Efficient Channel Equalization and Symbol Detection for MIMO OTFS Systems

Huiyang Qu, Guanghui Liu, *Senior Member, IEEE*, Muhammad Ali Imran, *Senior Member, IEEE*, Shan Wen, *Student Member, IEEE*, and Lei Zhang, *Senior Member, IEEE*

***Abstract*—The application of multiple-input multiple-output (MIMO) over orthogonal time frequency space (OTFS) modulation is envisioned to provide high-data-rate wireless transmission in high-mobility environments. However, in these communication scenarios, the multiple-dimensional interference, which can generate from space, delay and Doppler domains, challenges the channel equalization and symbol detection at the MIMO-OTFS receiver. To tackle this issue, we propose a time-space domain channel equalizer, relying on the mathematical least squares minimum residual algorithm, to remove the channel distortion on data symbols. The proposed channel equalizer adopts a recursion method to achieve symbol estimates, which can realize fast convergence by leveraging the sparsity of MIMO-OTFS channel matrix. Instead of directly remapping the equalized OTFS symbols into data bits, we develop an enhanced data detection (EDD) scheme to iteratively demodulate the superposed multi-antenna signal. The EDD can not only realize the linear-complexity interference cancellation, but also efficiently reap the spatial and multi-path diversities of MIMO-OTFS channel. The simulations show the proposed channel equalization and EDD algorithms enable the MIMO-OTFS receiver to robustly demodulate multi-stream 256-ary quadrature amplitude modulation symbols, under a maximum velocity of 550 km/h at 5.9 GHz carrier frequency.**

***Index Terms*—Multiple-input multiple-output (MIMO) orthogonal time frequency space (OTFS), channel equalization, enhanced data detection (EDD), interference cancellation, high mobility**

I. Introduction

Time- and frequency-domain selectivity is an inherent characteristic of wireless communication channels [1]. To address the frequency-domain selectivity induced by channel multi-path propagation, complex transceiver schemes are designed. For instance, by using multi-carrier modulation, such as cyclic prefix orthogonal frequency division multiplexing (CP-OFDM) [2], low complexity per-subcarrier equalization algorithms can be adopted. However, on the other hand, the channel time-domain selectivity caused by the mobility in high-speed vehicular communication networks [3], challenges the traditional multi-carrier modulation. Specifically, the Doppler effect involved by mobility is represented as the inter-carrier interference (ICI) in OFDM or its varieties [4]–[8], which poses a bottleneck to overall system performance.

In order to tackle the doubly-selective channels, a novel two-dimensional (2-D) modulation technique, namely orthogonal time frequency space (OTFS), was proposed in [9]. Since 2018, P. Raviteja and Yi Hong have shown its robust performance in high-mobility environments [10]–[12]. The core principle of OTFS is that it maps the modulated data symbols in a delay-Doppler (DD) plane, and employs the inverse symplectic finite Fourier transform (ISFFT) to spread them over the whole time-frequency (TF) grids. Then, the TF-domain data is modulated by a multi-carrier scheme, e.g., OFDM [13], [14]. Through this procedure, the transmitted data symbols can be equally impacted by the doubly-dispersive channel, thus mitigating the performance loss [15].

Similar to OFDM, multiple-input multiple-output (MIMO) can be combined with OTFS as well for further increasing the transmission throughput [16]–[22]. To achieve the robust data transmission in MIMO-OTFS system, efficient channel equalization and data detection algorithms are required at the receiver side. In some high-mobility scenarios, such as high-speed train running in a clear and open signal-transmission space [23], the channel scattering components are limited and the channel responses may appear as a few impulses in the Doppler frequency [24]. The existing channel equalization schemes, including linear schemes [20], [25]–[27] and non-linear schemes [19], [28]–[30], have been developed to efficiently recover the transmitted multi-stream OTFS symbols. For instance, in [25], [26], the authors suppose each propagation path of channel has only one Doppler frequency shift. Thus, the OTFS transmission matrix has a special block circulant structure, which enables the linear schemes to work at low computational cost by using matrix decomposing techniques. Although the widely-used linear schemes are easy to be implemented [20], [27], they can not achieve the satisfactory demodulation performance, in contrast to the message passing (MP) based non-linear methods in [10], [19], [28]–[30]. In addition, thanks to the 2-D sparsity of the DD channel, the number of connection nodes in the factor graph of the MP equalizer can be greatly reduced, which yields an acceptable computational complexity.

However, in a general communication scenario, it may involve a large number of scattering objects, like in vehicle to everything (V2X) communications [31]. As analyzed in [32], [33], scattering-abundant channel can induce not only the inter-symbol interference (ISI), but also the inter-Doppler interference (IDI) with relatively wide band in the Doppler

This work was supported by the National Natural Science Foundation of China under Grants 62071097 and U20A20184. (Corresponding author: Guanghui Liu).

Huiyang Qu, Guanghui Liu and Shan Wen are with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: guanghuiliu@uestc.edu.cn; {hyqu, shanwen}@std.uestc.edu.cn).

Muhammad Ali Imran and Lei Zhang are with the James Watt School of Engineering, University of Glasgow, Glasgow, G12 8QQ, U.K. (e-mail: {Muhammad.Imran, Lei.Zhang}@glasgow.ac.uk).

domain. Moreover, when the OTFS signal undergoes a MIMO channel, the additional interference will be involved in the space domain, generating inter-antenna interference (IAI) (i.e., interference by other streams in multi-user MIMO system). In such a time-varying MIMO channel with significant surroundings, it would cause a burden for the number of connection nodes in the factor graph of the MP equalizer, yielding an unacceptable computational cost. Thus, how to efficiently and robustly demodulate the received MIMO-OTFS signal with the multi-dimensional interference, i.e., IAI, IDI and ISI, is another challenge to be addressed.

To solve the above problem, in this paper, a time-space (TS) domain channel equalizer, relying on the mathematical least squares minimum residual (LSMR) algorithm, is developed for MIMO-OTFS systems. It is capable of achieving the comparable equalization performance to the linear minimum mean square error (LMMSE), but preserving much lower computational complexity. As a cascade of TS-domain channel equalizer, an enhanced data detector (EDD) is proposed in the delay-Doppler-space (DDS) domain (i.e., the OTFS symbols are demodulated in this domain). The EDD involves a layer-by-layer interference cancellation and signal-combined detection scheme, which can significantly improve the data demodulation performance in MIMO-OTFS system. Our contributions can be summarized as follows.

1) Different from the existing channel modeling assumed for OTFS, we consider a more general communication environment that each sub-channel has a continuous spectrum in the Doppler frequency. The MIMO-OTFS signal transmission over the continuous-Doppler-spread channel is modeled in the TS, frequency-space (FS), and DDS domains, respectively [1]. Our analysis on the structures of MIMO-OTFS channel matrices in these dimensions can provide useful insights for channel equalization and symbol detection.
2) We formulate the MIMO-OTFS channel equalization as a linear optimization problem, thereby proposing a TS-domain LSMR-based channel equalizer. Specifically, in this paper, one important conclusion can be achieved that performing the linear equalization schemes, such as zero forcing (ZF) or LMMSE, in the TS, FS, or DDS domain are equivalent. Thus, by leveraging the sparsity of TS-domain channel matrix, an LSMR-based equalizer is proposed to retrieve the OTFS symbols at each transmit stream. The developed equalizer can achieve fast convergence under high Doppler spread, which yields a much lower computational cost than the LMMSE.
3) Instead of directly adopting a hard or soft decision on symbol estimates after channel equalization, we design an EDD in DDS domain to iteratively demodulate the multi-antenna data symbols. Benefiting from the precise data symbol estimates from channel equalizer, the interference (i.e., IAI and ISI) can be perfectly reconstructed and eliminated in EDD. Also, the interference-eliminated signal components are further accumulated by using equal gain combining (EGC) and maximum ratio combining (MRC [33]) techniques, so as to efficiently reap the spatial and multi-path diversities of MIMO-OTFS channel. The simulation results verify the proposed receiver can robustly demodulate 256-ary quadrature amplitude modulation (QAM) symbols, in the case of the MIMO-OTFS signal transmission over high-mobility environments.

[1]Note that symbols modeled in different domains are the same but only represented in different forms in order to compare the corresponding equalization methods.

The remainder of this paper is organized as follows. The MIMO-OTFS transmission framework and channel matrix are analyzed in Sec. II. In Sec. III, the LSMR based TS-domain channel equalization is proposed. In Sec. IV, an enhanced iterative data detection is developed. The simulation results are presented in Sec. V and conclusions are given in Sec. VI.

Notations: In this paper, vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. The element in the l-th row and v-th column of matrix $\mathbf{Y}$ is defined as $Y(l,v)$. We use $(\cdot)^{-1}$, $(\cdot)^{T}$ and $(\cdot)^{H}$ to represent inverse, transpose and conjugate transpose of a matrix, respectively. The operators $\odot$, $\mathrm{vec}(\cdot)$, $\mathrm{invec}(\cdot)$ and $\mathbb{E}\{\cdot\}$ denote Hadamard product, the vectorizing of a matrix, invertorizing of a vector, and expectation operation, respectively. $\mathbb{C}$ stands for the set of complex numbers. The zeroth order Bessel function of the first kind is denoted as $J_0(\cdot)$. $\mathbf{I}_M$ and $\mathbf{F}_M$ denote the $M \times M$ identity matrix and the fast Fourier transform (FFT) matrix [34], respectively.

II. OTFS Transmission over Time-varying MIMO Channel

In this section, we firstly model the OTFS-based MIMO system, and study the characteristics of MIMO-OTFS channel matrix, from TS, FS, and DDS dimensions, respectively. The analysis helps us further investigate the linear equalization of time-varying MIMO channels.

A. MIMO-OTFS System Model

Fig. 1 demonstrates MIMO-OTFS transmission model, where the receiver is equipped with N_{R} antennas to simultaneously serve N_{T} OTFS data streams. Note that the setup can be used for both multi-user single antenna transmitters and a single transmitter with multi-antenna cases. In this section, we only analyze the basic MIMO-OTFS transmission architecture in Fig. 1, while the proposed demodulation scheme, including channel equalization and EDD modules, will be introduced in Sec. III and IV, respectively.

We use $X_i[m,n]$ to denote a QAM symbol that allocated at the m-th delay and n-th Doppler bin of the DD plane in transmit stream i, where $m \in [0, M-1], n \in [0, N-1], i \in [1, N_{\mathrm{T}}]$. The 2-D data block $\mathbf{X}_i \in \mathbb{C}^{M\times N}$, which contains MN QAM symbols, are transformed from the DD domain to TF domain by using ISFFT, i.e., $\mathbf{D}_i = \mathbf{F}_M \mathbf{X}_i \mathbf{F}_N^H, \mathbf{D}_i \in \mathbb{C}^{M\times N}$, as shown in Fig. 1. Subsequently, $\mathbf{D}_i$ is sent into a conventional TF modulator, OFDM, to produce the time-domain signal block, according to $\mathbf{S}_i = \mathbf{F}_M^H \mathbf{D}_i, \mathbf{S}_i \in \mathbb{C}^{M\times N}$. By column-wise reading the entries in

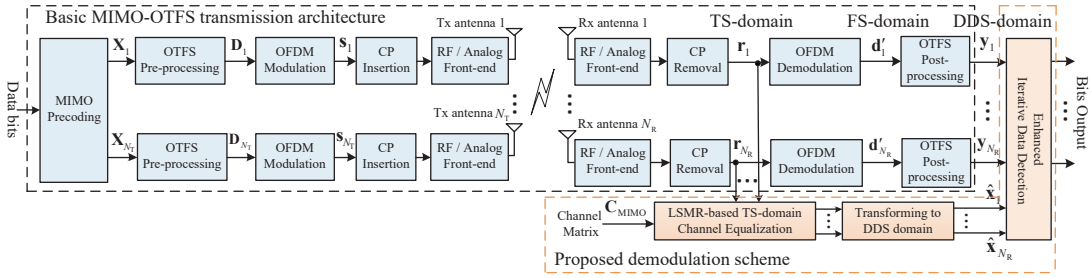


Fig. 1. The basic transmission architecture of MIMO-OTFS system and the proposed demodulation scheme.

$\mathbf{S}_i$, the data samples within one OTFS transmission block are expressed as $\mathbf{s}_i = \mathrm{vec}(\mathbf{S}_i), \mathbf{s}_i \in \mathbb{C}^{MN\times 1}$. To avoid the inter-block interference, a CP with length M_{cp} is inserted at the beginning of $\mathbf{s}_i$. In this paper, we consider only one CP for the whole OTFS transmission block, to achieve an improved spectrum efficiency.

We denote $h_{j,i}[k,l]$ as the time-varying channel response between transmit antenna i and receive antenna j, corresponding to the sampling time index k and delay tap $l, l \in [0, L-1]$. At receive antenna j, after CP removal, the received time-domain sample is denoted as $r_j[k]$. Omitting the additive noise effect [2], $r_j[k]$ can be written as

$$r_j[k] = \sum_{i=1}^{N_{\mathrm{T}}} \sum_{l=0}^{L-1} h_{j,i}[k,l]\, s_i[(k-l)_{MN}]. \tag{1}$$

Defining $\mathbf{r}_j = [r_j[0], r_j[1], \cdots, r_j[MN-1]]^T \in \mathbb{C}^{MN\times 1}$ as the received *time-domain* sequences in Fig. 1, we have

$$\mathbf{r}_j = \sum_{i=1}^{N_{\mathrm{T}}} \mathbf{C}_{j,i}\mathbf{s}_i, \tag{2}$$

where $\mathbf{C}_{j,i} \in \mathbb{C}^{MN\times MN}$ denotes the time-domain channel matrix between arbitrary transmit antenna i and receive antenna j, which follows $C_{j,i}(k1,k2) = h_{j,i}[k1,(k1-k2)_{MN}]$. Then, $\mathbf{r}_j$ is rearranged into a time-domain data block $\mathbf{R}_j \in \mathbb{C}^{M\times N}$ by $\mathbf{R}_j = \mathrm{invec}(\mathbf{r}_j)$. The TF-domain data block is obtained through FFT, which is expressed as $\mathbf{D'}_j = \mathbf{F}_M\mathbf{R}_j, \mathbf{D'}_j \in \mathbb{C}^{M\times N}$. In the end, the SFFT converts $\mathbf{D'}_j$ back into the DD domain according to $\mathbf{Y}_j = \mathbf{F}_M^H\mathbf{D'}_j\mathbf{F}_N, \mathbf{Y}_j \in \mathbb{C}^{M\times N}$.

To make the analysis and derivations clear, we propose the following two propositions to write the signal in matrix forms in frequency and DD domains.

Proposition 1: At receive antenna $j, j \in [1, N_{\mathrm{R}}]$, after OFDM modulation, the *frequency-domain* data vector $\mathbf{d'}_j$ in Fig. 1 can be expressed as

$$\mathbf{d'}_j = \sum_{i=1}^{N_{\mathrm{T}}} \mathbf{H}_{j,i}\mathbf{d}_i. \tag{3}$$

where $\mathbf{d'}_j = \mathrm{vec}(\mathbf{D'}_j) \in \mathbb{C}^{MN\times 1}$ and $\mathbf{d}_i = \mathrm{vec}(\mathbf{D}_i) \in \mathbb{C}^{MN\times 1}$ denote the received and transmitted frequency data symbols, respectively. These symbols are modulated at N OFDM symbols, each of which has M sucarriers. Matrix $\mathbf{H}_{j,i} \in \mathbb{C}^{MN\times MN}$ is the frequency-domain channel matrix between transmit antenna i and receive antenna j, i.e.,

$$\mathbf{H}_{j,i} \triangleq (\mathbf{I}_N \otimes \mathbf{F}_M)\,\mathbf{C}_{j,i}\left(\mathbf{I}_N \otimes \mathbf{F}_M^H\right). \tag{4}$$

Proof: See Appendix A. ■

The OTFS post-processing in Fig. 1, employs the SFFT to transform $\mathbf{d'}_j$ back into DD domain. Accordingly, we have the following transmission relation.

Proposition 2: The *DD-domain* received data vector at antenna j, i.e., $\mathbf{y}_j$ in Fig. 1, can be derived as

$$\mathbf{y}_j = \sum_{i=1}^{N_{\mathrm{T}}} \mathbf{G}_{j,i}\mathbf{x}_i, \tag{5}$$

where $\mathbf{x}_i \triangleq \mathrm{vec}\left(\mathbf{X}_i^T\right) \in \mathbb{C}^{MN\times 1}$ and $\mathbf{y}_i \triangleq \mathrm{vec}\left(\mathbf{Y}_i^T\right) \in \mathbb{C}^{MN\times 1}$ contains MN transmitted and received OTFS symbols, respectively. The DD-domain transmission matrix between transmit antenna i and receive antenna j is represented as

$$\mathbf{G}_{j,i} = \mathbf{\Pi}\,(\mathbf{F}_N \otimes \mathbf{I}_M)\,\mathbf{C}_{j,i}\left(\mathbf{F}_N^H \otimes \mathbf{I}_M\right)\mathbf{\Pi}^T, \tag{6}$$

where $\mathbf{\Pi}$ is a permutation matrix in (52). Note that introducing a permutation matrix in (6) enables the OTFS transmission matrix $\mathbf{G}_{j,i}$ to have a straightforward shape, which helps us better analyze the interference components. More details can be found in Appendix A.

Proof: See Appendix A. ■

B. Channel Model

In the existing MIMO-OTFS works, the channel responses are supposed to be sparse in the Doppler domain, i.e., the number of Doppler shifts is very limited due to the few channel scatters. However, this assumption may not be valid in a more general communication environment, e.g., V2X communications [31], where a large number of scattering objects can be involved. Hence, this paper exactly aims to handle such a scattering-abundant MIMO channel. To do this, we make the following assumptions.

Assumption 1: In the case of single user transmission but with multi-antennas, we suppose the sub-channels are independent to each other. The assumption is reasonable in

[2]In this section, we omit the effect of additive noise to focus on the transmission channel. As the noise samples are not correlated with data symbols, its effect will be investigated independently in Sec. V.

the case of multi-user MIMO with each user equipped with single antenna at the transmitters.

Assumption 2: In this paper, we use the classical Jakes' model [35] to represent the scattering-abundant channels, which has been widely employed in the wireless communications. For each tap, $h_{j,i}[k,l]$ follows Rayleigh distribution in time, of which the auto-correlation function is

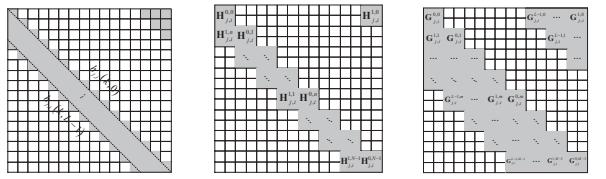
$$\mathbb{E}\left\{h_{j,i}[k_1,l]\,h^*_{j,i}[k_2,l]\right\} = \sigma^2_{j,i,l} J_0\left(2\pi f_d\left|k_1-k_2\right|T_s\right), \tag{7}$$

where $\sigma^2_{j,i,l}$ represents the channel power of the l-th tap between transmit antenna i and receive antenna j, and f_d is the maximum Doppler frequency. Accordingly, we have a "U-shaped" Doppler spectrum, i.e.,

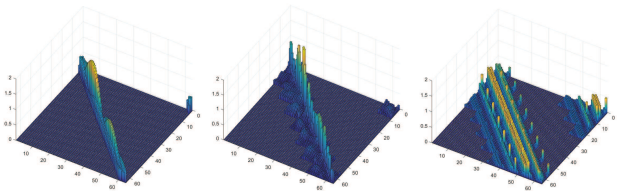
$$\varphi_{j,i,l}(v) = \begin{cases} \frac{1}{\pi\sqrt{f_d^2-v^2}}, & v\in[-f_d,f_d] \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

C. Analysis of Channel Matrix

In this subsection, we analyze the structures of sub-channel matrices from time, frequency and DD dimensions, i.e., $\mathbf{C}_{j,i}$, $\mathbf{H}_{j,i}$ and $\mathbf{G}_{j,i}$ in (2), (3) and (5), respectively. The analysis helps us propose efficient channel equalizer and data detector in Sec. III and IV, respectively.



(a) The structure of $\mathbf{C}_{j,i}$. (b) The structure of $\mathbf{H}_{j,i}$. (c) The structure of $\mathbf{G}_{j,i}$.



(d) An instance of $\mathbf{C}_{j,i}$. (e) An instance of $\mathbf{H}_{j,i}$. (f) An instance of $\mathbf{G}_{j,i}$.

Fig. 2. The illustrations of sub-channel matrices $\mathbf{C}_{j,i}$, $\mathbf{H}_{j,i}$ and $\mathbf{G}_{j,i}$, in the time, frequency, and DD domains, respectively. Parameters are set as $M=8$, $N=8$ and $L=4$. The CIRs of each sub-channel are obtained by the classical Jakes' channel simulator [36]. The maximum channel Doppler frequency f_d normalized to the OFDM subcarrier spacing Δf is 20%, corresponding to a velocity of 550 km/h at 5.9 GHz carrier frequency.

Firstly, let us recall the time-domain sub-channel matrix $\mathbf{C}_{j,i}$ in (2). It has a Toeplitz structure, but is not circulant due to the time-varying channel responses. Since the CP is only inserted at the beginning of the whole time-domain OTFS transmission block, instead of each OFDM symbol, $\mathbf{C}_{j,i}$ is not block diagonal, shown in Figs. 2 (a) and (d). Note that when considering a CP or zero-padding OTFS system [33], $\mathbf{C}_{j,i}$ has a block diagonal structure. In this case, the low-complexity time-domain channel equalization scheme [37], can be used to recover the OTFS symbols.

In the frequency domain, the sub-channel matrix $\mathbf{H}_{j,i}$ is obtained from (4), which can be further decomposed into a set of sub-matrices by using (3) and (4). Specifically, for the frequency-domain data symbols of the n-th OFDM symbol at receive antenna j, denoted as $\mathbf{d'}_j(n) = [D'_j[0,n], D'_j[1,n], \cdots, D'_j[M-1,n]]^T \in \mathbb{C}^{M\times 1}$, it follows

$$\mathbf{d'}_j(n) = \mathbf{H}^{0,n}_{j,i}\mathbf{d}_i(n) + \mathbf{H}^{1,n}_{j,i}\mathbf{d}_i(n-1), \tag{9}$$

where $\mathbf{H}^{0,n}_{j,i} \in \mathbb{C}^{M\times M}$ and $\mathbf{H}^{1,n}_{j,i} \in \mathbb{C}^{M\times M}$ are named as the ICI and ISI matries, respectively. They are achieved by

$$\mathbf{H}^{0,n}_{j,i} = \mathbf{F}_M \mathbf{C}^{0,n}_{j,i}\mathbf{F}^H_M; \quad \mathbf{H}^{1,n}_{j,i} = \mathbf{F}_M \mathbf{C}^{1,n}_{j,i}\mathbf{F}^H_M, \tag{10}$$

where $\mathbf{C}^{0,n}_{j,i} \in \mathbb{C}^{M\times M}$ and $\mathbf{C}^{1,n}_{j,i} \in \mathbb{C}^{M\times M}$ are lower and upper triangular matrices, respectively, composed by the CIR samples within the n-th OFDM duration, i.e,

$$\begin{aligned} C^{0,n}_{j,i}(m_1,m_2) &= h_{j,i}[(n-1)M+m_1, m_1-m_2] \\ C^{1,n}_{j,i}(m_1,m_2) &= h_{j,i}[(n-1)M+m_1, (m_1-m_2)_M]. \end{aligned} \tag{11}$$

Due to the large channel Doppler spread, the CIR values change quickly in Fig. 2(d), which yields the significant ICI in the frequency domain. Accordingly, the submatrices in the main diagonal of $\mathbf{H}_{j,i}$, i.e., $\mathbf{H}^{0,0}_{j,i}, \mathbf{H}^{0,1}_{j,i}, \cdots, \mathbf{H}^{0,n}_{j,i}, \cdots, \mathbf{H}^{0,N-1}_{j,i}$, contains the channel frequency responses and ICI, shown in Figs. 2(b) and (e). In this paper, since only one CP is inserted at the beginning of whole OTFS transmission frame, the ISI will be involved between the adjacent OFDM symbols. The ISI reveals as $\mathbf{H}^{1,0}_{j,i}, \mathbf{H}^{1,1}_{j,i}, \cdots, \mathbf{H}^{1,n}_{j,i}, \cdots, \mathbf{H}^{1,N-1}_{j,i}$. Note that if considering a CP or zero-padding OTFS transmission structure [33], there will be no ISI, indicating $\mathbf{H}^{1,n}_{j,i} = \mathbf{0}$.

Proposition 3: The DD-domain sub-channel matrix $\mathbf{G}_{j,i}$ in (6) is block-banded, of which the sub-matrix $\mathbf{G}^{l,m}_{j,i} \in \mathbb{C}^{N\times N}, l=0,\cdots,L-1; m=0,\cdots,M-1$, follows

$$\mathbf{G}^{l,m}_{j,i} = \left(\mathbf{F}_N \otimes \mathbf{a}^T_M(l)\right)\mathbf{C}_{j,i}\left(\mathbf{F}^H_N \otimes \mathbf{a}_M(m)\right), \tag{12}$$

where $\mathbf{a}_M(m)$ refers to the m-th column of $\mathbf{I}_M$.

Proof: See Appendix B. ■

According to **proposition** 3 and the analysis in [33], we can investigate the ISI and IDI generated in the delay and Doppler dimensions, respectively. Specifically, from Fig. 2(c), the ISI is contributed by the submatrices $\mathbf{G}^{l,m}_{i,j} \in \mathbb{C}^{N\times N}, l\neq 0, m=0,\cdots,M-1$. Meanwhile, the IDI is contributed by the off-diagonal elements of $\mathbf{G}^{0,m}_{i,j} \in \mathbb{C}^{N\times N}, m=0,\cdots,M-1$. Fig. 2(f) presents an instance of matrix $\mathbf{G}_{j,i}$, where the number of non-zero submatrices in each row or column direction is $L-1$, corresponding to the maximum channel delay spread [32].

Specially, facing a low or moderate-speed case, the channel is time-invariant during one OFDM symbol duration. As a result, matrix $\mathbf{G}_{j,i}$ has a circulant structure, which enables to design a low-complexity channel equalizer to recover the OTFS data symbols. In this paper, we focus on the high-mobility scenarios, where the channel varies even within

one OFDM symbol interval. Thus, the large Doppler spread induced IDI challenges the channel equalization [3].

III. Equalization of MIMO-OTFS Channel

In this section, we firstly derive a unified expression that can represent the TS, FS and DDS-domain transmission features of MIMO-OTFS. Note that symbols modeled in different domains are the same but at the different stages in the receiver as shown in Fig. 1, i.e., $\mathbf{r}_j$, $\mathbf{d}'_j$ and $\mathbf{y}_j$ are TS, FS and DDS-domain symbols, respectively. Then, we formulate a generic linear equalization of MIMO-OTFS channel, and conclude that performing the LMMSE scheme in the TS, FS, or DDS domain achieves the same performance. To avoid the shortages of the existing low-complexity LMMSE methods, i.e., limited application scenarios and inferior performance, we propose an efficient TS-domain MIMO-OTFS channel equalization scheme. It can simultaneously realize low computational cost and accurate symbol estimates.

A. The Linear Equalization

According to (2), and **propositions** 1–3, we can write a generic transmission model for MIMO-OTFS, i.e.,

$$\boldsymbol{\mu}_{\mathrm{MIMO}} = \boldsymbol{\Phi}_{\mathrm{MIMO}}\boldsymbol{\chi}_{\mathrm{MIMO}} + \boldsymbol{n}_{\mathrm{MIMO}}, \tag{13}$$

where $\boldsymbol{\chi}_{\mathrm{MIMO}} = \left[\boldsymbol{\chi}_1^T, \cdots, \boldsymbol{\chi}_{N_{\mathrm{T}}}^T\right] \in \mathbb{C}^{N_{\mathrm{T}}MN\times 1}$, $\boldsymbol{\mu}_{\mathrm{MIMO}} = \left[\boldsymbol{\mu}_1^T, \cdots, \boldsymbol{\mu}_{N_{\mathrm{R}}}^T\right] \in \mathbb{C}^{N_{\mathrm{R}}MN\times 1}$ and $\boldsymbol{n}_{\mathrm{MIMO}}$ denote the transmitted data symbols at N_{T} antennas, received data symbols at N_{R} antennas, and noise samples, respectively. Matrix $\boldsymbol{\Phi}_{\mathrm{MIMO}}$ represents a general MIMO channel matrix, of which the structure is

$$\boldsymbol{\Phi}_{\mathrm{MIMO}} = \begin{bmatrix} \boldsymbol{\Phi}_{1,1} & \cdots & \boldsymbol{\Phi}_{1,N_{\mathrm{T}}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Phi}_{N_{\mathrm{R}},1} & \cdots & \boldsymbol{\Phi}_{N_{\mathrm{R}},N_{\mathrm{T}}} \end{bmatrix}, \tag{14}$$

where $\boldsymbol{\Phi}_{j,i}$ denotes the sub-channel matrix between arbitrary transmit antenna i and receive antenna j.

Remark 1: Equ. (13) is a unified expression that can represent the TS, FS, or DDS-domain transmission features and the resulting interference, shown in (2), (3) and (5), by given the different meanings of $\boldsymbol{\Phi}_{j,i}$, $\boldsymbol{\chi}_i$ and $\boldsymbol{\mu}_i$. Specifically, the transmitted symbols in these three domain $\mathbf{s}_i$, $\mathbf{d}_i$ and $\mathbf{x}_i$ are related to the received symbols $\mathbf{r}_j$, $\mathbf{d}'_j$ and $\mathbf{y}_j$ in Fig. 1, respectively. For example, let $\boldsymbol{\chi}_i = \mathbf{x}_i$, $\boldsymbol{\mu}_j = \mathbf{y}_j$ and $\boldsymbol{\Phi}_{j,i} = \mathbf{G}_{j,i}$, we can achieve the DDS-domain MIMO-OTFS transmission in Fig. 1. By leveraging (13), a generic linear equalization scheme can be formulated in the following part.

By using the unified equation (13), channel equalization can be modeled as solving the following least square (LS) problem

$$\min_{\hat{\boldsymbol{\chi}}_{\mathrm{MIMO}}\in\mathbb{C}^{N_{\mathrm{T}}MN\times 1}} \left\|\boldsymbol{\Phi}_{\mathrm{MIMO}}\hat{\boldsymbol{\chi}}_{\mathrm{MIMO}} - \boldsymbol{\mu}_{\mathrm{MIMO}}\right\|_2. \tag{15}$$

[3]An effective approach to suppress the IDI can be found in [38]. It designs a transmitter window and enables the channel responses in the Doppler domain to become much sparser, which helps to reduce the channel equalization complexity.

A straightforward approach could be applied to (15) by using a linear equalizer $\mathbf{A} \in \mathbb{C}^{N_T MN\times N_T MN}$. For instance, the choice

$$\mathbf{A} = \left(\boldsymbol{\Phi}_{\mathrm{MIMO}}^H\boldsymbol{\Phi}_{\mathrm{MIMO}} + \lambda\mathbf{I}_{N_T MN}\right)^{-1}\boldsymbol{\Phi}_{\mathrm{MIMO}}^H, \tag{16}$$

gives an LMMSE equalizer if λ is the noise power, and the ZF equalizer if $\lambda = 0$. Thus, the MSE of the equalized data symbols at all transmit antennas is

$$\begin{aligned} \mathrm{MSE} &= \mathbb{E}\left\{\left\|\hat{\boldsymbol{\chi}}_{\mathrm{MIMO}} - \boldsymbol{\chi}_{\mathrm{MIMO}}\right\|^2\right\} \\ &= \mathbb{E}\left\{\left\|\mathbf{A}\boldsymbol{\mu}_{\mathrm{MIMO}} - \boldsymbol{\chi}_{\mathrm{MIMO}}\right\|^2\right\}. \end{aligned} \tag{17}$$

Corollary 1: Applying the LMMSE-based equalizer in the TS, FS, or DDS domain produces the same mean squared error (MSE) value.

Proof: See Appendix C. ■

According to **corollary 1**, the linear equalization schemes can be implemented in any one of the three domains without equalization performance loss. In practice, directly using the equalizer like (16) would be prohibitive in term of calculational complexity, since it scales as $\mathcal{O}\left(N_T^3M^3N^3\right)$ due to the matrix inversion.

The DDS-domain scheme: To tackle the calculational issue, the low-complexity DDS-domain equalizers were proposed in the existing MIMO-OTFS works [20], [27], by investigating the block circulant structure of the DDS-domain channel matrix. However, these low-complexity schemes are based on the assumption that the sub-channel appears as a few impulse responses in the DD domain. It does not fulfill the premise of channel model in this paper, as denoted in Sec. II-B. In other words, the above low-complexity LMMSE algorithms can not work in the continuous-Doppler-spread channels.

The FS-domain scheme: There can be another idea that adopting the low-complexity MIMO-OFDM equalizers in the FS domain of Fig. 1, such as "block banded equalizer", which has been widely used in the literature [34], [39], [40]. The "block banded equalizer" is achieved by approximating the "full band LMMSE" equalizer in (14) as

$$\mathbf{A}_{\mathrm{B-FS}} = \left(\mathbf{H}_{\mathrm{B-MIMO}}^H\mathbf{H}_{\mathrm{B-MIMO}} + \lambda\mathbf{I}_{N_T MN}\right)^{-1}\mathbf{H}_{\mathrm{B-MIMO}}^H \tag{18}$$

In (18), $\mathbf{H}_{\mathrm{B-MIMO}}$ is an approximation of $\mathbf{H}_{\mathrm{MIMO}}$, and it can be obtained by selecting the main diagonal, the Q superdiagonals and Q subdiagonals of each submatrices $\mathbf{H}_{j,i}^{0,n}, \mathbf{H}_{j,i}^{1,n}; n = 0, \cdots, N-1$ in (10). Although the "block banded equalizer" with relatively small value of Q can realize computational complexity reduction, it performs at the cost of losing the symbol estimation accuracy. This is because the operation of "Q-banded channel matrix" is essentially equivalent to approximate the time-varying channel by using a "Q-order" complex-exponential basis expansion model [41]. The ad-hoc approximation of the channel would cause a non-negligible modeling error, and further deteriorate the equalization performance.

B. The Proposed TS-domain Channel Equalization

Based on the analysis above, FS and DDS-domain channel equalizers are not suitable to be adopted for MIMO-OTFS.

In the following, we will propose to equalize the received data symbols in the TS domain, i.e., $\mathbf{r}_j$ in Fig. 1. According to the derivations and analysis in Sec. III-A, it is clear that the challenge of linear equalization is how to efficiently solve (15), especially when the number of antennas and length of transmission block are large. In mathematics, the LSMR is a conjugate gradient based iterative algorithm to solve large and sparse equation as (15). It is similar to the well-known least-squares QR decomposition (LSQR) approach [42], but has faster convergence speed and better numerical stability. Focusing on the MIMO-OTFS transmission in the TS domain, the CIR is indeed sparse, as described in Sec. II. Thus, a sparse TS-domain channel matrix $\mathbf{C}_{\rm MIMO}$ can be achieved, shown in Fig. 2. Taking advantage of the sparsity of TS-domain matrix, the iteration process in the LSMR converges quickly, yielding a much lower computational overhead than the conventional LMMSE method.

In this paper, we construct an LSMR based TS-domain channel equalizer in Fig. 1, to remove the channel distortions on the transmitted OTFS symbols $\mathbf{s}_{\rm MIMO}$. Specifically, according to (13) and (2), the TS-domain transmission can be formulated as

$$\mathbf{r}_{\rm MIMO} = \mathbf{C}_{\rm MIMO}\mathbf{s}_{\rm MIMO} + \boldsymbol{n}_{\rm MIMO}, \tag{19}$$

where $\mathbf{s}_{\rm MIMO} = \left[\mathbf{s}_1^T, \cdots, \mathbf{s}_{N_{\rm T}}^T\right]^T \in \mathbb{C}^{N_{\rm T}MN\times 1}$, $\mathbf{r}_{\rm MIMO} = \left[\mathbf{r}_1^T, \cdots, \mathbf{r}_{N_{\rm R}}^T\right]^T \in \mathbb{C}^{N_{\rm R}MN\times 1}$, and $\boldsymbol{n}_{\rm MIMO}$ denote the TS-domain transmitted data symbols at $N_{\rm T}$ antennas, received data symbols at $N_{\rm R}$ antennas, and noise samples, respectively. The TS-domain channel matrix $\mathbf{C}_{\rm MIMO}$ can be found in (70). According to (16), the TS-domain equalization matrix follows

$$\mathbf{A}_{\rm TS} = \left(\mathbf{C}_{\rm MIMO}^H\mathbf{C}_{\rm MIMO} + \lambda\mathbf{I}_{N_TMN}\right)^{-1}\mathbf{C}_{\rm MIMO}^H. \tag{20}$$

The LSMR algorithm aims to iteratively solve (20), of which the solving process is summarized in Algorithm I. It begins with the Golub-Kahan process (Bidiagonalization) [43]. In the i-th iteration, after the Bidiagonalization operation, the Krylov subspace $\mathbf{V}_i^{(k)} = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_i]$ is constructed to obtain $\mathbf{s}_{\rm MIMO}^i$ by solving the linear equation $\mathbf{s}_{\rm MIMO}^i = \mathbf{V}_i\mathbf{x}_i$. The estimated data vector $\mathbf{s}_{\rm MIMO}^i$ is further updated in the Krylov subspace. It will be seen in Sec. V that, after a few number of iterations, the LSMR can achieve a same MSE value to the LMMSE.

In the LSMR, the complexity order in each iteration is $\mathcal{O}\left(N_T^2MNL\right)$ [44]. For I iterations, the overall complexity order is $\mathcal{O}\left(N_T^2MNLI\right)$. In practice, the numbers of channel taps L and iteration times I are much smaller than the transmission block size, i.e., $LI \ll MN$. Thus, the overall complexity of the LSMR is significantly smaller than the complexity $\mathcal{O}\left(N_T^3M^3N^3\right)$ of the LMMSE.

Remark 2: When the TS-domain channel equalization scheme is implemented, shown in Fig. 1, the estimated TS data symbols $\hat{\mathbf{s}}_{\rm MIMO}$ can be directly transformed to DDS domain, generating the equalized OTFS symbols $\hat{\mathbf{x}}_{\rm MIMO}$. Then, $\hat{\mathbf{x}}_{\rm MIMO}$ can be remapped into data bits, thus accomplishing the demodulation process. Hence, it could be feasible to discard the OFDM and OTFS demodulation procedures in the basic transmission architecture of Fig. 1 for simplicity.

Algorithm 1: The LSMR based TS-domain channel equalization

Input: $\mathbf{C}_{\rm MIMO}$; $\mathbf{r}_{\rm MIMO}$
Output: $\hat{\mathbf{s}}_{\rm MIMO}$
1: **Initialization:**
1) Calculating parameters:
$\beta_1\mathbf{u}_1 = \mathbf{r}_{\rm MIMO}$, $\alpha_1\mathbf{v}_1 = \mathbf{C}_{\rm MIMO}^T\mathbf{r}_{\rm MIMO}$, $\mathbf{h}_1 = \mathbf{v}_1$, $\bar{\mathbf{h}}_0 = \mathbf{0}$,
$\mathbf{x}_0 = \mathbf{0}$, $\bar{\alpha}_1 = \alpha_1$, $\bar{\zeta}_1 = \alpha_1\beta_1$, $\rho_0 = 1$, $\bar{\rho}_0 = 1$, $\bar{c}_0 = 1$, $\bar{s}_0 = 1$
2: **for** $i = 1 : I$ **do**
3: Bidiagonalization:
$\beta_{i+1}\mathbf{u}_{i+1} = \mathbf{C}_{\rm MIMO}\mathbf{v}_i - \alpha_i\mathbf{u}_i$, $\beta_{i+1} = \|\mathbf{u}_{i+1}\|$,
$\alpha_{i+1} = \|\mathbf{v}_{i+1}\|$, $\alpha_{i+1}\mathbf{v}_{i+1} = \mathbf{C}_{\rm MIMO}^T\mathbf{u}_{i+1} - \beta_{i+1}\mathbf{v}_i$
4: Solving the LS equations in Krylov subspace:
$\rho_i = \sqrt{\bar{\alpha}_i^2 + \beta_{i+1}^2}$, $c_i = \bar{\alpha}_i/\rho_i$, $s_i = \beta_{i+1}/\rho_i$,
$\theta_{i+1} = s_i\alpha_{i+1}$, $\bar{\alpha}_{i+1} = c_i\alpha_{i+1}$, $\bar{\theta}_i = \bar{s}_{i-1}\rho_i$,
$\bar{\rho}_i = \sqrt{\left(\bar{c}_{i-1}\rho_i\right)^2 + \theta_{i+1}^2}$, $\bar{c}_i = \bar{c}_{i-1}\rho_i/\bar{\rho}_i$,
$\bar{s}_i = \theta_{i+1}/\rho_i$, $\zeta_i = \bar{c}_i\bar{\zeta}_i$, $\bar{\zeta}_{i+1} = -\bar{s}_i\bar{\zeta}_i$
5: Update $\mathbf{s}_{\rm MIMO}^i$:
$\mathbf{s}_{\rm MIMO}^i = \mathbf{s}_{\rm MIMO}^{i-1} + \left(\zeta_i/(\rho_i\bar{\rho}_i)\right)\bar{\mathbf{h}}_i$,
$\bar{\mathbf{h}}_i = \mathbf{h}_i - \left(\bar{\theta}_i\rho_i/(\rho_{i-1}\bar{\rho}_{i-1})\right)\bar{\mathbf{h}}_{i-1}$,
$\mathbf{h}_{i+1} = \mathbf{v}_{i+1} - \left(\theta_{i+1}/\rho_i\right)\mathbf{h}_i$,
$\mathbf{e}_i = \mathbf{r}_{\rm MIMO} - \mathbf{C}_{\rm MIMO}\mathbf{s}_{\rm MIMO}^i$
6: **end for**

However, we still retain the procedures since they are indispensable parts for the proposed EDD in the next section.

IV. Iteratively Enhanced Data Demodulation

A conventional demodulation strategy is to remap the equalized OTFS symbols into data bits by using hard or soft decision. However, in MIMO-OTFS system, such a "direct decision", comes with inferior demodulation performance. On the one hand, in the DDS domain (i.e., the OTFS symbols are exactly demodulated in this domain), the interference generates from multiple dimensions, i.e., space-domain IAI, delay-domain ISI, and Doppler-domain IDI. As shown in Fig. 3, the spatial-domain IAI can be regarded as the transmitted i-th signal stream undergoes channel j, i, where $i \neq j$, and received by antenna j. For the delay-domain ISI, it is induced by the sub-channel's multi-path propagation in the single-stream OTFS transmission. From Fig. 3, the ISI is contributed by the submatrices $\mathbf{G}_{j,i}^{l,0} \in \mathbb{C}^{N\times N}, l \neq 0$. Without interference cancellation, the demodulated data bits may not be reliable. In addition, there are significantly rich channel diversity gains in the MIMO-OTFS system, e.g., spatial and multi-path diversities. Nevertheless, the linear equalization schemes may not be capable of reaping these channel diversities, also yielding a poor demodulation performance.

Instead of using the "direct decision" above, we propose an enhanced data demodulation scheme in this section, to improve the performance of the MIMO-OTFS receiver. Our basic idea can be concluded as: 1) as shown in Fig. 1, we preserve both the equalized OTFS symbols $\hat{\mathbf{x}}_{\rm MIMO}$ and re-

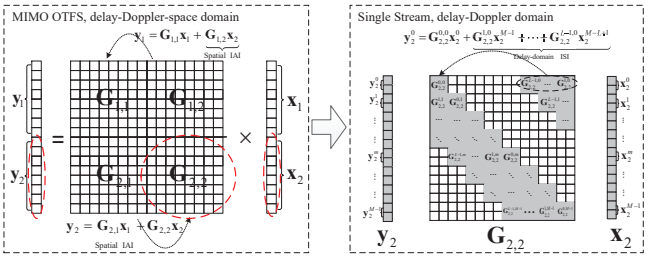


Fig. 3. Demonstration of multi-dimension interference in MIMO-OTFS system.

ceived data samples $\mathbf{y}_{\mathrm{MIMO}}$, and feed them into the enhanced iterative data detector; 2) by leveraging the equalized OTFS symbols of each transmit stream, the interference (i.e., IAI and ISI) can be reconstructed, and further removed from observation vector $\mathbf{y}_{\mathrm{MIMO}}$; 3) after interference cancellation, the "pure" signal components are accumulated together to resolve the transmitted OTFS symbols.

A. The Proposed Iterative Data Detector

The proposed EDD in the MIMO-OTFS receiver of Fig. 1, is detailed in Fig. 4, where the inputs come from the equalized OTFS symbols $\hat{\mathbf{x}}^{\mathrm{Ini}}_{\mathrm{MIMO}}$, received data samples $\mathbf{y}_{\mathrm{MIMO}}$ and DDS-domain channel matrix $\mathbf{G}_{\mathrm{MIMO}}$. As illustrated in Fig. 4, the EDD mainly consists of two-layer operations. The first layer is realized in the space dimension. In the k-th iteration, the IAI corresponding to transmit signal stream i at receive antenna antenna j is reconstructed by the estimated QAM symbols and channel matrix, according to $\sum_{i'=1,\neq i}^{N_{\mathrm{T}}} \mathbf{G}_{j,i}\hat{\mathbf{x}}_{i'}$. Then, the observation vector at receive antenna j, i.e., $\mathbf{y}_j$, will be updated through subtracting the resulting IAI: $\bar{\mathbf{y}}_{j,i} = \mathbf{y}_j - \sum_{i'=1,\neq i}^{N_T} \mathbf{G}_{j,i'}\hat{\mathbf{x}}_{i'}$, where $\bar{\mathbf{y}}_{j,i}$ can be treated as the "pure" signal of transmit stream i spreading at receive antenna j. After IAI elimination, the EGC is invoked to combine these "pure" signal components, which generates the accumulated signal vector $\bar{\mathbf{y}}_{\Sigma,i}$ and channel matrix $\mathbf{G}_{\Sigma,i}$. As a cascade of the IAI cancellation and EGC, the second layer is performed in the delay domain, involving the ISI cancellation and MRC. Similar to the first layer, the MRC scheme is adopted to combine the ISI-eliminated signal components, so as to collect the channel multi-path diversity.

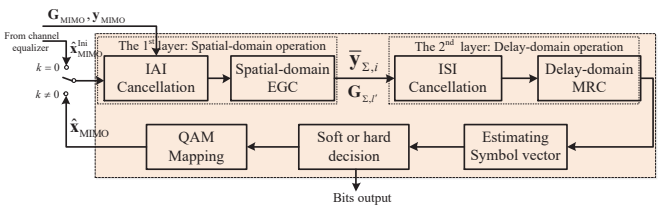


Fig. 4. The architecture of the iteratively enhanced data detector (EDD).

Specifically, the iteration process is summarized in Algorithm 2. In the initial step, the observation vector, initial data estimates from channel equalizer and channel matrix are denoted as $\mathbf{y}_{\mathrm{MIMO}}$, $\hat{\mathbf{x}}^{\mathrm{Ini}}_{\mathrm{MIMO}}$, and $\mathbf{G}_{\mathrm{MIMO}}$, respectively. The iteration process is described in following part.

Algorithm 2: Iteration procedure of the proposed data detector

Input: Channel matrix: $\mathbf{G}_{\mathrm{MIMO}}$; Observations: $\mathbf{y}_{\mathrm{MIMO}}$; Initial data estimates from channel equalizer: $\hat{\mathbf{x}}^{\mathrm{Ini}}_{\mathrm{MIMO}}$
Output: Estimated data bits
1: **for** Iteration times = 1 : K **do**
2: **IAI cancellation:** $\bar{\mathbf{y}}_{j,i} = \mathbf{y}_j - \sum_{i'=1,\neq i}^{N_T} \mathbf{G}_{j,i'}\hat{\mathbf{x}}_{i'}$.
3: **Spatial-domain EGC:** $\bar{\mathbf{y}}_{\Sigma,i} \triangleq \sum_{j=1}^{N_R} \bar{\mathbf{y}}_{j,i}$, $\mathbf{G}_{\Sigma,i} \triangleq \sum_{j=1}^{N_R} \mathbf{G}_{j,i}$.
4: **ISI cancellation:** use $\bar{\mathbf{y}}_{\Sigma,i}$ and $\mathbf{G}_{\Sigma,i'}$ to perform ISI cancellation according to (37).
5: **Delay-domain MRC** by (41).
6: Estimating data symbols by (42).
7: Soft or hard decision on symbol estimates.
8: QAM mapping.
9: **end for**

IAI cancellation: The OTFS symbol estimates, which come from the channel equalizer in Fig. 1 (in the initial iteration) or the QAM mapper in Fig. 4 (after the previous iteration), are denoted as $\hat{\mathbf{x}}_{\mathrm{MIMO}} = \left[\hat{\mathbf{x}}_1^T, \cdots, \hat{\mathbf{x}}_j^T, \cdots, \hat{\mathbf{x}}_{N_{\mathrm{T}}}^T\right]^T \in \mathbb{C}^{MNN_{\mathrm{T}}\times 1}$. At each receive antenna, we calculate the "pure" signal component from transmit antenna i as

$$\bar{\mathbf{y}}_{j,i} = \mathbf{y}_j - \sum_{i'=1,\neq i}^{N_T} \mathbf{G}_{j,i'}\hat{\mathbf{x}}_{i'}, \tag{30}$$

which also denotes the IAI cancellation. In (30), $\mathbf{y}_j$ is the observation vector at receive antenna j, $\hat{\mathbf{x}}_{i'}$ denotes the estimated symbol vector of transmit stream i', $\mathbf{G}_{j,i'}$ is the sub-channel matrix between transmit stream i' and receive antenna j, and $\bar{\mathbf{y}}_{j,i}$ refers to the spatial-impaired signal component of $\hat{\mathbf{x}}_i$ in receive antenna j after removing the spatial interference generated by other streams $\hat{\mathbf{x}}_{i'}, i' \neq i$.

Based on (5), the received DD-domain data vector at antenna j follows

$$\mathbf{y}_j = \sum_{i'=1,}^{N_T} \mathbf{G}_{j,i'}\mathbf{x}_{i'} + \mathbf{n}''_j, \tag{31}$$

where $\mathbf{n}''_j$ denotes the noise vector. Combing (30) and (31), we have

$$\bar{\mathbf{y}}_{j,i} = \mathbf{e}_j + \mathbf{G}_{j,i}\hat{\mathbf{x}}_i + \mathbf{n}_j, \tag{32}$$

where $\mathbf{e}_j$ denotes the accumulated error vector at receive antenna j, i.e.,

$$\mathbf{e}_j \triangleq \sum_{i'=1}^{N_T} \mathbf{G}_{j,i'}\left(\mathbf{x}_{i'} - \hat{\mathbf{x}}_{i'}\right). \tag{33}$$

At a relatively high signal-to-noise ratio (SNR) regime, such as SNR $\geq$ 20 dB, the entries in $\mathbf{e}_j$ gradually approach to zero as the number of iterations increases [4], which will be shown in Sec. V.

[4]However, there would be a particular phenomenon that $\mathbf{e}_j$ is always with large values no matter how many iterations are performed. It is known as error propagation, caused by the initial data estimation error from the channel equalizer. We will elaborate the issue in Sec. V.

Spatial-domain EGC: After IAI cancellation, we can obtain the "pure" signal components of transmit stream i spreading at receive antenna j, i.e., $\bar{\mathbf{y}}_{j,i}$. These "pure" signal components enable us to using EGC technique to equally combine the multi-antenna signal, so as to reap the channel spatial diversity. Relying on (32), for all receive antennas $j = 1, \cdots N_{\mathrm{R}}$, we have following equations corresponding to $\hat{\mathbf{x}}_i$, i.e.,

$$\begin{cases} \bar{\mathbf{y}}_{1,i} = \mathbf{G}_{1,i}\hat{\mathbf{x}}_i + \mathbf{e}_1 + \mathbf{n}_1, \\ \vdots \\ \bar{\mathbf{y}}_{N_{\mathrm{R}},i} = \mathbf{G}_{N_{\mathrm{R}},i}\hat{\mathbf{x}}_i + \mathbf{e}_{N_{\mathrm{R}}} + \mathbf{n}_{N_{\mathrm{R}}}. \end{cases} \tag{34}$$

The spatial-domain EGC is performed by equally combing these N_{R} equations in (34), i.e.,

$$\bar{\mathbf{y}}_{\Sigma,i} = \mathbf{G}_{\Sigma,i}\hat{\mathbf{x}}_i + \mathbf{e}_{\Sigma} + \mathbf{n}_{\Sigma}, \tag{35}$$

where $\bar{\mathbf{y}}_{\Sigma,i}$ denotes the accumulated signal at whole receive antennas $\bar{\mathbf{y}}_{\Sigma,i} \triangleq \sum_{j=1}^{N_R} \bar{\mathbf{y}}_{j,i}$, $\mathbf{G}_{\Sigma,i}$ is the spatial-gathered channel matrix, i.e., $\mathbf{G}_{\Sigma,i} \triangleq \sum_{j=1}^{N_R} \mathbf{G}_{j,i}$. In (35), $\mathbf{e}_{\Sigma}$ and $\mathbf{n}_{\Sigma}$ take the same form of $\bar{\mathbf{y}}_{\Sigma,i}$, which are denoted as the accumulated error, and noise at all receive antennas, respectively. From (34) to (35), the whole received "pure" signal components corresponding to the i-th transmit stream are aggregated. Accordingly, the channel spatial diversity is efficiently extracted.

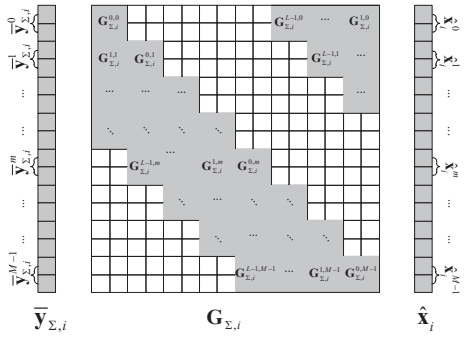


Fig. 5. The structure of $\mathbf{G}_{\Sigma,i}$, for the particular case $L = 4$.

So far, the spatial-domain operations of EDD, i.e., IAI cancellation and EGC in Fig. 4, have been performed in the first layer. Based on (35), an LS method can be used to acquire (or named as update) $\hat{\mathbf{x}}_i$. Nevertheless, we do not propose to do that. This is because: 1) it will introduce a highly computational cost due to the matrix inversion of an $MN \times MN$ matrix $\mathbf{G}_{\Sigma,i}$; 2) for each stream, the delay-domain ISI still exists. In contrast, the structure of matrix $\mathbf{G}_{\Sigma,i}$, illustrated in Fig. 5, enlightens us to decompose (35) as a set of sub-equations. These sub-equations can exactly represent the delay-and-Doppler domain transmission of each OTFS stream. Thus, by using these sub-equations, we can further design an ISI cancellation scheme, as well as the delay-domain MRC, to collect the channel multi-path diversity.

ISI cancellation: As shown in Fig. 4, the second layer is implemented in the delay domain, which contains the ISI cancellation and MRC operation. Similar to layer one, we firstly reconstruct the ISI, which corresponds to the transmitted symbols at a certain delay bin, by using the sub-matrices of $\mathbf{G}_{\Sigma,i}$ and sub-vectors of $\hat{\mathbf{x}}_i$. After subtracting the resulting ISI from observation vector $\bar{\mathbf{y}}_{\Sigma,i}$, the MRC technique is further employed to combine the ISI-eliminated signal. The MRC scheme can equivalently improve the detection SNR of the OTFS signal. As a result, it extracts the channel multi-path diversity. Specifically, we decompose the accumulated signal vector $\bar{\mathbf{y}}_{\Sigma,i}$ in (35) into M subvectors as

$$\bar{\mathbf{y}}_{\Sigma,i} = \left[\left(\bar{\mathbf{y}}^0_{\Sigma,i}\right)^T, \cdots, \left(\bar{\mathbf{y}}^m_{\Sigma,i}\right)^T, \cdots, \left(\bar{\mathbf{y}}^{M-1}_{\Sigma,i}\right)^T \right]^T.$$

The m-th subvector $\bar{\mathbf{y}}^m_{\Sigma,i} \in \mathbb{C}^{N\times 1}$ contains all Doppler-direction data symbols corresponding to the m-th delay bin, which follows

$$\bar{\mathbf{y}}^m_{\Sigma,i} = \mathbf{e}^m_{\Sigma} + \mathbf{n}^m_{\Sigma} + \mathbf{G}^{0,m}_{\Sigma,i}\hat{\mathbf{x}}^m_i + \sum_{l'=1}^{L-1} \mathbf{G}^{l',m}_{\Sigma,i}\hat{\mathbf{x}}^{(m-l')_M}_i, \tag{36}$$

where $\hat{\mathbf{x}}^m_i$ collects all estimated Doppler-direction data symbols corresponding to the m-th delay bin, i.e., $\hat{\mathbf{x}}^m_i = \left[\hat{X}_i[m,0], \cdots, \hat{X}_i[m,n], \cdots, \hat{X}_i[m,N-1]\right]^T$, and $\mathbf{G}^{l',m}_{\Sigma,i}$ denotes the sub-channel matrix in $\mathbf{G}_{\Sigma,i}$, which can be found in Fig. 5. Parameters $\mathbf{e}^m_{\Sigma}$ and $\mathbf{n}^m_{\Sigma}$ refer to error and noise sub-vectors, respectively, which take the same form of $\bar{\mathbf{y}}^m_{\Sigma,i}$. In the right-hand side of (36), the fourth term denotes the ISI components which corresponds to the m-th sub-block $\hat{\mathbf{x}}^m_i$. Accordingly, the ISI cancellation is performed by

$$\mathbf{G}^{0,m}_{\Sigma,i}\hat{\mathbf{x}}^m_i + \mathbf{e}^m_{\Sigma} + \mathbf{n}^m_{\Sigma} = \bar{\mathbf{y}}^m_{\Sigma,i} - \sum_{l'=1}^{L-1} \mathbf{G}^{l',m}_{\Sigma,i}\hat{\mathbf{x}}^{(m-l')_M}_i, \tag{37}$$

which can be rewritten as

$$\mathbf{G}^{0,m}_{\Sigma,i}\hat{\mathbf{x}}^m_i = \tilde{\mathbf{y}}^m_{\Sigma,i}. \tag{38}$$

In (38), $\tilde{\mathbf{y}}^m_{\Sigma,i}$ is

$$\tilde{\mathbf{y}}^m_{\Sigma,i} \triangleq \bar{\mathbf{y}}^m_{\Sigma,i} - \sum_{l'=1}^{L-1} \mathbf{G}^{l',m}_{\Sigma,i}\hat{\mathbf{x}}^{(m-l')_M}_i - \mathbf{e}^m_{\Sigma} - \mathbf{n}^m_{\Sigma}, \tag{39}$$

which can be treated as the m-th sub-block of the accumulated "pure" signal, after removing the ISI.

Delay-domain MRC: After ISI cancellation, shown in Fig. 1, we can obtain the "pure" observation signal components of the m-th sub-block, i.e., $\tilde{\mathbf{y}}^m_{\Sigma,i}$ in (39). As the maximum channel delay spread is L, these ISI-eliminated signal can be gathered from whole L propagation paths. Specifically, from (38) and (39), we can derive following L equations corresponding to $\hat{\mathbf{x}}^m_i$:

$$\begin{cases} \tilde{\mathbf{y}}^m_{\Sigma,i} = \mathbf{G}^{0,m}_{\Sigma,i}\hat{\mathbf{x}}^m_i, \\ \tilde{\mathbf{y}}^{(m+1)_M}_{\Sigma,i} = \mathbf{G}^{1,m}_{\Sigma,i}\hat{\mathbf{x}}^m_i, \\ \vdots \\ \tilde{\mathbf{y}}^{(m+L-1)_M}_{\Sigma,i} = \mathbf{G}^{L-1,m}_{\Sigma,i}\hat{\mathbf{x}}^m_i. \end{cases} \tag{40}$$

By leveraging the MRC technique, we combine these L equations in (40), to detect the OTFS symbol vector $\hat{\mathbf{x}}^m_i$. Specifically, we have

$$\mathbf{y}'^m_{\Sigma,i} = \bar{\mathbf{G}}^m_{\Sigma,i}\hat{\mathbf{x}}^m_i, \tag{41}$$

where

$$\mathbf{y'}_{\Sigma,i}^{m} \triangleq \left(\sum_{l'=0}^{L-1}\left(\mathbf{G}_{\Sigma,i}^{l',m}\right)^{H}\tilde{\mathbf{y}}_{\Sigma,i}^{(m+l')_M}\right),$$

$$\bar{\mathbf{G}}_{\Sigma,i}^{m} \triangleq \left(\sum_{l'=0}^{L-1}\left(\mathbf{G}_{\Sigma,i}^{l',m}\right)^{H}\mathbf{G}_{\Sigma,i}^{l',m}\right).$$

The right-hand side of (41) combines the whole signal branches of $\hat{\mathbf{x}}_i^m$ that spreads in the delay dimension, or rather, collects the submatrices in the column direction of $\mathbf{G}_{\Sigma,i'}$. By doing so, the accumulated signal components enable to extract the multi-path diversity, meanwhile improve the signal power equivalently. According to (41), $\hat{\mathbf{x}}_i^m$ is resolved by

$$\hat{\mathbf{x}}_i^m = \left(\bar{\mathbf{G}}_{\Sigma,i}^{m}\right)^{-1}\mathbf{y'}_{\Sigma,i}^{m}. \tag{42}$$

Remark 3: The proposed EDD is a layer-by-layer scheme. The serial-cascaded modules in Fig. 4, i.e., IAI-cancellation, EGC, ISI-cancellation, and MRC, are capable of realizing multi-dimensional interference elimination and channel diversity extraction. Also, the EDD in the MIMO-OTFS system, can independently cooperate with arbitrary channel equalizer, as the EDD only requires the initial equalized OTFS symbols to start.

Remark 4: In practice, the initial symbol estimates, obtained by the channel equalizer, are crucial to perform the iterative EDD. On the one hand, the interference in the MIMO-OTFS system is significant. With imprecise initial estimates, the residual error in (33) would be considerably large, which yields the error propagation and possibly results in a block error rate (BLER) platform. The FS-domain "block banded equalizer" [40] can work efficiently as long as the channel's time variations are not significant across the duration of one OFDM symbol period. However, since the OTFS signal always undergoes the highly doubly-dispersive channels, the "block banded equalizer" could not be effective due to the strong ICI. In contrast, our TS-domain LSMR channel equalizer can not only avoid the above drawbacks, but also help to reduce the number of iterations in the EDD, which will be shown in Sec. V.

B. Practical Implementation and Computational Complexity

For the practical implementation, it is not necessary to calculate each step in the EDD. For the spatial-domain operations, i.e., IAI cancellation and EGC in Fig. 4, the key outputs are $\bar{\mathbf{y}}_{\Sigma,i}$ and $\mathbf{G}_{\Sigma,i}$. 1) Calculating $\mathbf{G}_{\Sigma,i}$: matrix $\mathbf{G}_{\Sigma,i}$ is the accumulated channel matrix at all receive antennas corresponding to transmit antenna i. It requires N^2MLN_R complex additions (CAs) for each transmit stream. For the whole N_T transmit streams, the total complexity in terms of CAs is $N^2N_T^2ML$. 2) Calculating $\bar{\mathbf{y}}_{\Sigma,i}$: From (31) and (32), a much simpler method can be used to acquire $\bar{\mathbf{y}}_{\Sigma,i}$, i.e., by defining an error vector $\boldsymbol{\varepsilon}_{\mathrm{MIMO}} = \left[\boldsymbol{\varepsilon}_1^T, \cdots, \boldsymbol{\varepsilon}_{N_R}^T\right]^T$ as

$$\boldsymbol{\varepsilon}_{\mathrm{MIMO}} = \mathbf{y}_{\mathrm{MIMO}} - \mathbf{G}_{\mathrm{MIMO}}\hat{\mathbf{x}}_{\mathrm{MIMO}}. \tag{43}$$

In this way, (32) can be rewritten as $\bar{\mathbf{y}}_{j,i} = \boldsymbol{\varepsilon}_j + \mathbf{G}_{j,i}\hat{\mathbf{x}}_i$. Further, (35) follows

$$\bar{\mathbf{y}}_{\Sigma,i} = \boldsymbol{\varepsilon}_{\Sigma} + \mathbf{G}_{\Sigma,i}\hat{\mathbf{x}}_i, \tag{44}$$

where

$$\boldsymbol{\varepsilon}_{\Sigma} \triangleq \sum\nolimits_{j=1}^{N_R}\boldsymbol{\varepsilon}_j. \tag{45}$$

In the right-hand side of (44), $\boldsymbol{\varepsilon}_{\Sigma}$ is obtained from (43) and (45), which needs $2MNN_R$ CAs and $N^2N_T^2ML$ complex multiplications (CMs). For the second term, i.e., $\mathbf{G}_{\Sigma,i}\hat{\mathbf{x}}_i$, the computational complexity has been counted in calculating (43).

To realize the delay-domain MRC based detection, it only requires following operations. 3) Obtaining $\tilde{\mathbf{y}}_{\Sigma,i}^m$: from (39), calculating $\tilde{\mathbf{y}}_{\Sigma,i}^m$ requires $(L+2)N$ CAs and N^2L CMs. 4) Updating $\hat{\mathbf{x}}_i^m$: there is no need to directly use matrix inversion technique in (42) to achieve $\hat{\mathbf{x}}_i^m$, since $\bar{\mathbf{G}}_{\Sigma,i}^m$ in (42) is a circulant matrix. Thus, the FFT can be used in the practical implementation. Specifically, $\bar{\mathbf{G}}_{\Sigma,i}^m$ can be decomposed as

$$\bar{\mathbf{G}}_{\Sigma,i}^{m} = \mathbf{F}_N\boldsymbol{\Lambda}_{\Sigma,i}^{m}\mathbf{F}_N^H, \tag{46}$$

where $\mathbf{F}_N$ demotes $N\times N$ FFT matrix, and $\boldsymbol{\Lambda}_{\Sigma,i}^m$ is a diagonal matrix. The diagonal elements can be achieved by applying the FFT to the first row of $\bar{\mathbf{G}}_{\Sigma,i}^m$. As a result, (42) can be rewritten as:

$$\hat{\mathbf{x}}_i^m = \mathbf{F}_N\left(\boldsymbol{\Lambda}_{\Sigma,i}^{m}\right)^{-1}\mathbf{F}_N^H\mathbf{y'}_{\Sigma,i}^{m}. \tag{47}$$

Based on the above derivations, updating $\hat{\mathbf{x}}_i^m$ can be simplified as follows:

Step 1: Applying IFFT to $\mathbf{y'}_{\Sigma,i}^m$, which requires the complexity order of $\mathcal{O}\left(N\log\left(N\right)\right)$.

Step 2: Calculation the diagonal elements in $\boldsymbol{\Lambda}_{\Sigma,i}^m$, which requires the complexity order of $\mathcal{O}\left(N\log\left(N\right)\right)$.

Step 3: Element-wise division in the operation $\left(\boldsymbol{\Lambda}_{\Sigma,i}^{m}\right)^{-1}\mathbf{F}_N^H\mathbf{y'}_{\Sigma,i}^{m}$, which requires N times complex divisions.

Step 4: Using FFT to the results in step 3, which requires the complexity order of $\mathcal{O}\left(N\log\left(N\right)\right)$.

Overall, updating $\hat{\mathbf{x}}_i^m$ requires the complexity order of $\mathcal{O}\left(N\log\left(N\right)\right)$. According to the above analysis and the derivations, implementing the proposed EDD requires the complexity order of $\mathcal{O}\left(N^2N_T^2MLK\right)$.

V. Simulations and Discussions

In this section, the demodulation performance is evaluated for the MIMO-OTFS system, of which the parameters are set according to Table I. We assume each channel tap follows a Rayleigh distribution with an exponentially decaying power delay profile [34], and no spatial correlation between antennas. In all simulations, we consider UE's velocity of 550 km/h at 5.9 GHz carrier frequency, corresponding to a maximum channel Doppler frequency f_d of 3000 Hz. To avoid the detection performance loss introduced by channel estimation errors, the channel responses are assumed as perfectly known at the receiver side.

A. Channel Equalization Performance

We firstly investigate the channel equalization performance by evaluating the MSE of the equalized OTFS symbols, i.e.,

$$\mathrm{MSE} = \mathbb{E}\left\{\left\|\hat{\mathbf{x}}_{\mathrm{MIMO}}^{\mathrm{Ini}} - \mathbf{x}_{\mathrm{MIMO}}\right\|^2\right\}. \tag{48}$$

TABLE I
SYSTEM SIMULATION PARAMETERS

Parameter	Value
Num. of subcarriers (M)	32
Num. of symbols (N)	32
Num. of channel taps (L)	5
Transmission mode	2×2 MIMO, 3×4 MIMO
Carrier frequency (f_c)	5.9 GHz
Subcarrier spacing (Δf)	15 KHz
Maximum user's velocity	550 km/h
Modulation scheme	QPSK, 16-QAM, 64-QAM, 256-QAM
Channel encoder	$[1, 5/7]_8$ convolutional code
Channel coding rate	1/2
Channel Doppler spectrum	"U-shaped" spectrum [31]
Num. of simulated OTFS blocks	1×10^5

The proposed TS-domain equalizer is compared with the widely-used frequency-domain "block banded equalizer" [34], which provides a good trade off between equalization performance and calculational complexity by selecting the main diagonal, the Q superdiagonals and Q subdiagonals of each frequency-domain channel matrix.

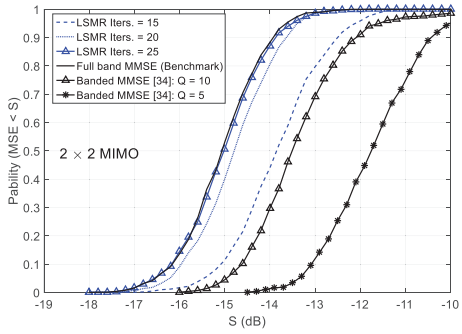


Fig. 6. The CDF of the MSE in 2×2 MIMO-OTFS case, where SNR = 20 dB.

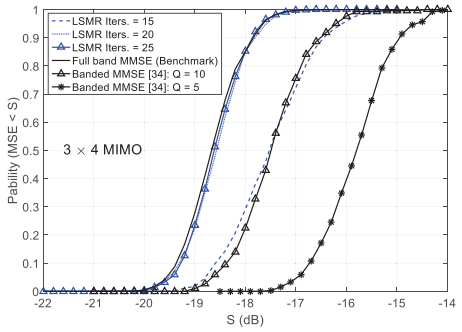


Fig. 7. The CDF of the MSE in 3×4 MIMO case, where SNR = 20 dB. The maximum UE's velocity is 550 km/h.

Figs. 6 and 7 demonstrate the cumulative density function (CDF) of the MSE in 2×2 MIMO and 3×4 MIMO transmission cases, respectively, where SNR = 20 dB, and label S(dB) in X-axis denotes the threshold value. The CDF curve "full band LMMSE", which is achieved by using the full-banded LMMSE method in (16), is plotted as a benchmark. As we can see, it requires 25 and 20 iterations for the LSMR, respectively in 2×2 and 3×4 MIMO transmission, to achieve the comparable equalization performance to the full-banded LMMSE. Also, our scheme significantly outperforms the frequency-domain "block banded equalizer" [34].

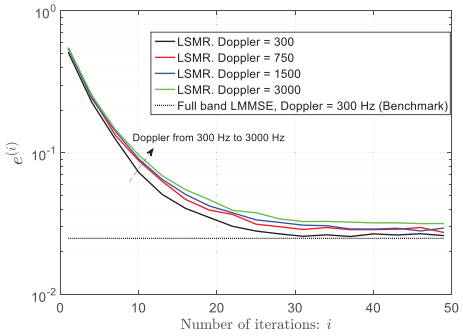


Fig. 8. The comparison of the equalization error $e^{(i)}$ with the number of iterations i, where SNR = 20 dB.

To evaluate the convergence speed of LSMR, we have compared $e^{(i)}$ versus number of iterations i in Fig. 8. Note that

$$e^{(i)} = \mathbb{E}\left\{\left\|\hat{\mathbf{s}}_{\mathrm{MIMO}}^{(i)} - \mathbf{s}_{\mathrm{MIMO}}\right\|^2\right\}$$

denotes the estimation error after the i-th iteration, where $\hat{\mathbf{s}}_{\mathrm{MIMO}}^{(i)}$ represents the estimated symbol vector after the i-th iteration. In Fig. 8, the black dash line denotes the estimation error corresponding to the full band LMMSE algorithm, which is regarded as a threshold to evaluate the convergence of the LSMR. As we can see, after 25 iterations, the value of $e^{(i)}$ gradually approaches to a stationary level for all considered Doppler spread cases. Also, the smaller channel Doppler spread is, the quicker convergence rate can be obtained. Finally, even if the channel matrix may be ill-conditioned when the Doppler is 3000 Hz [45], the LSMR is still capable of acquiring the relatively accurate and robust numerical solutions.

B. Demodulation Performance

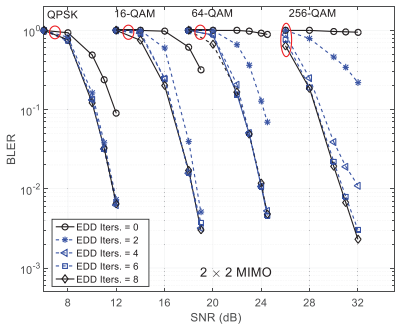


Fig. 9. The BLER comparison of SNR with different iterations in the EDD.

Fig. 9 compares the BLER versus SNR with different iteration times in the proposed EDD, where f_d is 3000 Hz. The receiver's architecture is constructed by the LSMR-based channel equalizer and the EDD[5], and the transmission mode

[5]The number of iterations in LSMR is chosen as 20, to achieve a tradeoff between the computational complexity and equalization performance.

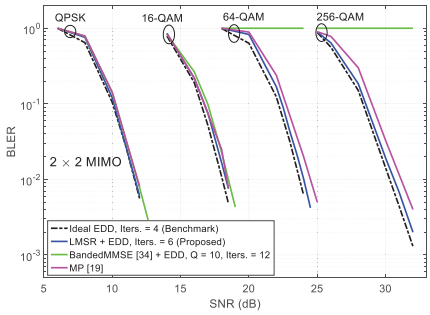


Fig. 10. The BLER versus SNR in 2 × 2 MIMO transmission case.

is 2×2 MIMO. When Iters. = 0, the EDD is not invoked. The receiver only adopting the LSMR-based channel equalizer can not demodulate the high-order QAM signaling. By leveraging the EDD scheme, the BLER platform decreases quickly with few iterations. For all QAM constellations, a solid BLER performance can be achieved after six iterations.

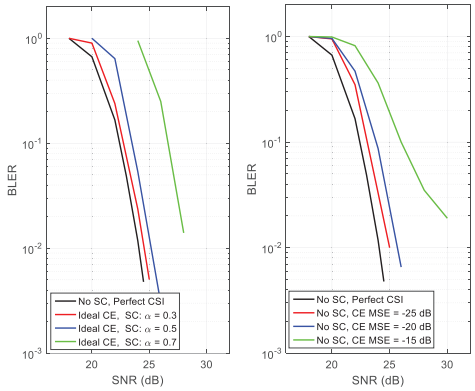
Fig. 10 demonstrate the BLER comparison of SNR with different demodulation schemes in 2×2 MIMO transmission case. We plot the BLER curve "Ideal EDD", which is obtained by feeding the perfect initial symbol estimates into the EDD, to verify the demodulation gap involved by the channel equalization error. It can be seen from Fig. 10 that, the SNR loss caused by our scheme is negligible in contrast to scheme "Ideal EDD". Additionally, the other state-of-the-art detection methods, i.e., the MP based symbol detection [19] and the frequency-domain "block banded equalizer" with EDD, denoted as "BandedMMSE [34] + EDD", are compared. For the MP, the maximum iterations are 10, 30, 50 and 100, respectively for QPSK, 16-QAM, 64-QAM, and 256-QAM signaling. For low-order QAM signaling, the receivers adopting these two detectors can realize the comparable demodulation performance to ours. For high-order QAM signaling, the proposed receiver achieves 1 dB SNR gain over the MP-based detector at the BLER level of 10^{-2}. In contrast, the receiver "BandedMMSE [34] + EDD, $Q = 10$, Iters. = 12" fails to retrieve the high-order QAM symbols due to the intolerable channel equalization error.

Fig. 11(a) demonstrates the BLER comparison of SNR with different spatial correlation (SC) coefficients, where α denotes the SC coefficient for transmitter and receiver antennas. As we can see, the proposed detector is robust to the small or moderate SC degree, e.g., $\alpha = 0.3/0.5$. Even with moderate SC degree, e.g., $\alpha = 0.5$, the SNR gap between cases "No SC, Perfect SCI" and "Ideal CE, SC: $\alpha = 0.5$" is about 1 dB, at BLER = 10^{-2}. Additionally, when the antenna's correlation is relatively strong, i.e., $\alpha = 0.7$, there is a relatively large SNR loss compared with ideal case.

Fig. 11(b) illustrates the influence of imperfect channel stae information to the demodulation performance. Here, we introduce error n_e in the channel estimation (CE) using the model

$$h'_{j,i}[k,l] = h_{j,i}[k,l] + n_e.$$

The MSE of CE, i.e., the variance of n_e is considered as –25, –20 or –15 dB. When MSE $\leq$ –20 dB, the demodulation SNR loss introduced by CE error is less than 1.5 dB, at the BLER level of 10^{-2}. When the CE error is significant, e.g., MSE = – 15 dB, more transmitting power will be consumed to achieve a satisfied demodulation performance.



(a) The BLER versus SNR with different spatial correlation coefficients. (b) The BLER versus SNR with different channel estimation error.

Fig. 11. The demodulation performance with non-ideal issues.

C. Computational Complexity

We conclude the computational load of different methods in Table II. According to [10], [19], implementing the MP requires the complexity order of $\mathcal{O}\left(N^2N_T^2MLAI_{MP}\right)$. In the proposed receiver architecture (LSMR + EDD), the complexity orders for LSMR and EDD are $\mathcal{O}\left(N_T^2MNLI\right)$ and $\mathcal{O}\left(N^2N_T^2MLK\right)$, respectively. Thus, the overall complexity order is $\mathcal{O}\left(N^2N_T^2MLK\right)$. In the same way, the overall complexity orders for "Full-band MMSE + EDD" and "Banded-MMSE + EDD" are $\mathcal{O}\left(M^3N^3N_T^3\right)$ and $\mathcal{O}\left(MN^2N_T^3Q^3\right)$, respectively. From Table 1, as the size of modulation alphabet A increases, the implementation complexity of MP increases significantly. In our proposal, i.e., LSMR + EDD, the implementation complexity is independent of the size of modulation alphabet, thus achieving a constant computational

TABLE II
THE COMPARISON OF COMPUTATIONAL COMPLEXITY FOR DIFFERENT DETECTION METHODS

Equalization	Detection	Overall Complexity	2 × 2 MIMO OTFS			
			QPSK	16-QAM	64-QAM	256-QAM
Full-band MMSE	EDD	$\mathcal{O}\left(M^3N^3N_T^3\right)$	$\mathcal{O}\left(8\times10^9\right)$			
Banded MMSE [34]		$\mathcal{O}\left(MN^2N_T^3Q^3\right)$	$\mathcal{O}\left(2\times10^8\right)$			
LSMR		$\mathcal{O}\left(N^2N_T^2MLK\right)$	$\mathcal{O}\left(4\times10^6\right)$			
MP [19]		$\mathcal{O}\left(N^2N_T^2MLAI_{MP}\right)$	$\mathcal{O}\left(2\times10^7\right)$	$\mathcal{O}\left(3\times10^8\right)$	$\mathcal{O}\left(2\times10^9\right)$	$\mathcal{O}\left(10^{10}\right)$

load for all considered QAM signaling. Also, as the number of iterations in EDD is quite smaller than that in the MP, i.e., $K \ll I_{MP}$, the overall computational load of the proposed architecture is much lower than the MP.

VI. Conclusions

In this paper, we studied the orthogonal time frequency space (OTFS) for multiple-input multiple-output (MIMO) system with the focus on channel equalization and symbol detection. Specifically, through analyzing the MIMO-OTFS transmission architecture, we found that the time-space (TS) domain channel matrix has a sparse structure. By leveraging the sparsity of channel matrix, a least squares minimum residual (LSMR) based TS-domain channel equalizer was proposed to remove the channel distortions on OTFS symbols. Based on the channel equalization, we developed an enhanced data detection (EDD) scheme to improve the demodulation performance of MIMO-OTFS receiver. The simulations demonstrated that the receiver adopting the proposed LSMR and EDD, can robustly demodulate the superposed high-order quadrature amplitude modulation (QAM) symbols, e.g., 256-QAM, in the case of multiple-stream OTFS signal transmission over high-mobility scenarios.

Appendix A
Proof of Propositions 1 and 2

At receive antenna j, the frequency-domain data vector $\mathbf{d'}_j$ is

$$\mathbf{d'}_j = \sum_{i=1}^{N_{\mathrm{T}}} \left(\mathbf{I}_N \otimes \mathbf{F}_M\right) \mathbf{C}_{j,i} \left(\mathbf{I}_N \otimes \mathbf{F}_M^H\right) \mathbf{d}_i, \tag{49}$$

where $\mathbf{d}_i = \mathrm{vec}\left(\mathbf{D}_i\right) \in \mathbb{C}^{MN\times 1}$ contains the frequency data symbols modulated at N OFDM symbols, each of which has M sucarriers. Defining the frequency-domain channel matrix between transmit antenna i and receive antenna j as $\mathbf{H}_{j,i} \triangleq \left(\mathbf{I}_N \otimes \mathbf{F}_M\right) \mathbf{C}_{j,i} \left(\mathbf{I}_N \otimes \mathbf{F}_M^H\right)$, (49) follows

$$\mathbf{d'}_j = \sum_{i=1}^{N_{\mathrm{T}}} \mathbf{H}_{i,j} \mathbf{d}_i. \tag{50}$$

The DD-domain data vector at receive antenna j is expressed as

$$\bar{\mathbf{y}}_j = \sum_{i=1}^{N_{\mathrm{T}}} \left(\mathbf{F}_N \otimes \mathbf{I}_M\right) \mathbf{C}_{j,i} \left(\mathbf{F}_N^H \otimes \mathbf{I}_M\right) \bar{\mathbf{x}}_i, \tag{51}$$

where $\bar{\mathbf{x}}_i \triangleq \mathrm{vec}\left(\mathbf{X}_i\right) \in \mathbb{C}^{MN\times 1}$ collects all QAM symbols modulated at the DD plane of the i-th stream. The format of (51) is not preferable in demonstrating the DD-domain transmissions. Thus, to better understand the OTFS transmission in this paper, we rewrite (51) into (5), by introducing a permutation matrix $\mathbf{\Pi} \in \mathbb{C}^{MN\times MN}$, i.e.,

$$\mathbf{\Pi} = \left[\begin{array}{ccc} \left(\mathbf{I}_N \otimes \mathbf{a}_M^T\left(0\right)\right)^T & \cdots & \left(\mathbf{I}_N \otimes \mathbf{a}_M^T\left(M-1\right)\right)^T \end{array} \right]^T, \tag{52}$$

where $\mathbf{a}_M\left(m\right)$ refers to the m-th column of $\mathbf{I}_M$, as well as defining $\mathbf{x}_i \triangleq \mathrm{vec}\left(\mathbf{X}_i{}^T\right) = \mathbf{\Pi}\mathrm{vec}\left(\mathbf{X}_i\right)$ and $\mathbf{y}_i \triangleq \mathrm{vec}\left(\mathbf{Y}_i{}^T\right) = \mathbf{\Pi}\mathrm{vec}\left(\mathbf{Y}_i\right)$. It follows $\mathbf{y}_j = \sum_{i=1}^{N_{\mathrm{T}}} \mathbf{G}_{j,i}\mathbf{x}_i$, where $\mathbf{G}_{j,i} \in \mathbb{C}^{MN\times MN}$ denotes the DD-domain transmission matrix between transmit antenna j and receive antenna i, i.e.,

$$\mathbf{G}_{j,i} = \mathbf{\Pi}\left(\mathbf{F}_N \otimes \mathbf{I}_M\right) \mathbf{C}_{j,i} \left(\mathbf{F}_N^H \otimes \mathbf{I}_M\right) \mathbf{\Pi}^H. \tag{53}$$

■

Appendix B
Proof of Proposition 3

The time-domain channel matrix $\mathbf{C}_{j,i}$ in (53), can be expressed as

$$\mathbf{C}_{j,i} = \sum_{l=0}^{L-1} \mathbf{B}_{MN}^l \mathrm{diag}\left(\mathbf{h}_{j,i}^l\right), \tag{54}$$

where $\mathbf{B}_{MN} \in \mathbb{C}^{MN\times MN}$ is a forward cyclic shift matrix [46]. For simplicity, we denote the CIR vector of the l-th path between transmit antenna i and receive antenna j as $\mathbf{h}_{j,i}^l = \left\{h_{j,i}\left[M_{\mathrm{cp}}, l\right], \cdots, h_{j,i}\left[M_{\mathrm{cp}}+MN-1, l\right]\right\}$. Inserting (54) into (53), we have (57). Define the (m', m)-th submatrix in (57) as $\mathbf{G}_{j,i}^{m',m} \in \mathbb{C}^{N\times N}; m', m \in [0, M-1]$, which follows

$$\mathbf{G}_{j,i}^{m',m} = \left(\mathbf{F}_N \otimes \mathbf{a}_M^T\left(m'\right)\right) \mathbf{C}_{j,i} \left(\mathbf{F}_N^H \otimes \mathbf{a}_M\left(m\right)\right). \tag{57}$$

Inserting (54) into (55), the submatrix $\mathbf{G}_{i,j}^{m',m}$ is rewritten as

$$\begin{aligned} \mathbf{G}_{j,i}^{m',m} &= \sum_{l=0}^{L-1} \left(\mathbf{F}_N \otimes \mathbf{a}_M^T\left(m'\right)\right) \mathbf{B}_{MN}^l \mathrm{diag}\left(\mathbf{h}_{j,i}^l\right) \left(\mathbf{F}_N^H \otimes \mathbf{a}_M\left(m\right)\right) \\ &= \sum_{l=0}^{L-1} \left(\mathbf{F}_N \otimes \mathbf{a}_M^T\left(\left(m'-l\right)_M\right)\right) \mathrm{diag}\left(\mathbf{h}_{j,i}^l\right) \left(\mathbf{F}_N^H \otimes \mathbf{a}_M\left(m\right)\right). \end{aligned} \tag{58}$$

Supposing the maximum channel delay spread $L < M$, we have the following conclusions [47]:

1) if and only if $\left(m'-l\right)_M = m$, $\mathbf{G}_{j,i}^{m',m} \neq \mathbf{0}$;

2) When $\left(m'-l\right)_M = m$, $\mathbf{G}_{j,i}^{m',m}$ is a circulant matrix with its first row in (59). ■

$$\mathbf{G}_{j,i} = \left[\begin{array}{ccccc} \left(\mathbf{F}_N \otimes \mathbf{a}_M^T\left(0\right)\right) \mathbf{C}_{j,i} \left(\mathbf{F}_N^H \otimes \mathbf{a}_M\left(0\right)\right) & \cdots & \cdots & \left(\mathbf{F}_N \otimes \mathbf{a}_M^T\left(0\right)\right) \mathbf{C}_{j,i} \left(\mathbf{F}_N^H \otimes \mathbf{a}_M\left(M-1\right)\right) \\ \left(\mathbf{F}_N \otimes \mathbf{a}_M^T\left(1\right)\right) \mathbf{C}_{j,i} \left(\mathbf{F}_N^H \otimes \mathbf{a}_M\left(0\right)\right) & \cdots & \cdots & \left(\mathbf{F}_N \otimes \mathbf{a}_M^T\left(1\right)\right) \mathbf{C}_{j,i} \left(\mathbf{F}_N^H \otimes \mathbf{a}_M\left(M-1\right)\right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\mathbf{F}_N \otimes \mathbf{a}_M^T\left(M-1\right)\right) \mathbf{C}_{j,i} \left(\mathbf{F}_N^H \otimes \mathbf{a}_M\left(0\right)\right) & \cdots & \cdots & \left(\mathbf{F}_N \otimes \mathbf{a}_M^T\left(M-1\right)\right) \mathbf{C}_{j,i} \left(\mathbf{F}_N^H \otimes \mathbf{a}_M\left(M-1\right)\right) \end{array} \right] \tag{57}$$

$$\mathbf{G}_{j,i}^{m',m}\left(0,:\right) = \frac{1}{\sqrt{N}} \mathbf{F}_N \left\{h_{j,i}\left[M_{\mathrm{cp}}+m', \left(m'-m\right)_M\right], \cdots, h_{j,i}\left[M_{\mathrm{cp}}+\left(N-1\right)M+m', \left(m'-m\right)_M\right]\right\}. \tag{59}$$

APPENDIX C
PROOF OF COROLLARY 1

Time-space domain: We begin with the TS-domain MIMO-OTFS transmission. According to (2) and (13), the MIMO transmission is formulated as

$$\mathbf{r}_{\mathrm{MIMO}} = \mathbf{C}_{\mathrm{MIMO}}\mathbf{s}_{\mathrm{MIMO}} + \boldsymbol{n}_{\mathrm{MIMO}}, \tag{69}$$

where $\mathbf{r}_{\mathrm{MIMO}} = \left[\mathbf{r}_1^T, \cdots, \mathbf{r}_{N_{\mathrm{R}}}^T\right]^T \in \mathbb{C}^{N_{\mathrm{R}}MN\times 1}$ is the TS-domain received data symbols at all antennas, $\mathbf{s}_{\mathrm{MIMO}} = \left[\mathbf{s}_1^T, \cdots, \mathbf{s}_{N_{\mathrm{T}}}^T\right]^T \in \mathbb{C}^{N_{\mathrm{T}}MN\times 1}$ represent the transmitted TS-domain signal at all antennas, $\boldsymbol{n}_{\mathrm{MIMO}}$ denotes the additive noise vector and TS-domain channel matrix follows

$$\mathbf{C}_{\mathrm{MIMO}} = \begin{bmatrix} \mathbf{C}_{1,1} & \cdots & \mathbf{C}_{1,N_{\mathrm{T}}} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{N_{\mathrm{R}},1} & \cdots & \mathbf{C}_{N_{\mathrm{R}},N_{\mathrm{T}}} \end{bmatrix}. \tag{70}$$

According to (16), the TS-domain equalization matrix is derived as

$$\mathbf{A}_{\mathrm{TS}} = \left(\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{C}_{\mathrm{MIMO}} + \lambda\mathbf{I}_{N_TMN}\right)^{-1}\mathbf{C}_{\mathrm{MIMO}}^H. \tag{71}$$

Thus, the MSE can be written in (72).

Frequency-space domain: We denote $\mathbf{d'}_{\mathrm{MIMO}} = \left[\mathbf{d'}_1^T, \cdots, \mathbf{d'}_{N_{\mathrm{R}}}^T\right]^T$ and $\mathbf{d}_{\mathrm{MIMO}} = \left[\mathbf{d}_1^T, \cdots, \mathbf{d}_{N_{\mathrm{T}}}^T\right]^T$ as the FS-domain received and transmitted data symbols for all antennas, respectively. By leveraging (49), we have $\mathbf{d'}_{\mathrm{MIMO}} = \left(\mathbf{I}_{N_{\mathrm{R}}} \otimes \left(\mathbf{I}_N \otimes \mathbf{F}_N\right)\right)\mathbf{r}_{\mathrm{MIMO}}; \mathbf{d}_{\mathrm{MIMO}} = \left(\mathbf{I}_{N_{\mathrm{T}}} \otimes \left(\mathbf{I}_N \otimes \mathbf{F}_N\right)\right)\mathbf{s}_{\mathrm{MIMO}}$. Thus, the FS-domain transmission is expressed as

$$\mathbf{d'}_{\mathrm{MIMO}} = \mathbf{H}_{\mathrm{MIMO}}\mathbf{d}_{\mathrm{MIMO}} + \boldsymbol{n'}_{\mathrm{MIMO}}, \tag{73}$$

where $\mathbf{H}_{\mathrm{MIMO}}$ denotes the FS-domain channel matrix, i.e.,

$$\mathbf{H}_{\mathrm{MIMO}} = \left(\mathbf{I}_{N_{\mathrm{R}}} \otimes \left(\mathbf{I}_N \otimes \mathbf{F}_N\right)\right)\mathbf{C}_{\mathrm{MIMO}}\left(\mathbf{I}_{N_{\mathrm{T}}} \otimes \left(\mathbf{I}_N \otimes \mathbf{F}_N\right)\right)^H, \tag{74}$$

and the noise vector $\boldsymbol{n'}_{\mathrm{MIMO}}$ follows $\boldsymbol{n'}_{\mathrm{MIMO}} = \left(\mathbf{I}_{N_{\mathrm{R}}} \otimes \left(\mathbf{I}_N \otimes \mathbf{F}_N\right)\right)\boldsymbol{n}_{\mathrm{MIMO}}$. From (16), the FS-domain equalization matrix is

$$\mathbf{A}_{\mathrm{FS}} = \left(\mathbf{H}_{\mathrm{MIMO}}^H\mathbf{H}_{\mathrm{MIMO}} + \lambda\mathbf{I}_{N_TMN}\right)^{-1}\mathbf{H}_{\mathrm{MIMO}}^H, \tag{75}$$

and the MSE refers to (76). For simplicity, we define $\mathbf{K}_1 = \mathbf{I}_{N_{\mathrm{R}}} \otimes \left(\mathbf{I}_N \otimes \mathbf{F}_N\right)$ and $\mathbf{K}_2 = \mathbf{I}_{N_{\mathrm{T}}} \otimes \left(\mathbf{I}_N \otimes \mathbf{F}_N\right)$. Note that $\mathbf{K}_1$ and $\mathbf{K}_2$ are unitary matrices. By leveraging (74) – (76), we have

$$\begin{aligned}
&\mathrm{Tr}\left\{\left(\mathbf{A}_{\mathrm{FS}}\mathbf{d'}_{\mathrm{MIMO}} - \mathbf{d}_{\mathrm{MIMO}}\right)\left(\mathbf{A}_{\mathrm{FS}}\mathbf{d'}_{\mathrm{MIMO}} - \mathbf{d}_{\mathrm{MIMO}}\right)^H\right\} \\
&= \mathrm{Tr}\left\{\mathbf{K}_2\left(\left(\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{C}_{\mathrm{MIMO}} + \lambda\mathbf{I}_{N_{\mathrm{T}}MN}\right)^{-1}\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)\right. \\
&\qquad \left.\left(\left(\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{C}_{\mathrm{MIMO}} + \lambda\mathbf{I}_{N_{\mathrm{T}}MN}\right)^{-1}\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)^H\mathbf{K}_2^H\right\} \\
&= \mathrm{Tr}\left\{\left(\left(\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{C}_{\mathrm{MIMO}} + \lambda\mathbf{I}_{N_{\mathrm{T}}MN}\right)^{-1}\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)\right. \\
&\qquad \left.\left(\left(\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{C}_{\mathrm{MIMO}} + \lambda\mathbf{I}_{N_{\mathrm{T}}MN}\right)^{-1}\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)^H\right\} \\
&= \mathrm{Tr}\left\{\left(\mathbf{A}_{\mathrm{TS}}\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)\left(\mathbf{A}_{\mathrm{TS}}\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)^H\right\}
\end{aligned} \tag{76}$$

According to (79), we can achieve $\mathrm{MSE}_{\mathrm{FS}} = \mathrm{MSE}_{\mathrm{TS}}$.

Delay-Doppler-space domain: For simplicity, we directly formulate the DDS-domain MIMO-OTFS transmission as

$$\mathbf{y}_{\mathrm{MIMO}} = \mathbf{G}_{\mathrm{MIMO}}\mathbf{x}_{\mathrm{MIMO}} + \boldsymbol{n''}_{\mathrm{MIMO}}, \tag{81}$$

where $\mathbf{y}_{\mathrm{MIMO}} = \left(\mathbf{I}_{N_{\mathrm{R}}} \otimes \left(\mathbf{\Pi}\left(\mathbf{F}_{\mathbf{N}} \otimes \mathbf{I}_{\mathbf{M}}\right)\right)\right)\mathbf{r}_{\mathrm{MIMO}}$ and $\mathbf{x}_{\mathrm{MIMO}} = \left(\mathbf{I}_{N_{\mathrm{T}}} \otimes \left(\mathbf{\Pi}\left(\mathbf{F}_{\mathbf{N}} \otimes \mathbf{I}_{\mathbf{M}}\right)\right)\right)\mathbf{s}_{\mathrm{MIMO}}$ denote the received and transmitted DDS-domain data vector, respectively. Vector $\boldsymbol{n''}_{\mathrm{MIMO}} = \left(\mathbf{I}_{N_{\mathrm{R}}} \otimes \left(\mathbf{\Pi}\left(\mathbf{F}_{\mathbf{N}} \otimes \mathbf{I}_{\mathbf{M}}\right)\right)\right)\boldsymbol{n}_{\mathrm{MIMO}}$ is the noise vector. The DDS-domain channel matrix $\mathbf{G}_{\mathrm{MIMO}} = \left(\mathbf{I}_{N_{\mathrm{R}}}\left(\mathbf{\Pi}\left(\mathbf{F}_{\mathbf{N}} \otimes \mathbf{I}_{\mathbf{M}}\right)\right)\right)\mathbf{C}_{\mathrm{MIMO}}\left(\mathbf{I}_{N_{\mathrm{R}}}\left(\Pi\left(\mathbf{F}_N \otimes \mathbf{I}_M\right)\right)\right)^H$. The MSE of the DDS-domain symbol estimates is denoted as (82), where $\mathbf{A}_{\mathrm{DDS}} = \left(\mathbf{G}_{\mathrm{MIMO}}^H\mathbf{G}_{\mathrm{MIMO}} + \lambda\mathbf{I}_{N_TMN}\right)^{-1}\mathbf{G}_{\mathrm{MIMO}}^H$. In similar, we define $\mathbf{P}_1 = \left(\mathbf{I}_{N_{\mathrm{R}}} \otimes \left(\mathbf{\Pi}\left(\mathbf{F}_{\mathbf{N}} \otimes \mathbf{I}_{\mathbf{M}}\right)\right)\right)$ and $\mathbf{P}_2 = \left(\mathbf{I}_{N_{\mathrm{T}}} \otimes \left(\mathbf{\Pi}\left(\mathbf{F}_{\mathbf{N}} \otimes \mathbf{I}_{\mathbf{M}}\right)\right)\right)$. According to the above, we have

$$\begin{aligned}
&\mathrm{Tr}\left\{\left(\mathbf{A}_{\mathrm{DDS}}\mathbf{y}_{\mathrm{MIMO}} - \mathbf{x}_{\mathrm{MIMO}}\right)\left(\mathbf{A}_{\mathrm{DDS}}\mathbf{y}_{\mathrm{MIMO}} - \mathbf{x}_{\mathrm{MIMO}}\right)^H\right\} \\
&= \mathrm{Tr}\left\{\mathbf{P}_2\left(\left(\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{C}_{\mathrm{MIMO}} + \lambda\mathbf{I}_{N_{\mathrm{T}}MN}\right)^{-1}\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)\right. \\
&\qquad \left.\left(\left(\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{C}_{\mathrm{MIMO}} + \lambda\mathbf{I}_{N_{\mathrm{T}}MN}\right)^{-1}\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)^H\mathbf{P}_2^H\right\} \\
&= \mathrm{Tr}\left\{\left(\left(\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{C}_{\mathrm{MIMO}} + \lambda\mathbf{I}_{N_{\mathrm{T}}MN}\right)^{-1}\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)\right. \\
&\qquad \left.\left(\left(\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{C}_{\mathrm{MIMO}} + \lambda\mathbf{I}_{N_{\mathrm{T}}MN}\right)^{-1}\mathbf{C}_{\mathrm{MIMO}}^H\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)^H\right\} \\
&= \mathrm{Tr}\left\{\left(\mathbf{A}_{\mathrm{TS}}\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)\left(\mathbf{A}_{\mathrm{TS}}\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)^H\right\}.
\end{aligned} \tag{82}$$

Based on (83), we have $\mathrm{MSE}_{\mathrm{DDS}} = \mathrm{MSE}_{\mathrm{TS}}$.

■

REFERENCES

[1] L. Zhang, A. Farhang, G. Feng, and O. Onireti, *Radio Access Network Slicing and Virtualization for 5G Vertical Industries*. Wiley, 2020. [Online]. Available: https://books.google.co.uk/books?id=-V0MEAAAQBAJ

[2] B. Muquet, Z. Wang, G. B. Giannakis, M. De Courville, and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions?" *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 2136–2148, Dec. 2002.

[3] P. Wang, B. Di, H. Zhang, K. Bian, and L. Song, "Cellular V2X communications in unlicensed spectrum: Harmonious coexistence with VANET in 5G systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5212–5224, Aug. 2018.

[4] H. Qu, G. Liu, Y. Wang, Q. Chen, C. Yi, and J. Peng, "A time-domain approach to channel estimation and equalization for the SC-FDM system," *IEEE Trans. Broadcast.*, vol. 65, no. 4, pp. 713–726, Dec. 2019.

[5] L. Zhang, A. Ijaz, P. Xiao, M. M. Molu, and R. Tafazolli, "Filtered OFDM systems, algorithms, and performance analysis for 5G and beyond," *IEEE Trans. Commun.*, vol. 66, no. 3, pp. 1205–1218, Mar. 2017.

$$\mathrm{MSE}_{\mathrm{TS}} = \frac{1}{N_{\mathrm{T}}MN}\mathbb{E}\left\{\mathrm{Tr}\left\{\left(\mathbf{A}_{\mathrm{TS}}\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)\left(\mathbf{A}_{\mathrm{TS}}\mathbf{r}_{\mathrm{MIMO}} - \mathbf{s}_{\mathrm{MIMO}}\right)^H\right\}\right\}. \tag{72}$$

$$\mathrm{MSE}_{\mathrm{FS}} = \frac{1}{N_{\mathrm{T}}MN}\mathbb{E}\left\{\mathrm{Tr}\left\{\left(\mathbf{A}_{\mathrm{FS}}\mathbf{d'}_{\mathrm{MIMO}} - \mathbf{d}_{\mathrm{MIMO}}\right)\left(\mathbf{A}_{\mathrm{FS}}\mathbf{d'}_{\mathrm{MIMO}} - \mathbf{d}_{\mathrm{MIMO}}\right)^H\right\}\right\}. \tag{76}$$

$$\mathrm{MSE}_{\mathrm{DDS}} = \frac{1}{N_{\mathrm{T}}MN}\mathbb{E}\left\{\mathrm{Tr}\left\{\left(\mathbf{A}_{\mathrm{DDS}}\mathbf{y}_{\mathrm{MIMO}} - \mathbf{x}_{\mathrm{MIMO}}\right)\left(\mathbf{A}_{\mathrm{DDS}}\mathbf{y}_{\mathrm{MIMO}} - \mathbf{x}_{\mathrm{MIMO}}\right)^H\right\}\right\}, \tag{82}$$

[6] L. Zhang, P. Xiao, A. Zafar, A. U. Quddus, and R. Tafazolli, "FBMC system: an insight into doubly dispersive channel impact," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 3942–3956, May 2017.

[7] X. Zhang, L. Zhang, P. Xiao, D. Ma, J. Wei, and Y. Xin, "Mixed numerologies interference analysis and inter-numerology interference cancellation for windowed OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7047–7061, Aug. 2018.

[8] G. Liu, S. V. Zhidkov, H. Li, L. Zeng, and Z. Wang, "Low-complexity iterative equalization for symbol-reconstruction-based OFDM receivers over doubly selective channels," *IEEE Trans. Broadcast.*, vol. 58, no. 3, pp. 390–400, Sept. 2012.

[9] 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 (WCNC)*, Mar. 2017.

[10] P. Raviteja, P. K. T., H. Yi, and V. Emanuele, "Interference cancellation and iterative detection for orthogonal time frequency space modulation," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6501–6515, Aug. 2018.

[11] P. Raviteja, K. T. Phan, and Y. Hong, "Embedded pilot-aided channel estimation for OTFS in delay-Doppler channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4906–4917, May 2019.

[12] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Effective diversity of OTFS modulation," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 249–253, Nov. 2019.

[13] A. Rezazadehreyhani, A. Farhang, M. Ji, R. R. Chen, and B. Farhang-Boroujeny, "Analysis of discrete-time MIMO OFDM-based orthogonal time frequency space modulation," in *Proc. IEEE International Conference on Communications (ICC)*, May 2018.

[14] W. Shen, L. Dai, J.-p. An, P. Fan, and R. W. Heath, "Channel estimation for orthogonal time frequency space (OTFS) massive MIMO," *IEEE Trans. Signal Process.*, vol. 67, no. 16, pp. 4204–4217, Aug. 2019.

[15] S. Li, J. Yuan, W. Yuan, Z. Wei, B. Bai, and D. W. K. Ng, "Performance analysis of coded OTFS systems over high-mobility channels," *IEEE Trans. Wireless Commun.*, *Early Access*, 2021.

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

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

[18] R. M. Augustine, G. Surabhi, and A. Chockalingam, "Space-time coded OTFS modulation in high-Doppler channels," in *Proc. IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, May 2019.

[19] M. K. Ramachandran and A. Chockalingam, "MIMO-OTFS in high-Doppler fading channels: signal detection and channel estimation," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2018.

[20] G. Surabhi and A. Chockalingam, "Low-complexity linear equalization for $2\times$ 2 MIMO-OTFS signals," in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Aug. 2020, pp. 1–5.

[21] Y. Liu, S. Zhang, F. Gao, J. Ma, and X. Wang, "Uplink-aided high mobility downlink channel estimation over massive MIMO-OTFS system," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 9, pp. 1994–2009, 2021.

[22] M. Li, S. Zhang, F. Gao, P. Fan, and O. A. Dobre, "A new path division multiple access for the massive MIMO-OTFS networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 903–918, 2021.

[23] C. X. Wang, A. Ghazal, B. Ai, Y. Liu, and P. Fan, "Channel measurements and models for high-speed train communication systems: A survey." *IEEE Commun. Surv. Tut.*, vol. 18, no. 2, pp. 974–987, Dec. 2016.

[24] R. Hadani, S. Rakib, A. Molisch, C. Ibars, A. Monk, M. Tsatsanis, J. Delfeld, A. Goldsmith, and R. Calderbank, "Orthogonal time frequency space (otfs) modulation for millimeter-wave communications systems," in *Proc. IEEE MTT-S International Microwave Symposium (IMS)*, Oct. 2017.

[25] G. Surabhi and A. Chockalingam, "Low-complexity linear equalization for OTFS modulation," *IEEE Commun. Lett.*, vol. 24, no. 2, pp. 330–334, Feb. 2020.

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

[27] P. Singh, A. Gupta, H. B. Mishra, and R. Budhiraja, "Low-complexity ZF/MMSE receivers for MIMO-OTFS systems with imperfect CSI," *arXiv preprint arXiv:2010.04057*, 2020.

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

[29] V. Khammammetti and S. K. Mohammed, "OTFS-based multiple-access in high Doppler and delay spread wireless channels," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 528–531, Apr. 2019.

[30] L. Li, Y. Liang, P. Fan, and Y. Guan, "Low complexity detection algorithms for OTFS under rapidly time-varying channel," in *Proc. IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, May 2019.

[31] 3GPP, *Study on LTE-based V2X Services*, Rel. 14, TR 36.885 (V14.0.0) Std., 2016.

[32] H. Qu, G. Liu, L. Zhang, S. Wen, and M. A. Imran, "Low-complexity symbol detection and interference cancellation for OTFS system," *IEEE Trans. Commun.*, vol. 69, no. 3, pp. 1524–1537, Mar. 2021.

[33] T. Thaj and E. Viterbo, "Low complexity iterative rake decision feedback equalizer for zero-padded OTFS systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 15 606–15 622, Dec. 2020.

[34] F. Rottenberg, X. Mestre, F. Horlin, and J. Louveaux, "Efficient equalization of time-varying channels in MIMO OFDM systems," *IEEE Trans. Signal Process.*, vol. 67, no. 21, pp. 5583–5595, Nov. 2019.

[35] W. C. Jakes and D. C. Cox, *Microwave mobile communications*. Wiley-IEEE Press, 1994.

[36] C. Xiao, Y. R. Zheng, and N. C. Beaulieu, "Second-order statistical properties of the WSS Jakes' fading channel simulator," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 888–891, June 2002.

[37] S. S. Das, V. Rangamgari, S. Tiwari, and S. C. Mondal, "Time domain channel estimation and equalization of CP-OTFS under multiple fractional dopplers and residual synchronization errors," *IEEE Access*, vol. 9, pp. 10 561–10 576, Dec 2020.

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

[39] A. Gorokhov and J.-P. Linnartz, "Robust OFDM receivers for dispersive time-varying channels: Equalization and channel acquisition," *IEEE Trans. Commun.*, vol. 52, no. 3, pp. 572–583, Apr. 2004.

[40] P. Schniter, "Low-complexity equalization of OFDM in doubly selective channels," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1002–1011, Apr. 2004.

[41] K. Fang, L. Rugini, and G. Leus, "Low-complexity block turbo equalization for OFDM systems in time-varying channels," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5555–5566, Nov. 2008.

[42] C. C. Paige and M. A. Saunders, "LSQR: An algorithm for sparse linear equations and sparse least squares," *ACM Trans. Mathemat. Software*, vol. 8, no. 1, pp. 43–71, 1982.

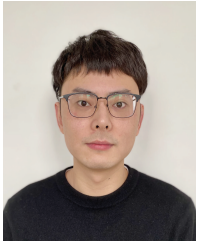
[43] G. Golub and W. Kahan, "Calculating the singular values and pseudo-inverse of a matrix," *Journal Soc. Indus. Appl. Math. Ser. B Number. Anal.*, vol. 2, no. 2, pp. 205–224, 1965.

[44] D. C.-L. Fong and M. Saunders, "LSMR: An iterative algorithm for sparse least-squares problems," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2950–2971, Oct. 2011.

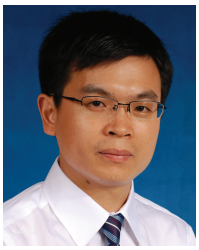
[45] H. Qu, G. Liu, L. Zhang, M. A. Imran, and S. Wen, "Low-dimensional subspace estimation of continuous-doppler-spread channel in OTFS systems," *IEEE Trans. Commun.*, 2021, Early Access.

[46] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Practical pulse-shaping waveforms for reduced-cyclic-prefix OTFS," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 957–961, Jan. 2019.

[47] G. H. Golub and C. F. Van Loan, *Matrix Computations, 4th ed.* Johns Hopkins Univ. press, 2013.



Huiyang Qu received the B.Sc. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China in 2015. He is currently working toward the Ph.D. degree at UESTC. His research interests include channel estimation and equalization, 5G and beyond waveform design and mmWave communication system.



Guanghui Liu (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in electronic engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2002 and 2005, respectively. In 2005, he joined Samsung Electronics, South Korea, as a Senior Engineer. In 2009, he became an Associate Professor with the School of Electronics Engineering, UESTC, where he has been a Full Professor since 2014 and is currently with the School of Information and Communication Engineering. His general research interests include digital signal processing and telecommunications, with emphasis on digital video processing and transmission, B5G/6G waveform design, and AI for physical layer techniques. He has published over 100 papers in peer-reviewed journals and conferences, and filed over 60 patents (6 U.S. granted patents) in the above areas. He was a recipient of the Natural Science Award and the Science and Technology Progress Award, both from Ministry of Education of China in 2015.

Muhammad Ali Imran (Senior Member, IEEE) is Dean University of Glasgow UESTC and a Professor of Wireless Communication Systems with research interests in fundamental communication theory, self-organised networks and the wireless sensor systems. He heads the Communications, Sensing and Imaging (CSI) research group at University of Glasgow and is Director of Centre for Educational Development and Innovation. He is an Affiliate Professor at the University of Oklahoma, USA; Adjunct Research Professor at Ajman University, UAE and a visiting Professor at 5G Innovation Centre, University of Surrey, UK. He has over 20 years of combined academic and industry experience with several leading roles in multi-million pounds funded projects. He has filed 15 patents; has authored/co-authored over 400 journal and conference publications; has authored 2 books, edited 8 books and authored more than 30 book chapters; has successfully supervised over 40 postgraduate students at Doctoral level. He has been a consultant to international projects and local companies in the area of self-organised networks. He is a Fellow of IET and a Senior Fellow of HEA.

Shan Wen (Student member, IEEE) received the B.Sc. degree in Electronic and Information Engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2017, where he is currently working toward the Ph.D. degree. His research interests include faster-than-Nyquist signaling, high-capacity communications, waveform design, channel estimation and (iterative) equalization, MIMO-FTN system.

Lei Zhang (Senior Member, IEEE) is a Senior Lecturer at the University of Glasgow. He has academia and industry combined research experience on wireless communications and networks, and distributed systems for IoT, blockchain, autonomous systems. His 20 patents are granted/filed in 30+ countries/regions. He published 3 books, and 150+ papers in peer-reviewed journals, conferences and edited books. Dr. Zhang is an associate editor of IoT Journal, IEEE Wireless Communications Letters and Digital Communications and Networks, and a guest editor of IEEE JSAC. He received the IEEE ComSoc TAOS Technical Committee Best Paper Award 2019 and IEEE ICEICT'21 Best Paper Award. Dr. Zhang is the founding Chair of IEEE Special Interest Group on Wireless Blockchain Networks in Cognitive Networks Technical Committee (TCCN). He delivered tutorials in IEEE ICC'20, IEEE PIMRC'20, IEEE Globecom'21, IEEE VTC'21 Fall, IEEE ICBC'21 and EUSIPCO'21