mathbf{s} , the associated pilots are set as $\{s_k^q = 1, k = 0, \Delta, \dots, (N_p - 1)\Delta\}$ for the q -th symbol. After N -point inverse discrete Fourier transform (IDFT) and adding the CP, the signals are mapped to the carrier before they are transmitted. Then, they are transmitted over the cascaded multipath channel and are received by the BS. The BS demodulates them, removes the CP and performs discrete Fourier transform (DFT) based demodulation of the received signals. As we have known, the received signals are given by the circular convolution of the channel and the transmitted signals in the delay-domain [39], which can be transformed into the multiplication operator in the FD. The cascaded channel (5) in the time-delay domain may be rewritten as $\mathbf{H}^q(\tau) = \sum_{l=1}^L \mathbf{H}^q(\tau_l) \delta(\tau - \tau_l)$, where $\mathbf{H}^q(\tau_l)$ denotes the channel at delay τ_l . After sampling, it is discretized into N points $\{\mathbf{H}^q(n)\}_{n=0}^{N-1}$ corresponding to N delay taps with $\mathbf{H}^q(n) = \mathbf{H}^q(\tau_l)$, $n = \tau_l/T_s$ and $\mathbf{H}^q(n) = 0$ otherwise. Then upon applying the DFT to $\{\mathbf{H}^q(n)\}_{n=0}^{N-1}$ for obtaining the cascaded channel at the k -th subcarrier, we have

$$\mathbf{H}_k^q = \sum_{n=0}^{N-1} \mathbf{H}^q(n) e^{-j2\pi \frac{nk}{N}} = \sum_{l=0}^L \mathbf{H}^q(\tau_l) e^{-j2\pi \frac{\tau_l k}{NT_s}}. \quad (7)$$

Furthermore, the RIS RCs are frequency-insensitive within a certain bandwidth [19]–[21]. Hence, the signal $\mathbf{y}_k^q \in \mathbb{C}^{M_b \times 1}$ received at the k -th subcarrier is described as

$$\mathbf{y}_k^q = \mathbf{H}_k^q \boldsymbol{\theta}^q s_k^q + \mathbf{n}_k^q = \mathbf{H}_k^q \boldsymbol{\theta}^q + \mathbf{n}_k^q, \quad (8)$$

where $\mathbf{n}_k^q \in \mathbb{C}^{M_b \times 1}$ and $\boldsymbol{\theta}^q = [\theta_1^q, \dots, \theta_{m_r}^q, \dots, \theta_{M_r}^q]^T \in \mathbb{C}^{M_r \times 1}$ denote the additive white Gaussian noise (AWGN) and RIS RC, respectively. The coefficient $\theta_{m_r}^q = \alpha_{m_r}^q e^{j\beta_{m_r}^q}$, where $\alpha_{m_r}^q \in [0, 1]$ and $\beta_{m_r}^q \in [0, 2\pi)$ represent the amplitude and phase shift of the m_r -th RIS element. We set $\alpha_{m_r} = 1, \forall m_r \in \{1, \dots, M_r\}$ for maximizing the reflected signal power and simplifying the hardware complexity.

Observe from (7) and (8) that the cascaded channel at a specific subcarrier frequency is composed of the paths at all delay taps, where each delay tap is associated with multiple paths. A possible way to decouple the paths is to exploit the delay-domain. When applying the N_p -point IDFT to the received pilots $\{\mathbf{y}_k^q, k = 0, \Delta, \dots, (N_p - 1)\Delta\}$ for characterizing their relationship in the delay-domain, we have

$$\tilde{\mathbf{y}}^q(d) = \sum_{s_d=1}^{S_d} c_{s_d} e^{j2\pi f_{s_d} t^q} \mathbf{a}_{M_b}(\varphi_{s_d}^B) \mathbf{a}_{M_r}^H(\varphi_{s_d}^R) \boldsymbol{\theta}^q + \tilde{\mathbf{n}}^q(d), \quad (9)$$

where $d \in \{0, \dots, D-1\}$ denotes the delay taps. To expound further, we have $\tilde{\mathbf{y}}^q(d) = \sum_{i=1}^{N/N_p} \mathbf{y}^q(d + iN_p)$ and $\tilde{\mathbf{n}}^q(d) = \sum_{i=1}^{N/N_p} \mathbf{n}^q(d + iN_p)$, where $\{\mathbf{y}^q(n), \mathbf{n}^q(n)\}_{n=0}^{N-1}$ represent the received signals and noise. The proof of (9) is provided in Appendix A.

Observe from (9) that the signal received at a certain delay tap is only associated with the channels having the same delay. In other words, the delay-domain expression of (9) serendipitously decouples the mixed paths having different delays. Again, there are also multiple paths at each delay tap, which have different amplitudes and phases in the received signal. This phenomenon is totally different from the single-path case. Therefore, we have to decouple the paths at each delay tap. If we succeed in estimating the parameters $c_{s_l}, f_{s_l}, \varphi_{s_l}^B$ and $\varphi_{s_l}^R$ at each delay tap, then the overall cascaded CIR in (5) can be obtained.

III. PROPOSED ESTIMATION FOR LARGE BS RECEIVER ARRAYS

In this section, we study the scenario, when large RAs are employed at the next-generation BSs, where the high angular resolution can be exploited for distinguishing the UL paths arriving with arbitrary angles at a high degree of accuracy [42]–[44]. Recall that we have separated the coupled paths having different delays and derived the expression (9). The angular-domain is further exploited for separating the paths within a specific delay tap. To simplify our notation, we omit the sign of delay taps d but retain the superscript q for indicating symbol index hereafter in this section. Therefore, the expression (9) can be recast as

$$\tilde{\mathbf{y}}^q = \sum_{s=1}^S c_s e^{j2\pi f_s t^q} \mathbf{a}_{M_b}(\varphi_s^B) \mathbf{a}_{M_r}^H(\varphi_s^R) \boldsymbol{\theta}^q + \tilde{\mathbf{n}}^q. \quad (10)$$

Since there are four sets of parameters $\{c_s, f_s, \varphi_s^B, \varphi_s^R\}_{s=1}^S$ to be estimated, we propose a two-stage estimation strategy, which is depicted in Fig.3, where a total of Q symbols are mapped to a specific transmission frame. The training subframe is used for performing channel estimation, which is divided into two stages: Stage I having Q_1 OFDM symbols and Stage II of Q_2 OFDM symbols. The estimation strategy contains three steps: 1) we estimate the AoAs $\{\varphi_s^B\}_{s=1}^S$ at the BS and decouple the Doppler frequency using the OFDM symbols in Stage I, where the RIS RCs are randomly generated⁴; 2) we estimate the Doppler shifts using the OFDM symbols in Stage II, where the RIS RCs remain unchanged; 3) we perform Doppler compensation for each single path and estimate the equivalent AoDs at the RIS and the complex channel gains using the OFDM symbols received in Stage I again.

A. BS AoAs Estimation and Doppler Decoupling

Firstly, we estimate the AoAs at the BS by exploiting the fine angular resolution of large arrays. Let $\tilde{c}_s^q = c_s e^{j2\pi f_s t^q} \mathbf{a}_{M_r}^H(\varphi_s^R) \boldsymbol{\theta}^q$ and the expression (10) can be rewritten in a compact form as

$$\tilde{\mathbf{y}}^q = \sum_{s=1}^S \tilde{c}_s^q \mathbf{a}_{M_b}(\varphi_s^B) + \tilde{\mathbf{n}}^q. \quad (11)$$

⁴This is an appealing technique of reducing the mutual coherence among columns of the sensing matrix in the expression (23). There are some design methods, which specifically aim for improving the performance, hence the motivated readers might like to refer to [38], [45], [46] for further inspiration.

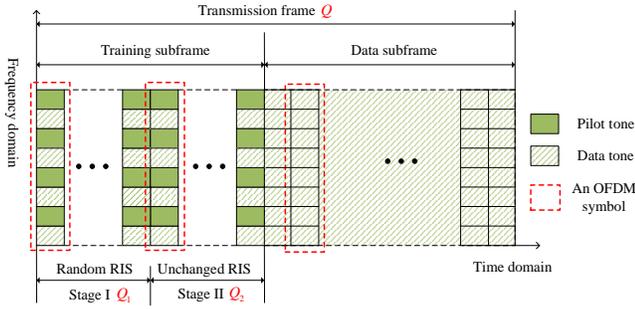


Fig. 3. Transmission strategy for large BS receiver arrays.

It is assumed that the number of paths corresponding to each delay tap is far less than the number of antennas at BS, i.e. $S \ll M_b$. Hence the estimation can be formulated as a typical CS problem, which can be solved by various algorithms, such as convex relaxation exemplified by the least absolute shrinkage and selection operator (LASSO), greedy algorithms like orthogonal matching pursuit (OMP) and iterative Bayesian algorithms like sparse Bayesian learning (SBL) [47]. Generally speaking, the convex algorithms perform well in terms of sparse recovery performance. However, they are more complex [48]. The greedy algorithms are less complex and provide performance guarantees, especially when the number of non-zero entries of the sparse vector is small [48]. The Bayesian algorithms are based on the assumption of a particular prior distribution of the sparse vector. They update the parameters iteratively, which is usually time-consuming [47]. In our settings, the number of paths at a particular delay tap is usually low due to the sparsity of mmWave channels. To elaborate briefly, the number of non-zero entries in the sparse vector in the expressions of (13), (21) and (25) is small, which supports low-complexity implementations under a certain performance guarantee of OMP. Therefore, we opted for the classic OMP as our CS-based recovery algorithms, which is also widely used in the existing literatures [22], [31].

Constructing the dictionary matrix based on a series of steering vectors at the BS as $\mathbf{A}^{M_b} = \left[\mathbf{a}_{M_b} \left(-\frac{1}{2}\right), \mathbf{a}_{M_b} \left(-\frac{1}{2} + \frac{1}{M_b}\right), \dots, \mathbf{a}_{M_b} \left(-\frac{1}{2} + \frac{M_b-1}{M_b}\right) \right] \in \mathbb{C}^{M_b \times M_b}$, so the expression (11) can be formulated as

$$\tilde{\mathbf{y}}^q = \mathbf{A}^{M_b} \tilde{\mathbf{c}}^q + \tilde{\mathbf{n}}^q, \quad (12)$$

for which the angular resolution is $1/M_b$, where $\tilde{\mathbf{c}}^q \in \mathbb{C}^{M_b \times 1}$ denotes the sparse vector associated with S non-zero entries corresponding to $\{c_s e^{j2\pi f_s t^q} \mathbf{a}_{M_r}^H(\varphi_s^R) \boldsymbol{\theta}^q\}_{s=1}^S$. Furthermore, the estimated AoAs can be expressed as $\hat{\varphi}_s^B = -\frac{1}{2} + \frac{i_b-1}{M_b}$, $i_b \in \mathcal{I}^q$, where \mathcal{I}^q denotes the index set of non-zero entries in $\tilde{\mathbf{c}}^q$. Upon stacking the signals $\{\tilde{\mathbf{y}}^q\}_{q=1}^{Q_1}$ received in Stage I we arrive at

$$\tilde{\mathbf{y}} = \tilde{\mathbf{A}} \tilde{\mathbf{c}} + \tilde{\mathbf{n}}, \quad (13)$$

where

$$\tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{A}^{M_b} & 0 & \dots & 0 \\ 0 & \mathbf{A}^{M_b} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{A}^{M_b} \end{bmatrix}, \quad (14)$$

$\tilde{\mathbf{y}} = \left[(\tilde{\mathbf{y}}^1)^T, (\tilde{\mathbf{y}}^2)^T, \dots, (\tilde{\mathbf{y}}^{Q_1})^T \right]^T \in \mathbb{C}^{Q_1 M_b \times 1}$, $\tilde{\mathbf{c}} = \left[(\tilde{\mathbf{c}}^1)^T, (\tilde{\mathbf{c}}^2)^T, \dots, (\tilde{\mathbf{c}}^{Q_1})^T \right]^T \in \mathbb{C}^{Q_1 M_b \times 1}$, and $\tilde{\mathbf{n}} = \left[(\tilde{\mathbf{n}}^1)^T, (\tilde{\mathbf{n}}^2)^T, \dots, (\tilde{\mathbf{n}}^{Q_1})^T \right]^T \in \mathbb{C}^{Q_1 M_b \times 1}$ denote the sensing matrix, observation vector, sparse vector and noise, respectively. Recall that we have assumed that the BS AoAs remain constant during the transmission frame of Fig.3, which indicates that these Q_1 sparse vectors share a common sparsity support, i.e. $\mathcal{I}^1 = \dots = \mathcal{I}^{Q_1} = \mathcal{I}$. Therefore, the AoAs estimation problem can be solved by the multiple measurement vector (MMV)-based OMP algorithm [31].

Once we obtained the index set \mathcal{I} , we can derive the estimate of non-zero entries in $\tilde{\mathbf{c}}^q$ by applying the least squared (LS) algorithm, which is presented as

$$\hat{\mathbf{c}}^q = \left(\left(\mathbf{A}_{\mathcal{I}}^{M_b} \right)^H \mathbf{A}_{\mathcal{I}}^{M_b} \right)^{-1} \left(\mathbf{A}_{\mathcal{I}}^{M_b} \right)^H \tilde{\mathbf{y}}^q \in \mathbb{C}^{S \times 1}, \quad (15)$$

where $\mathbf{A}_{\mathcal{I}}^{M_b} = [\mathbf{A}^{M_b}]_{:, \mathcal{I}}$ denotes the submatrix composed of the specific columns associated with the non-zero index set corresponding to the estimated AoAs. So far, the cascaded paths associated with different Doppler shifts at a single delay tap have been decoupled.

B. Doppler Shift Estimation

In the previous step, we have drawn a distinction among the paths at each delay tap and acquired the estimation of coefficients, which are represented by

$$[\hat{\mathbf{c}}^q]_s = c_s e^{j2\pi f_s t^q} \mathbf{a}_{M_r}^H(\varphi_s^R) \boldsymbol{\theta}^q + [\hat{\mathbf{n}}^q]_s, \quad s = 1, 2, \dots, S, \quad (16)$$

where $\hat{\mathbf{n}}^q = \left(\left(\mathbf{A}_{\mathcal{I}}^{M_b} \right)^H \mathbf{A}_{\mathcal{I}}^{M_b} \right)^{-1} \left(\mathbf{A}_{\mathcal{I}}^{M_b} \right)^H \tilde{\mathbf{n}}^q$ denotes the estimation error. We can find that the difference of $[\hat{\mathbf{c}}^q]_s$ between successive symbols lies in the exponential term containing the Doppler shift, when the RIS RCs remain unchanged.

Therefore, we set $\boldsymbol{\theta}^1 = \dots = \boldsymbol{\theta}^{Q_2} = \boldsymbol{\theta}$ in Stage II of Fig.3 and combine the first and last $Q_2 - 1$ OFDM symbols for mitigating the effect of Doppler estimation error, which are shown as

$$\begin{aligned} r_1 &= \sum_{q=Q_1+1}^{Q-1} [\hat{\mathbf{c}}^q]_s \\ &= \sum_{q=Q_1+1}^{Q-1} c_s e^{j2\pi f_s (q-1)(T+T_{cp})} \mathbf{a}_{M_r}^H(\varphi_s^R) \boldsymbol{\theta}^q + \sum_{q=Q_1+1}^{Q-1} [\hat{\mathbf{n}}^q]_s, \end{aligned} \quad (17)$$

$$\begin{aligned}
r_2 &= \sum_{q=Q_1+2}^Q [\hat{\mathbf{c}}^q]_s \\
&= e^{j2\pi f_s(T+T_{cp})} \sum_{q=Q_1+1}^{Q-1} c_s e^{j2\pi f_s(q-1)(T+T_{cp})} \mathbf{a}_{M_r}^H(\varphi_s^R) \boldsymbol{\theta}^q + \sum_{q=Q_1+2}^Q [\hat{\mathbf{n}}^q]_s,
\end{aligned} \tag{18}$$

whose signal-to-noise (SNR) can be expressed as

$$\begin{aligned}
\text{SNR} &= \frac{|c_s \mathbf{a}_{M_r}^H(\varphi_s^R) \boldsymbol{\theta}^q|^2 \left| \sum_{q=Q_1+1}^{Q-1} e^{j2\pi f_s(q-1)(T+T_{cp})} \right|^2}{\mathbb{E} \left(\left| \sum_{q=Q_1+1}^{Q-1} [\hat{\mathbf{n}}^q]_s \right|^2 \right)} \\
&= \frac{|c_s \mathbf{a}_{M_r}^H(\varphi_s^R) \boldsymbol{\theta}^q|^2 \left| \sum_{q=Q_1+1}^{Q-1} e^{j2\pi f_s(q-1)(T+T_{cp})} \right|^2}{Q_2 - 1},
\end{aligned} \tag{19}$$

which can be improved by appropriately selecting the value of Q_2 , and we can get $\frac{r_2}{r_1} \approx e^{j2\pi f_s(T+T_{cp})}$. Therefore, the estimation of Doppler shift of the s -th path is approximated as

$$\hat{f}_s = \frac{\angle(r_2/r_1)}{2\pi(T+T_{cp})}. \tag{20}$$

C. Equivalent AoDs and Complex Gains Estimation

Now that we have obtained the estimated Doppler shift of each path, we can perform Doppler compensation for mitigating the impact of the Doppler effect and estimate both the equivalent AoDs at the RIS and the cascaded complex gains. As for the s -th path, we multiply both sides of (16) by $e^{-j2\pi \hat{f}_s t^q}$ and acquire the adjusted coefficients, which is expressed as

$$[\bar{\mathbf{c}}^q]_s \approx c_s \mathbf{a}_{M_r}^H(\varphi_s^R) \boldsymbol{\theta}^q + [\bar{\mathbf{n}}^q]_s. \tag{21}$$

Since the number of elements M_r at the RIS is usually large in typical application settings, it possesses a sufficiently fine angular resolution for distinguishing the equivalent AoDs at the RIS. Likewise, in Subsection III-A, we can resort to CS-based algorithms for estimating the equivalent AoDs φ_s^R and gains c_s by making use of the signals in Stage I relying on the random RIS RCs. Let us now define the dictionary matrix associated with the steering vectors at the RIS as

$$\begin{aligned}
\mathbf{A}^{M_r} &= \left[\mathbf{a}_{M_r}^* \left(-\frac{1}{2} \right), \mathbf{a}_{M_r}^* \left(-\frac{1}{2} + \frac{1}{M_r} \right), \dots, \mathbf{a}_{M_r}^* \left(-\frac{1}{2} + \frac{M_r - 1}{M_r} \right) \right] \\
&\in \mathbb{C}^{M_r \times M_r}
\end{aligned} \tag{22}$$

with an angular resolution of $1/M_r$. After collecting Q_1 adjusted coefficients, we have

$$\bar{\mathbf{c}}_s = \Xi \mathbf{c}_s + \bar{\mathbf{n}}_s, \tag{23}$$

where $\bar{\mathbf{c}}_s = [[\bar{\mathbf{c}}^1]_s, \dots, [\bar{\mathbf{c}}^{Q_1}]_s]^T$, $\bar{\mathbf{n}}_s = [[\bar{\mathbf{n}}^1]_s, \dots, [\bar{\mathbf{n}}^{Q_1}]_s]^T$ and the sensing matrix obeys $\Xi = [\boldsymbol{\theta}^1, \boldsymbol{\theta}^2, \dots, \boldsymbol{\theta}^{Q_1}]^T \mathbf{A}^{M_r} \in \mathbb{C}^{Q_1 \times M_r}$. The problem can be readily solved by the classic OMP algorithm, since there is only a single non-zero entry in the

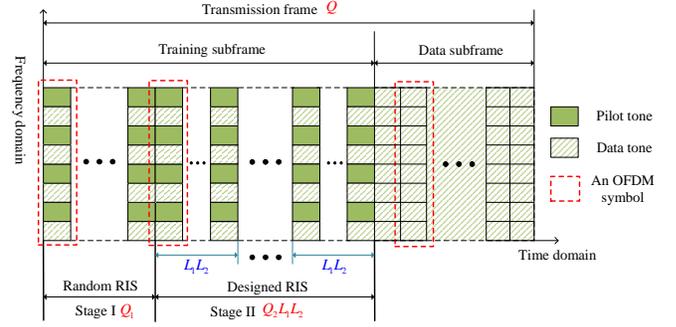


Fig. 4. Transmission strategy for small BS receiver arrays.

sparse vector \mathbf{c}_s , whose index is denoted as i_r . Hence, the corresponding equivalent AoD and the complex gain can be calculated as $\hat{\varphi}_s^R = -\frac{1}{2} + \frac{i_r - 1}{M_r}$ and $\hat{c}_s = \left(\left([\Xi]_{:,i_r} \right)^H [\Xi]_{:,i_r} \right)^{-1} \left([\Xi]_{:,i_r} \right)^H \bar{\mathbf{c}}_s$, respectively.

After obtaining the parameters $\left\{ \hat{c}_s, \hat{f}_s, \hat{\varphi}_s^B, \hat{\varphi}_s^R \right\}_{s=1}^{S_d}$, $\forall d \in \{0, 1, \dots, D-1\}$, the overall cascaded channel can be recovered according to expression (5).

IV. PROPOSED ESTIMATION STRATEGY FOR SMALL BS RECEIVER ARRAYS

The powerful massive MIMO technology incurs high energy consumption and hardware complexity, which may however be significantly reduced under the assistance of RISs [7]. Accordingly in this section, we discuss the estimation of the cascaded channel for small BS RA arrays. Initially a single antenna is considered and the algorithm then can be extended to multiple antennas by parallel processing, whose complexity is still acceptable owing to using only a few BS RAs. The notation d for the delay tap is placed in the subscript and expression (9) can be rewritten as

$$\tilde{\mathbf{y}}_d^q = \sum_{s_d=1}^{S_d} c_{s_d} e^{j2\pi f_{s_d} t^q} \mathbf{a}_{M_r}^H(\varphi_{s_d}^R) \boldsymbol{\theta}^q + \tilde{\mathbf{n}}_d^q. \tag{24}$$

It is not feasible to separate the paths associated with a given delay tap by exploiting the angular resolution of the BS array, since it is not fine enough for a few antennas. Hence the algorithm of Section III is unsuitable, which motivates us to conceive a new estimation method. The extra degree of freedom provided by RISs captures our attention, whose reflection pattern can be delicately designed in support of channel estimation [20].

Therefore, we propose a modified estimation scheme, as shown in Fig.4. It contains three steps: 1) we coarsely estimate the equivalent AoDs $\hat{\varphi}_{s_d}^R$ by neglecting the phase shift imposed by the Doppler effect using the OFDM symbols received in Stage I, where the RIS RCs are randomly generated; 2) we estimate the Doppler shift using the OFDM symbols received in Stage II, where the RIS RCs are designed to be a specific pattern associated with $\hat{\varphi}_{s_d}^R$; 3) we then proceed by refining the estimates of the equivalent AoDs and complex gains with the aid of the estimated Dopplers using the OFDM symbols in Stage I again.

A. Preliminary Estimation of the Equivalent AoDs

The RIS-based passive beamforming possesses high angular resolution given its numerous elements, which can be exploited for separating the different cascaded paths having different equivalent AoDs at the RIS. However, by observing expression (24), we find that the term containing the Doppler frequency is time-varying, hence it is challenging to organize the problem into a CS-based form. Here, we temporarily assume that this term is time-invariant in Stage I, which has a trivial impact on the estimation of angles over a considerable range of Doppler shift, which will be expounded on in Section V.

Thus, by adopting the same dictionary matrix as shown in (22) and collecting Q_1 OFDM symbols received at the BS, we have

$$\tilde{\mathbf{y}}_d = \Xi \tilde{\mathbf{c}}_d + \tilde{\mathbf{n}}_d, \quad (25)$$

where $\tilde{\mathbf{y}}_d = [\tilde{y}_d^1, \tilde{y}_d^2, \dots, \tilde{y}_d^{Q_1}]^T \in \mathbb{C}^{Q_1 \times 1}$, $\tilde{\mathbf{n}}_d = [\tilde{n}_d^1, \tilde{n}_d^2, \dots, \tilde{n}_d^{Q_1}]^T \in \mathbb{C}^{Q_1 \times 1}$ and Ξ is consistent with the one in (23). Furthermore, $\tilde{\mathbf{c}}_d$ is the sparse vector with S_d non-zero entries, each of which approximately equals to $c_{s_d} e^{j2\pi f_{s_d} t} \approx c_{s_d} e^{j2\pi f_{s_d} t^1} \approx \dots \approx c_{s_d} e^{j2\pi f_{s_d} t^{Q_1}}$. The problem may then be solved by OMP as well. After getting the index set of non-zero entries \mathcal{I} , we could obtain a rough estimate of the equivalent AoDs of the s_d -th path in the d -th delay tap, which is calculated as $\hat{\varphi}_{s_d}^R = -\frac{1}{2} + \frac{i_{s_d}-1}{M_r}$, where $i_{s_d} \in \mathcal{I}$, $s_d = 1, \dots, S_d$.

B. Doppler Estimation by Designing the RIS Reflection Pattern

Then, it becomes feasible to utilize the OFDM symbols in Stage II to perform Doppler estimation, provided that the paths associated with each delay tap can be separated. Specifically, the steering vectors at the RIS associated with different equivalent AoDs are asymptotically orthogonal, which can be exploited to design the RIS reflection pattern for mitigating the interference of other paths and to extract a certain path.

Accordingly, we set $\boldsymbol{\theta}^q = \mathbf{a}_{M_r}(\hat{\varphi}_{s_d}^R)$ during the q -th OFDM symbol time. Then the received signal will be expressed as

$$\begin{aligned} \tilde{\mathbf{y}}_d^q &= \sum_{s'_d=1}^{S_d} c_{s'_d} e^{j2\pi f_{s'_d} t^q} \mathbf{a}_{M_r}^H(\varphi_{s'_d}^R) \mathbf{a}_{M_r}(\hat{\varphi}_{s_d}^R) + \tilde{n}_d^q \\ &= \underbrace{c_{s_d} e^{j2\pi f_{s_d} t^q} \mathbf{a}_{M_r}^H(\varphi_{s_d}^R) \mathbf{a}_{M_r}(\hat{\varphi}_{s_d}^R)}_{\text{desired signal}} \\ &\quad + \underbrace{\sum_{s'_d \neq s_d} c_{s'_d} e^{j2\pi f_{s'_d} t^q} \mathbf{a}_{M_r}^H(\varphi_{s'_d}^R) \mathbf{a}_{M_r}(\hat{\varphi}_{s_d}^R)}_{\text{interference}} + \tilde{n}_d^q \\ &\stackrel{(b)}{\approx} c_{s_d} e^{j2\pi f_{s_d} t^q} M_r + \tilde{n}_d^q, \end{aligned} \quad (26)$$

where (b) holds due to the orthogonality of steering vectors, Hence the amplitude of the desired signal is much higher than that of the interference imposed by other paths. In other words, when the passive RIS beam is aligned with the estimated equivalent AoD of the s_d -th cascaded path, most of the signal

energy received at the BS is gleaned from the s_d -th path, which helps us to separate the coupled paths.

At this point we can estimate the Doppler shift of each individual path by calculating the phase difference of signals received in different OFDM symbols. If we use the same method as for the large array counterpart of Section III, Q_2 consecutive OFDM symbols are used for estimating each equivalent AoD $\hat{\varphi}_{s_d}^R$. The corresponding pilot overhead will become $L_1 L_2 Q_2$, since there are $L_1 L_2$ paths in total considering all delay taps, where Q_2 tends to be a dozen or so (verified in Section V), which is too high to be acceptable. Thus, we propose an interleaved method for estimating the Doppler shifts, which is depicted in Fig.4. The OFDM symbols in Stage II of Fig.4 are divided into Q_2 blocks and each block contains $L_1 L_2$ OFDM symbols. We set $\boldsymbol{\theta}^q = \mathbf{a}_{M_r}^H(\hat{\varphi}_{s_d}^R)$ for $q = Q_1 + \sum_{d'=1}^{d-1} S_{d'} + s_d + (q_2 - 1)L_1 L_2$, $q_2 = 1, 2, \dots, Q_2$. This means that the RIS RCs are set to the same value associated with a particular estimated equivalent AoD with an interval of $L_1 L_2$ OFDM symbols. Upon summing up the first and last $Q_2 - 1$ signals received at the d -th delay tap we get

$$r_{1,d} = \sum_{q_2=1}^{Q_2-1} \tilde{y}_d^q = \sum_{q_2=1}^{Q_2-1} c_{s_d} e^{j2\pi f_{s_d} (q-1)(T+T_{cp})} M_r + \sum_{q_2=1}^{Q_2-1} \tilde{n}_d^q, \quad (27)$$

$$\begin{aligned} r_{2,d} &= \sum_{q_2=2}^{Q_2} \tilde{y}_d^q \\ &= e^{j2\pi f_{s_d} L_1 L_2 (T+T_{cp})} \sum_{q_2=1}^{Q_2-1} c_{s_d} e^{j2\pi f_{s_d} (q-1)(T+T_{cp})} M_r + \sum_{q_2=2}^{Q_2} \tilde{n}_d^q. \end{aligned} \quad (28)$$

Due to the reasons argued in Section III-B, the estimated Doppler shift of the s_d -th path at the d -th delay tap approximately becomes

$$\hat{f}_{s_d} = \frac{\angle(r_{2,d}/r_{1,d})}{2\pi L_1 L_2 (T + T_{cp})}. \quad (29)$$

We found that the estimation error of the interleaved method is lower than that of the consecutive one when adopting the same value of Q_2 , since the denominator of (29) is larger than that in (20). Therefore, our proposed interleaved Doppler estimation strategy can achieve the same precision with less OFDM symbols, which substantially reduces the overhead.

C. Refined Estimation of the Equivalent AoDs and Complex Gains

Recall that we have obtained the rough estimates of the equivalent AoDs at the RIS in the first step by neglecting the phase shift induced by the Doppler, which can then be refined by exploiting the estimated Doppler shifts in the second step.

We resort to the signals $\{\tilde{y}_d^q\}_{q=1}^{Q_1}$ received in Stage I again, from which the expression (24) is reformulated as

$$\tilde{\mathbf{y}}_d = \mathbf{\Omega}_d \mathbf{c}_d + \tilde{\mathbf{n}}_d, \quad (30)$$

where $\mathbf{\Omega}_d$ is shown at the top of the next page and

$$\mathbf{\Omega}_d = \begin{bmatrix} e^{j2\pi\hat{f}_1 t^1} \mathbf{a}_{M_r}^H(\varphi_1^R) \boldsymbol{\theta}^1 & \dots & e^{j2\pi\hat{f}_{s_d} t^1} \mathbf{a}_{M_r}^H(\varphi_{s_d}^R) \boldsymbol{\theta}^1 & \dots & e^{j2\pi\hat{f}_{s_d} t^1} \mathbf{a}_{M_r}^H(\varphi_{S_d}^R) \boldsymbol{\theta}^1 \\ e^{j2\pi\hat{f}_1 t^2} \mathbf{a}_{M_r}^H(\varphi_1^R) \boldsymbol{\theta}^2 & \dots & e^{j2\pi\hat{f}_{s_d} t^2} \mathbf{a}_{M_r}^H(\varphi_{s_d}^R) \boldsymbol{\theta}^2 & \dots & e^{j2\pi\hat{f}_{s_d} t^2} \mathbf{a}_{M_r}^H(\varphi_{S_d}^R) \boldsymbol{\theta}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ e^{j2\pi\hat{f}_1 t^{Q_1}} \mathbf{a}_{M_r}^H(\varphi_1^R) \boldsymbol{\theta}^{Q_1} & \dots & e^{j2\pi\hat{f}_{s_d} t^{Q_1}} \mathbf{a}_{M_r}^H(\varphi_{s_d}^R) \boldsymbol{\theta}^{Q_1} & \dots & e^{j2\pi\hat{f}_{s_d} t^{Q_1}} \mathbf{a}_{M_r}^H(\varphi_{S_d}^R) \boldsymbol{\theta}^{Q_1} \end{bmatrix} \in \mathbb{C}^{Q_1 \times S_d}, \quad (31)$$

$\tilde{\mathbf{y}}_d = [\tilde{y}_d^1, \tilde{y}_d^2, \dots, \tilde{y}_d^{Q_1}]^T$, $\tilde{\mathbf{n}}_d = [\tilde{n}_d^1, \tilde{n}_d^2, \dots, \tilde{n}_d^{Q_1}]^T$, $\mathbf{c}_d = [c_1, c_2, \dots, c_{S_d}]^T \in \mathbb{C}^{S_d \times 1}$. The optimal equivalent AoDs $\Psi = \{\varphi_{s_d}^R\}_{s_d=1}^{S_d}$ and the equivalent complex gains $\{c_{s_d}\}_{s_d=1}^{S_d}$ are the ones that make both sides of the equation (30) as similar as possible. Therefore, the optimization problem can be described as

$$\{\Psi, \mathbf{c}_d\} = \arg \min_{\Psi, \mathbf{c}_d} \|\tilde{\mathbf{y}}_d - \mathbf{\Omega}_d \mathbf{c}_d\|_2^2, \quad (32)$$

$$s.t. \varphi_{s_d}^R \in \left[-\frac{1}{2}, \frac{1}{2}\right), \forall s_d = 1, \dots, S_d. \quad (33)$$

In order to find the two groups of most appropriate parameters, we refer to the classic alternative optimization method. Firstly, we fix \mathbf{c}_d and apply the popular gradient descent (GD) method for minimizing the objective function of (32) subject to the constraints of (33). Let us now set the derivative of the objective function with respect to $\varphi_{s_d}^R$ as follows:

$$(g_{s_d}^R)^{n-1} = \frac{\partial \|\tilde{\mathbf{y}}_d - (\mathbf{\Omega}_d)^{n-1} (\mathbf{c}_d)^{n-1}\|_2^2}{\partial \varphi_{s_d}^R}, \quad (34)$$

where the superscript n represents the n -th iteration. The detailed derivation is provided in Appendix B. Thus, we can update the estimate of $\varphi_{s_d}^R$ as

$$(\varphi_{s_d}^R)^n = (\varphi_{s_d}^R)^{n-1} - \eta (g_{s_d}^R)^{n-1}, \quad (35)$$

where η denotes the step size. Additionally, note that since the value of $\varphi_{s_d}^R$ is restricted to the range $[-1/2, 1/2)$, it is necessary to perform the following operation

$$(\varphi_{s_d}^R)^n = \begin{cases} (\varphi_{s_d}^R)^n - 1, & (\varphi_{s_d}^R)^n \geq 1/2 \\ (\varphi_{s_d}^R)^n + 1, & (\varphi_{s_d}^R)^n < -1/2 \\ (\varphi_{s_d}^R)^n, & \text{otherwise} \end{cases}. \quad (36)$$

After acquiring the estimate of $\varphi_{s_d}^R$, we can update the values of the equivalent complex gain based on the LS estimator, which is formulated as

$$(\mathbf{c}_d)^n = \left(((\mathbf{\Omega}_d)^n)^H (\mathbf{\Omega}_d)^n \right)^{-1} ((\mathbf{\Omega}_d)^n)^H \tilde{\mathbf{y}}_d. \quad (37)$$

The initial values of the equivalent AoDs are set to the ones estimated in the first step, hence we have $(\varphi_{s_d}^R)^0 = \hat{\varphi}_{s_d}^R$. Accordingly, $(\mathbf{\Omega}_d)^0$ and $(\mathbf{c}_d)^0$ can be derived from the expressions (31) and (37). After a number of iterations, more precise estimates of the equivalent AoD and of the complex gain of each path can be obtained.

TABLE II
SYSTEM PARAMETERS IN THE SIMULATIONS

Parameter	Values
Carrier frequency f_c	30 GHz
Subcarrier spacing B	240 KHz
The number of subcarriers N	256
The number of pilot subcarriers N_p	16
Maximum cascaded delay τ_D	250 ns
Velocity of the user	80 km/h
Doppler spread D_s	4500 Hz
The number of RIS elements M_r	100
The number of paths of BS-RIS link L_1	3
The number of paths of RIS-User link L_2	5
The number of OFDM symbols in relevance interval	300

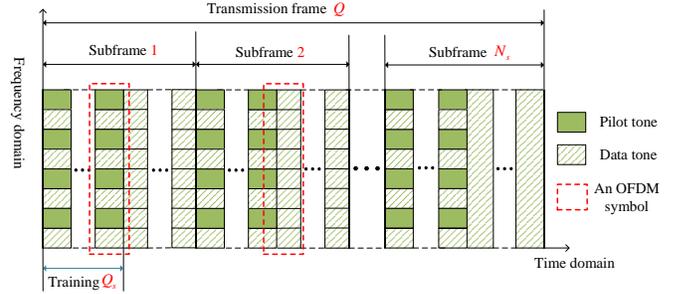


Fig. 5. Transmission strategy proposed in [22].

V. SIMULATION RESULTS

In this section, we will numerically characterize the channel estimation performance of the proposed methods for both scenarios. The overall parameter settings for our simulations are presented in Table II [22], [49], [50]. Therefore, the OFDM symbol duration is $T_s = \frac{1}{\Delta f} = 4.2 \mu s$. In IoV scenarios, the vehicles usually move at speeds between 50-100 km/h [51], thus the maximum Doppler shift will be 1400-2800 Hz when communicating at a carrier frequency of 30 GHz. The simulation results are mainly recorded at the speed of 80 km/h, so the Doppler spread is about $D_s = 4500$ Hz and the channel coherence time is $T_c = \frac{1}{4D_s} \approx 55.6 \mu s$. As to the delay spread for IoV, it is shown in [52] that the vehicles' blocking effect does not significantly affect the DS. So we set the maximum cascaded delay τ_D to be around 250 ns [53], which results in the coherence bandwidth of $W_c = \frac{1}{2\tau_D} = 2000$ KHz [35]. Besides, the physical AoAs and AoDs are generated randomly from $[-\pi/2, \pi/2)$ and the complex gains are also random and normalized to 1.

For comparison, we consider the channel estimation algorithm proposed in [22] as the benchmark, which takes advantage of the sparsity of the channel in the angular domain, but the authors did not take the Doppler effect into consideration.

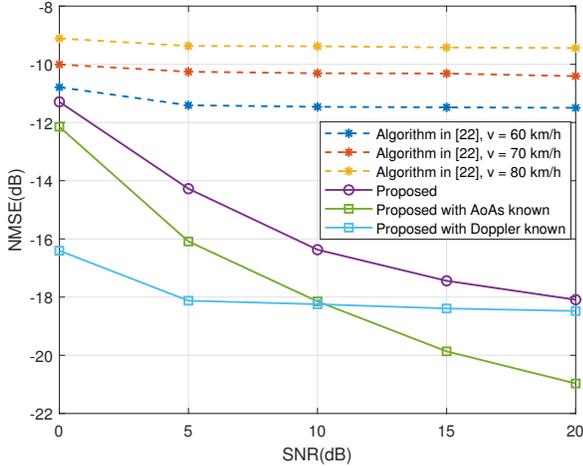


Fig. 6. Channel estimation performance against SNR.

The transmission strategy of [22] is presented in Fig.5, where the total transmission frame is divided into N_s subframes. During each of these, the first Q_s OFDM symbols are used for performing channel estimation and the remaining OFDM symbols are used for pure data. In Table III, we compare the computational complexity of the proposed algorithms and the benchmark.

The normalized mean squared error (NMSE) of the cascaded channel estimation can be expressed as

$$\text{NMSE}_{\mathbf{H}} = \frac{\mathbb{E} \left(\sum_{q=1}^Q \sum_{d=0}^{D-1} \left\| \hat{\mathbf{H}}^q(\tau_d) - \mathbf{H}^q(\tau_d) \right\|_2^2 \right)}{\mathbb{E} \left(\sum_{q=1}^Q \sum_{d=0}^{D-1} \left\| \mathbf{H}^q(\tau_d) \right\|_2^2 \right)}, \quad (38)$$

where $\hat{\mathbf{H}}^q(\tau_d)$ denotes the estimate of the cascaded channel at the d -th delay tap during the q -th OFDM symbol. Likewise, the NMSE of the Doppler shifts, AoAs, equivalent AoDs and complex gains are defined similarly to (38).

A. Performance of the Algorithm Proposed for Large BS Arrays

In this subsection, we investigate the performance of our channel estimation method proposed for large BS arrays. The number of RAs M_b at the BS is set to 30. The other parameters are consistent with the descriptions above. In addition to [22], we also compare the proposed algorithm to two other baselines, one for perfectly known AoAs at the BS and the other for known Doppler shifts.

Fig.6 shows the channel estimation performance vs. SNR. In this simulation, $Q_1 = 18$ symbols are utilized in Stage I for estimating the equivalent angles and complex gains with the aid of randomly generated RIS RCs and $Q_2 = 12$ symbols are used in Stage II for estimating the Doppler shifts using the unchanged RIS reflection pattern. As for the method in [22], the number of subframes is set to $N_s = 3$, where the number of training symbols within each subframe is $Q_s = 30$. In order to compare the performance gain of our proposed algorithm over the traditional one, we simulate the performance of the

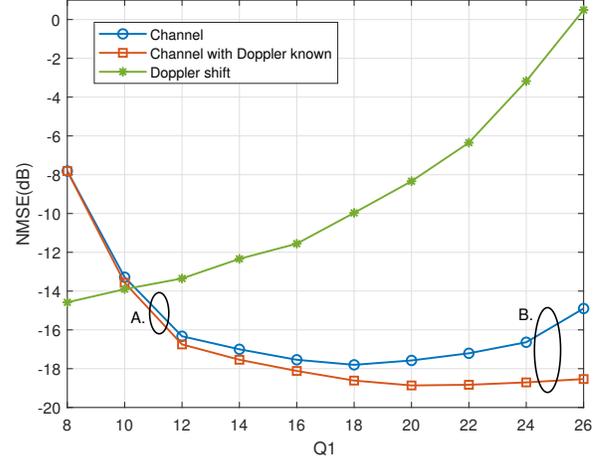


Fig. 7. Channel estimation performance against the number of OFDM symbols in Stage I with $Q = 30$.

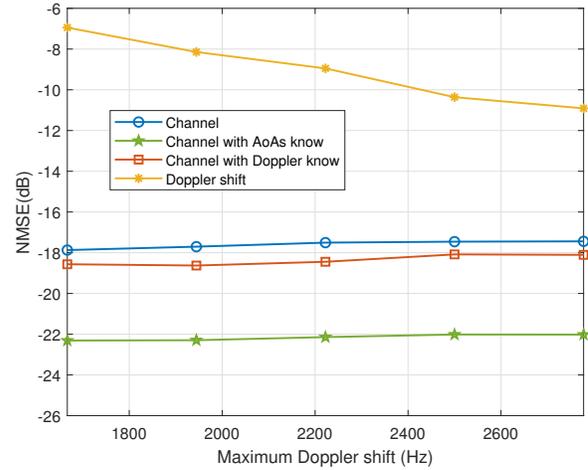
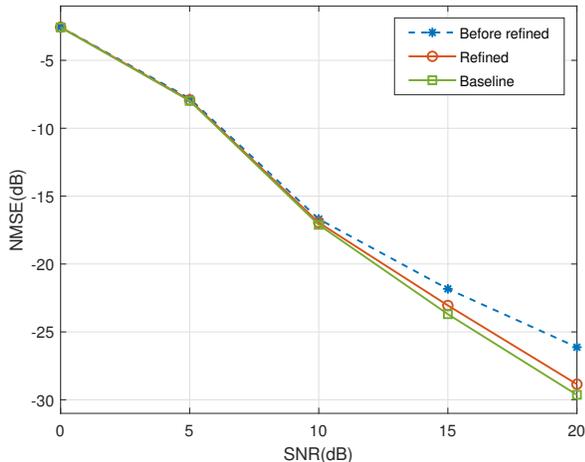


Fig. 8. Channel estimation performance against the maximum Doppler shift with SNR=15dB.

traditional method for user velocity of 60 km/h, 70 km/h and 80 km/h. The results are shown in dashed lines from top to bottom, respectively. The solid line labelled as *Proposed* represents the NMSE of the estimated channel for $v = 80$ km/h based on our strategy proposed in Section III. The curve labelled as *Proposed with AoAs known* represents the channel estimation results, when the AoAs are known as a prior at BS, while the curve labelled as *Proposed with Doppler known* is the one, when Doppler shift of each path is known. Upon increasing the velocity (maximum Doppler shift), the performance of the traditional algorithm in [22] degrades significantly due to ignoring the Doppler effect. By contrast, our proposed method performs well as a benefit of Doppler compensation. Furthermore, we can find that the eroded accuracy of Doppler shift estimation at low SNRs is the main limitation of the channel estimation performance. However, the main limitation becomes the accuracy of angle estimation at high SNRs. This is consistent with the method introduced in Subsection III-B, where the approximation accuracy in (20)

TABLE III
COMPLEXITY COMPARISON OF DIFFERENT ESTIMATION ALGORITHMS

Algorithm	Complexity
Benchmark in [22]	$\mathcal{O}(L_1 L_2 (N_s Q_s)^2 M_b^2 + L_1 L_2 M_r N_s Q_s)$
Proposed algorithm for large BS array	$\mathcal{O}(L_1 L_2 Q_1^2 M_b^2 + L_1 L_2 Q_2 + L_1 L_2 M_r Q_1)$
Proposed algorithm for small BS array	$\mathcal{O}(Q_1 M_r L_1 L_2 + L_1 L_2 Q_2 + N_{\text{iter}}(M_r Q_1 + Q_1 (L_1 L_2)^2))$



(a) Angle estimation performance against SNR.

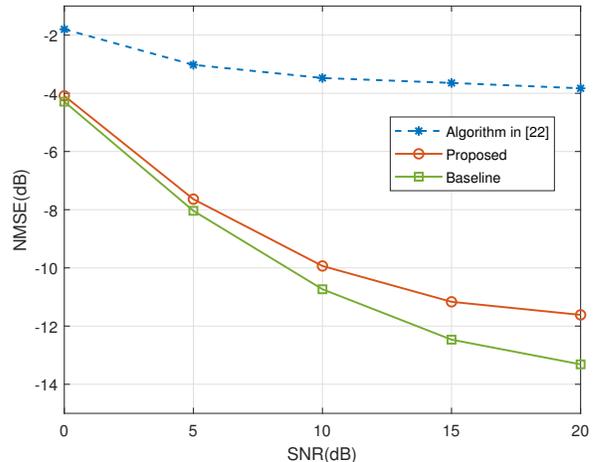
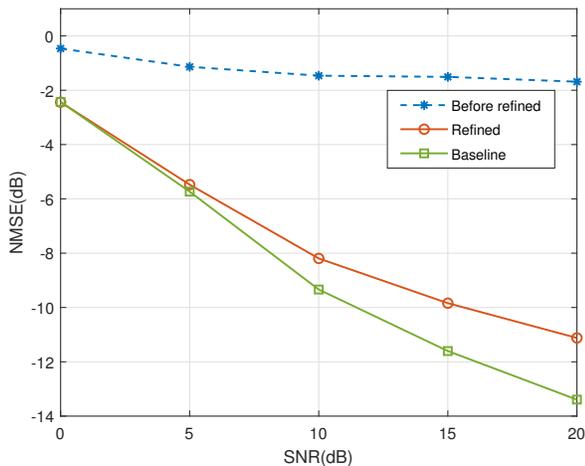


Fig. 10. Channel estimation performance against SNR.



(b) Gain estimation performance against SNR.

Fig. 9. The effect of preliminary estimation in Section IV.

critically depends on the SNR. As for the pilot overhead, the number of training symbols for our proposed method is 30, which is far less than the overhead of the traditional one given by $N_s Q_s = 90$.

In Fig.7, we investigate the optimal sharing of the number of OFDM training symbols between Stage I and Stage II. As described before, Q_1 symbols are used in Stage I to estimate the AoAs, while the equivalent AoDs at the RIS as well as the complex gains and the Q_2 symbols are used in Stage II to estimate the Doppler frequency shifts. The number of symbols $Q = Q_1 + Q_2$ in the training subframe is fixed. The curve labelled as *Doppler shift* represents the Doppler shift estimation performance and the curve labelled

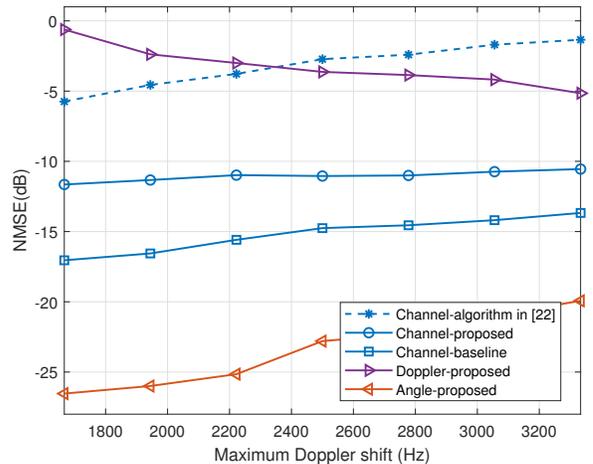


Fig. 11. Channel estimation performance against the maximum Doppler shift with SNR=15dB.

as *Channel* represents the NMSE of the estimated cascaded channel. We observe in Fig.7 that the channel estimation performance first improves and then degrades upon increasing Q_1 . When $Q_1 \leq 18$, the increased number of measurements in Stage I boosts the accuracy of the channel angles and gains, which improves the overall channel estimation performance. However, when $Q_1 > 18$, the reduction in the number of symbols Q_2 used for estimating the Doppler shifts in Stage II leads to a reduced Doppler estimation accuracy, where the NMSE of Doppler estimation is higher than -10dB. The poor Doppler estimation accuracy will degrade the entire channel estimation performance. Thus, with the increase of Q_1 , the

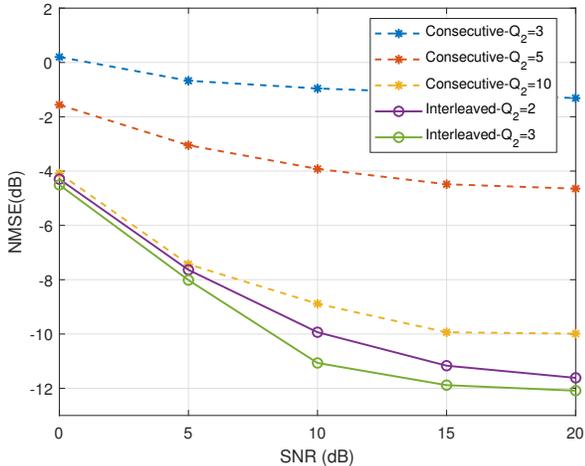


Fig. 12. Channel estimation performance with different RIS design patterns.

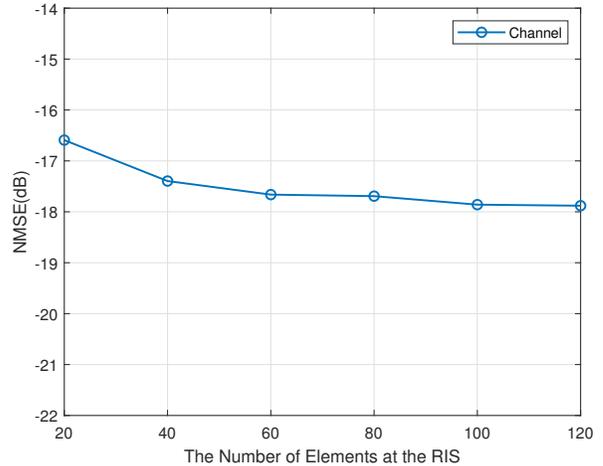


Fig. 14. Channel estimation performance against the number of antennas at the RIS.

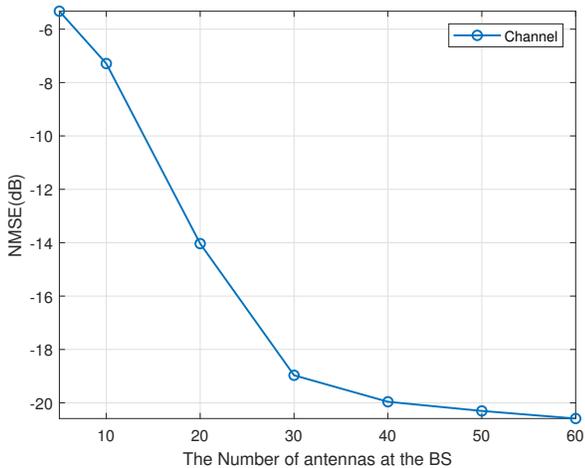


Fig. 13. Channel estimation performance against the number of antennas at the BS.

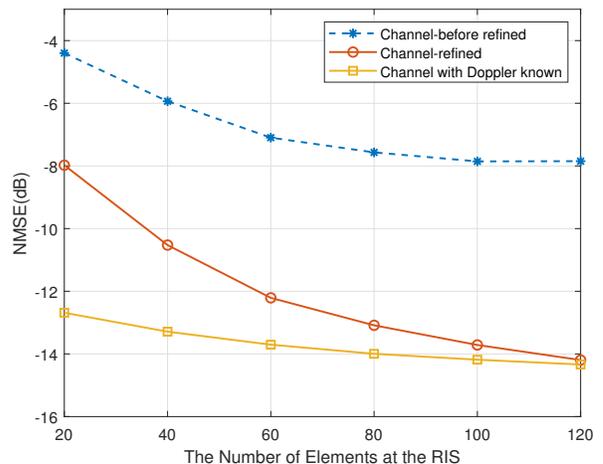


Fig. 15. Channel estimation performance against the number of antennas at the RIS.

NMSE performance first increases and then decreases. We also simulate the baseline when the channel estimator relies on perfectly known Doppler frequency shifts. Compared to points A and B in Fig.7, the gap between the proposed and the baseline solutions widens with the increase of Q_1 . More explicitly, the angle and gain estimation performance is the main factor affecting the overall channel estimation performance when Q_1 is small, while the impact of Doppler estimation performance prevails, when Q_1 is large.

In addition to that, we evaluate the performance of the algorithm for different maximum Doppler shifts spanning from about 1700 Hz to 2700 Hz (the corresponding speed is 60-80 km/h), which is shown in Fig.8. The curves are defined similarly to Fig.7. Upon increasing f_m , the estimation performance of the cascaded channel still remains stable, yielding more precise Doppler estimation. On the one hand, this verifies that our algorithm proposed for large BS arrays in high-mobility scenarios is robust to the Doppler effect. On the other hand, combined with Fig.7, we may conclude that the

estimation of Doppler has little impact on the overall channel estimation, provided that the accuracy of Doppler estimation is within a certain limit.

B. Performance of the Algorithm Proposed for Small BS Arrays

In this subsection, we present simulation results for characterizing the performance of the algorithm proposed in Section IV for small BS arrays. We choose the estimation performance with known Doppler shifts as our baseline. The value of Q_1 is set to 20 as well. Recall that we first roughly estimate the equivalent AoDs at the RIS by neglecting the Doppler effect and then refine the estimate by taking the Doppler shifts into consideration in the third step.

To elaborate, Fig.9 (a) and (b) show the estimation performance of the equivalent AoDs at the RIS and complex gains against SNR, respectively. The curve labelled as *Before refined* represents the coarse estimate in the first step, while and the curve labelled as *Refined* represents the estimate after

adjusting them with the aid of the Doppler in the third step. Observe that the angle estimate is similar before and after refining. However, the estimate of the equivalent complex gains improves substantially after adjustment, along with the Doppler estimation.

Fig.10 presents the performance of the cascaded channel's estimation for a user velocity of 80 km/h. In the algorithm of [22], we set $Q_s = 20$ and $N_s = 3$. Compared to the traditional one, our proposed method significantly improves the performance and approaches the baseline.

Furthermore, the MMSE performance vs. the maximum Doppler shift is investigated in Fig.11. Observe that the estimate of the equivalent AoDs at the RIS degrades upon increasing the maximum Doppler shifts. However, the estimated NMSE of the cascaded channel, although slightly deteriorated, remains generally low, which is a benefit of the accurate Doppler estimation. Our proposed algorithm is relatively robust to a wide range of Doppler shifts.

Fig.12 compares the performance of the pair of RIS reflection patterns designed. The curves *Consecutive* represent estimating the Doppler shift by calculating the phase difference of $Q_2 - 1$ consecutive symbols, while the curve *Interleaved* is the one adopting the interleaved mode. The number of training symbols used for estimating channels at all delay taps is $Q_1 + Q_2 L_1 L_2$. As for the consecutive processing method, 170 OFDM symbols with $Q_2 = 10$ are required for achieving a relative accurate estimation, which may be deemed excessive compared to all the 300 OFDM symbols in the relevance interval. As for the interleaved design introduced in Section IV-B, the number of training symbols may be reduced to 50 with $Q_2 = 2$, which is 16% of the number of total symbols during the relevance interval.

C. The Impact of Channel Dimension on Estimation Performance

The channel dimension is determined by the number of antennas at the BS and the number of elements at the RIS. When the BS array is large, we can refer to the algorithm proposed in Section III, which separates the paths by exploiting the angular selectivity of the BS array. However, we refer to the algorithm of Section IV, when the angular resolution is insufficient. For a small BS array scenario, the number of antennas at the BS does not impact the channel estimation performance due to parallel processing mechanism at BS. We first simulate the channel estimation performance versus the number of antennas at the BS for large BS array scenarios by applying the algorithm of Section III. As verified by Fig.13, the channel estimation performance becomes better upon increasing the number of antennas at the BS. This is because a large BS array provides a higher angular resolution for distinguishing the arriving paths.

Then we investigate the channel estimation performance versus the number of elements at the RIS for both the large BS array scenario and the small BS array scenario. Fig.14 shows the estimation performance for large BS array scenarios by using the algorithm proposed in Section III. We observe that the RIS dimension only has a modest impact on the estimation performance. This is because the number of RIS

elements mainly affects the passive beamforming gain, when conducting channel estimation in Section III, which is usually large enough [6], [7].

However, the RIS dimension has a significant impact on the channel estimation performance using the algorithm proposed for small BS arrays in Section IV. Fig.15 shows the channel estimation performance versus the number of RIS elements. As the number of RIS elements increases, the performance of the overall channel estimation gradually improves and gets closer to the baseline when the Doppler frequency shifts are known. Recall that we estimate the Doppler frequency shifts by exploiting the asymptotical orthogonality of the steering vectors of the RIS array. Upon increasing the RIS size, the orthogonality of the RIS array is enhanced, and the Doppler estimation becomes more accurate, which improves the overall channel estimation performance.

VI. CONCLUSIONS

In this paper, we have studied the channel estimation issues of RIS-assisted dispersive, high-Doppler IoV systems in two-hop channels. In order to separate the Doppler shifts associated with each path, we solved the problem in the delay-domain and proposed a pair of strategies for both large and small BS arrays. The high angular resolution of large BS arrays was utilized to separate paths at each delay tap but small arrays have limited angular resolution. Hence we adopted an interleaved method to design the reflection pattern of the RIS to help distinguish the paths. After Doppler compensation, the performance of channel estimation improved significantly. Furthermore, due to the angular sparsity, CS-based algorithms were applied for estimating the angles and gains, which remarkably reduced the overhead. Our simulation results demonstrated that our proposed methods overperform their traditional counterparts.

APPENDIX A

For better illustration, we stack N received signals as $\mathbf{Y}^t = [\mathbf{y}(0), \dots, \mathbf{y}(n), \dots, \mathbf{y}(N-1)]$, where $\mathbf{y}(n) = \mathbf{y}(nT_s)$ is the n -th sampling point. We then apply the DFT to \mathbf{Y}^t for obtaining the corresponding signals $\mathbf{Y}^f = \mathbf{Y}^t \mathbf{W}^N$ in the frequency domain, where $\mathbf{Y}^f = [\mathbf{y}_0, \dots, \mathbf{y}_N]$ and \mathbf{W}^N denotes the N -point DFT matrix. Recall that we uniformly choose N_p subcarriers to carry pilots, so the N_p pilot signals received at the BS are denoted as $\mathbf{Y}_{N_p}^f = [\mathbf{y}_0, \mathbf{y}_\Delta, \dots, \mathbf{y}_{(N_p-1)\Delta}]$. In order to obtain the delay-domain signals, we then apply the N_p -point IDFT to $\mathbf{Y}_{N_p}^f$, which is expressed as

$$\tilde{\mathbf{Y}} = \frac{1}{N_p} \mathbf{Y}_{N_p}^f (\mathbf{W}^{N_p})^H, \quad (39)$$

where $\tilde{\mathbf{Y}} = [\tilde{\mathbf{y}}(0), \dots, \tilde{\mathbf{y}}(d), \dots, \tilde{\mathbf{y}}(N_p-1)]$ with $\tilde{\mathbf{y}}(d)$ denoting the signal at the d -th delay tap, which can be further

expressed as

$$\begin{aligned}
\tilde{\mathbf{y}}(d) &= \frac{1}{N_p} \sum_{p=0}^{N_p-1} \mathbf{y}_{p \frac{N}{N_p}} \cdot e^{j2\pi \frac{dp}{N_p}} \\
&\stackrel{(a)}{=} \frac{1}{N_p} \sum_{p=0}^{N_p-1} \left(\sum_{n=0}^{N-1} \mathbf{y}(n) e^{-j2\pi \frac{pn}{N_p}} \right) e^{j2\pi \frac{dp}{N_p}} \\
&= \sum_{n=0}^{N-1} \mathbf{y}(n) \frac{1}{N_p} \frac{1 - e^{-j2\pi(n-d)}}{1 - e^{-j2\pi \frac{n-d}{N_p}}} \\
&\stackrel{(b)}{=} \sum_{i=0}^{\frac{N}{N_p}-1} \mathbf{y}(d + iN_p) \tag{40}
\end{aligned}$$

where (a) is derived by the DFT from time-domain signals to the corresponding subcarrier-related frequency-domain signals, and (b) holds true because we have

$$\frac{1}{N_p} \frac{1 - e^{-j2\pi(n-d)}}{1 - e^{-j2\pi \frac{n-d}{N_p}}} = \begin{cases} 1, & n = d + iN_p, \quad i \in \mathbb{Z} \\ 0, & \text{others,} \end{cases} \tag{41}$$

According to the expression (8), the same IDFT is applied to the right hand side, which will be formulated as

$$\tilde{\mathbf{\Gamma}} = \frac{1}{N_p} \mathbf{\Gamma} (\mathbf{W}^{N_p})^H + \tilde{\mathbf{n}}, \tag{42}$$

where $\tilde{\mathbf{\Gamma}} = [\tilde{\gamma}(0), \dots, \tilde{\gamma}(d), \dots, \tilde{\gamma}(N_p - 1)]$, $\mathbf{\Gamma} = [\gamma_0, \gamma_\Delta, \dots, \gamma_{(N_p-1)\Delta}]$ with $\gamma_k = \mathbf{H}_k \boldsymbol{\theta}$ and $\tilde{\mathbf{n}} = 1/N_p [\mathbf{n}_0, \dots, \mathbf{n}_{(N_p-1)\Delta}] (\mathbf{W}^{N_p})^H$. Applying the same formula as (40) and because the maximum number of delay taps D in the cascaded channel is no higher than N_p , $\tilde{\gamma}(d)$ is given by

$$\tilde{\gamma}(d) = \mathbf{H}(d) \boldsymbol{\theta} + \tilde{\mathbf{n}}(d), \tag{43}$$

where we have $\mathbf{H}(d) = \mathbf{H}(dT_s) = \sum_{s_d=1}^{S_d} c_{s_d} e^{j2\pi f_{s_d} t} \mathbf{a}_{M_b}(\varphi_{s_d}^B) \mathbf{a}_{M_r}^H(\varphi_{s_d}^R)$. Therefore, the expression (9) is arrived at.

APPENDIX B

Upon unfolding the objective function (32) we get

$$\|\tilde{\mathbf{y}}_d - \mathbf{\Omega}_d \mathbf{c}_d\|_2^2 = \tilde{\mathbf{y}}_d^H \tilde{\mathbf{y}}_d - \tilde{\mathbf{y}}_d^H \mathbf{\Omega}_d \mathbf{c}_d - \mathbf{c}_d^H \mathbf{\Omega}_d^H \tilde{\mathbf{y}}_d + \mathbf{c}_d^H \mathbf{\Omega}_d^H \mathbf{\Omega}_d \mathbf{c}_d. \tag{44}$$

Then the partial derivative of $\varphi_{s_d}^R$ may be calculated as

$$\begin{aligned}
g_{s_d}^R &= \frac{\partial \|\tilde{\mathbf{y}}_d - \mathbf{\Omega}_d \mathbf{c}_d\|_2^2}{\partial \varphi_{s_d}^R} \\
&= -\tilde{\mathbf{y}}_d^H \frac{\partial \mathbf{\Omega}_d}{\partial \varphi_{s_d}^R} \mathbf{c}_d - \mathbf{c}_d^H \frac{\partial \mathbf{\Omega}_d^H}{\partial \varphi_{s_d}^R} \tilde{\mathbf{y}}_d + \mathbf{c}_d^H \frac{\partial \mathbf{\Omega}_d^H \mathbf{\Omega}_d}{\partial \varphi_{s_d}^R} \mathbf{c}_d, \tag{45}
\end{aligned}$$

where

$$\frac{\partial \mathbf{\Omega}_d}{\partial \varphi_{s_d}^R} = \begin{bmatrix} \mathbf{0}_{s_d-1}^T & \epsilon_{s_d}^1 & \mathbf{0}_{S_d-s_d}^T \\ \mathbf{0}_{s_d-1}^T & \epsilon_{s_d}^2 & \mathbf{0}_{S_d-s_d}^T \\ \vdots & \vdots & \vdots \\ \mathbf{0}_{s_d-1}^T & \epsilon_{s_d}^{Q_1} & \mathbf{0}_{S_d-s_d}^T \end{bmatrix}, \tag{46}$$

which means that all the columns in the partial derivative matrix are given by zero vectors, except for the s_d -th column. The q -th element of the s_d -th column is

$$\epsilon_{s_d}^q = \sum_{m_r=1}^{M_r} j2\pi(m_r - 1) \theta_{m_r}^q e^{j2\pi f_{s_d} t^q} e^{j2\pi(m_r-1)\varphi_{s_d}^R}, \tag{47}$$

where $\theta_{m_r}^q = [\boldsymbol{\theta}^q]_{m_r}$. Moreover, we have $\frac{\partial \mathbf{\Omega}_d^H}{\partial \varphi_{s_d}^R} = \left(\frac{\partial \mathbf{\Omega}_d}{\partial \varphi_{s_d}^R} \right)^H$ and

$$\frac{\partial \mathbf{\Omega}_d^H \mathbf{\Omega}_d}{\partial \varphi_{s_d}^R} = \frac{\partial \mathbf{\Omega}_d^H}{\partial \varphi_{s_d}^R} \mathbf{\Omega}_d + \mathbf{\Omega}_d^H \frac{\partial \mathbf{\Omega}_d}{\partial \varphi_{s_d}^R}. \tag{48}$$

Therefore, combining the expressions (46-48) with the formula (45), the partial derivative of $\varphi_{s_d}^R$ can be obtained.

REFERENCES

- [1] H. Zhou, W. Xu, J. Chen, and W. Wang, "Evolutionary V2X Technologies Toward the Internet of Vehicles: Challenges and Opportunities," *Proceedings of the IEEE*, vol. 108, no. 2, pp. 308–323, Feb. 2020.
- [2] D. C. N. et al., "6G Internet of Things: A Comprehensive Survey," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 359–383, Jan. 2022.
- [3] K. Z. G. et al., "Millimeter-Wave Communication for Internet of Vehicles: Status, Challenges, and Perspectives," *IEEE Internet Things J.*, vol. 7, no. 9, pp. 8525–8546, Sept. 2020.
- [4] C. X. et al., "OTFS-Aided RIS-Assisted SAGIN Systems Outperform Their OFDM Counterparts in Doubly Selective High-Doppler Scenarios," *IEEE Internet Things J.*, vol. 10, no. 1, pp. 682–703, Jan. 2023.
- [5] M. D. Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. d. Rosny, and S. Tretyakov, "Smart Radio Environments Empowered by Reconfigurable Intelligent Surfaces: How It Works, State of Research, and The Road Ahead," *J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.
- [6] C. Pan, H. Ren, K. Wang, J. F. Kolb, M. ElKashlan, M. Chen, M. D. Renzo, Y. Hao, J. Wang, A. Swindlehurst, X. You, and L. Hanzo, "Reconfigurable Intelligent Surfaces for 6G Systems: Principles, Applications, and Research directions," *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 14–20, Jun. 2021.
- [7] Q. Wu and R. Zhang, "Intelligent Reflecting Surface Enhanced Wireless Network via Joint Active and Passive Beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.
- [8] Z. Peng, T. Li, C. Pan, H. Ren, W. Xu, and M. D. Renzo, "Analysis and Optimization for RIS-Aided Multi-Pair Communications Relying on Statistical CSI," *IEEE Trans. Veh. Technol.*, vol. 70, no. 4, pp. 3897–3901, Apr. 2021.
- [9] S. Gong, C. Xing, X. Zhao, S. Ma, and J. An, "Unified IRS-Aided MIMO Transceiver Designs via Majorization Theory," *IEEE Trans. Signal Process.*, vol. 69, pp. 3016–3032, May. 2021.
- [10] A. U. Makarfi, K. M. Rabie, O. Kaiwartya, X. Li, and R. Kharel, "Physical Layer Security in Vehicular Networks with Reconfigurable Intelligent Surfaces," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, 2020, pp. 1–6.
- [11] Y. Zhu, B. Mao, and N. Kato, "Intelligent Reflecting Surface in 6G Vehicular Communications: A Survey," *IEEE Open J. Veh. Technol.*, vol. 3, pp. 266–277, 2022.
- [12] J. Wang, W. Zhang, X. Bao, T. Song, and C. Pan, "Outage analysis for intelligent reflecting surface assisted vehicular communication networks," *Proc. IEEE Glob. Commun. Conf.*, vol. 3, pp. 1–6, 2020.
- [13] Y. Chen, Y. Wang, and L. Jiao, "Robust transmission for reconfigurable intelligent surface aided millimeter wave vehicular communications with statistical CSI," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 928–944, Feb. 2022.
- [14] T. Dhruvakumar and A. Chaturvedi, "Reflecting surfaces for beyond line-of-sight coverage in millimeter wave vehicular networks," *Veh. Commun.*, vol. 33, pp. 1–4, 2022.
- [15] D. Mishra and H. Johansson, "Channel Estimation and Low-complexity Beamforming Design for Passive Intelligent Surface Assisted MISO Wireless Energy Transfer," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19)*, May. 2019, pp. 4659–4663.

- [16] K. Ardah, S. Gharekhloo, A. L. F. de Almeida, and M. Haardt, "TRICE: A Channel Estimation Framework for RIS-Aided Millimeter-Wave MIMO Systems," *IEEE Signal Process. Lett.*, vol. 28, pp. 513–517, Feb. 2021.
- [17] C. Hu, L. Dai, S. Han, and X. Wang, "Two-Timescale Channel Estimation for Reconfigurable Intelligent Surface Aided Wireless Communications," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7736–7747, Nov. 2021.
- [18] X. Wei, D. Shen, and L. Dai, "Channel Estimation for RIS Assisted Wireless Communications—Part II: An Improved Solution Based on Double-Structured Sparsity," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1403–1407, 2021.
- [19] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent Reflecting Surface Meets OFDM: Protocol Design and Rate Maximization," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4522–4535, Jul. 2020.
- [20] B. Zheng and R. Zhang, "Intelligent Reflecting Surface-Enhanced OFDM: Channel Estimation and Reflection Optimization," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 518–522, Apr. 2020.
- [21] B. Zheng, C. You, and R. Zhang, "Intelligent Reflecting Surface Assisted Multi-User OFDMA: Channel Estimation and Training Design," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8315–8329, Dec. 2020.
- [22] Z. Wan, Z. Gao, and M. S. Alouini, "Broadband Channel Estimation for Intelligent Reflecting Surface Aided mmwave Massive MIMO Systems," in *Proc. IEEE Int. Conf. Commun. (IEEE ICC'20)*, Jun. 2020, pp. 1–6.
- [23] Y. Liu, S. Zhang, F. Gao, J. Tang, and O. A. Dobre, "Cascaded Channel Estimation for RIS Assisted mmwave MIMO Transmissions," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 2065–2069, Sept. 2021.
- [24] W. Wu, H. Wang, W. Wang, and R. Song, "Doppler mitigation method aided by reconfigurable intelligent surfaces for high-speed channels," *IEEE Wireless Commun. Letter.*, vol. 11, no. 3, pp. 627–631, Mar. 2022.
- [25] S. Sun and H. Yan, "Channel estimation for reconfigurable intelligent surface-assisted wireless communications considering Doppler effect," *IEEE Wireless Commun. Letter.*, vol. 10, no. 4, pp. 790–794, Apr. 2021.
- [26] 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.
- [27] C. Xu, J. An, T. Bai, L. Xiang, S. Sugiura, R. G. Maunder, L.-L. Yang, and L. Hanzo, "Reconfigurable Intelligent Surface Assisted Multi-Carrier Wireless Systems for Doubly Selective High-Mobility Ricean Channels," *IEEE Trans. Veh. Technol.*, vol. 71, no. 4, pp. 4023–4041, Apr. 2022.
- [28] C. Xu, J. An, T. Bai, S. Sugiura, R. G. Maunder, Z. Wang, L.-L. Yang, and L. Hanzo, "Channel Estimation for Reconfigurable Intelligent Surface Assisted High-Mobility Wireless Systems," *IEEE Trans. Veh. Technol.*, pp. 1–16, Sept. 2022, *early access*.
- [29] S. Gao, X. Cheng, and L. Yang, "Estimating Doubly-Selective Channels for Hybrid mmwave Massive MIMO Systems: A Doubly-Sparse Approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 5703–5715, Sept. 2020.
- [30] W. Zhou and W. H. Lam, "Channel estimation and data detection for OFDM systems over fast-fading and dispersive channels," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1381–1392, Mar. 2010.
- [31] S. Ma, W. Shen, J. An, and L. Hanzo, "Wideband Channel Estimation for IRS-Aided Systems in the Face of Beam Squint," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6240–6253, Oct. 2021.
- [32] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
- [33] H. Franz and M. Gerald, *Wireless Communications Over Rapidly Time-Varying Channels*, 1st ed. Academic Press, Inc., 2011.
- [34] R. Steele and L. Hanzo, *Mobile Radio Communications*. Wiley-IEEE Press, 1999.
- [35] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [36] W. Guo, W. Zhang, P. Mu, F. Gao, and H. Lin, "High-mobility wideband massive MIMO communications: Doppler compensation analysis and scaling laws," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3177–3191, Jun. 2019.
- [37] S. Srivastava, R. K. Singh, A. K. Jagannatham, and L. Hanzo, "Delay-Doppler and Angular Domain 4D-Sparse CSI Estimation in OTFS Aided MIMO Systems," *IEEE Trans. Veh. Technol.*, pp. 1–6, Sept. 2022, *early access*.
- [38] X. Ge, W. Shen, C. Xing, L. Zhao, and J. An, "Training Beam Design for Channel Estimation in Hybrid mmwave MIMO Systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7121–7134, Sept. 2022.
- [39] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications with MATLAB*. Wiley-IEEE Press, Aug. 2010.
- [40] W. Zhou and W.-H. Lam, "Channel estimation and data detection for OFDM systems over fast-fading and dispersive channels," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1381–1892, Mar. 2010.
- [41] Y. Choi, P. Voltz, and F. Cassara, "On channel estimation and detection for multicarrier signals in fast and selective rayleigh fading channels," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1375–1387, Aug. 2001.
- [42] C. Xing, S. Wang, S. Chen, S. Ma, H. V. Poor, and L. Hanzo, "Matrix-Monotonic Optimization – part I: Single-variable optimization," *IEEE Trans. Signal Process.*, vol. 69, pp. 738–754, Nov. 2021.
- [43] C. Xing, S. Wang, S. Chen, S. Ma, H. V. Poor, and L. Hanzo, "Matrix-Monotonic Optimization – part II: Multi-Variable Optimization," *IEEE Trans. Signal Process.*, vol. 69, pp. 179–194, Nov. 2021.
- [44] C. Xing, S. Ma, and Y. Zhou, "Matrix-Monotonic Optimization for MIMO Systems," *IEEE Trans. Signal Process.*, vol. 63, pp. 334–348, Jan. 2015.
- [45] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory.*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.
- [46] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Trans. Inf. Theory.*, vol. 57, no. 7, pp. 4680–4688, Jul. 2011.
- [47] Y. Arjoune, N. Kaabouch, H. E. Ghazi, and A. Tamtaoui, "Compressive sensing: Performance comparison of sparse recovery algorithms," in *Proc. IEEE 7th Annu. Comput. Commun. Workshop Conf.*, 2017, pp. 1–7.
- [48] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.
- [49] A. Taha, M. Alrabeiah, , and A. Alkhateeb, "Enabling large intelligent surfaces with compressive sensing and deep learning," *IEEE Access*, vol. 9, pp. 44 304–44 321, 2021.
- [50] 3GPP, "Physical channels and modulation," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.211, 01 2023, version 17.4.0.
- [51] H. Zhou, W. Xu, J. Chen, and W. Wang, "Evolutionary V2X technologies toward the Internet of Vehicles: Challenges and opportunities," *Proc. IEEE*, vol. 108, no. 2, pp. 308–323, Feb. 2020.
- [52] J.-J. Park, J. Lee, J. Liang, K.-W. Kim, K.-C. Lee, , and M.-D. Kim, "Millimeter wave vehicular blockage characteristics based on 28 GHz measurements," in *Proc. 2017 IEEE 86th Vehicular Technology Conf.*, 2017, pp. 1–5.
- [53] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 38.901, 03 2022, version 17.0.0.



Jiaxin Chen received the B.S. degree in electronic engineering from Beijing Institute of Technology, Beijing, China, in 2021. She is currently pursuing the M.S. degree with the School of Information and Electronics, Beijing Institute of Technology. Her main research interests include channel estimation, reconfigurable intelligent surface, orthogonal time frequency space (OTFS) modulation and compressive sensing.



Wenqian Shen received the B.S. degree from Xi'an Jiaotong University, Shaanxi, China in 2013 and the Ph.D. degree from Tsinghua University, Beijing, China in 2018. She is currently an associate professor with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China. Her research interests include massive MIMO and mmWave/THz communications. She has published several journal and conference papers in IEEE Transaction on Signal Processing, IEEE Transaction on Communications, IEEE Transaction on Vehicular Technology, IEEE ICC, etc. She has won the IEEE Best Paper Award at the IEEE ICC 2017.



Shixun Luo received the B. Eng. degree in 2013 and Ph.D. degree in 2020, both from the School of Information and Electronics, Beijing Institute of Technology, China, in communications engineering. Since Oct. 2020, she has been with the School of Cyberspace Science and Technology, Beijing Institute of Technology, China, where she is currently Research Engineer. Her research interests include satellite communication, UAV networks, and physical-layer security, especially spread spectrum signal processing.



Siqi Ma received the Ph.D. degree from the Beijing Institute of Technology, Beijing with the School of Information and Electronics., in 2022. His current research interests include massive MIMO, mmWave/THz communications, energy-efficient communications, intelligent reflecting surface and networks.



Chengwen Xing (Member, IEEE) received the B.E. degree from Xidian University, Xi'an, China, in 2005 and the Ph.D. degree from The University of Hong Kong, Hong Kong, China, in 2010. Since September 2010, he has been with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China, where he is currently a Full Professor. From September 2012 to December 2012, he was a Visiting Scholar with the University of Macau, Macau, China. His current research interests include machine learning, statistical signal

processing, convex optimization, multivariate statistics, array signal processing and high-frequency band communication systems.



Lajos Hanzo (<http://www-mobile.ecs.soton.ac.uk>, https://en.wikipedia.org/wiki/Lajos_Hanzo) (FIEEE'04) received Honorary Doctorates from the Technical University of Budapest and Edinburgh University. He is a Foreign Member of the Hungarian Science-Academy, Fellow of the Royal Academy of Engineering (FREng), of the IET, of EURASIP and holds the IEEE Eric Sumner Technical Field Award.