Wideband Channel Estimation for IRS-Aided Systems in the Face of Beam Squint

Siqi Ma, Wenqian Shen, Jianping An, Lajos Hanzo, *Fellow, IEEE*

***Abstract*—Intelligent reflecting surfaces (IRSs) improve both the bandwidth and energy efficiency of wideband communication systems by using low-cost passive elements for reflecting the impinging signals with adjustable phase shifts. To realize the full potential of IRS-aided systems, having accurate channel state information (CSI) is indispensable, but it is challenging to acquire, since these passive devices cannot carry out transmit/receive signal processing. The existing channel estimation methods conceived for wideband IRS-aided communication systems only consider the channel's frequency selectivity, but ignore the effect of beam squint, despite its severe performance degradation. Hence we fill this gap and conceive wideband channel estimation for IRS-aided communication systems by explicitly taking the effect of beam squint into consideration. We demonstrate that the mutual correlation function between the spatial steering vectors and the cascaded two-hop channel reflected by the IRS has two peaks, which leads to a pair of estimated angles for a single propagation path, due to the effect of beam squint. One of these two estimated angles is the frequency-independent 'actual angle', while the other one is the frequency-dependent 'false angle'. To reduce the influence of false angles on channel estimation, we propose a twin-stage orthogonal matching pursuit (TS-OMP) algorithm, where the path angles of the cascaded two-hop channel reflected by the IRS are obtained in the first stage, while the propagation gains and delays are obtained in the second stage. Moreover, we propose a bespoke pilot design by exploiting the specific the characteristics of the mutual correlation function and cross-entropy theory for achieving an improved channel estimation performance. Our simulation results demonstrate the superiority of the proposed channel estimation algorithm and pilot design over their conventional counterparts.**

***Index Terms*—Wideband channel estimation, intelligent reflecting surface (IRS), reconfigurable intelligent surface (RIS), beam squint.**

I. Introduction

The intelligent reflecting surface (IRS) aided communication concept, which is also named as reconfigurable intelligent surface (RIS), has emerged as a promising solution for next-generation systems [1]–[3], which is capable of achieving improved bandwidth and energy efficiency at a low cost. In contrast to traditional amplify-and-forward (AF) relaying, IRS-aided systems use low-cost passive elements for reflecting the incident signals with adjustable phase shifts [4]. By appropriately designing the phase shifts of IRS elements and the cascaded two-hop channel reflected through an IRS, we can improve the link quality [5], [6]. Hence apart from active beamforming at the base station (BS), passive beamforming has also been harnessed at the IRS in narrow-band scenarios in [7]–[11]. Specifically, Wu *et al.* [7] proposed a semidefinite relaxation (SDR) based method for maximizing the spectral efficiency, while Huang *et al.* [8] advocated a gradient descent approach and sequential fractional programming method for maximizing the energy efficiency. As a further development, Feng *et al.* [10] conceived a deep reinforcement learning based framework for solving the non-convex optimization problem of passive beamforming design at the IRS. Moreover, a geometric mean decomposition-based method [12], a convex optimization-based technique [13] and block coordinate descent iterative algorithms [14] were proposed for joint passive and active beamforming in wideband IRS-aided systems. However, all these contributions relied on the idealized simplifying assumption that the channel state information (CSI) is perfectly known, even though in reality channel estimation is quite challenging in IRS-aided systems. This is because the reflective elements are passive devices, which cannot perform active transmit/receive signal processing. Since there is a paucity of literature on this challenging subject, we conceive an efficient channel estimation method for IRS-aided systems.

Despite the paucity of related solutions, some narrow-band schemes have been disseminated in [15]–[18]. Specifically, Mishra *et al.* [15] proposed an 'on-off' state control based channel estimation method, where only a single element of a reflecting surface was switched on, while all other elements remained off at each time slot. In this way, the channel reflected through the activated element can be estimated without interference from the signals reflected by all other elements of the IRS. As a further development, Wang *et al.* [16] proposed a three-phase channel estimation method for the uplink of IRS-aided multiuser systems, which exploited the fact that the IRS elements reflect the signals arriving from different users to the BS via the same IRS to BS channel. Lin [17] *et al.* proposed a Lagrange optimization based channel estimation strategy for minimizing the mean-squared error of channel estimation. However, the above methods require a large number of measurements for distinguishing the large number of reflective elements at the IRS. To circumvent this problem, Wang *et al.* [18] formulated the channel estimation of IRS-aided systems as a sparse signal recovery problem by exploiting the channel's angular-domain sparsity. Then, they

This work was supported in part by the the National Key Research and Development Program of China under No. 2019YFB1803200, and by the National Natural Science Foundation of China (NSFC) under Grant 61901034. L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/P034284/1 and EP/P003990/1 (COALESCE) as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028). (*Corresponding author: Wenqian Shen.)*

S. Ma, W. Shen, J. An are with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (E-mails: sqma@bit.edu.cn; wshen@bit.edu.cn; an@bit.edu.cn). S. Ma is also with Shaanxi Key Laboratory of Information Communication Network and Security, Xian University of Posts & Telecommunications, Xian, Shaanxi 710121, China.

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).

applied popular compressed sensing (CS) algorithms, such as the basis pursuit and simultaneous orthogonal matching pursuit techniques for recovering the channel at a much reduced number of measurements.

To expound a little further, some authors have also extended the realm of channel estimation solutions from narrow-band to wideband scenarios [19]–[22]. Specifically, Zheng *et al.* [19], [20] formulated a wideband channel estimation problem for IRS-aided orthogonal frequency-division multiplexing (OFDM) systems and proposed a least-square (LS) based method for estimating frequency-selective fading channels. To reduce the overhead of channel training, Yang *et al.* [22] proposed an IRS element-grouping method, where each group consists of a set of adjacent IRS elements that share a common reflecting coefficient. Wan [21] *et al.* designed a CS-based wideband channel estimation method by assuming that the path-angle of the BS to IRS channel is known. Then, they utilized a distributed orthogonal matching pursuit algorithm for estimating the IRS to user channel. These wideband channel estimation methods [19]–[22] do indeed consider the frequency-selectivity of the wideband channel, but they ignore the effect of beam squint. The challenge is that the beam squint of wideband systems will lead to a pair of distinct angles of the cascaded two-hop channel reflected through the IRS for the same propagation path, which makes traditional channel estimation methods ineffective. To the best of our knowledge, this is the first contribution proposing channel estimation for IRS-aided systems by considering the effect of beam squint. Hence the main contributions of this paper are summarized as follows:

- We formulate the estimation problem of the reflected channel, which is defined as the cascaded BS-to-IRS and IRS-to-user channel, in the face of beam squint. Specifically, the cascaded channel is characterized by the equivalent angles, gains and delays of the associated propagation paths. In contrast to the conventional channel model, where the spatial steering vectors are frequency-independent, we consider the realistic frequency-dependent steering vectors and the extended angular range of the steering vectors of the reflected channel.
- We study the effect of beam squint on the channel estimation of IRS-aided communication systems and propose a twin-stage OMP (TS-OMP) algorithm, which is robust to beam squint. We can readily search for the peak of the correlation function between the spatial steering vectors and the reflected channel, which determines the equivalent angles. However, due to the extended angular range of the steering vector, a pair of angles will be returned for a single propagation path. One of the two estimated angles is the frequency-independent actual angle, while the other one is the frequency-dependent false angle, which may substantially erode the performance of channel estimation. To eliminate the influence of false angles, we propose a TS-OMP algorithm. Specifically, in the first stage, we exploit the fact that the false angles are slightly different for the different subcarriers, while the actual angles remain the same for all subcarriers. Then, we propose a block-sparse method for estimating the actual equivalent angles, while mitigating the influence of the subcarrier-dependent false angles. In the second stage, based on the equivalent angles estimated in the first stage, we propose an OMP based method for calculating the gains and delays of the equivalent paths having different equivalent angles.
- Finally, we propose the corresponding pilot design, where a pair of pilot design requirement are considered. Firstly, in order to eliminate the interference imposed by false angles in the first stage, we have to reduce the accumulated values corresponding to the false angles of the mutual correlation function across the different subcarriers. This means that the false angles at the pilot-subcarriers corresponding to a certain path should be separated from each other. Secondly, a high grade of orthogonality is preferred among the columns of the measurement matrix for the estimation of channel gains and delays in the second stage. Based on the above-mentioned pair of requirements, we propose a cross-entropy based pilot design method, which improves the cascaded BS-IRS-user channel estimation performance. Finally, our simulation results demonstrate that the proposed channel estimation algorithm and pilot design combination outperforms its conventional counterpart.

	[16]	[18]	[19]	[20]	[21]	Our Proposed Method
Narrowband Scenario	✓	✓				
Wideband Scenario			✓	✓	✓	✓
LS/MMSE	✓		✓	✓		
CS Based		✓			✓	✓
Beam Squint						✓

TABLE I
CONTRASTING OUR CONTRIBUTIONS TO THE STATE-OF-THE-ART

To clarify our contributions clearly, we contrast the proposed channel estimation method against the state-of-the-art in Table I.

The rest of this paper is organized as follows. In Section II, we present the system model of IRS-aided wideband systems considering the effect of beam squint. In Section III, we analyze the effect of beam squint on estimation of BS-IRS-user channel. In Section IV, we propose a TS-OMP method for estimating the channel, which is robust to beam squint, while in Section V, we propose a pilot design based on cross-entropy. In Section VI, our simulation results are provided, followed by our conclusion in Section VII.

Notations: We use the following notations throughout the paper. We let a, $\mathbf{a}$, $\mathbf{A}$ represent the scalar, vector, and matrix respectively; $\mathrm{vec}\{\cdot\}$ denotes the vectorization of a matrix and $\mathrm{ivec}\{\cdot\}$ denotes the invectorization of a vector; $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$, and $(\cdot)^{-1}$ denote the transpose, conjugate transpose, and inverse of a matrix, respectively; $[\mathbf{a}]_{m:n}$ denotes the m-th to n-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 operator $\circ$, $\otimes$ and $*$ represent the Hadamard-product, Kronecker product and convolution, respectively; Finally, $\mathbf{0}$ denotes the zero matrix, $\mathbf{I}$ denotes identity matrix, and $\mathbf{1}_{m,n}$

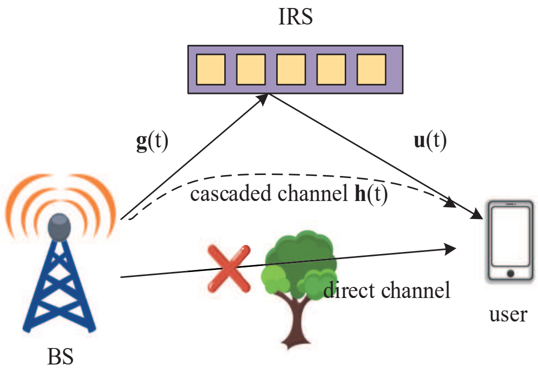


Fig. 1. IRS-aided communication system

denotes all 1 matrix of size $m \times n$.

II. System Model With Beam Squint

As depicted in Fig. 1, we consider an IRS-aided wideband OFDM system, where both the BS and the user have a single antenna- [1], while the IRS is an $(M \times 1)$-element uniform linear array (ULA) [23] [24]. Let $\mathbf{g}(t) \in \mathbb{C}^{M\times 1}$ denote the channel impulse response (CIR) spanning from the BS to the IRS, and $\mathbf{u}(t) \in \mathbb{C}^{M\times 1}$ denote the CIR of the link spanning from the IRS to the user. In this paper, we neglect the direct path between the BS and the user. If however the direct path is non-negligible, we can estimate it by turning off the IRS and using traditional channel estimation methods [18]. Based on this assumption, the direct channel can be neglected in the IRS-aided channel model. OFDM having N_p subcarriers is adopted for combating the multipath effects. We define the transmission bandwidth as W, resulting in W/N_p subcarriers. We assume that the cyclic prefix (CP) is longer than the maximum multipath delay and all the complex-valued channel envelopes remain approximately constant within the channel's coherence time [20].

Assume furthermore that there are L_1 propagation paths between the BS and the IRS, where $\tau_{l_1,m}^{\mathrm{TR}}$ denotes the time delay of the l_1-th path spanning from the BS to the m-th IRS elements, where $l_1 \in \{1, 2, \cdots, L_1\}$ and $m \in \{1, 2, \cdots, M\}$. Thus, the CIR between the BS and the m-th IRS element can be expressed as [25]

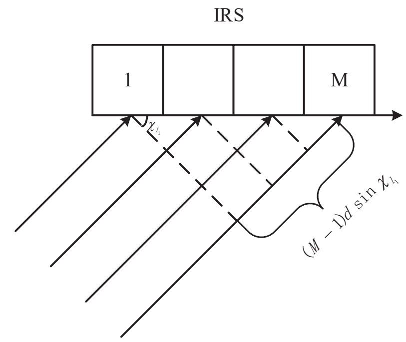
$$g_m(t) = \sum_{l_1=1}^{L_1} \overline{\alpha}_{l_1} e^{-j2\pi f_c \tau_{l_1,m}^{\mathrm{TR}}} \delta(t - \tau_{l_1,m}^{\mathrm{TR}}), \tag{1}$$

where f_c is the carrier frequency and $\overline{\alpha}_{l_1}$ is the complex path gain of the l_1-th path. We define the channel vector $\mathbf{g}(t) = [g_1(t), \cdots, g_M(t)]^{\mathrm{T}} \in \mathbb{C}^{M\times 1}$. Similar to the definition of $g_m(t)$, the CIR between the m-th IRS element and the user can be expressed as

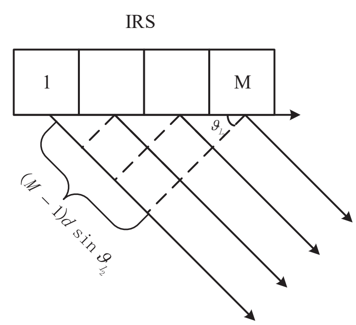
$$u_m(t) = \sum_{l_2=1}^{L_2} \overline{\beta}_{l_2} e^{-j2\pi f_c \tau_{l2,m}^{\mathrm{RR}}} \delta(t - \tau_{l_2,m}^{\mathrm{RR}}), \tag{2}$$

[1]The channel estimation method proposed in this paper can be readily extended to multiple-antenna aided BSs and multiple users relying on single antennas by using orthogonal pilots. The details are shown in Appendix A.

where L_2 and $\overline{\beta}_{l_2}$ denote the number of paths between the IRS as well as the user and the complex gain of the l_2-th path, $l_2 \in \{1, \cdots, L_2\}$. We denote the delay of the l_2-th path from the m-th IRS element to the user by $\tau_{l2,m}^{\mathrm{RR}}$ and define the channel vector by $\mathbf{u}(t) = [u_1(t), \cdots, u_M(t)]^{\mathrm{T}} \in \mathbb{C}^{M\times 1}$.



(a) Illustrate AOA χ_{l_1}.



(b) Illustrate DOA ϑ_{l_2}.

Fig. 2. The illustration of AOA and DOA

Let us express the signal reflected by the m-th IRS element as

$$\begin{aligned} r_m(t) &= \theta_m g_m(t) * s(t) \\ &= \theta_m \sum_{l_1=1}^{L_1} \overline{\alpha}_{l_1} e^{-j2\pi f_c \tau_{l_1,m}^{\mathrm{TR}}} s(t - \tau_{l_1,m}^{\mathrm{TR}}), \end{aligned} \tag{3}$$

where $\boldsymbol{\theta} = [\theta_1, \cdots, \theta_m, \cdots, \theta_M]^T = [\varepsilon_1 e^{j\phi_1}, \cdots, \varepsilon_m e^{j\phi_m}, \cdots, \varepsilon_M e^{j\phi_M}]^T$ represent the reflection coefficients of the IRS, where $\phi_m \in [0, 2\pi)$ and $\varepsilon_m \in [0, 1]$ are the phase shift and amplitude reflection coefficient of the m-th IRS element, respectively. To maximize the signal power reflected by the IRS, we set $\varepsilon_m = 1, \forall_m \in M$ [7]. Moreover, $s(t)$ is the signal transmitted by the BS.

Then, the signal $r_m(t)$ is reflected by the IRS to the user and the received signal is expressed as

$$\begin{aligned} y_m(t) &= u_m(t) * r_m(t) + n_m(t) \\ &= \sum_{l_2=1}^{L_2} \overline{\beta}_{l_2} e^{-j2\pi f_c \tau_{l_2,m}^{\mathrm{RR}}} r_m(t - \tau_{l_2,m}^{\mathrm{RR}}) + n_m(t), \end{aligned} \tag{4}$$

where $n_m(t)$ is the additive complex Gaussian noise. By substituting (3) into (4), we have

$$\begin{aligned} y_m(t) = \theta_m \sum_{l_1=1}^{L_1}\sum_{l_2=1}^{L_2} \overline{\alpha}_{l_1}\overline{\beta}_{l_2} e^{-j2\pi f_c \tau^{\mathrm{TR}}_{l_1,m}} e^{-j2\pi f_c \tau^{\mathrm{RR}}_{l_2,m}} \\ \cdot s(t-\tau^{\mathrm{TR}}_{l_1,m}-\tau^{\mathrm{RR}}_{l_2,m}) + n_m(t) \\ = \theta_m h_m(t) * s(t) + n_m(t), \end{aligned} \tag{5}$$

where CIR of the m-th cascaded BS-IRS-user element channel is written as

$$\begin{aligned} h_m(t) = \sum_{l_1=1}^{L_1}\sum_{l_2=1}^{L_2} \overline{\alpha}_{l_1}\overline{\beta}_{l_2} e^{-j2\pi f_c \tau^{\mathrm{TR}}_{l_1,m}} e^{-j2\pi f_c \tau^{\mathrm{RR}}_{l_2,m}} \\ \cdot \delta(t-\tau^{\mathrm{TR}}_{l_1,m}-\tau^{\mathrm{RR}}_{l_2,m}). \end{aligned} \tag{6}$$

For the sake of simplicity, we define $\tau^{\mathrm{TR}}_{l_1} = \tau^{\mathrm{TR}}_{l_1,1}$ and $\tau^{\mathrm{RR}}_{l_2} = \tau^{\mathrm{RR}}_{l_2,1}$. Based on the assumption that IRS array size is much smaller than the distance between the BS and the IRS as well as the IRS-user distance, the path delay $\tau^{\mathrm{TR}}_{l_1,m}$ and $\tau^{\mathrm{RR}}_{l_2,m}$ can be expressed as

$$\tau^{\mathrm{TR}}_{l_1,m} = \tau^{\mathrm{TR}}_{l_1} + (m-1)\frac{d\sin\chi_{l_1}}{c} = \tau^{\mathrm{TR}}_{l_1} + (m-1)\frac{\varphi^{\mathrm{TR}}_{l_1}}{f_c}, \tag{7}$$

$$\tau^{\mathrm{RR}}_{l_2,m} = \tau^{\mathrm{RR}}_{l_2} - (m-1)\frac{d\sin\vartheta_{l_2}}{c} = \tau^{\mathrm{RR}}_{l_2} - (m-1)\frac{\varphi^{\mathrm{RR}}_{l_2}}{f_c}, \tag{8}$$

where χ_{l_1} denotes the angle of arrival (AOA) of the l_1-th path from the BS to the IRS, as shown in Fig. 2 (a). Similarly, ϑ_{l_2} denotes the angle of departure (AOD) of the l_2-th path from the IRS to the user, as shown in Fig. 2 (b). We define $\varphi^{\mathrm{TR}}_{l_1} = \frac{d\sin\chi_{l_1}}{\lambda_c}$ and $\varphi^{\mathrm{RR}}_{l_2} = \frac{d\sin\vartheta_{l_2}}{\lambda_c}$ as the normalized AOA and AOD, respectively, where λ_c represents the carrier wavelength. For the typical half-wavelength element spacing of $d = \lambda_c/2$, we have $\varphi^{\mathrm{TR}}_{l_1} \in [-1/2, 1/2)$, $\varphi^{\mathrm{RR}}_{l_2} \in [-1/2, 1/2)$.

By substituting (7) and (8) into (6), we have

$$\begin{aligned} h_m(t) = \sum_{l_1=1}^{L_1}\sum_{l_2=1}^{L_2} \underbrace{\overline{\alpha}_{l_1} e^{-j2\pi f_c \tau^{TR}_{l_1}}}_{\alpha_{l_1}} \underbrace{\overline{\beta}_{l_2} e^{-j2\pi f_c \tau^{RR}_{l_2}}}_{\beta_{l_2}} \\ e^{-j2\pi(m-1)\varphi^{TR}_{l_1}} e^{j2\pi(m-1)\varphi^{RR}_{l_2}} \delta(t-\tau^{TR}_{l_1,m}-\tau^{RR}_{l_2,m}), \end{aligned} \tag{9}$$

where $\alpha_{l_1} = \overline{\alpha}_{l_1} e^{-j2\pi f_c \tau^{TR}_{l_1}}$ and $\beta_{l_2} = \overline{\beta}_{l_2} e^{-j2\pi f_c \tau^{RR}_{l_2}}$. By applying the continuous time Fourier transform to (9), the frequency-domain (FD) cascaded channel corresponding to the m-th element of IRS can be written as

$$\begin{aligned} h_m(f) &= \int_{-\infty}^{+\infty} h_m(t) e^{-j2\pi f t} dt \\ &= \sum_{l_1=1}^{L_1}\sum_{l_2=1}^{L_2} \alpha_{l_1}\beta_{l_2} e^{-j2\pi(m-1)\left(\varphi^{\mathrm{TR}}_{l_1}-\varphi^{\mathrm{RR}}_{l_2}\right)} e^{-j2\pi f\left(\tau^{\mathrm{TR}}_{l_1,m}+\tau^{\mathrm{RR}}_{l_2,m}\right)} \\ &= \sum_{l_3=1}^{L_1L_2} c^{\mathrm{C}}_{l_3} e^{-j2\pi(m-1)\varphi^{\mathrm{C}}_{l_3}\left(1+\frac{f}{f_c}\right)} e^{-j2\pi f \tau^{\mathrm{C}}_{l_3}}, \end{aligned} \tag{10}$$

where the equivalent delay $\tau^{\mathrm{C}}_{l_3}$, angle $\varphi^{\mathrm{C}}_{l_3}$ and the complex gain $c^{\mathrm{C}}_{l_3}$ of the cascaded BS-IRS-user channel can be defined as

$$\tau^{\mathrm{C}}_{l_3} = \tau^{\mathrm{TR}}_{l_1} + \tau^{\mathrm{RR}}_{l_2}, \tag{11}$$

$$\varphi^{\mathrm{C}}_{l_3} = \varphi^{\mathrm{TR}}_{l_1} - \varphi^{\mathrm{RR}}_{l_2}, \tag{12}$$

$$c^{\mathrm{C}}_{l_3} = \alpha_{l_1}\beta_{l_2}, \tag{13}$$

where $l_3 \in \{1, 2, \cdots, L_1L_2\}$ and the range of $\varphi^{\mathrm{C}}_{l_3}$ is $(-1, 1)$. Let us discuss the range of $\varphi^{\mathrm{C}}_{l_3}$ in the next section in detail. Note that if we can obtain the parameters of $\tau^{\mathrm{C}}_{l_3}, \varphi^{\mathrm{C}}_{l_3}$ and $c^{\mathrm{C}}_{l_3}$, the cascaded FD channel can be recovered according to (10).

Based on the convolution theorem and equation (5), the FD signal received by the user through the m-th IRS element can be expressed as

$$y_m(f) = \theta_m h_m(f) s(f) + n_m(f), \tag{14}$$

where $n_m(f)$ is the additive white Gaussian noise (AWGN). The FD signal $s(f)$ denotes the pilot symbol transmitted by the BS. Without loss of generality, we let $s(f) = 1$. Thus, the FD signal received by the user can be expressed as

$$\begin{aligned} y(f) &= \sum_{m=1}^{M} y_m(f) \\ &= \boldsymbol{\theta}^{\mathrm{T}}\mathbf{h}(f) + n(f), \end{aligned} \tag{15}$$

where $n(f) \sim CN(0, \delta^2)$ and δ^2 is the noise power. The cascaded channel vector $\mathbf{h}(f) = [h_1(f), h_2(f), \cdots, h_M(f)]^{\mathrm{T}}$ can be rewritten as

$$\mathbf{h}(f) = \sum_{l_3=1}^{L_1L_2} c^{\mathrm{C}}_{l_3}\mathbf{a}\left(\left(1+\frac{f}{f_c}\right)\varphi^{\mathrm{C}}_{l_3}\right) e^{-j2\pi f \tau^{\mathrm{C}}_{l_3}}, \tag{16}$$

where

$$\begin{aligned} &\mathbf{a}((1+\frac{f}{f_c})\varphi^{\mathrm{C}}_{l_3}) = \\ &[1, e^{-j2\pi\left(1+\frac{f}{f_c}\right)\varphi^{\mathrm{C}}_{l_3}}, \cdots, e^{-j2\pi m\left(1+\frac{f}{f_c}\right)\varphi^{\mathrm{C}}_{l_3}}, \\ &\cdots, e^{-j2\pi(M-1)\left(1+\frac{f}{f_c}\right)\varphi^{\mathrm{C}}_{l_3}}]^{\mathrm{T}} \end{aligned} \tag{17}$$

is the spatial steering vector. In contrast to the conventional channel model [26], [27], the spatial steering vector is dependent on frequency and has the following property: $\lim_{M\to\infty}\frac{1}{M}\mathbf{a}^{\mathrm{H}}\left(\left(1+\frac{f}{f_c}\right)\varphi_1\right)\mathbf{a}\left(\left(1+\frac{f}{f_c}\right)\varphi_2\right) = \delta\left(\varphi_1-\varphi_2\right)$, which is termed as the angular orthogonality property. The proof is similar to that in [28].

The channel model derived in (10) can be directly extended to the uniform planar array (UPA) scenario. For an IRS of $M_x \times M_y$ elements, we have

$$\begin{aligned} h_{m_x,m_y}(f) = \sum_{l_1=1}^{L_1}\sum_{l_2=1}^{L_2} \alpha_{l_1}\beta_{l_2} e^{-j2\pi(m_x-1)\left(u^{\mathrm{TR}}_{l_1}-u^{\mathrm{RR}}_{l_2}\right)\left(1+\frac{f}{f_c}\right)} \\ \times e^{-j2\pi(m_y-1)\left(v^{\mathrm{TR}}_{l_1}-v^{RR}_{l_2}\right)\left(1+\frac{f}{f_c}\right)} e^{-j2\pi f\left(\tau^{TR}_{l_1}+\tau^{\mathrm{RR}}_{l_2}\right)}, \end{aligned} \tag{18}$$

where $h_{m_x,m_y}(f)$ denotes the cascaded FD channel corresponding to the (m_x, m_y)-th IRS element. We have $u^{\mathrm{x}}_l = \frac{d\cos(\overline{u}^{\mathrm{x}}_l)}{\lambda_c}$ and $v^{\mathrm{x}}_l = \frac{d\cos(\overline{u}^{\mathrm{x}}_l)\sin(\overline{v}^{\mathrm{x}}_l)}{\lambda_c}$, where $\overline{u}^{\mathrm{x}}_l$ and $\overline{v}^{\mathrm{x}}_l$ denote the elevational angle and azimuth angle, respectively, $l \in \{l_1, l_2\}$, $\mathrm{x} \in \{\mathrm{RR}, \mathrm{TR}\}$.

III. Effect Of Beam Squint on channel estimation of IRS-Aided systems

In contrast to the conventional wideband channel model [26], [27], where the steering vector is independent of frequency, in this paper, we consider the frequency-dependent steering vectors of (16) for wideband IRS-aided systems. We observe from (16) that the spatial angle $\left(1+\frac{f}{f_c}\right)\varphi_{l_3}^{\mathrm{C}}$ has different angular spread over different subcarriers, which is referred to as "beam squint" [28]–[30]. The effect of beam squint will change the equivalent angle range $\varphi_{l_3}^{\mathrm{C}}$ of the cascaded channel, hence the angular range of the steering vectors of the reflected channel is no longer equivalent to $[-1/2, 1/2)$ [23] and we should consider the extended angular range of steering vectors spanning the interval of $(-1, 1)$, which will be discussed as follows.

According to (16) and the angular orthogonality property of steering vectors, one can get the estimation of the equivalent angle $\varphi_{l_3}^{\mathrm{C}}$ by finding the peak of mutual correlation function $\Gamma^f(x)$ between steering vectors and the cascaded channel, which can be expressed as

$$\begin{aligned}\Gamma^f(x) &= \mathbf{a}^{\mathrm{H}}\left(\left(1+\frac{f}{f_c}\right)x\right)\mathbf{h}(f)\\ &= \sum_{l_3=1}^{L_1L_2}\sum_{m=1}^{M} c_{l_3} e^{-j2\pi f\tau_{l_3}^{\mathrm{C}}} e^{-j2\pi(m-1)\left(1+\frac{f}{f_c}\right)(x-\varphi_{l_3}^{\mathrm{C}})}\\ &= \sum_{l_3=1}^{L_1L_2} c_{l_3} e^{-j2\pi f\tau_{l_3}^{\mathrm{C}}} \frac{\sin\left(\pi M\left(1+\frac{f}{f_c}\right)(x-\varphi_{l_3}^{\mathrm{C}})\right)}{\sin\left(\pi\left(1+\frac{f}{f_c}\right)(x-\varphi_{l_3}^{\mathrm{C}})\right)},\end{aligned} \tag{19}$$

where the search range is $x \in (-1, 1)$, since $\varphi_{l_3}^{\mathrm{C}} \in (-1, 1)$. We consider a single path of the channel $\mathbf{h}(f)$ for better elaborating on the effect of beam squint on the range of the equivalent angle $\varphi_{l_3}^{\mathrm{C}}$. Without loss of generality, we consider the path of $l_3 = 1$. Then equation of (19) can be simplified as

$$\Gamma^f(x) = c_1 e^{-j2\pi f\tau_1^{\mathrm{C}}} \frac{\sin\left(\pi M\left(1+\frac{f}{f_c}\right)(x-\varphi_1^{\mathrm{C}})\right)}{\sin\left(\pi\left(1+\frac{f}{f_c}\right)(x-\varphi_1^{\mathrm{C}})\right)}, \tag{20}$$

where the range of φ_1^{C} is $(-1, 1)$.

For the conventional wideband channel model disregarding the effect of beam squint, the term of $\frac{f}{f_c}$ is ignored in (20). Hence, the steering vector is expressed as $\mathbf{a}(\varphi_{l_3}^{\mathrm{C}}) = \left[1, e^{-j2\pi\varphi_{l_3}^{\mathrm{C}}} \ \cdots, e^{-j2\pi(M-1)\varphi_{l_3}^{\mathrm{C}}}\right]^{\mathrm{T}}$ and for the single path of $l_3 = 1$, $\Gamma^f(x) = c_1 e^{-j2\pi f\tau_1^{\mathrm{C}}} \frac{\sin\left(\pi M\left(x-\varphi_1^{\mathrm{C}}\right)\right)}{\sin\left(\pi\left(x-\varphi_1^{\mathrm{C}}\right)\right)}$. Since the value of M is usually large, the function $\frac{1}{M}\left|\frac{\sin(\pi M\xi)}{\sin(\pi\xi)}\right|$ only has nonzero value when $\xi \in \mathbb{Z}$. Thus, we can estimate the equivalent angle $\varphi_{l_3}^{\mathrm{C}}$ of the cascaded channel $\mathbf{h}(f)$ by finding the peak of the function $\Gamma^f(x)$. When we search for x between -1 and 1, the function $\Gamma^f(x)$ reaches its peak at $x = \varphi_1^{\mathrm{C}}$, while the other peak at $x = \varphi_1^{\mathrm{C}} + 1$ ($\varphi_1^{\mathrm{C}} < 0$) or $x = \varphi_1^{\mathrm{C}} - 1$ ($\varphi_1^{\mathrm{C}} > 0$). Then, we will get two estimated angles, one of which is the actual angle φ_1^{C} and the other one is separated from the actual angle by 1, which is termed as false angle. The false angle is frequency-independent, and the steering vector $\mathbf{a}(\varphi_1^{\mathrm{C}})$ of the actual angle and steering vector $\mathbf{a}(\varphi_1^{\mathrm{C}} \pm 1)$ of false angle are equivalent. Thus, it is easy to understand that searching across the range of $x \in (-1, 1)$ is equivalent to $x \in [-\frac{1}{2}, \frac{1}{2})$, and the equivalent angle estimation is formulated as

$$x = \begin{cases} \varphi_1^{\mathrm{C}} & \varphi_1^{\mathrm{C}} \in [-\frac{1}{2}, \frac{1}{2}) \\ \varphi_1^{\mathrm{C}} - 1 & \varphi_1^{\mathrm{C}} \in [\frac{1}{2}, 1) \\ \varphi_1^{\mathrm{C}} + 1 & \varphi_1^{\mathrm{C}} \in (-1, -\frac{1}{2}], \end{cases} \tag{21}$$

where $x \in [-\frac{1}{2}, \frac{1}{2})$. Explicitly, we only have to consider the range of equivalent angles in the steering vector as $[-\frac{1}{2}, \frac{1}{2})$ instead of $(-1, 1)$. Similar conclusions can be found in [23].

However, when we consider the effect of beam squint, the range of equivalent angles in the steering vector $\mathbf{a}\left(\left(1+\frac{f}{f_c}\right)x\right)$ can no longer be equivalent to $[-\frac{1}{2}, \frac{1}{2})$. That is because the false angles become frequency-dependent. Specifically, when we search for x between -1 and 1, we will get the actual peak at $x = \varphi_1^{\mathrm{C}}$ and the false peak at $x = \varphi_1^{\mathrm{C}} + \frac{f_c}{f+f_c}$ ($-1 < \varphi_1^{\mathrm{C}} \leq 0$) or $x = \varphi_1^{\mathrm{C}} - \frac{f_c}{f+f_c}$ ($0 < \varphi_1^{\mathrm{C}} < 1$), which is frequency-dependent. Thus, we will get two estimated angles, one of which is the actual angle of φ_1^{C} and the other is the false angle of $\varphi_1^{\mathrm{C}} \pm \frac{f_c}{f+f_c}$. The false angle is separated from the actual angle by $\frac{f_c}{f+f_c}$. Thus, the squint of false angle over all subcarriers is $\frac{f_c}{0+f_c} - \frac{f_c}{W+f_c}$, which is independent of the specific equivalent angle φ_1^{C}.

The false angle has a grave impact on channel estimation. We will elaborate by considering an example, where there is a single path with equivalent angle of $\varphi_1^{\mathrm{C}} = -1/6$ in cascaded FD channel $\mathbf{h}(f)$. The angular-domain index x is selected from -1 to 1, the carrier frequency is $f_c = 10$GHz, the number of subcarrier is $N_p = 128$ and the system bandwidth is $W = 500$MHz. Thus, the subcarrier frequency is $f = \frac{n_p}{N_p}W$, where $n_p \in \{0, 1, \cdots, N_p - 1\}$. The function $\Gamma^f(x)$ at subcarriers $30, 60, 90, 120$ is shown in Fig. 3. We observe that there are two peaks over the angular range of $(-1, 1)$, which correspond to the index of the actual angle and the index of false angle. The index of the actual angle remains $-\frac{1}{6}$ for these four subcarriers, as well as for all other subcarriers. By contrast, the index of the false angle varies from about 0.78 to 0.82 between subcarriers 30 and 120. Furthermore, we note that the peak of the false angle is quite comparable to that of the actual angle, which will seriously interfere with the estimation of the actual angle, since we cannot readily distinguish the actual angle and false angle by searching for the peak of the mutual correlation function. This motivates us to propose an efficient estimation method of the cascaded FD channel $\mathbf{h}(f)$, which is robust to beam squint.

IV. IRS-Aided Channel Estimation

A. Overview of the Proposed Channel Estimation Method

In this section, we propose a TS-OMP based method for estimating the parameters of equivalent angles, delays and gains of cascaded channel. The philosophy of proposed method is as follows:

- In the first stage, we propose a block-sparse processing based method for estimating the equivalent angles of

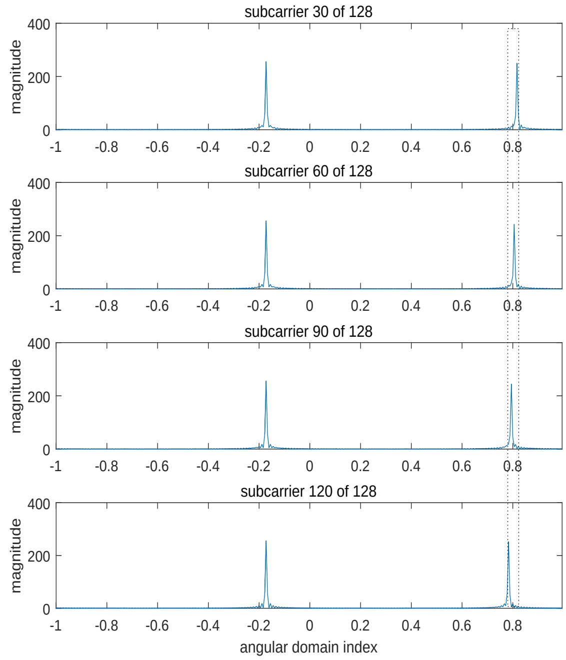


Fig. 3. The mutual correlation function $\Gamma^f(x)$ between steering vectors $\mathbf{a}\left((1+\frac{f}{f_c})x\right)$ and the cascaded FD channel $h(f)$ at different subcarriers, for $\varphi_1^{\mathrm{C}}=-1/6$, $M=256$, $N_p=128$, $f_c=$10GHz, $W=$ 500MHz.

the cascaded channel. As shown in Fig. 3, we find that the index of the actual angle remains the same for all subcarriers, while that of the false angle is different for the different subcarriers. Thus, if we accumulate the function $\Gamma^f(\varphi_1^{\mathrm{C}})$ across the different subcarriers, where φ_1^{C} denotes the actual angle of path $l_3=1$, the accumulated value $\Gamma^f(\varphi_1^{\mathrm{C}})$ corresponding to the actual angle will be quite high. By contrast, when we accumulate the function $\Gamma^f(\varphi_1^{\mathrm{F}})$ across the subcarriers, where $\varphi_1^{\mathrm{F}}=\varphi_1^{\mathrm{C}}\pm\frac{f_c}{f+f_c}$ denotes the false angle corresponding to the path of $l_3=1$, the value $\Gamma^f(\varphi_1^{\mathrm{F}})$ for the angle φ_1^{F} accumulated over the subcarriers does not increase as fast as the accumulated value of $\Gamma^f(\varphi_1^{\mathrm{C}})$. Therefore, we propose to accumulate the mutual correlation function $\Gamma^f(x)$ over the subcarriers for eliminating the influence of the false angle.

- In the second stage, based on the estimated equivalent angles, we propose an OMP based method for estimating the equivalent path delays and gains of the cascaded channel. Since there may be multiple paths having different delays for the same equivalent angle, an appropriate stopping condition is needed.

B. Equivalent Angle Estimation for the Cascaded Channel

Observe from (16), that there may be multiple paths having the same equivalent angle among the L_1L_2 number of cascaded channel paths. Thus, (15) can be rewritten as

$$\begin{aligned} y(f) &= \sum_{l_3=1}^{L_1L_2}\boldsymbol{\theta}^{\mathrm{T}}\mathbf{a}\left(\left(1+\frac{f}{f_c}\right)\varphi_{l_3}^{\mathrm{C}}\right)c_{l_3}e^{-j2\pi f\tau_{l_3}^{\mathrm{C}}}+n(f)\\ &= \sum_{i=1}^{N_a}\boldsymbol{\theta}^{\mathrm{T}}\mathbf{a}\left(\left(1+\frac{f}{f_c}\right)\overline{\varphi}_i\right)\sum_{j_i=1}^{J_i}\overline{c}_{j_i}e^{-j2\pi f\overline{\tau}_{j_i}}+n(f), \end{aligned} \tag{22}$$

where $\overline{c}_{j_i}$ and $\overline{\tau}_{j_i}$ denote the equivalent gain and delay of the j_i-th cascaded channel path, $j_i=\{1,2,\cdots,J_i\}$ and J_i denotes the number of cascaded channel paths having the same equivalent angle $\overline{\varphi}_i$, where $i=\{1,2,\cdots,N_a\}$ and N_a denotes the number of different values of equivalent angles $\varphi_{l_3}^{\mathrm{C}}$ in the cascaded channel. Thus, we have $\sum_i^{N_a}J_i=L_1L_2$.

In the n_p-th subcarrier, i.e., $f=\frac{n_pW}{N_p}$, the signal received by the user can be expressed as

$$y(n_p)=\boldsymbol{\theta}^{\mathrm{T}}\mathbf{A}(n_p)\mathbf{z}(n_p)+n(n_p), \tag{23}$$

where we denote $y(\frac{n_pW}{N_p})$ and $n(\frac{n_pW}{N_p})$ as $y(n_p)$ and $n(n_p)$ for simplicity. The sparse vector $\mathbf{z}(n_p)\in\mathbb{C}^{N_d\times 1}$ has N_a none-zero elements. The matrix $\mathbf{A}(n_p)\in\mathbb{C}^{M\times N_d}$ is the dictionary matrix composed of N_d steering vectors, which can be expressed as

$$\begin{aligned} \mathbf{A}(n_p)=&[\mathbf{a}((-1)(1+\frac{n_pW}{N_pf_c})),\mathbf{a}((-1+\frac{2}{N_d})(1+\frac{n_pW}{N_pf_c})),.\\ &\cdots,.\mathbf{a}((1-\frac{2}{N_d})(1+\frac{n_pW}{N_pf_c}))]. \end{aligned} \tag{24}$$

In order to estimate the equivalent angles $\{\overline{\varphi}_i\}_{i=1}^{N_a}$ of the cascaded channel, we need multiple sets of observations at the user. We consider N_s OFDM symbols and assume that the IRS reflection coefficients are reconfigured during different OFDM symbols. Thus, the signal received by the user in the n_p-th subcarrier after N_s OFDM symbols can be expressed as

$$\begin{aligned} \mathbf{y}(n_p)&=\boldsymbol{\Theta}\mathbf{A}(n_p)\mathbf{z}(n_p)+\mathbf{n}(n_p)\\ &=\mathbf{F}(n_p)\mathbf{z}(n_p)+\mathbf{n}(n_p), \end{aligned} \tag{25}$$

where $\mathbf{y}(n_p)=[y^1(n_p),y^2(n_p),\cdots,y^{N_s}(n_p)]^{\mathrm{T}}$ and the element $y^{N_s}(n_p)$ denotes the signal received during the N_s-th OFDM symbol. The matrix $\boldsymbol{\Theta}=[\boldsymbol{\theta}^1,\boldsymbol{\theta}^2,\cdots,\boldsymbol{\theta}^{N_s})]^{\mathrm{T}}\in\mathbb{C}^{N_s\times M}$ and the vector $\boldsymbol{\theta}^{N_s}\in\mathbb{C}^{M\times 1}$ denotes the reflection coefficient vector during the N_s-th OFDM symbol. The noise vector in (25) is $\mathbf{n}(n_p)=[n^1(n_p),n^2(n_p),\cdots,n^{N_s}(n_p)]^{\mathrm{T}}$ and the element $n^{N_s}(n_p)$ represents the Gaussian noise during the N_s-th OFDM symbol. While we have matrix $\mathbf{F}(n_p)=\boldsymbol{\Theta}\mathbf{A}(n_p)\in\mathbb{C}^{N_s\times N_d}$. To elaborate, equation (25) represents a sparse signal recovery problem, where $\mathbf{y}(n_p)$ is the observation vector, $\mathbf{F}(n_p)$ is the measurement matrix and $\mathbf{z}(n_p)$ is the sparse vector to be recovered. By using classic OMP algorithm, we can get the non-zero elements' indices in the vector $\mathbf{z}(n_p)$, which are expressed as $\mathcal{I}^a=\{\mathcal{I}_i^{\mathrm{a}}\}_{i=1}^{i=N_a}$. The equivalent angles $\overline{\varphi}_i$ can be estimated as $\overline{\varphi}_i=-1+\frac{2(\mathcal{I}_i^{\mathrm{a}}-1)}{N_d}$.

However, as discussed in Section III, we have to combine several subcarriers for suppressing the deleterious influence of false angles. Collecting the signal $\mathbf{y}(n_p)$ received in N_{P1}

subcarriers where N_{P1} is the number of pilot subcarriers, we have

$$\overline{\mathbf{y}} = \overline{\mathbf{F}}\overline{\mathbf{z}} + \overline{\mathbf{n}}, \tag{26}$$

where

$$\overline{\mathbf{F}} = \begin{bmatrix} \mathbf{F}(n_1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{F}(n_2) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{F}(n_{N_{P1}}) \end{bmatrix} \in \mathbb{C}^{N_s N_{P1} \times N_{P1} N_d}, \tag{27}$$

and $\overline{\mathbf{y}} = [\mathbf{y}(n_1)^{\mathrm{T}}, \mathbf{y}(n_2)^{\mathrm{T}}, \cdots, \mathbf{y}(n_{N_{P1}})^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{N_s N_{P1} \times 1}$, $\overline{\mathbf{z}} = [\mathbf{z}(n_1)^{\mathrm{T}}, \mathbf{z}(n_2)^{\mathrm{T}}, \cdots, \mathbf{z}(n_{N_{P1}})^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{N_{P1} N_d \times 1}$, $\overline{\mathbf{n}} = [\mathbf{n}(n_1)^{\mathrm{T}}, \mathbf{n}(n_2)^{\mathrm{T}}, \cdots, \mathbf{n}(n_{N_{P1}})^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{N_s N_{P1} \times 1}$. Equation (26) also represents a sparse signal recovery problem. Since the actual angle is the same for all subcarriers, the vector $\overline{\mathbf{z}}$ exhibits inherent block sparsity. By exploiting the block-sparsity, we propose a TS-OMP algorithm for eliminating the effect of false angles and get an accurate estimate of the cascaded channel angles, which is summarized in Algorithm 1.

In Step 1 of Algorithm 1, we apply an elementary transformation to (26) for locating those particular elements of vector $\overline{\mathbf{z}}$ that represent the same equivalent angle in different subcarriers, which are then lumped together.. Specifically, this operation can be expressed as $[\widetilde{\mathbf{z}}]_{(n_d-1)N_{P1}+n_p} = [\overline{\mathbf{z}}]_{(n_p-1)N_d+n_d}$, where $n_p = \{1, 2, \cdots, N_{P1}\}$ and $n_d = \{1, 2, \cdots, N_d\}$. To ensure that equation (26) holds, we apply a similar transformation to the columns of matrix $\overline{\mathbf{F}}$, which can be expressed as $[\widetilde{\mathbf{F}}]_{:,(n_d-1)N_{P1}+n_p} = [\overline{\mathbf{F}}]_{:,(n_p-1)N_d+n_d}$. Thus, (26) can be rewritten as $\overline{\mathbf{y}} = \widetilde{\mathbf{F}}\widetilde{\mathbf{z}} + \overline{\mathbf{n}}$. In this way, the N_{P1} adjacent elements of vector $\widetilde{\mathbf{z}}$ are either all zero elements or all non-zero elements. In Step 3, we estimate the indices of non-zero elements of vector $\widetilde{\mathbf{z}}$. Traditionally, we can obtain the indices by finding the maximum value of objective function (OF), which can be expressed as $\mathcal{I}_i^a = \underset{\mathcal{I}_i^a \in \{1,2,\cdots,N_d\}}{\arg\max} \sum_{k=1}^{N_{P1}} \left\| ([\widetilde{\mathbf{F}}_{b_1}]_{:,t})^{\mathrm{H}} \mathbf{r}_a^{i-1} \right\|_2^2$, where $t = (\mathcal{I}_i^a - 1)N_{P1} + k$ and the expression $\mathbf{r}_a^{i-1}$ can be found in Step 8, while $\widetilde{\mathbf{F}}_{b_1}$ denotes the buff matrix processing in Step 6, where we set $\widetilde{\mathbf{F}}_{b_1} = \widetilde{\mathbf{F}}$ in Step 1 of Algorithm 1. We find that there are many zeros elements in each column of matrix $\widetilde{\mathbf{F}}$. To reduce the complexity, we remove those items of $\mathbf{r}_a^{i-1}$ that interact with the zero elements of $\mathbf{F}_{b_1}$. The OF can be rewritten as $\mathcal{I}_i^a = \underset{\mathcal{I}_i^a \in \{1,2,\cdots,N_d\}}{\arg\max} \sum_{k=1}^{N_{P1}} \left\| ([\widetilde{\mathbf{F}}_{b_1}]_{j:q,t})^{\mathrm{H}} [\mathbf{r}_a^{i-1}]_{j:q} \right\|_2^2$, where $j = (k-1)N_s + 1$, $q = kN_s$ and $t = (\mathcal{I}_i^a - 1)N_{\mathrm{P1}} + k$. Repeat Steps 3 to 9 of Algorithm 1 until the stopping criterion is met, delivering the index set $\mathcal{I}^a = \{\mathcal{I}_i^a\}_{i=1}^{N_a}$.

Then, we can use least squares (LS) estimator to estimate the non-zero elements of vector $\mathbf{z}(n_p)$, which can be expressed as

$$\mathrm{LS}: \quad [\mathbf{z}(n_p)]_{\mathcal{I}^a} = \left([\mathbf{F}(n_p)]_{:,\mathcal{I}^a}\right)^{\dagger} \mathbf{y}(n_p), \tag{28}$$

where

$$([\mathbf{F}(n_p)]_{:,\mathcal{I}^a})^{\dagger} = \left[([\mathbf{F}(n_p)]_{:,\mathcal{I}^a})^{\mathrm{H}} [\mathbf{F}(n_p)]_{:,\mathcal{I}^a}\right]^{-1} ([\mathbf{F}(n_p)]_{:,\mathcal{I}^a})^{\mathrm{H}} \tag{29}$$

denotes the pseudo-inverse of $[\mathbf{F}(n_p)]_{:,\mathcal{I}^a}$.

C. Delay- and Gain- Estimation of the Cascaded Channel

According to (22) and (23), the i-th non-zero element of vector $\mathbf{z}(n_p)$ can be expressed as

$$[\mathbf{z}(n_p)]_{\mathcal{I}_i^a} = \sum_{j_i=1}^{J_i} \overline{c}_{j_i} e^{-j2\pi \frac{n_p W}{N_P} \overline{\tau}_{j_i}}. \tag{30}$$

Upon collecting $[\mathbf{z}(n_p)]_{\mathcal{I}_i^a}$ of pilot-subcarriers, we have

$$\mathbf{z}^{\mathcal{I}_i^a} = \sum_{j_i=1}^{J_i} \overline{c}_{j_i} \overline{\mathbf{b}}\left(\overline{\tau}_{j_i}\right), \tag{31}$$

where $\mathbf{z}^{\mathcal{I}_i^a} = \left[[\mathbf{z}(n_1)]_{\mathcal{I}_i^a}, [\mathbf{z}(n_2)]_{\mathcal{I}_i^a}, \cdots, [\mathbf{z}(n_{N_{P1}})]_{\mathcal{I}_i^a}\right]^{\mathrm{T}}$ and $\overline{\mathbf{b}}\left(\overline{\tau}_{j_i,i}\right)$ can be expressed as

$$\overline{\mathbf{b}}\left(\overline{\tau}_{j_i}\right) = \left[e^{-j2\pi \frac{n_1 W}{N_P} \overline{\tau}_{j_i}}, e^{-j2\pi \frac{n_2 W}{N_P} \overline{\tau}_{j_i}}, \cdots, e^{-j2\pi \frac{n_{N_{P1}} W}{N_P} \overline{\tau}_{j_i}}\right]^{\mathrm{T}}. \tag{32}$$

Equation (31) can also be formulated as a sparse signal recovery problem, which can be expressed as

$$\mathbf{z}^{\mathcal{I}_i^a} = \mathbf{B}\overline{\mathbf{c}}^{\mathcal{I}_i^a}, \tag{33}$$

where the sparse vector $\overline{\mathbf{c}}^{\mathcal{I}_i^a} \in \mathbb{C}^{N_\tau \times 1}$ has J_i non-zero elements and $\mathbf{B} = [\mathbf{b}(0), \mathbf{b}(1), \cdots, \mathbf{b}(N_\tau - 1)] \in \mathbb{C}^{N_{P1} \times N_\tau}$ is the dictionary matrix for the delay domain composed of N_τ steering vectors. The vector $\mathbf{b}(k)$ of equation(33) can be expressed as

$$\mathbf{b}(k) = [e^{-j2\pi \frac{W}{N_p} \frac{kT_\tau}{N_\tau} n_1}, e^{-j2\pi \frac{W}{N_p} \frac{kT_\tau}{N_\tau} n_2}, \cdots, e^{-j2\pi \frac{W}{N_p} \frac{kT_\tau}{N_\tau} n_{N_{P1}}}]^{\mathrm{T}}, \tag{34}$$

where T_τ is the maximum channel delay in the cascaded channel. In this way, we can utilize our OMP-based method to estimate the vector $\overline{\mathbf{c}}^{\mathcal{I}_i^a}$. The non-zero elements of vector $\overline{\mathbf{c}}^{\mathcal{I}_i^a}$ correspond to the channel gain $\overline{c}_{j_i}$ and the index corresponds to the channel delay $\overline{\tau}_{j_i}$. The exact details are shown in Algorithm 1.

According to (22), the FD channel response of the $\mathbf{h}(n_p)$ of the n_p-th subcarrier in (16) can be expressed as

$$\mathbf{h}(f) = \sum_{i=1}^{N_d} \mathbf{a}\left(\left(1 + \frac{f}{f_c}\right)\overline{\varphi}_i\right) \sum_{j_i=1}^{J_i} \overline{c}_{j_i} e^{-j2\pi f \overline{\tau}_{j_i}}. \tag{35}$$

Thus, after obtain the parameters of $\{\overline{\varphi}_i, \overline{c}_{j_i}, \overline{\tau}_{j_i}\}$, where $i \in \{1, 2, \cdots, N_a\}$ and $j_i = \{1, 2, \cdots, J_i\}$, the cascaded FD channel $h(f)$ of all the subcarriers can be obtained according to (35).

V. Pilot Design

In this section, we propose a cross-entropy based pilot design for improving the estimation performance of the channel's equivalent angles, delays and gains. The pilot design guidelines have to consider both the first stage and the second stage of the proposed TS-OMP algorithm, which are as follows:

- In the first stage of the TS-OMP algorithm, we estimate the equivalent angles of cascaded FD channel. We define

Input: Received signal $\overline{\mathbf{y}}$. Measurement matrix $\overline{\mathbf{F}}$ and dictionary matrix $\mathbf{B}$.
Initialization: $\mathbf{r}_a^0 = \overline{\mathbf{y}}$, $\mathcal{I}^a = [\,]$, $\widetilde{\mathbf{F}}_{b_2} = [\,]$ and iteration counter $i = 1$.

Perform elementary transformation: $[\widetilde{\mathbf{F}}]_{:,(n_d-1)N_{P1}+n_p} = [\overline{\mathbf{F}}]_{:,(n_p-1)N_d+n_d}$ and $[\widetilde{\mathbf{z}}]_{(n_a-1)N_{P1}+n_p} = [\overline{\mathbf{z}}]_{(n_p-1)N_a+n_a}$. We set $\widetilde{\mathbf{F}}_{b_1} = \widetilde{\mathbf{F}}$.
repeat
  Estimate angular support:
$$\mathcal{I}_i^a = \underset{\mathcal{I}_i^a \in \{1,2,\cdots,N_d\}}{\arg\max} \sum_{k=1}^{N_{P1}} \left\| \left([\widetilde{\mathbf{F}}_{b_1}]_{j:q,t}\right)^{\mathrm{H}} [\mathbf{r}_a^{i-1}]_{j:q} \right\|_2^2,$$
  where $j = (k-1)N_s + 1$, $q = kN_s$ and $t = (\mathcal{I}_i^a - 1)N_{\mathrm{P}1} + k$.
  Update index set: $\mathcal{I}^a = \mathcal{I}^a \cup \{\mathcal{I}_i^a\}$.
  Expand matrix: $\widetilde{\mathbf{F}}_{b_2} = \widetilde{\mathbf{F}}_{b_2} \cup [\widetilde{\mathbf{F}}_{b_1}]_{:,(\mathcal{I}_i^a-1)N_{P1}+1:\mathcal{I}_i^a N_{P1}}$.
  Update matrix: $[\widetilde{\mathbf{F}}_{b_1}]_{:,(\mathcal{I}_i^a-1)N_{P_1}+1:\mathcal{I}_i^a N_{P_1}} = \mathbf{0}$.
  Calculate the value $\mathbf{z}_b$:
$$\mathbf{z}_b = \left((\widetilde{\mathbf{F}}_{b_2})^{\mathrm{H}} \widetilde{\mathbf{F}}_{b_2}\right)^{-1} (\widetilde{\mathbf{F}}_{b_2})^{\mathrm{H}} \overline{\mathbf{y}}.$$
  Update residual: $\mathbf{r}_a^i = \overline{\mathbf{y}} - \widetilde{\mathbf{F}}_{b_2} \mathbf{z}_b$.
  Set $i = i + 1$.
until $\frac{\|\mathbf{r}_a^{i-2} - \mathbf{r}_a^{i-1}\|_2^2}{\|\overline{\mathbf{y}}\|_2^2} \le \zeta$
Exploit LS estimator to calculate vector $\overline{\mathbf{z}}$ according to (28).
for $1 \le i \le N_d$ **do**
  Initialization: $\mathbf{B}_{b_1} = \mathbf{B}$, $\mathbf{B}_{b_2} = [\,]$, $\mathbf{R}_\tau^0 = \mathbf{z}_{P1}^{\mathcal{I}_i^a}$, $\mathbf{T}_{\mathbf{b}} = [\,]$ and the counter $t = 1$.
  repeat
    Estimate delay support:
$$k = \underset{k \in \{1,2,\cdots,N_\tau\}}{\arg\max} \left\| \left([\mathbf{B}_{b_1}]_{:,k}\right)^{\mathrm{H}} \mathbf{r}_\tau^{t-1} \right\|_2^2$$
    Update index set: $\mathbf{T}_b = \mathbf{T}_b \cup \{k\}$.
    Expand matrix: $\mathbf{B}_{b_2} = \mathbf{B}_{b_2} \cup [\mathbf{B}_{b_1}]_{:,k}$.
    Update matrix: $[\mathbf{B}_{b_1}]_{:,k} = \mathbf{0}$.
    Calculate the value $\overline{\mathbf{c}}_{\mathcal{I}_i^a}$:
$$\overline{\mathbf{c}}_{\mathcal{I}_i^a} = \left((\mathbf{B}_{b_2})^{\mathrm{H}} \mathbf{B}_{b_2}\right)^{-1} (\mathbf{B}_{b_2})^{\mathrm{H}} \mathbf{z}^{\mathcal{I}_i^a}.$$
    Update residual: $\mathbf{r}_\tau^t = \mathbf{z}^{\mathcal{I}_i^a} - \mathbf{B}_{b_2} \overline{\mathbf{c}}_{\mathcal{I}_i^a}$.
    Set $t = t + 1$.
  until $\frac{\|\mathbf{r}_\tau^{t-2} - \mathbf{r}_\tau^{t-1}\|_2^2}{\left\|\mathbf{z}_{P1}^{\mathcal{I}_i^a}\right\|_2^2} \le \zeta$
  Output the set $\mathbf{T}_b$ and vector $\overline{\mathbf{c}}_{\mathcal{I}_i^a}$.
  We have the channel angle $\overline{\varphi}_i = -1 + \frac{2(\mathcal{I}_i^{\mathrm{a}} - 1)}{N_a}$, the delay $\overline{\tau}_{j_i} = \{\frac{[\mathbf{T}_b]_{j_i} T_\tau}{N_\tau} | j_i \in \{1, 2, \cdots, J_i\}\}$ and the channel gain $\overline{c}_{j_i} = [\overline{\mathbf{c}}_{\mathcal{I}_i^a}]_{j_i}$.
end for
Output: $\overline{\varphi}_i, \overline{c}_{j_i}, \overline{\tau}_{j_i}$

Algorithm 1: The proposed TS-OMP algorithm

$v_{n_1} = \varphi_1^{\mathrm{C}} \pm \frac{f_{n_1}}{f_{n_1} + f_c}$ and $v_{n_{N_{P1}}} = \varphi_1^{\mathrm{C}} \pm \frac{f_{n_{N_{P1}}}}{f_{n_{N_{P1}}} + f_c}$ as the false angle corresponding to the actual angle φ_1^{C} at the n_1-th and $n_{N_{P1}}$-th subcarrier, where $f_{n_1} = \frac{n_1 W}{N_P}$ and $f_{n_{N_{P1}}} = \frac{n_{N_{P1}} W}{N_P}$ denotes the frequency of the n_1-th and $n_{N_{P1}}$-th subcarrier, respectively. To eliminate the interference imposed by the false angle, we should decrease the accumulated values of function $\Gamma^f(\varphi_{l_3}^{\mathrm{F}})$ over the pilot- subcarriers, as mentioned in Section III. This means that the false angle $\varphi_{l_3}^{\mathrm{F}} = \varphi_{l_3}^{\mathrm{C}} \pm \frac{f_c}{f + f_c}$ corresponding to a certain actual angle $\varphi_{l_3}^{\mathrm{C}}$ at different subcarriers should be at a different angular index in Fig. 3. We have defined the resolution of angular-domain dictionary matric as $\frac{2}{N_d}$ in (24). Thus, our first guideline is formulated as $v_{n_1} - v_{n_{N_{P1}}} > \frac{2}{N_d}$.

- In the second stage of our TS-OMP algorithm, we formulate the problem of estimating the equivalent channel delays and gains as a sparse signal recovery problem. Naturally, a high grade of orthogonality of the measurement matrix columns is preferred for sparse signal recovery [31] [32]. Hence the second guideline of pilot design should ensure that the measurement matrix $\mathbf{B}$ in (33), which depends on the pilot-subcarriers, has near-orthogonal columns.

As for our first guideline of pilot design, we have

$$\frac{f_c}{f_c + f_{n_1}} - \frac{f_c}{f_c + f_{n_{N_{P1}}}} \ge \frac{2}{N_d}. \tag{36}$$

Thus, the span $D = f_{n_{N_{P1}}} - f_{n_1}$ of pilots has to satisfy

$$D \ge \frac{f_{n_1}(f_c + \frac{N_d}{2} f_c) + f_c^2 - f_{n_1}(\frac{N_d}{2} - 1) - f_{n_1}^2}{(\frac{N_d}{2} - 1) f_c + f_{n_1}}. \tag{37}$$

As for our second guideline of pilot design, we have to design $\mathbf{B}$ so that $\mathbf{B}^{\mathrm{H}}\mathbf{B}$ becomes an approximately identity matrix,

$$\mathbf{B}^{\mathrm{H}}\mathbf{B} \approx N_{\mathrm{P}1} \mathbf{I}_{N_\tau}, \tag{38}$$

where $\mathbf{I}_{N_\tau}$ is the identity matrix of size $N_\tau \times N_\tau$. Thus, our pilot design problem can formulated as

$$\min_{\mathbf{B}} \mu(\mathbf{B}) = \left\|\mathbf{B}^{\mathrm{H}}\mathbf{B} - N_{\mathrm{P}1}\mathbf{I}_{N_\tau}\right\|_2^2 \quad \text{s.t.} \quad (37), \tag{39}$$

where the measurement matrix $\mathbf{B}$ is dependent of the subcarrier index of pilots $\{n_1, n_2, \cdots, n_{N_{P1}}\}$. The optimization problem (39) can then be solved by exhaustive search. However, the complexity of exhaustive search is excessive. For example, if the number of subcarriers is 128 and the number of pilots is 6, we need 5.4×10^9 searches to find the optimal pilots. Inspired-by the cross-entropy method of [33], we hence propose the pilot design of Algorithm 2. The design criterion of the proposed algorithm is to get minimum value of $\mu\mathbf{B}$ under the constraint of (37).

In Step 1 of Algorithm 2, we define the probability vector $\mathbf{P}_{\mathrm{B}}^i \in \mathbb{C}^{N_P \times 1}$ of the i-th iteration, where the n_p-th element of $\mathbf{P}_{\mathrm{B}}^i$ denotes the probability that the n_p-th element of $\mathbf{d}_{n_c}$ is equal to 1. The n_p-th element of $\mathbf{d}_{n_c} \in \mathbb{C}^{N_{\mathrm{P}} \times 1}$

Input: The number of candidates N_c, the number of elites N_e and the number of iterations N_{iter}. The number of subcarriers $N_{\rm P}$ and the number of pilots $N_{\rm P1}$.
Initialization: $P_{\rm B}^0 = \frac{N_{\rm P1}}{N_{\rm P}} \times \mathbf{1}_{N_{\rm P},1}$ and iteration counter $i = 0$.
for $0 \leq i \leq N_{iter} - 1$ **do**
1. Randomly generate N_c candidate vectors $\{\mathbf{d}_{n_c}\}_{n_c=1}^{N_c}$ according to $\mathbf{P}_{\rm B}^i$ and constraint of (37), where $\mathbf{d}_{n_c} \in \{0,1\}^{N_{\rm P}\times 1}$. Generate the N_c candidate measurement matrix $\{\mathbf{B}_{n_c}\}_{n_c=1}^{N_c}$ according to $\mathbf{d}_{n_c}$.
2. Calculate the objective function: $\mu(\mathbf{B}_{n_c}) = \left\|\mathbf{B}_{n_c}^{\rm H}\mathbf{B}_{n_c} - N_{\rm P1}\mathbf{I}_{N_\tau}\right\|_2^2$
3. Sort the objective function $\mu(\mathbf{B}_{n_c})$ in ascending order: $\mu(\mathbf{B}_{d_{e,1}}) \leq \mu(\mathbf{B}_{d_{e,2}}) \leq \cdots \leq \mu(\mathbf{B}_{d_{e,N_c}})$;
4. $\mathbf{P}_{\rm B}^{i+1} = (1-\overline{w})\mathbf{P}_{\rm B}^i + \frac{\overline{w}}{N_e}(\mathbf{d}_{d_{e,1}} + \mathbf{d}_{d_{e,2}} + \cdots + \mathbf{d}_{d_{e,N_e}})$;
5. $i = i + 1$;

end
Output: Designed pilot (the index of none-zero elements in $\mathbf{d}_{d_{e,1}}$).

Algorithm 2: Pilot design

indicates whether the pilot found occupies the n_p-th subcarrier. Specifically, if the n_p-th element of vector $\mathbf{d}_{n_c}$ is equal to 1, then the n_p-th subcarrier is a pilot. Otherwise, if the n_p-th element of vector $\mathbf{d}_{n_c}$ is equal to 0, the n_p-th subcarrier is not a pilot. We utilize the probability matrix $\mathbf{P}_{\rm B}^i \in \mathbb{C}^{N_P\times 1}$ to generate N_c candidate vectors expressed as $\mathbf{d}_{n_c} \in \{0,1\}^{N_P\times 1}$. Then, we check whether the candidate pilots satisfy (37). If not, we regenerate the pilots until the number of candidates reaches N_c. we then use the candidate pilots chosen to generate candidate measurement matrix $\{\mathbf{B}_{n_c}\}_{n_c=1}^{N_c}$ according to (34). In Step 2 of Algorithm 2, we calculate the OF: $\mu(\mathbf{B}_{n_c}) = \left\|\mathbf{B}_{n_c}^{\rm H}\mathbf{B}_{n_c} - N_{\rm P1}\mathbf{I}_{N_\tau}\right\|_2^2$. In Step 3 of Algorithm 2, we sort the OF $\mu(\mathbf{B}_{n_c})$ in ascending order and retain the first N_e elements. Then, we select the corresponding N_e candidates $\{\mathbf{B}_{n_c}\}_{n_c=d_{e,1}}^{n_c=d_{e,N_e}}$ as elites and record their indices as $\{d_{e,1}, d_{e,2}, \cdots, d_{e,N_e}\}$. In Step 4 of Algorithm 2, we update the probability matrix as $\mathbf{P}_{\rm B}^{i+1} = (1-\overline{w})\mathbf{P}_{\rm B}^i + \frac{\overline{w}}{N_e}(\mathbf{d}_{d_{e,1}} + \mathbf{d}_{d_{e,2}} + \cdots + \mathbf{d}_{d_{e,N_e}})$, where $\overline{w}$ is smooth parameter and we set as 0.8 in this paper. After N_{iter} iterations, we arrive at the pilot subcarrier index set of $\{n_1, \cdots, n_p, \cdots, n_{N_{P1}}\}$, where n_p is the index of non-zero elements in $\mathbf{d}_{d_{e,1}}$. The design principle of the parameter N_c, N_e and N_{iter} is an open question. Usually, the number of elites N_e is smaller than the number of candidates N_c. In the simulation, we will show the effect of different values of N_c, N_e and N_{iter} on the performance of the proposed Algorithm 2.

VI. Simulation Results

In this section, we present our simulation results for characterizing the performance of the proposed channel estimation and pilot design method. We assume that the IRS elements of our ULA have a half-wavelength spacing. The carrier frequency is $f_c = 20$ GHz, The AOA is $\varphi_{l_1}^{\rm TR} \in [-\frac{\pi}{2}, \frac{\pi}{2})$ and the DOA is $\varphi_{l_2}^{\rm RR} \in [-\frac{\pi}{2}, \frac{\pi}{2})$. The maximum delay spread of the cascaded channel is 200 ns. The signal bandwidth is $W = 510$ MHz and the other parameteters are set to $M = 256$, $N_p = 128$, $N_c = 100$, $N_e = 20$ and $L_1L_2 = 6$. We define the normalized mean-squared error (NMSE) of the cascaded channel at all subcarriers as

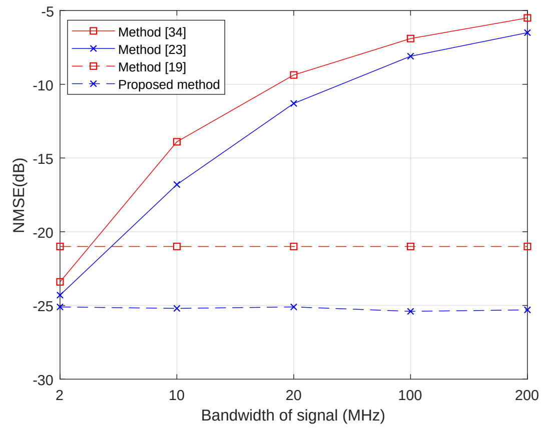


Fig. 4. The NMSE performance of cascaded channel estimation against the system bandwidth. The number of IRS elements $M = 256$. The number of subcarriers $N_p = 128$. The frequency $f_c = 20$ GHz and the pilot overhand is 5%.

$$\mathrm{NMSE}_{\mathbf{h}} = \frac{E\left(\sum_{n_p=1}^{N_p}\left\|\widehat{\mathbf{h}}(\frac{n_pW}{N_P}) - \mathbf{h}(\frac{n_pW}{N_P})\right\|_2^2\right)}{E\left(\sum_{n_p=1}^{N_p}\left\|\mathbf{h}(\frac{n_pW}{N_P})\right\|_2^2\right)}, \tag{40}$$

where $\widehat{\mathbf{h}}(\frac{n_pW}{N_P})$ denotes the estimate of $\mathbf{h}(\frac{n_pW}{N_P})$.Furthermore, we define the NMSE of vector $\mathbf{z}(n_p)$ in (25) and $\overline{\mathbf{c}}^{\mathcal{I}_i^a}$ in (33) as

$$\mathrm{NMSE}_{\mathbf{z}} = \frac{E\left(\sum_{n_p=1}^{N_P}\|\widehat{\mathbf{z}}(n_p) - \mathbf{z}(n_p)\|_2^2\right)}{E\left(\sum_{n_p=1}^{N_P}\|\mathbf{z}(n_p)\|_2^2\right)}, \tag{41}$$

$$\mathrm{NMSE}_{\overline{\mathbf{c}}} = \frac{E\left(\sum_{i=1}^{N_d}\left\|\widehat{\mathbf{c}}^{I_i^a} - \overline{\mathbf{c}}^{I_i^a}\right\|_2^2\right)}{E\left(\sum_{i=1}^{N_d}\left\|\overline{\mathbf{c}}^{I_i^a}\right\|_2^2\right)}, \tag{42}$$

where $\widehat{\mathbf{c}}^{I_i^a}$ and $\widehat{\mathbf{z}}(n_p)$ denote the estimate of $\overline{\mathbf{c}}^{I_i^a}$ and $\mathbf{z}(n_p)$.

Fig. 4 shows the NMSE performance of the cascaded IRS channel vs. the bandwidth. The curves representing *Method [23]* and *Method [34]* characterize the NMSE performance of the conventional compressive sensing based method proposed in [23] and [34], where the range of equivalent angles is assume to be $[-1/2, 1/2)$ and the effect of the beam squint is ignored. The curve representing *Method [19]* characterizes a LS estimator used for estimating the channel $\boldsymbol{h}(f)$ in (15). In the simulations, we set the number of measurements to 35 for

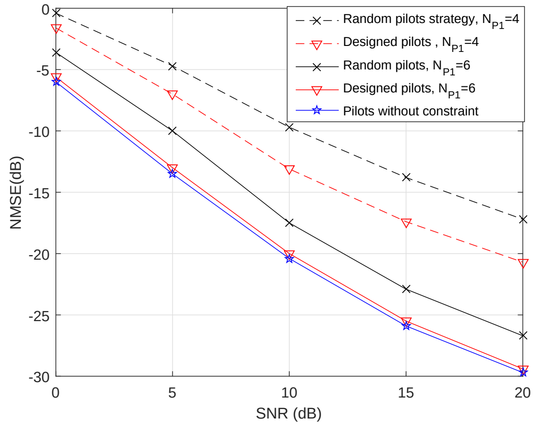


Fig. 5. The NMSE vs. SNR performance of cascaded channel gains and delays estimation. The number of IRS elements $M = 256$. The number of subcarriers $N_p = 128$. The frequency $f_c = 20$ GHz and the signal bandwidth is 510 MHz.

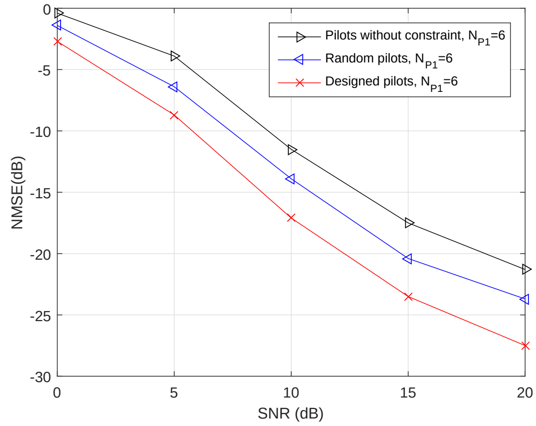


Fig. 7. The NMSE vs. SNR performance of cascaded channel estimation. The number of IRS elements $M = 256$. The number of subcarriers $N_p = 128$. The frequency $f_c = 20$ GHz and the signal bandwidth is 510 MHz.

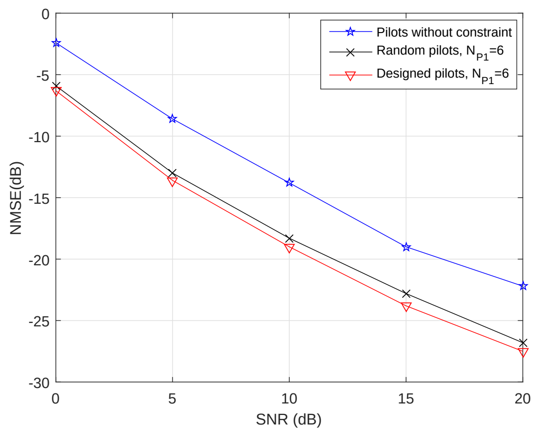


Fig. 6. The NMSE vs. SNR performance of cascaded channel angles estimation. The number of IRS elements $M = 256$. The number of subcarriers $N_p = 128$. The frequency $f_c = 20$ GHz and the signal bandwidth is 510 MHz.

the methods of [23], [34] and also for our proposed method, which is defined as the number of OFDM symbols required for channel estimation. Furthermore, we set the number of measurements to 1000 for the method of [19]. We observe that when the system bandwidth is small, the conventional method and the proposed method have a similar NMSE performance. However, as the bandwidth grows, the effect of beam squint becomes more severe, which degrades the performance of the traditional channel estimation method proposed in [23] and [34]. Furthermore, we find that although both the method of [19] and our proposed method are robust to the effect of beam squint, the number of measurements required by our proposed method is much lower than that of the method [19].

Fig. 5 shows the impact of different pilot designs on the channel estimation of delays and gains vs. SNR. The Y-axis represents the NMSE of vector $\bar{\mathbf{c}}^{\mathcal{I}_i^a}$. In Fig. 5, we estimate the channel delays and gains with known equivalent angles. The curve *Designed pilots,* $N_{P1} = 6$ and *Designed pilots,* $N_{P1} = 4$ rely on the pilots obtained by Algorithm 2, respectively. The curve *Random pilots,* $N_{P1} = 6$ and *Random pilots,* $N_{P1} = 4$ correspond to the randomly selected pilots, where the number of pilots is 6 and 4, respectively. The curve *Pilots without constraint* denotes the pilots whose subcarrier indices obtained by Algorithm 2, except that our first pilot design guideline is ignored. Fig. 5 shows that the proposed pilot design substantially improves the channel gain and delay estimation performance, and the constraint of (37) has a minor effect on the delay and gain estimation performance of the second stage.

Fig. 6 shows the impact of different pilot designs on the equivalent angle estimation vs. SNR. The Y-axis represents the NMSE of vector $\mathbf{z}(n_p)$. We consider the range of equivalent angles as $(-1, 1)$. Fig. 6 shows that the proposed pilot design substantially improve the algorithm's equivalent angles estimation performance. However, since the curve *Pilots without constraint* does not consider the first guideline, it represents poor performance. Observe in Fig. 5 and Fig. 6, that the proposed pilot design obtains both good equivalent angle as well as gain and delay estimation performance, simultaneously.

Fig. 7 shows the impact of different pilot designs on the estimation of cascaded channel against SNR. The definition of curves in Fig. 7 is the same as those in Fig. 6. Fig. 7 shows that the proposed pilot design benefically improves the algorithm's cascaded channel estimation performance.

Fig. 8 shows the NMSE performance comparison of the proposed TS-OMP algorithm and of the conventional methods [23] [34] [19] vs. the SNR. The curve *Proposed TS-OMP and designed pilots* represents the NMSE of the cascaded channel estimated by the proposed TS-OMP algorithm based on our designed pilots. The curves representing *Method [23] and random pilots*, *Method [34] and random pilots* and *Method [19] and random pilots* characterize the NMSE of the cascaded channel estimated by the conventional IRS channel estimation methods proposed in [23] [34] [19] based on random pilots. The curve representing *Ideal solution* represents the NMSE of channel estimation relying on known equivalent angles and delays of the cascaded channel. In the simulation, we set the number of measurements to 35 for the methods of

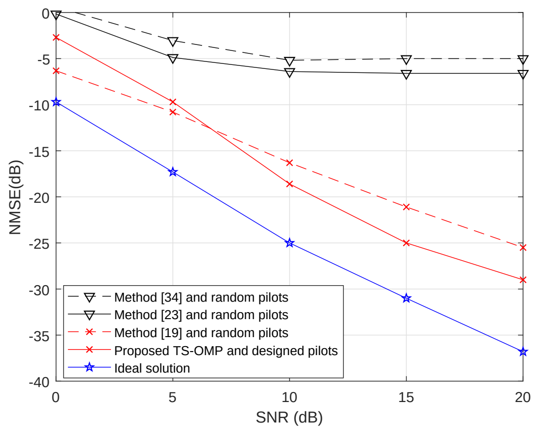


Fig. 8. The NMSE vs. SNR performance of cascaded channel estimation. The number of IRS elements $M = 256$. The frequency $f_c = 20$ GHz and the signal bandwidth is 510 MHz. The pilot overhand is 5%.

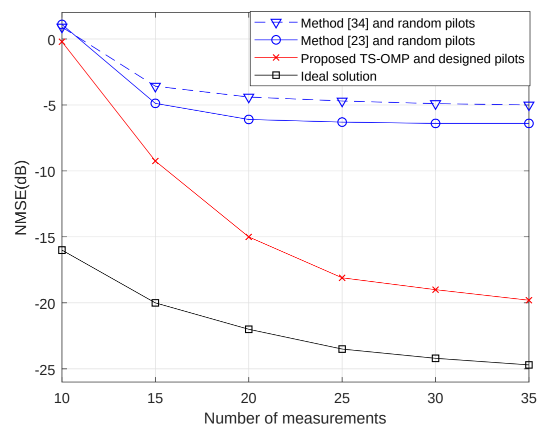


Fig. 9. The NMSE performance of cascaded channel estimation against the number of measurements. The number of IRS elements $M = 256$. The frequency $f_c = 20$ GHz and the signal bandwidth is 510 MHz. The SNR is 10dB. The pilot overhand is 5%.

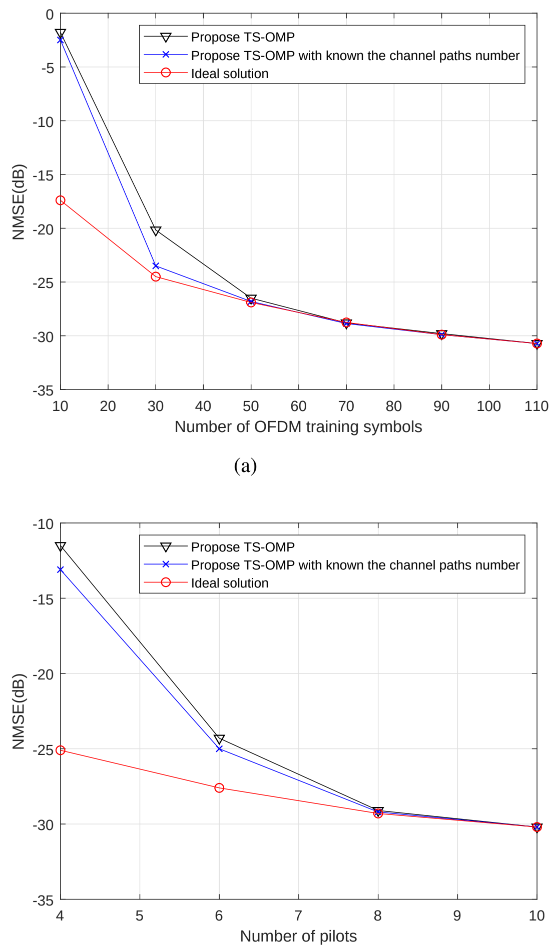
[23], [34] and also for our proposed method. Furthermore, we set the number of measurements to 1000 for the method of [19]. We observe that the proposed TS-OMP algorithm and cross-entropy based pilot design substantially improve the channel estimation performance in IRS-aided wideband systems. Furthermore, we find that although both the method of [19] and our proposed method have a similar NMSE performance, the number of measurements required by our proposed method is much lower than that of the method [19]. Observe from Algorithm 1 that the complexity of the proposed TS-OMP algorithm mainly comes from the angular and delay support estimation in Step 3 and 15. The method proposed in [23] and [34] also have to estimate the support. However, the proposed TS-OMP algorithm removes a large number of zero-valued elements of measurement matrix $\overline{\mathbf{F}}$ in (26) when we estimate the support. Therefore, the dimension of the measurement matrix is reduced. In this way, the complexity of the proposed TS-OMP algorithm is lower than that of the solutions in [23] and [34]. Due to the fact that the number of OFDM training symbols N_s required by [19] is much higher than the proposed TS-OMP, we may conclude that the complexity of [19] is higher than that of TS-OMP.

Fig. 9 shows the NMSE performance of the proposed channel estimation method vs. the number of measurements, which is defined as the number of OFDM symbols required for channel estimation. The SNR is 10dB. The curves representing *Method [23] and random pilots* and representing *Method [34] and random pilots* show that we cannot achieve the accurate channel estimation, not even for a high training overhead of OFDM symbols, which is due to the effect of beam squint.

Fig. 10 show the NMSE performance of the cascaded channel estimation against the number of OFDM training symbols and the number of pilots. In the simulation, we set that the number of pilots to 10 for Fig. 10 (a) and the number of OFDM training symbols to 100 for Fig. 10 (b). The SNR is 10dB. The curve '*Proposed TS-OMP with known channel path number*' represents the case, where we know the number of cascaded channel paths, which means that the stopping criterion is ideal. Observe from Fig. 10 (a) and Fig. 10 (b) that the proposed TS-OMP algorithm may attain a similar NMSE performance to that of the ideal solutions when the number of measurements N_s and the number of pilots N_{P1} are very high.

Fig. 11 shows the symbol error rate (SER) against SNR. We use 16-QAM transmitted symbols. The curve *Random phase* represents the SER performance, where the IRS phases are chosen randomly. The curve *Method [23] and designed phase* represents the SER performance when the CSI is estimated by the method proposed in [23]. In the simulation, we use the IRS phase parameter design method proposed in [14], which formulates the problem of phase parameter design as a MSE minimization problem and utilizes the classical iterative block coordinate descent (BCD) algorithm to solve it. The curve *Designed phase and designed pilots* represents the SER performance, where the IRS phase is designed when the CSI is estimated by the proposed TS-OMP algorithm based on our designed pilots. The curve *Designed phase and random pilots* represents the SER performance, where the IRS phase is designed when the CSI is estimated by the proposed TS-OMP algorithm based on random pilots. The curve *Designed phase and perfect CSI* represents the SER performance, where the IRS phase is based on the perfect CSI. The curve *Random phase* and *Designed phase and perfect CSI* shows that the appropriate design of the IRS phase substantially reduces the SER. The curve *Designed phase and perfect CSI* and *Designed phase and designed pilots* shows that the proposed TS-OMP algorithm based on our pilot design achieves similar SER to the perfect CSI scenario.

Fig. 12 and Fig. 13 show the effect of the parameters N_c, N_e and N_{iter} on the performance of Algorithm 2. In Fig. 12, we set the number of iterations to $N_{iter} = 20$ and the number of subscarriers to $N_p = 128$. The optimal solution is found by exhaustive search. We find that the performance of the proposed Algorithm 2 is improved as N_c increases, and the performance is not sensitive to the parameter N_e. Observe from Fig. 13 that performance of the proposed Algorithm 2 is improved as N_{iter} increases and converges



(a)

(b)

Fig. 10. The NMSE performance of cascaded channel estimation. (a) The NMSE performance of cascaded channel estimation against the number of OFDM training symbols. (b) The NMSE performance of cascaded channel estimation against the number of pilots

towards the optimal solution, when N_{iter} is sufficiently large. Our proposed Algorithm 2 is a cross-entropy based method, whose convergence is not effected by the constraint [35]. Thus, the proposed Algorithm 2 is capable of finding the globally optimal solution with a sufficiently high probability, when N_{iter} and N_c are large. The details of the proof can be found in [36].

VII. Conclusions

Cascaded channel estimation of IRS-aided systems have been investigated in wideband scenarios considering the effect of beam squint. The cascaded channel is characterized by the equivalent angles, gains and delays of propagation paths. We demonstrated that it is hard to distinguish between the actual angle and false angle, which is occurred due to the effect of beam squint. We proposed a TS-OMP algorithm, which can estimate the cascaded channel with small training overhand. Moreover, to further improve the channel estimation performance, we propose a pilot design method based on cross-entropy theory. Our simulation results confirm that the proposed TS-OMP algorithm and pilot design method achieve better performance than their conventional counterparts.

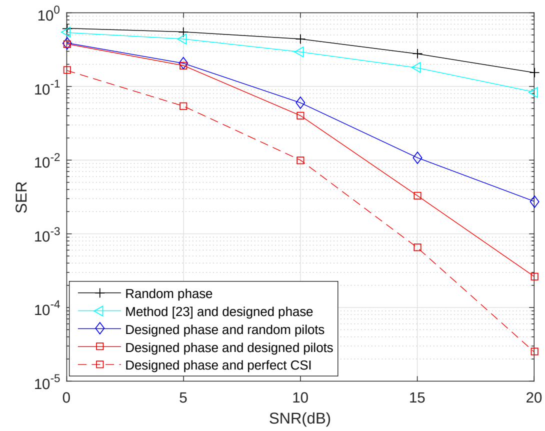


Fig. 11. The SER performance against SNR. The number of IRS elements $M = 256$. The number of subcarriers $N_p = 128$. The frequency $f_c = 20$ GHz and the signal bandwidth is 510 MHz. The pilot overhand is 5%.

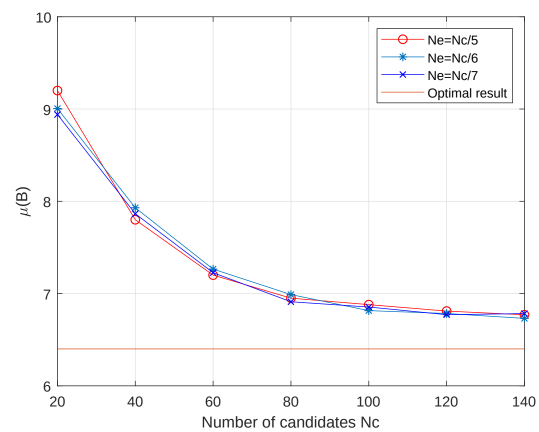


Fig. 12. The value of $\mu(\mathbf{B})$ against the number of candidates N_c. The number of subcarriers $N_p = 128$. The frequency $f_c = 20$ GHz and the signal bandwidth is 510 MHz. The number of pilots N_{P1} is 6. The number of iterations N_{iter} is 20.

Appendix A

We consider an IRS-aided wideband OFDM system having N_P subcarriers, where the BS has N antennas and the user has a single antenna. We assign N_{P1} subcarriers as pilots in each OFDM symbol and transmit N OFDM symbols in a frame. The n-th BS antenna transmits pilots at the n_p-th subcarrier, which is expressed as $\mathbf{s}_n(f) \in \mathbb{C}^{N\times 1}$. According to (15), the FD signal $\mathbf{y}(f) \in \mathbb{C}^{N\times 1}$ received by the user from the N-antenna BS is expressed as

$$\mathbf{y}(f) = \sum_{n=1}^{N} \boldsymbol{\theta}^{\mathrm{T}} \mathbf{h}_n^{\mathrm{H}}(f)\mathbf{s}_n(f) + \mathbf{n}(f), \tag{43}$$

where $\mathbf{h}_n(f)$ denotes the frequency-domain (FD) cascaded channel of n-th antenna installed at the BS. The vector $\mathbf{n}(f)$ is Gaussian noise. We adopt orthogonal pilots $\mathbf{s}_n(f)$ for the different BS antennas, i.e., $|\mathbf{s}_n^{\mathrm{H}}(f)\mathbf{s}_{n'}(f)| = 1$, when $n = n'$

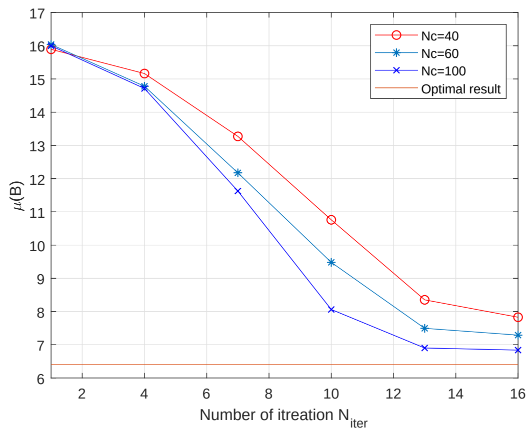


Fig. 13. The value of $\mu(\mathbf{B})$ against the number of iterations N_{iter}. The number of subcarriers $N_p = 128$. The frequency $f_c = 20$ GHz and the signal bandwidth is 510 MHz. The number of pilots N_{P1} is 6.

and $|\mathbf{s}_n^{\mathrm{H}}(f)\mathbf{s}_{n'}(f)| = 0$ when $n \neq n'$. Thus, multiply $\mathbf{s}_n^*(f)$ both sides of (43), we have

$$y_n(f) = \boldsymbol{\theta}^{\mathrm{T}}\mathbf{h}_n(f) + n(f), \tag{44}$$

where $n(f) = \mathbf{n}^{\mathrm{T}}(f)\mathbf{s}_n^*(f)$ follows the complex Gaussian distribution. Our task is to estimate the FD cascaded channel $\mathbf{h}_n(f)$ from (44), and (44) is similar to (15). In this way, the proposed method to estimate FD cascaded channel $\mathbf{h}(f)$ for BS equipped with one antenna can be directly utilized to estimate FD cascaded channel $\mathbf{h}_n(f)$.

References

[1] J. Chen, Y. Liang, Y. Pei, and H. Guo, "Intelligent reflecting surface: A programmable wireless environment for physical layer security," *IEEE Access*, vol. 7, pp. 82 599–82 612, 2019.

[2] L. Dai, B. Wang, M. Wang, X. Yang, J. Tan, S. Bi, S. Xu, F. Yang, Z. Chen, M. D. Renzo, C. Chae, and L. Hanzo, "Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, pp. 45 913–45 923, 2020.

[3] Y. Han, W. Tang, S. Jin, C. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical csi," *IEEE Trans. Veh Technol.*, vol. 68, pp. 8238–8242, Oct. 2019.

[4] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, pp. 106–112, Jan. 2020.

[5] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, Aug. 2019.

[6] E. Basar, "Reconfigurable intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G," *IEEE Trans. on Commun.*, pp. 1–1, Feb. 2020.

[7] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, pp. 5394–5409, Nov. 2019.

[8] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, pp. 4157–4170, Aug. 2019.

[9] B. Di, H. Zhang, L. Li, L. Song, Y. Li, and Z. Han, "Practical hybrid beamforming with finite-resolution phase shifters for reconfigurable intelligent surface based multi-user communications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4565–4570, Apr. 2020.

[10] K. Feng, Q. Wang, X. Li, and C. Wen, "Deep reinforcement learning based intelligent reflecting surface optimization for MISO communication systems," *IEEE Wireless Commun. Lett.*, pp. 1–1, Jan. 2020.

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

[12] K. Ying, Z. Gao, S. Lyu, Y. Wu, H. Wang, and M. Alouini, "GMD-based hybrid beamforming for large reconfigurable intelligent surface assisted millimeter-wave massive MIMO," *IEEE Access*, vol. 8, pp. 19 530–19 539, 2020.

[13] S. Zhang and R. Zhang, "Capacity characterization for intelligent reflecting surface aided MIMO communication," *IEEE J. Sel. Areas Commun.*, pp. 1–1, 2020.

[14] H. Li, R. Liu, M. Liy, Q. Liu, and X. Li, "IRS-enhanced wideband MU-MISO-OFDM communication systems," in *Proc. IEEE WCNC*, 2020, pp. 1–6.

[15] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surfaceassisted MISO wireless energy transfer," in *Proc. ICASSP*, May. 2019, pp. 1–4.

[16] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Tran. Wireless Commun.*, pp. 1–1, 2020.

[17] J. Lin, G. Wang, R. Fan, T. A. Tsiftsis, and C. Tellambura, "Channel estimation for wireless communication systems assisted by large intelligent surfaces," *arXiv preprint arXiv:1911.02158*, 2019.

[18] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems," *IEEE Signal Process. Lett.*, vol. 27, pp. 905–909, 2020.

[19] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," *IEEE Wireless Commun. Lett.*, pp. 1–1, Dec. 2019.

[20] B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface assisted multi-user OFDMA: Channel estimation and training design," *arXiv preprint arXiv:2003.00648*, 2020.

[21] Z. Wan, Z. Gao, and M. S. Alouini, "Broadband channel estimation for intelligent reflecting surface aided mmWave massive MIMO systems," in *Proc. IEEE ICC*, 2020, pp. 1–6.

[22] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization," *IEEE Trans. Commun.*, pp. 1–1, Mar. 2020.

[23] J. Chen, Y. Liang, R. Fan, H. V. Cheng, and W. Yu, "Channel estimation for reconfigurable intelligent surface aided multi-user MIMO systems," *arXiv preprint arXiv:1912.03619*, 2019.

[24] X. Ma, Z. Chen, W. Chen, Z. Li, Y. Chi, C. Han, and S. Li, "Joint channel estimation and data rate maximization for intelligent reflecting surface assisted TeraHertz MIMO communication systems," *IEEE Access*, vol. 8, pp. 99 565–99 581, 2020.

[25] M. Jian, F. Gao, Z. Tian, S. Jin, and S. Ma, "Angle-domain aided UL/DL channel estimation for wideband mmwave massive MIMO systems with beam squin," *IEEE Trans. wireless Commun.*, vol. 18, no. 7, pp. 3515–3527, Jul. 2019.

[26] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen, and R. S. Blum, "Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 35, pp. 1524–1538, Jul. 2017.

[27] X. Ma, F. Yang, S. Liu, J. Song, and Z. Han, "Design and optimization on training sequence for mmWave communications: A new approach for sparse channel estimation in massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1486–1497, Jul. 2017.

[28] X. Gao, L. Dai, S. Zhou, A. M. Sayeed, and L. Hanzo, "Wideband beamspace channel estimation for millimeter-wave MIMO systems relying on lens antenna arrays," *IEEE Transactions on Signal Processing*, vol. 67, no. 18, pp. 4809–4824, Jul. 2019.

[29] Y. Chen, D. Chen, T. Jiang, and L. Hanzo, "Channel-covariance and angle-of-departure aided hybrid precoding for wideband multiuser millimeter wave MIMO systems," *IEEE Tran. Commun.*, vol. 67, no. 12, pp. 8315–8328, 2019.

[30] K. Venugopal, N. Gonzlez-Prelcic, and R. W. Heath, "Optimal frequency-flat precoding for frequency-selective millimeter wave channels," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5098–5112, Aug. 2019.

[31] M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5695–5702, Dec. 2007.

[32] J. M. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1395–1408, Jul. 2009.

[33] S. Ma, W. Shen, J. An, and L. Hanzo, "Antenna array diagnosis for millimeter-wave mimo systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4585–4589, Apr. 2020.

[34] X. Ma, Z. Chen, W. Chen, Z. Li, Y. Chi, C. Han, and S. Li, "Joint channel estimation and data rate maximization for intelligent reflecting surface assisted Terahertz MIMO communication systems," *IEEE Access*, May 2020.
[35] P. Minvielle, E. Tantar, A. Tantar, and P. Berisset, "Sparse antenna array optimization with the cross-entropy method," *IEEE Trans. Antennas Propag.*, vol. 59, no. 8, pp. 2862–2871, Aug. 2011.
[36] A. Costaa, O. D. Jonesb, and D. Kroesec, "Convergence properties of the cross-entropy method for discrete optimization," *Opera. Research Lett.*, pp. 573–580, Sep. 2007.

Siqi Ma received the B.E. degree from the Beijing Institute of Technology, Beijing, China, in 2015, where he is currently pursuing the Ph.D. degree with the School of Information and Electronics. His current research interests include massive MIMO, mmWave/THz communications, energy-efficient communications, intelligent reflecting surface and networks.

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. 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.

Jianping An (M'08) received the B.E. degree from Information Engineering University in 1987, and the M.S. and Ph.D. degrees from Beijing Institute of Technology, in 1992 and 1996, respectively. Since 1996, he has been with the School of Information and Electronics, Beijing Institute of Technology, where he now holds the post of Full Professor. From 2010 to 2011, he was a Visiting Professor at University of California, San Diego. He has published more than 150 journal and conference articles and holds (or co-holds) more than 50 patents. He has received various awards for his academic achievements and the resultant industrial influences, including the National Award for Scientific and Technological Progress of China (1997) and the Excellent Young Teacher Award by the China's Ministry of Education (2000). Since 2010, he has been serving as a Chief Reviewing Expert for the Information Technology Division, National Scientific Foundation of China. Prof. An's current research interest is focused on digital signal processing theory and algorithms for communication systems.

Lajos Hanzo (http:www-mobile.ecs.soton.ac.uk, https:en.wikipedia.orgwikiLajos_Hanzo) (FIEEE'04, Fellow of the Royal Academy of Engineering F(REng), of the IET and of EURASIP), received his Master degree and Doctorate in 1976 and 1983, respectively from the Technical University (TU) of Budapest. He was also awarded the Doctor of Sciences (DSc) degree by the University of Southampton (2004) and Honorary Doctorates by the TU of Budapest (2009) and by the University of Edinburgh (2015). He is a Foreign Member of the Hungarian Academy of Sciences and a former Editor-in-Chief of the IEEE Press. He has served several terms as Governor of both IEEE ComSoc and of VTS. He has published 1900+ contributions at IEEE Xplore, 19 Wiley-IEEE Press books and has helped the fast-track career of 123 PhD students. He holds the Chair of Telecommunications and directs the research of Next-Generation Wireless at the University of Southampton, UK