Channel Estimation for Orthogonal Time Frequency Space (OTFS) Massive MIMO

Wenqian Shen, Linglong Dai, Shuangfeng Han, Chih-Lin I, and Robert W. Heath, Jr., *Fellow, IEEE*

***Abstract*—Orthogonal time frequency space (OTFS) modulation outperforms orthogonal frequency division multiplexing (OFDM) in high-mobility scenarios. One challenge for OTFS massive MIMO is downlink channel estimation due to the required high pilot overhead. In this paper, we propose a 3D structured orthogonal matching pursuit (3D-SOMP) algorithm based channel estimation technique. First, we show that the OTFS MIMO channel exhibits 3D structured sparsity: normal sparsity along the delay dimension, block sparsity along the Doppler dimension, and burst sparsity along the angle dimension. Based on the 3D structured channel sparsity, we then formulate the downlink channel estimation problem as a sparse signal recovery problem. Simulation results show that the proposed 3D-SOMP algorithm can achieve accurate channel state information with low pilot overhead.**

I. Introduction

One goal of future wireless communications (the emerging 5G or beyond 5G) is to support reliable communications in high-mobility scenarios, such as on high-speed railways with a speed of up to 500 km/h [1] or on vehicles with a speed of up to 300 km/h [2]. The dominant modulation technique for 4G and the emerging 5G is orthogonal frequency division multiplexing (OFDM). In high-mobility scenarios, OFDM may experience significant inter-carrier interference (ICI) due to the large Doppler spread of time-variant channels. ICI will severely degrade the performance of OFDM systems when traditional transceivers are used. To cope with ICI, some modifications of the traditional OFDM were proposed at the cost of more complicated transceiver design [3]–[7].

Orthogonal time frequency space (OTFS) is an alternative to OFDM to tackle the time-variant channels [8]–[10]. Leveraging the basis expansion model for the channel [11], OTFS converts the time-variant channels into the *time-independent* channels in the delay-Doppler domain. Accordingly, the information bearing data is multiplexed into the roughly constant channels in the delay-Doppler domain. OTFS with massive multiple-input multiple-output (MIMO) can further increase the spectrum efficiency. Such benefits require that the downlink channel state information (CSI) is known at the transmitter to design the transmit beamforming vectors [12]. With a large number of base station (BS) antennas in OTFS massive MIMO systems, downlink channel estimation is challenging due to the required high pilot overhead.

In OTFS systems, the BS uses pilots that are transmitted in the delay-Doppler domain for channel estimation [13]. An impulse-based method was proposed for single-input single-output (SISO) systems, where an impulse in the delay-Doppler domain was transmitted as the training pilots. It was extended to OTFS MIMO systems by transmitting several impulses with proper guard between adjacent impulses to distinguish different BS antennas [14]. An alternative method using PN sequences in the delay-Doppler domain as the training pilots was proposed for OTFS SISO channel estimation in [15]. In that method, three coefficients of channels, namely, delay shift, Doppler shift, and fade coefficient are estimated. Then the delay-Doppler channel can be calculated accordingly. The existing OTFS channel estimation techniques can not be directly extended to OTFS massive MIMO systems, since a large number of antennas are required to be distinguished by transmitting orthogonal pilots, which will lead to high pilot overhead.

In this paper, we propose a 3D structured orthogonal matching pursuit (3D-SOMP) algorithm based downlink channel estimation technique for OTFS massive MIMO systems. We present the discrete-time formulation of OTFS systems and show that the OTFS massive MIMO channel exhibits a delay-Doppler-angle 3D structured sparsity. The channel estimator makes use of the training pilots that are transmitted in the delay-Doppler domain. Decomposing the channel based on its structure, we formulate the channel estimation problem as a sparse signal recovery problem. To solve this problem, we propose a 3D-SOMP algorithm, which can achieve accurate CSI with low pilot overhead.

Notation: Boldface capital letters stand for matrices and lower-case letters stand for column vectors. The transpose, conjugate, conjugate transpose, and inverse of a matrix are denoted by $(\cdot)^{\mathrm{T}}$, $(\cdot)^*$, $(\cdot)^{\mathrm{H}}$ and $(\cdot)^{-1}$, respectively. $\odot$ is the Hadamard product operator. $\|\mathbf{s}\|$ is the ℓ_2-norm of the vector $\mathbf{s}$. $\mathbf{\Psi}^{\dagger} = (\mathbf{\Psi}^{\mathrm{H}}\mathbf{\Psi})^{-1}\mathbf{\Psi}^{\mathrm{H}}$ is the Moore-Penrose pseudo-inverse of matrix $\mathbf{\Psi}$.

II. System Model

In this section, we present the discrete-time formulation of OTFS modulation and OTFS demodulation in SISO systems. Then, we describe an extension of OTFS into massive MIMO systems.

W. Shen is with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: wshen@bit.edu.cn). L. Dai is with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: daill@tsinghua.edu.cn). S. Han and C. I are with the Green Communication Research Center, China Mobile Research Institute, Beijing 100053, China (e-mails: hanshuangfeng,icl@chinamobile.com). R. W. Heath Jr. is with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712-1687, USA (e-mail: rheath@utexas.edu). R. W. Heath Jr. is also on the Technical Advisory Board of Cohere Technologies, which developed OTFS. The terms of this arrangement have been reviewed and approved by The University of Texas at Austin in accordance with its policy on objectivity in research.

978-1-5386-8088-9/19/$31.00 ©2019 IEEE

A. OTFS SISO Modulation

We consider the OTFS SISO architecture as commonly assumed [8]–[10]. A quadrature amplitude modulated (QAM) data sequence of length MN is first rearranged into a 2D data block. This is called a 2D OTFS frame $\mathbf{X}^{\mathrm{DD}} \in \mathbb{C}^{M\times N}$ in the delay-Doppler domain, where M and N are the numbers of resource units along the delay dimension and Doppler dimension. OTFS modulation at the transmitter is composed of a pre-processing block and a traditional frequency-time modulator such as OFDM or filter bank multicarrier (FBMC). The pre-processing block is realized by using an inverse symplectic finite Fourier transform (ISFFT) and a transmit windowing function. The ISFFT of $\mathbf{X}^{\mathrm{DD}}$ is [8]

$$\mathbf{X}^{\mathrm{ISFFT}} = \mathbf{F}_{\mathrm{M}}\mathbf{X}^{\mathrm{DD}}\mathbf{F}_{\mathrm{N}}^{\mathrm{H}}, \tag{1}$$

where $\mathbf{F}_{\mathrm{M}} \in \mathbb{C}^{M\times M}$ and $\mathbf{F}_{\mathrm{N}} \in \mathbb{C}^{N\times N}$ are discrete Fourier transform (DFT) matrices. A transmit windowing matrix $\mathbf{W}^{\mathrm{tx}} \in \mathbb{C}^{M\times N}$ multiplies $\mathbf{X}^{\mathrm{ISFFT}}$ element-wise to produce the 2D block in the frequency-time domain $\mathbf{X}^{\mathrm{FT}} \in \mathbb{C}^{M\times N}$ as

$$\mathbf{X}^{\mathrm{FT}} = \mathbf{X}^{\mathrm{ISFFT}} \odot \mathbf{W}^{\mathrm{tx}}. \tag{2}$$

There are several uses of the windowing matrix such as randomizing the phases of the transmitted symbols to eliminate the inter-cell interference [13]. In this paper, we assume a trivial window for simple expression, i.e., $\mathbf{W}^{\mathrm{tx}}$ is a matrix of all ones.

Assuming an OFDM modulator, the M-point inverse DFT (IDFT) is applied on each column of $\mathbf{X}^{\mathrm{FT}}$ to obtain the 2D transmit signal block $\mathbf{S} \in \mathbb{C}^{M\times N}$, i.e.,

$$\mathbf{S} = \mathbf{F}_{\mathrm{M}}^{\mathrm{H}}\mathbf{X}^{\mathrm{FT}}, \tag{3}$$

where $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_N]$. Each column vector $\mathbf{s}_i \in \mathbb{C}^{M\times 1}$ can be regarded as an OFDM symbol. Note that N OFDM symbols $\{\mathbf{s}_i\}_{i=1}^{N}$ occupy the bandwidth $M\Delta f$ and have the duration NT, where Δf and T are the subcarrier spacing and symbol duration. By combing (1)-(3),

$$\mathbf{S} = \mathbf{X}^{\mathrm{DD}}\mathbf{F}_{\mathrm{N}}^{\mathrm{H}}. \tag{4}$$

To avoid inter-symbol interference between blocks, the OFDM modulator usually adds cyclic prefix (CP) for each OFDM symbol $\mathbf{s}_i$ via a CP addition matrix $\mathbf{A}_{\mathrm{CP}} \in \mathbb{C}^{(M+N_{\mathrm{CP}})\times M}$ [9] with N_{CP} being the length of CP. By reading the 2D transmit signal block $\mathbf{S}$ column-wise, the 1D transmit signal $\mathbf{s} \in \mathbb{C}^{(M+N_{\mathrm{CP}})N\times 1}$ is

$$\mathbf{s} = \mathrm{vec}\{\mathbf{A}_{\mathrm{CP}}\mathbf{S}\}. \tag{5}$$

B. OTFS SISO Demodulation

At the receiver, the κ-th element r_κ of the received signal $\mathbf{r} \in \mathbb{C}^{(M+N_{\mathrm{CP}})N\times 1}$ after the time-variant channel $h_{\kappa,\ell}$ with length $L+1$ is expressed as

$$r_\kappa = \sum_{\ell=0}^{L} h_{\kappa,\ell}s_{\kappa-\ell} + v_\kappa, \tag{6}$$

where $s_{\kappa-\ell}$ is the $(\kappa-\ell)$-th element of the transmit signal $\mathbf{s}$ and v_κ is the additive Gaussian noise at the receiver.

The OTFS demodulation at the receiver consists of a traditional frequency-time demodulator such as the OFDM or FBMC demodulator and a post-processing block. Specifically, assuming an OFDM demodulator, the received signal $\mathbf{r}$ is first rearranged as a matrix $\mathbf{R}$ of size $(M+N_{\mathrm{CP}})\times N$. Each column vector of $\mathbf{R}$ can be regarded as a received OFDM symbol including CP. Then, the OFDM demodulator removes the CP by multiplying $\mathbf{R}$ with a CP removal matrix $\mathbf{R}_{\mathrm{CP}} \in \mathbb{C}^{M\times(M+N_{\mathrm{CP}})}$ [9] to obtain the OFDM symbols $\mathbf{R}_{\mathrm{CP}}\mathbf{R}$ without CPs. Applying the M-point DFT on each OFDM symbol (without CP), we obtain the received 2D block $\mathbf{Y}^{\mathrm{FT}} \in \mathbb{C}^{M\times N}$ in the frequency-time domain as

$$\mathbf{Y}^{\mathrm{FT}} = \mathbf{F}_{\mathrm{M}}\mathbf{R}_{\mathrm{CP}}\mathbf{R}. \tag{7}$$

The post-processing block is realized by a receive windowing matrix $\mathbf{W}^{\mathrm{rx}} \in \mathbb{C}^{M\times N}$ and the SFFT. The receive windowing matrix $\mathbf{W}^{\mathrm{rx}}$ multiplies $\mathbf{Y}^{\mathrm{FT}}$ element-wise, i.e.,

$$\mathbf{Y}^{\mathrm{FT,W}} = \mathbf{Y}^{\mathrm{FT}} \odot \mathbf{W}^{\mathrm{rx}}. \tag{8}$$

Then, the SFFT is applied for $\mathbf{Y}^{\mathrm{FT,W}}$ to obtain the 2D data block $\mathbf{Y}^{\mathrm{DD}} \in \mathbb{C}^{M\times N}$ in the delay-Doppler domain as

$$\mathbf{Y}^{\mathrm{DD}} = \mathbf{F}_{\mathrm{M}}^{\mathrm{H}}\mathbf{Y}^{\mathrm{FT,W}}\mathbf{F}_{\mathrm{N}}. \tag{9}$$

Generally, the receive window is matched with the transmit window, i.e., $\mathbf{W}^{\mathrm{rx}} = \mathbf{W}^{\mathrm{tx}*}$ [13]. By combing (7)-(9),

$$\mathbf{Y}^{\mathrm{DD}} = \mathbf{R}_{\mathrm{CP}}\mathbf{R}\mathbf{F}_{\mathrm{N}}. \tag{10}$$

The received data block $\mathbf{Y}^{\mathrm{DD}}$ is given by the phase compensated 2D periodic convolution of the transmit data block $\mathbf{X}^{\mathrm{DD}}$ with the delay-Doppler channel impulse response (CIR) $\mathbf{H}^{\mathrm{DD}} \in \mathbb{C}^{M\times N}$ as shown in the following Lemma 1.

Lemma 1: We denote the $(\ell+1, k+1+\frac{N}{2})$-th element of $\mathbf{Y}^{\mathrm{DD}}$ and $\mathbf{X}^{\mathrm{DD}}$ as $Y_{\ell,k}^{\mathrm{DD}}$ and $X_{\ell,k}^{\mathrm{DD}}$, where $\ell = 0, 1, \cdots, M-1$ and $k = -\frac{N}{2}, \cdots, 0, \cdots, \frac{N}{2}-1$. Then $Y_{\ell,k}^{\mathrm{DD}}$ is given by

$$Y_{\ell,k}^{\mathrm{DD}} \overset{N\to\infty}{=} \sum_{\ell'=0}^{M-1}\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} X_{\ell',k'}^{\mathrm{DD}}H_{\ell-\ell',k-k'}^{\mathrm{DD}}w_{\ell,k-k'} + V_{\ell,k}^{\mathrm{DD}}, \tag{11}$$

where $w_{\ell,k-k'} = e^{j2\pi\frac{\ell(k-k')}{N(M+N_{\mathrm{CP}})}}$. $V_{\ell,k}^{\mathrm{DD}}$ is the additive noise in the delay-Doppler domain. $H_{\ell,k}^{\mathrm{DD}}$ is the $(\ell+1, k+1+\frac{N}{2})$-th element of the delay-Doppler CIR $\mathbf{H}^{\mathrm{DD}}$ and

$$H_{\ell,k}^{\mathrm{DD}} = \sum_{i=1}^{N} h_{(i-1)(M+N_{\mathrm{CP}})+1,(\ell)_M}e^{-j2\pi(i-1)\frac{k}{N}}, \tag{12}$$

where $(\ell)_M$ is the remainder after division of ℓ by M. Note that $H_{\ell,k}^{\mathrm{DD}} = H_{\ell+M,k+N}^{\mathrm{DD}}$, thus (11) can be regarded as periodic convolution.

Proof: See [16]. ■

We observe from (11) that the transmit data $X_{\ell',k'}^{\mathrm{DD}}$ in the delay-Doppler domain experiences roughly constant channel $H_{\ell,k}^{\mathrm{DD}}$ in the delay-Doppler domain. Since each transmit data $X_{\ell',k'}^{\mathrm{DD}}$ in the delay-Doppler domain is expanded onto the whole frequency-time domain as shown in (1), it can exploit

the full diversity of the frequency-time channel. As a result, OTFS has improved performance over the traditional OFDM especially in high-mobility scenarios [8]–[10].

C. OTFS Massive MIMO

We explain how OTFS works in massive MIMO systems to further increase the spectrum efficiency by using multi-user MIMO. The BS is equipped with $N_{\rm t}$ antennas to simultaneously serve U single-antenna users. Downlink precoding is performed to eliminate the inter-user interference. For example, the zero-forcing Tomlinson-Harashima precoding is adopted in [12]. After precoding, the transmit data block $\mathbf{X}^{\rm DD}$ in the delay-Doppler domain will be modulated through the OTFS modulation and transmitted at $N_{\rm t}$ antennas. At the user side, the received signal is demodulated through the OTFS demodulation. The main challenge for OTFS massive MIMO is the downlink channel estimation due to the required high pilot overhead. Next, we will focus on the downlink channel estimation in massive MIMO systems.

III. Proposed 3D-SOMP Based Channel Estimation in OTFS Massive MIMO Systems

In this section, we first show the 3D structured sparsity of OTFS MIMO channels. Then, we formulate the downlink channel estimation problem as a sparse signal recovery problem and solve it through a 3D-SOMP algorithm.

A. 3D Structured Sparsity of Delay-Doppler-angle Channel

We consider the downlink time-variant channel consisting of $N_{\rm p}$ dominant propagation paths. Each dominant path is composed of $N_{\rm s}$ subpaths. The s_i-th subpath of the i-th dominant path has a complex path gain α_{s_i} and Doppler frequency ν_{s_i}. The delays of all subpaths of the i-th dominant path can be regarded as the same τ_i [17]. We denote the physical AoD of the s_i-th subpath as θ_{s_i}. When a typical uniform linear array (ULA) of antennas is considered, the spatial angle associated with θ_{s_i} is defined as $\psi_{s_i} = \frac{d}{\lambda}\sin\theta_{s_i}$ [18], where d is the antenna spacing and λ is the wavelength of the carrier frequency. Typically, $d = \lambda/2$ and $\theta_{s_i} \in [-\pi/2, \pi/2)$, thus $\psi_{s_i} \in [-1/2, 1/2)$. The time-variant channel associated with the $(p+1)$-th antenna ($p = 0, 1, \cdots, N_{\rm t}-1$) can be expressed as [18]

$$h_{\kappa,\ell,p} = \sum_{i=1}^{N_{\rm p}} \sum_{s_i=1}^{N_{\rm s}} \alpha_{s_i} e^{j2\pi\nu_{s_i}\kappa T_{\rm s}} {\rm p_{rc}}(\ell T_{\rm s} - \tau_i) e^{-j2\pi p\psi_{s_i}}, \quad (13)$$

where ${\rm p_{rc}}(\tau)$ is the band-limited pulse shaping filter response evaluated at τ and $T_{\rm s} = \frac{1}{M\Delta f}$ is the system sampling interval. Based on (12), we express the delay-Doppler CIR of the $(p+1)$-th antenna (which is referred to as delay-Doppler-space CIR $H^{\rm DDS}_{\ell,k,p}$ in OTFS massive MIMO systems, where ℓ, k and p correspond to the delay, Doppler and spatial index) as follows

$$H^{\rm DDS}_{\ell,k,p} = \sum_{n=1}^{N} h_{(n-1)(M+N_{\rm CP})+1,(\ell)_M,p} e^{-j2\pi(n-1)\frac{k}{N}}. \quad (14)$$

The delay-Doppler-angle channel $H^{\rm DDA}_{\ell,k,r}$ is defined by applying inverse DFT for $H^{\rm DDS}_{\ell,k,p}$ along the space-dimension p as

$$H^{\rm DDA}_{\ell,k,r} \triangleq \sum_{p=0}^{N_{\rm t}-1} H^{\rm DDS}_{\ell,k,p} e^{j2\pi\frac{rp}{N_{\rm t}}}, \quad (15)$$

where $r = -\frac{N_{\rm t}}{2}, \cdots, 0, \cdots, \frac{N_{\rm t}}{2}-1$ is the angle index. By combining (13)-(15),

$$H^{\rm DDA}_{\ell,k,r} = \sum_{i=1}^{N_{\rm p}} \sum_{s_i=1}^{N_{\rm s}} \beta_{s_i} \Upsilon_N(\nu_{s_i} NT - k) \times {\rm p_{rc}}((\ell)_M T_{\rm s} - \tau_i)\, \Upsilon_{N_{\rm t}}(r - \psi_{s_i} N_{\rm t}), \quad (16)$$

where $\beta_{s_i} = \alpha_{s_i} e^{j2\pi\nu_{s_i}T_{\rm s}}$, $\Upsilon_N(x) \triangleq \sum_{n=1}^{N} e^{j2\pi\frac{x}{N}(n-1)} = \frac{\sin(\pi x)}{\sin(\pi\frac{x}{N})} e^{j\pi\frac{x(N-1)}{N}}$ and $T = (M + N_{\rm CP})T_{\rm s}$.

The function $\Upsilon_N(x)$ has the following characteristic: $|\Upsilon_N(x)| \approx 0$ when $|x| \gg 1$ [19]. Therefore, $H^{\rm DDA}_{\ell,k,r}$ has non-neglectable elements only if $\ell \approx \tau_i M\Delta f$, $k \approx \nu_{s_i} NT$, and $r \approx \psi_{s_i} N_{\rm t}$. We arrange $H^{\rm DDA}_{\ell,k,r}$ into a 3D tensor $\mathcal{H} \in \mathbb{C}^{M\times N\times N_{\rm t}}$. Since the Doppler-spread ν_{s_i} of the channel is smaller than the system bandwidth, the 3D channel $\mathcal{H}$ is block-sparse along the Doppler dimension, where the unique non-zero block is centered around $k = 0$ [20], [21]. Since the angle-spread of a dominant path is limited, the 3D channel $\mathcal{H}$ is burst-sparse along the angle dimension with $N_{\rm p}$ non-zero bursts [16], [20]. The difference between the burst-sparse and the traditional block-sparse is that the start position of the non-zero bursts are not necessarily to be $\{1, 1+D, 1+2D, \cdots\}$ where D is the length of non-zero blocks [22]. To sum up, the 3D channel tensor $\mathcal{H}$ is sparse along the delay dimension, block-sparse along the Doppler dimension, and burst-sparse along the angle dimension. This 3D structured sparsity can be used to estimate the CSI with low pilot overhead.

B. Formulation of Downlink Channel Estimation

We denote the length of pilots along the Doppler dimension and the delay dimension as N_ν and M_τ. We propose to use complex Gaussian random sequences as the training pilots. To avoid interference between pilots and data caused by the 2D periodic convolution in the delay-Doppler domain, guard intervals are required. Note that the delay-Doppler-angle channel $\mathcal{H}$ in OTFS massive MIMO systems has finite support $[0 : M_{\max} - 1]$ along the delay dimension and $\left[-\frac{N_{\max}}{2} : \frac{N_{\max}}{2} - 1\right]$ along the Doppler dimension [13]. The length of guard intervals should be $M_{\rm g} \geq M_{\max} - 1$ along the delay dimension and $\frac{N_{\rm g}}{2} \geq \frac{N_{\max}}{2} - 1$ along the Doppler dimension. To reduce the overall pilot overhead in OTFS massive MIMO systems, we propose the non-orthogonal pilot pattern, i.e., the pilots transmitted at different antennas completely overlap in the delay-Doppler domain, but the training sequences at different antennas are independently generated.

The training pilots in the delay-Doppler domain at the $(p+1)$-th antenna are denoted as $x_{\ell,k,p}$ with $\ell = 0, 1, \cdots, M_\tau - 1$, $k = -\frac{N_\nu}{2}, \cdots, 0, \cdots, \frac{N_\nu}{2} - 1$, and $p = 0, 1, \cdots, N_{\rm t} - 1$.

According to (11), the received pilots $y_{\ell,k}$ in the delay-Doppler domain at the user side can be expressed as

$$y_{\ell,k} = \sum_{p=0}^{N_{\rm t}-1} \sum_{\ell'=0}^{M_{\rm g}-1} \sum_{k'=-\frac{N_{\rm g}}{2}}^{\frac{N_{\rm g}}{2}-1} w_{\ell-\ell',k'} H^{\rm DDS}_{\ell',k',p} x_{\ell-\ell',k-k',p} + v_{\ell,k}, \tag{17}$$

where $w_{\ell-\ell',k'} = e^{j2\pi \frac{(\ell-\ell')k'}{N(M+N_{\rm CP})}}$. Using (15), we have

$$H^{\rm DDS}_{\ell,k,p} = \sum_{r=-\frac{N_{\rm t}}{2}}^{\frac{N_{\rm t}}{2}-1} H^{\rm DDA}_{\ell,k,r} e^{-j2\pi \frac{rp}{N_{\rm t}}}. \tag{18}$$

By substituting (18) into (17) and expressing $z_{\ell-\ell',k-k',r} = \sum_{p=0}^{N_{\rm t}-1} e^{-j2\pi \frac{rp}{N_{\rm t}}} x_{\ell-\ell',k-k',p}$, we have

$$y_{\ell,k} = \sum_{r=-\frac{N_{\rm t}}{2}}^{\frac{N_{\rm t}}{2}-1} \sum_{\ell'=0}^{M_{\rm g}-1} \sum_{k'=-\frac{N_{\rm g}}{2}}^{\frac{N_{\rm g}}{2}-1} w_{\ell-\ell',k'} H^{\rm DDA}_{\ell',k',r} z_{\ell-\ell',k-k',r} + v_{\ell,k}. \tag{19}$$

To simplify the expression, we rewrite (19) into the vector-matrix form. We arrange $y_{\ell,k}$ and $H^{\rm DDA}_{\ell',k',r}$ into column vectors $\mathbf{y} \in \mathbb{C}^{M_\tau N_\nu \times 1}$ and $\mathbf{h}_r \in \mathbb{C}^{M_{\rm g} N_{\rm g} \times 1}$. As a result, (19) can be rewritten in vector-matrix form

$$\mathbf{y} = \sum_{r=-\frac{N_{\rm t}}{2}}^{\frac{N_{\rm t}}{2}-1} \mathbf{W} \odot \mathbf{Z}_{{\rm c},r} \mathbf{h}_r + \mathbf{v}, \tag{20}$$

where $\mathbf{Z}_{{\rm c},r} \in \mathbb{C}^{M_\tau N_\nu \times M_{\rm g} N_{\rm g}}$ is the 2D periodic convolution matrix with the $(\ell N_\nu + k + N_\nu/2 + 1, \ell' N_{\rm g} + k' + N_{\rm g}/2 + 1)$-th element of $\mathbf{Z}_{{\rm c},r}$ being equal to $z_{\ell-\ell',k-k',r}$. $\mathbf{W} \in \mathbb{C}^{M_\tau N_\nu \times M_{\rm g} N_{\rm g}}$ is a matrix with the $(\ell N_\nu + k + N_\nu/2 + 1, \ell' N_{\rm g} + k' + N_{\rm g}/2 + 1)$-th element being $w_{\ell-\ell',k'}$. By denoting $\mathbf{\Psi} = \left[\mathbf{W} \odot \mathbf{Z}_{{\rm c},-\frac{N_{\rm t}}{2}}, \cdots, \mathbf{W} \odot \mathbf{Z}_{{\rm c},0}, \cdots, \mathbf{W} \odot \mathbf{Z}_{{\rm c},\frac{N_{\rm t}}{2}-1}\right]$ and $\mathbf{h} = \left[\mathbf{h}^{\rm T}_{-\frac{N_{\rm t}}{2}}, \cdots, \mathbf{h}^{\rm T}_0, \cdots, \mathbf{h}^{\rm T}_{\frac{N_{\rm t}}{2}-1}\right]^{\rm T}$, (20) can be expressed as a compressive sensing problem

$$\mathbf{y} = \mathbf{\Psi}\mathbf{h} + \mathbf{v}. \tag{21}$$

Note that $\mathbf{h}$ can be inversely vectorized to obtain a truncated delay-Doppler-angle channel $\mathcal{H}_{\rm g} \in \mathbb{C}^{M_{\rm g} \times N_{\rm g} \times N_{\rm t}}$, i.e., $\mathcal{H}_{\rm g} = {\rm invec}\{\mathbf{h}\}$, which is composed of the non-zero part of $\mathcal{H}$ with $\ell = 0, 1, \cdots, M_{\rm g} - 1$, $k = -\frac{N_{\rm g}}{2}, \cdots, 0, \cdots, \frac{N_{\rm g}}{2} - 1$, and $r = -\frac{N_{\rm t}}{2}, \cdots, 0, \cdots, \frac{N_{\rm t}}{2} - 1$. In the next subsection, we propose a 3D-SOMP algorithm to recover the channel vector $\mathbf{h}$ (or the truncated 3D channel $\mathcal{H}_{\rm g}$) in (21).

C. 3D-SOMP Algorithm

The proposed 3D-SOMP algorithm is presented in **Algorithm 1**. Different from the traditional OMP algorithm, to use the 3D structured sparsity of $\mathbf{h}$ (or $\mathcal{H}_{\rm g}$), we rearrange the correlation vector $\mathbf{e}$ as a tensor $\mathcal{E} \in \mathbb{C}^{M_{\rm g} \times N_{\rm g} \times N_{\rm t}}$ as in step 6. For the sake of presentation, we first introduce some notations of a N-dimensional ($N \geq 3$) tensor $\mathcal{M} \in \mathbb{C}^{I_1 \times I_2 \times, \cdots, \times I_N}$.

Algorithm 1 Proposed 3D-SOMP Algorithm

1: **Input:**
   1) Measurements $\mathbf{y}$; 2) Sensing matrix $\mathbf{\Psi}$
2: **Initialization:**
   $i = 0$, $\Omega = \emptyset$, $\mathbf{h}^{(i)} = \mathbf{0}$, $\mathbf{r} = \mathbf{y} - \mathbf{\Psi}\mathbf{h}^{(i)}$
3: **for** $i \leq N_{\rm p}$ **do**
4: $\quad i = i + 1$
5: $\quad \mathbf{e} = \mathbf{\Psi}^{\rm H}\mathbf{r}$
6: $\quad \mathcal{E} = {\rm invec}\{\mathbf{e}\}$
7: $\quad e_\tau(m) = \|\mathbf{E}_{(1)}(m,:)\|$
8: $\quad m_\tau^{(i)} = \arg\max_m e_\tau(m)$
9: $\quad e_\nu(n) = \|\mathcal{E}(m_\tau^{(i)}, n, :)\|$
10: $\quad n_\nu^{(i)} = \arg\min_n \left\|\mathbf{e}_\nu\left(\frac{N_{\rm g}}{2} - n : \frac{N_{\rm g}}{2} + n - 1\right)\right\|$, s.t. $\|\mathbf{e}_\nu\left(\frac{N_{\rm t}}{2} - n : \frac{N_{\rm t}}{2} + n - 1\right)\| \geq \epsilon \|\mathbf{e}_\nu\|$
11: $\quad \Lambda_\nu^{(i)} = \left\{\frac{N_{\rm g}}{2} - n_\nu^{(i)}, \cdots, \frac{N_{\rm g}}{2}, \cdots, \frac{N_{\rm g}}{2} + n_\nu^{(i)} - 1\right\}$
12: $\quad e_\theta(r) = \left\|\mathcal{E}\left(m_\tau^{(i)}, \Lambda_\nu^{(i)}, r\right)\right\|$
13: $\quad \mathbf{d}_\theta = \mathbf{L}^{\rm H}\mathbf{e}_\theta$
14: $\quad g_\theta(r) = \|\mathbf{D}_\theta(r,:)\|$
15: $\quad p_s = \arg\max_r g_\theta(r)$
16: $\quad \Lambda_\theta^{(i)} = \{p_s, p_s + 1, \cdots, p_s + D - 1\}$
17: $\quad \Omega = \Omega \cup (m_\tau^{(i)}, \Lambda_\nu^{(i)}, \Lambda_\theta^{(i)})$
18: $\quad \mathbf{h}^{(i)}|_\Omega = \mathbf{\Psi}_\Omega^\dagger \mathbf{y}$, $\mathbf{h}^{(i)}|_{\Omega^{\rm c}} = 0$
19: $\quad \mathbf{r} = \mathbf{y} - \mathbf{\Psi}\mathbf{h}^{(i)}$
20: **end for**
21: **Output:**
   Recovered channel vector $\hat{\mathbf{h}} = \mathbf{h}^{(N_{\rm p})}$.

The mode-n fiber is obtained by fixing all indexes but the n-th index of $\mathcal{M}$, i.e., $\mathcal{M}(i_1, i_2, \cdots, i_{n-1}, :, i_{n+1}, \cdots, i_N)$. The slice is obtained by fixing all but two indexes of $\mathcal{M}$. Finally, the mode-n unfolding matrix $\mathbf{M}_{(n)} \in \mathbb{C}^{I_n \times I_1 I_2 \cdots I_{n-1} I_{n+1} \cdots I_N}$ is obtained by arranging all the mode-n fibers as the columns of $\mathbf{M}_{(n)}$.

Our proposed 3D-SOMP algorithm identifies the 3D support (indexes of non-zero elements) of $\mathcal{H}_{\rm g}$ in an one-by-one fashion. The support of each dominant path of the channel is estimated in each iteration. In the i-th iteration, the algorithm starts by obtaining the mode-1 unfolding matrix $\mathbf{E}_{(1)} \in \mathbb{C}^{M_{\rm g} \times N_{\rm g} N_{\rm t}}$ of tensor $\mathcal{E}$. By calculating the ℓ_2-norm of row vectors of $\mathbf{E}_{(1)}$, the correlation vector $\mathbf{e}_\tau \in \mathbb{C}^{M_{\rm g} \times 1}$ along the delay dimension is obtained in step 7. Thus, the delay-dimension index $m_\tau^{(i)}$ of the i-th dominant path can be obtain by finding the largest element of $\mathbf{e}_\tau$ as in step 8.

Then, the user fixes the delay-dimension index $m_\tau^{(i)}$ and focuses on the slice $\mathcal{E}(m_\tau^{(i)}, :, :) \in \mathbb{C}^{N_{\rm g} \times N_{\rm t}}$. The correlation vector $\mathbf{e}_\nu \in \mathbb{C}^{N_{\rm g} \times 1}$ along the Doppler dimension is obtained in step 9. Note that the truncated 3D channel $\mathcal{H}_{\rm g}$ is block-sparse along the Doppler dimension. The length of the unique non-zero block is estimated in step 10, such that the Doppler-dimension support $\Lambda_\nu^{(i)}$ of the i-th dominant path is obtained as in step 11.

Finally, we focus on $\mathcal{E}\left(m_\tau^{(i)}, \Lambda_\nu^{(i)}, :\right)$ to obtain the angle-

dimension support of the i-th dominant path. Similarly, the angle-dimension correlation vector $\mathbf{e}_\theta \in \mathbb{C}^{N_\mathrm{t}\times 1}$ is obtained in step 12. Since the truncated 3D channel $\mathcal{H}_\mathrm{g}$ is burst-sparse along the angle dimension. The length of the non-zero burst is assumed as D. The user needs to estimate the start position of the non-zero burst. The burst sparsity is first transformed into the traditional block sparsity through a lifting transformation method following [22]. In this method, a burst-sparse vector of size $N_\mathrm{t}\times 1$ is connected to a block-sparse vector with a higher diemnsion $N_\mathrm{t}D\times 1$ via a lifting matrix $\mathbf{L}\in\{0,1\}^{N_\mathrm{t}\times N_\mathrm{t}D}$. The start position of the non-zero burst in the burst-sparse vector is corresponding to the support of the non-zero block in the higher-dimensional block-sparse vector. The $((p-1)D+q)$-th column of $\mathbf{L}$ ($p=1,2,\cdots,N_\mathrm{t}$ and $q=1,2,\cdots,D$) only has one non-zero element 1 at location $p\oplus q$ where

$$p\oplus q=\begin{cases} p+q, & \text{if } p+q\le N_\mathrm{t},\\ p+q-N_\mathrm{t}, & \text{if } p+q>N_\mathrm{t}.\end{cases} \tag{22}$$

To transform the burst sparsity of the truncated 3D channel $\mathcal{H}_\mathrm{g}$ along the angle dimension into the traditional block sparsity, the angle-dimension correlation vector $\mathbf{e}_\theta$ is modified by the lifting matrix $\mathbf{L}$ as in step 13. Then we can obtain the start position p_s of the non-zero burst in step 14-15. Therefore, the angle-dimension support corresponding to the i-th dominant path can be obtained as $\Lambda_\theta^{(i)}=\{p_s,p_s+1,\cdots,p_s+D-1\}$ in step 16.

After obtaining the delay-Doppler-angle 3D support $\Omega=\Omega\cup\left(m_\tau^{(i)},\Lambda_\nu^{(i)},\Lambda_\theta^{(i)}\right)$ in step 17, the user can partially estimate the channel $\mathbf{h}^{(i)}$ through the LS as in step 18. Finally, the residual measurements is computed by subtracting the contribution of $\mathbf{h}^{(i)}$ in step 19.

According to CS theory [23], [24], the length of measurements $M_\tau N_\nu$ is $\propto S\log(L)$, where S and L are the sparsity level and length of the sparse vector $\mathbf{h}$. According to the problem formulation in the last subsection, $S=N_{\max}N_\mathrm{p}D$ and $L=N_\mathrm{g}M_\mathrm{g}N_\mathrm{t}$. Therefore, the pilot overhead $M_\tau N_\nu$ of our proposed channel estimation technique is $\propto N_{\max}N_\mathrm{p}D\log(N_\mathrm{g}M_\mathrm{g}N_\mathrm{t})$. Note that the number of dominant paths is usually small, e.g., $N_\mathrm{p}=6$ [20]. Since the angle spread of a dominant path is usually limited, the length of non-zero block D along the angle dimension is usually much smaller than the number of BS antennas N_t, e.g., $D\approx N_\mathrm{t}/10$ [20]. The lengths of guard intervals N_g and M_g can be set as $N_{\max}$ and $M_{\max}$. For the traditional impulse based channel estimation technique (extended to OTFS massive MIMO systems [14]), the pilot overhead is $\propto N_\mathrm{t}N_{\max}M_{\max}$. Therefore, the required pilot overhead of the proposed 3D-SOMP based channel estimation is much lower than that of the previously proposed impulse based channel estimation.

IV. Simulation Results

In this section, we investigate the performance of the proposed 3D-SOMP based channel estimation technique, in terms of the normalized mean square error (NMSE) of channel estimation. The traditional impulse based channel estimation technique extended to MIMO systems is presented as a benchmark [14]. We also present the NMSE of the traditional OMP based channel estimation technique for comparison, where the OMP algorithm is used to recover $\mathbf{h}$ in (21). We simulate the standardized spatial channel model in 3GPP considering the urban macro cell environment [17]. The detailed system parameters are summarized in Table I. We define the pilot overhead ratio η as the ratio between the number of resource units for pilot transmission and the number of total resource units in the delay-Doppler domain.

TABLE I
System parameters for simulation

Parameter	Values
Carrier frequency (GHz)	2.15
Duplex mode	FDD
Subcarrier spacing (kHz)	15
Cyclic prefix duration (us)	16.6
FFT size	1024
Transmission bandwidth (# of resource blocks)	50
Size of a OTFS frame (M,N)	$(600,12)$
# of BS antennas	$8\sim 64$
# of user antennas	1
Channel model:3GPP standardized channel model	Urban macro cell
# of dominant channel paths	6
# of sub-paths per dominant path	20
User velocity (km/h)	360

In Fig. 1, we show the NMSE performance comparison against the pilot overhead ratio η. The number of BS antennas is 16 and the SNR is 5 dB. We observe that to achieve that same NMSE performance, the required pilot overhead of the proposed 3D-SOMP based channel estimation technique is smaller than that of the traditional impulse based channel estimation technique. This is because that non-orthogonal pilot pattern is used in the proposed channel estimation technique to reduce the pilot overhead. Moreover, the proposed 3D-SOMP based channel estimation technique outperforms the traditional OMP based technique when the same pilot overhead is considered. This is because that the proposed 3D-SOMP algorithm uses the 3D structured sparsity of the delay-Doppler-angle channel in OTFS massive MIMO systems.

In Fig. 2, we present the NMSE performance comparison against the number of BS antennas N_t. The pilot overhead ratio is set as 50% and the SNR is 5 dB. We observe that the NMSE performance of the traditional impulse based channel estimation technique severely degrades (NMSE is larger than 10^{-1}) when the the number of BS antennas increases larger than 8. This is due to insufficient pilot overhead when the number of BS antennas is large. Interference from adjacent impulses degrades its NMSE performance. On the contrary, the proposed 3D-SOMP based channel estimation technique works well with a large number of BS antennas. Moreover, the proposed 3D-SOMP based channel estimation technique outperforms the traditional OMP based channel estimation technique for the considered numbers of BS antennas.

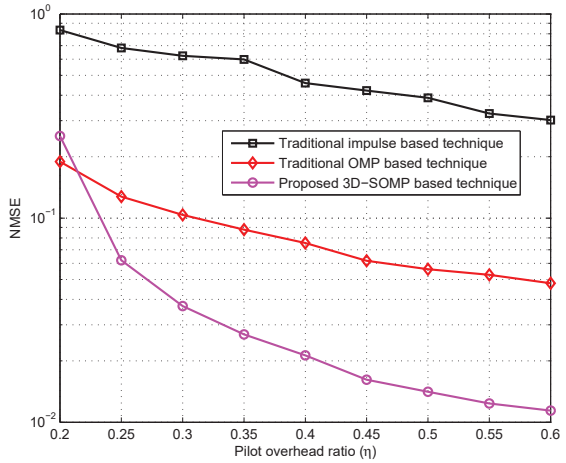

Fig. 1. The NMSE performance comparison against the pilot overhead ratio η. The number of BS antennas is 16 and the SNR is 5 dB.

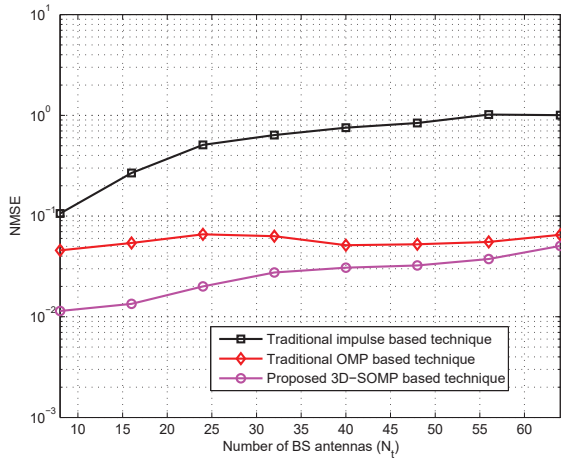

Fig. 2. The NMSE performance comparison against the number of BS antennas. The pilot overhead ratio is 50% and the SNR is 5 dB.

V. Conclusions

In this paper, we studied the OTFS modulation for massive MIMO systems for the first time with the focus on channel estimation. Specifically, we transformed the time-variant massive MIMO channels into the delay-Doppler-angle 3D channel in OTFS massive MIMO systems. We found that the 3D channel is structured sparse, i.e., sparse along the delay dimension, block-sparse along the Doppler dimension, and burst-sparse along the angle dimension. By using the 3D structured sparsity, we proposed a 3D-SOMP algorithm, which can achieve accurate CSI with low pilot overhead.

References

[1] B. Ai, X. Cheng, T. Krner, Z. D. Zhong, K. Guan, R. S. He, L. Xiong, D. W. Matolak, D. G. Michelson, and C. Briso-Rodriguez, "Challenges toward wireless communications for high-speed railway," *IEEE Trans. Intell. Transport. Syst.*, vol. 15, no. 5, pp. 2143–2158, Oct. 2014.

[2] J. Choi, V. Va, N. Gonzalez-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath, "Millimeter-wave vehicular communication to support massive automotive sensing," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 160–167, Dec. 2016.

[3] A. F. Molisch, M. Toeltsch, and S. Vermani, "Iterative methods for cancellation of intercarrier interference in OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 2158–2167, Jul. 2007.

[4] G. Leus, S. Zhou, and G. B. Giannakis, "Orthogonal multiple access over time- and frequency-selective channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1942–1950, Aug. 2003.

[5] Z. Wang, S. Zhou, G. B. Giannakis, C. R. Berger, and J. Huang, "Frequency-domain oversampling for zero-padded OFDM in underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 37, no. 1, pp. 14–24, Jan. 2012.

[6] X.-G. Xia, "Precoded and vector OFDM robust to channel spectral nulls and with reduced cyclic prefix length in single transmit antenna systems," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1363–1374, Aug. 2001.

[7] T. Ebihara and G. Leus, "Doppler-resilient orthogonal signal-division multiplexing for underwater acoustic communication," *IEEE J. Oceanic Eng.*, vol. 41, no. 2, pp. 408–427, Apr. 2016.

[8] R. Hadani, S. Rakib, M. Tsatsanis, A. Monk, A. J. Goldsmith, A. F. Molisch, and R. Calderbank, "Orthogonal time frequency space modulation," in *Proc. IEEE Wireless Communications and Networking Conference (IEEE WCNC'17)*, Mar. 2017, pp. 1–6.

[9] A. Farhang, A. RezazadehReyhani, L. E. Doyle, and B. Farhang-Boroujeny, "Low complexity modem structure for OFDM-based orthogonal time frequency space modulation," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 344–347, Jun. 2018.

[10] R. Hadani, S. Rakib, S. Kons, M. Tsatsanis, A. Monk, C. Ibars, J. Delfeld, Y. Hebron, A. J. Goldsmith, A. F. Molisch, and R. Calderbank, "Orthogonal time frequency space modulation," *arXiv preprint arXiv:1808.00519*, 2018.

[11] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE*, vol. 86, no. 10, pp. 1969–1986, Oct. 1998.

[12] R. Hadani and A. Monk, "OTFS: A new generation of modulation addressing the challenges of 5G," *arXiv preprint arXiv:1802.02623*, 2018.

[13] A. Monk, R. Hadani, M. Tsatsanis, and S. Rakib, "OTFS-orthogonal time frequency space," *arXiv preprint arXiv:1608.02993*, Aug. 2016.

[14] M. K. Ramachandran and A. Chockalingam, "MIMO-OTFS in high-doppler fading channels: Signal detection and channel estimation," *arXiv preprint arXiv:1805.02209*, 2018.

[15] K. Murali and A. Chockalingam, "On OTFS modulation for high-doppler fading channels," *arXiv preprint arXiv:1802.00929*, Feb. 2018.

[16] W. Shen, L. Dai, J. An, P. Fan, and R. Heath, "Channel estimation for orthogonal time frequency space (OTFS) massive MIMO," *Submitted to the IEEE Trans. Signl. Proc.*

[17] J. Salo, G. Del Galdo, J. Salmi, P. Kysti, M. Milojevic, D. Laselva, and C. Schneider, "MATLAB implementation of the 3GPP Spatial Channel Model (3GPP TR 25.996)," Jan. 2005. [Online]. Available: http://www.tkk.fi/Units/Radio/scm/.

[18] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Top. Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[19] X. Gao, L. Dai, S. Han, C. L. I, and X. Wang, "Reliable beamspace channel estimation for millimeter-wave massive MIMO systems with lens antenna array," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6010–6021, Sep. 2017.

[20] "Spatial channel model for multiple input multiple output (MIMO) simulations," *3GPP TR 25.996, V12.0.0 (2014-09)*.

[21] L. Dai, B. Wang, M. Peng, and S. Chen, "Hybrid precoding-based millimeter-wave massive MIMO-NOMA with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 1, pp. 131–141, Jan. 2019.

[22] A. Liu, V. K. N. Lau, and W. Dai, "Exploiting burst-sparsity in massive MIMO with partial channel support information," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7820–7830, Nov. 2016.

[23] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[24] L. Dai, Z. Wang, and Z. Yang, "Spectrally efficient time-frequency training OFDM for mobile large-scale MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 251–263, Feb. 2013.