Alternative threshold-based channel estimation and message-passing-based symbol detection in MIMO-OTFS systems using superimposed pilots

Rabah Ouchikh, Abdeldjalil Aissa El Bey, Thierry Chonavel, Mustapha Djeddou

Alternative threshold-based channel estimation and message-passing-based symbol detection in MIMO-OTFS systems using superimposed pilots

Rabah Ouchikh[a], Abdeldjalil Aïssa-El-Bey[b], Thierry Chonavel[b], Mustapha Djeddou[a]

[a] *Laboratoire Télécommunications, Ecole Militaire Polytechnique, Bordj El-Bahri, Algeria.*
[b] *IMT Atlantique, Lab-STICC, UMR CNRS 6285, F-29238 Brest, France.*

Abstract

In this work, we investigate the challenging problem of channel estimation in high-mobility environments for advanced mobile communication systems (5G and beyond). First, we propose an iterative algorithm for channel estimation and symbol detection in the delay-Doppler domain for multiple-input multiple-output orthogonal time frequency space (OTFS) system. The proposed algorithm is based on a superimposed pilot pattern to improve the spectral efficiency of the system. It iterates between data-aided channel estimation and message-passing-aided data detection. The channel estimation step is based on a threshold method. This step considers interference-plus-noise caused by the data symbols and the additive noise to adapt the threshold at each iteration. The data detection step is based on an adapted version of the message-passing algorithm proposed in the literature for uncoded OTFS. Then, to improve the channel estimation efficiency, we suggest an interference cancellation scheme executed at each iteration of the proposed algorithm. Finally, we compare the computational complexity and the achieved performance in terms of normalized mean square error of channel estimation, bit error rate, and spectral efficiency against five state-of the-art methods.

Keywords: MIMO, OTFS, channel estimation, iterative algorithm, superimposed pilot.

1. Introduction

The new generations of mobile communications are awaited to provide reliable high-speed communications even in high-mobility environments (high-speed trains, UAV communications, etc.) [1]. However, the orthogonal frequency division multiplexing (OFDM), which is the most popular modulation waveform deployed in fifth generation (5G) communication systems [2], see its performance degraded in high-mobility scenarios [3]. The orthogonal time-frequency space (OTFS) modulation, recently proposed in [4, 5], is one of the promising techniques for 5G and beyond. OTFS remains robust in channels where the Doppler shift is very high (high-mobility scenario). This new waveform incorporates the information symbols in the delay-Doppler (DD) domain, unlike OFDM where the data symbols are incorporated in the time-frequency domain [6, 7].

Multiple-input multiple-output (MIMO) can also be used with OTFS and benefit from the spatial diversity to further increase the transmission rate [8–12]. To ensure robust data transmission in a MIMO-OTFS system, efficient channel estimation (CE) and data detection algorithms are required.

The proposal of a CE scheme requires to make a compromise between the spectral efficiency (SE) of the system, the computational complexity, and the performance in terms of normalized mean square error (NMSE) of CE and bit error rate (BER).

Several CE schemes in the delay-Doppler domain have been proposed. These algorithms can be classified into three categories depending on the used data-transmission architecture. The first category includes conventional schemes using a super-frame architecture., i.e.: the first frame is used only for the transmission of pilots and the other frames for the transmission of data symbols. The second category includes embedded schemes carrying pilots and data symbols on the same frame. Although these schemes use guard symbols between the pilots and data symbols to avoid interference, they provide improved SE over conventional schemes. The third category of algorithms consists of superimposing pilots and data symbols without using guard intervals. These schemes further improve the SE but have the drawback of increasing the computational complexity, due mainly to the use of iterative schemes to manage the interference between the pilots and data symbols.

In the present paper, we propose an iterative algorithm for CE and data detection for MIMO-OTFS systems. To guarantee high SE, we opted for a superimposed pilot pattern. The proposed algorithm iterates between data-aided CE and message-passing-aided data detection. The CE step is based on a threshold method that accounts for interference-plus-noise caused by the data symbols and the additive noise to adapt the threshold at each iteration. The data detection step is based on an adapted version of the message-passing (MP) algorithm [13]. The main original features of this manuscript are synthesized as follows:

(i) We proposed an iterative algorithm for CE and symbol detection and a superimposed pilot pattern for MIMO-OTFS systems. The CE step is based on a simple threshold method and the symbol detection step is based on a message-passing algorithm.

(ii) We derive an optimal threshold expression for CE. Unlike classical methods [3] (where the threshold assumes the existence of additive noise only), in our case, the determination of the threshold takes into account the additive noise and the interference caused by data symbols.

(iii) In order to improve the CE efficiency, we propose an interference cancellation scheme executed at each iteration of the designed algorithm.

(iv) We compare the computational complexity and the performance of the proposed algorithm against five

state-of-the-art methods, namely: conventional pilot aided (CPA) [14], embedded pilot (EP) [3], row-group orthogonal matching pursuit (RG-OMP), row-group Bayesian learning (RG-BL) [15], and block sparse Bayesian learning with block reorganization (BSBL-BR) [16], in terms of CE-NMSE, BER, and SE.

The remainder of this manuscript is structured as follows. Related works are given in Section 2. Section 3 is dedicated to the description of the MIMO-OTFS system model. The details of the proposed algorithm for CE and data detection are given in Section 4. The complexity analysis and the performance of the proposed scheme are given in Section 5.

2. Related works

Several CE algorithms in the DD domain for OTFS modulation have been proposed in the literature. They can be classified into three groups, based on their transmission scheme.

2.1. Conventional pilot aided (CPA) design

This group utilizes methods that allocate an entire frame for pilot transmission, which involves using the initial OTFS frame to estimate the channel and the following frames for transmitting data [14, 17–21]. In [18], a method is proposed for compressed sensing in 2D Turbo that models the support matrix of the DD channel using a Markov random field and the Bernoulli Gaussian distribution. Reference [19] presents an uplink DD-CE algorithm for OTFS multiple access systems, formulating the CE problem as a sparse recovery problem and solving it using OMP and modified subspace pursuit algorithms. Reference [17] proposes two low-complexity algorithms for CE in OTFS with fractional delay and Doppler. The first (Modified Maximum Likelihood CE) is based on joint 2D estimation of the DD shift of a path. The second (two-step method) decouples the joint estimation into two separate 1D estimation steps. Reference [20] suggests the OMP with Binary-division Refinement algorithm for CS-based OTFS-CE, which allows for continuous parameter estimation in the DD domain. Finally, reference [21] develops a pseudo-random noise pilot-based CE scheme in the DD domain, leveraging the sparse DD impulse response that reflects the physical geometry of the wireless channel. In [14], a CPA design for MIMO-OTFS CE in the DD domain has been proposed. A super-frame architecture has been adopted., i.e.: the first frame is used for CE and the following frames are used for symbols transmission. The CE scheme uses impulses at some fixed locations in the DD domain as pilots for estimation. This algorithm achieves close to oracle performance with low complexity but suffers from SE degradation due to the super-frame architecture.

2.2. Embedded pilot (EP) scheme

This second group includes the incorporation of both pilots and data symbols in a single OTFS frame by incorporating guard intervals around the pilots to prevent data interference during the CE process [3, 15, 16, 22–30, 33]. Authors in [3] propose an embedded pilot-aided CE for OTFS in the DD channels. An arrangement of pilot, data symbols, and guard intervals in the same frame followed by a CE algorithm has been adopted for single-input single-output OTFS (SISO-OTFS) systems. Then, an extension to MIMO-OTFS systems has been also suggested in [3]. Thanks to the guard intervals, there is no interference on received pilot contributions caused by data symbols. Therefore, a simple threshold technique is used for CE. Compared to the CPA design, the performance and computational complexity achieved by this scheme are almost the same and there is an improvement of the SE. However, due to the increase in pilot overhead, this improvement remains small especially in MIMO-OTFS and multi-user (MU) MIMO-OTFS (MU-MIMO-OTFS) systems. In [15], a sparse-channel model has been proposed for reducing the pilot overhead of MIMO-OTFS systems. To estimate the channel parameters in the DD domain, pilots are arranged in the time-frequency domain. This reduces pilot overhead, pre-processing complexity and training period. The resultant DD domain channel state information (CSI) becomes simultaneously row and group (RG)-sparse. With the benefit of this property, an OMP-based RG-OMP technique has been suggested, complemented in a practical way by an improved RG-BL framework based on Bayesian learning (BL). Even if these two proposed techniques are complex compared to the previous schemes, they reach good performance in terms of CE NMSE and BER with a considerable improvement of the SE. In [16], an algorithm for MIMO-OTFS CE has been proposed, where the CE problem is formulated as a block sparse signal recovery problem. It is solved by the suggested BSBL-BR method. Compared to RG-OMP and RG-BL, BSBL-BR is less efficient but has low computational complexity and better SE. In [31], a new rearrangement of pilots has been proposed for OTFS. Afterwards, an OTFS pilot sequence optimization is formulated and solved via a particle swarm optimization algorithm. The CE problem is formulated as a sparse signal recovery problem and solved using the classical OMP algorithm. Thanks to the optimized pilot sequence, the accuracy of CE is improved. The SE is also improved compared to the previous schemes due to the reduced proportion of guard symbols. In [32], the CE problem for OTFS is formulated as a sparse signal recovery problem based on a reduced guard symbols pilot pattern. Several algorithms are designed to solve this problem, namely: Iterative two-stage, Reweighted iterative two-stage, Monte Carlo Markov chain, and soft Bayesian pursuit (SoBaP). These algorithms offer a good compromise between performance and complexity and remain applicable in high-mobility scenarios and low-latency applications. The suggested scheme offers a SE close to that of [31]. In [33], a joint CE and data detection (JCEDD) algorithm for hybrid reconfigurable intelligent surface aided millimeter wave OTFS systems has been proposed. In this method, the channel gain and the data

symbols are simultaneously obtained using the MP algorithm. The channel parameters (the Doppler shift, the channel sparsity, and the channel variance) are updated through expectation-maximization algorithm. This method achieves a trade-off between the estimation and detection performance, the power gap, and the pilot overhead.

2.3. Superimposed pilots design

This group includes a superimposed pilot scheme, where pilots and data symbols are spread across the DD domain [34–36]. In [34], a data-aided CE algorithm is proposed for a superimposed pilot and data transmission scheme to enhance the SE. Initially, the channel is coarsely estimated based on pilot symbols. Then, an iterative process is used for soft detection of data symbols and channel estimate to improve CE and data detection. In [35], pilots and data symbols are overlaid in the DD grid, spreading the pilot energy across the entire domain, resulting in a higher SE than previous systems that used guard intervals. In [36], an iterative algorithm for CE and data detection has been suggested. This sparse CE problem is reformulated as a specific marginalization of maximum a posteriori (MAP) problem and is resolved using a Bayesian approach based on the variational mean-field approximation via the variational Bayesian expectation maximization (VB-EM) algorithm.

3. System model

In this section, we discuss the basic concepts of a MIMO-OTFS transceiver following the description introduced in [14]. The parameters of the MIMO-OTFS system considered in this work, along with, their physical significations are shown in Table 1. The block diagram of the system is given in Fig 1.

Table 1: MIMO-OTFS system parameters.

Parameters	**Physical meaning**
N, M	Doppler bins, delay bins
N_t, N_r	Number of Tx antennas, number of Rx antennas
G_t, G_r	Transmit pulse matrix, Receive pulse matrix
Δt, Δf	Time slot duration, subcarrier spacing
$T = N\Delta t$, $M\Delta f$	Frame duration, frame bandwidth
k_ν, l_τ	Maximum Doppler tap, maximum delay tap
f_c, t_c	Carrier frequency, coherence time

At the transmitter side, for the p^{th} antenna ($p = 1 : N_t$), the 2D grid of symbols in the DD domain $\boldsymbol{X}_{DD}^{(p)} \in \mathbb{C}^{M\times N}$ is formed by mapping a sequence of information bits into MN symbols (QAM, PSK, etc.). Then, $\boldsymbol{X}_{DD}^{(p)}$ is transformed into a 2D time-frequency grid $\boldsymbol{X}_{TF}^{(p)} \in \mathbb{C}^{M\times N}$ via an inverse symplectic finite Fourier transform (ISFFT) as follows:

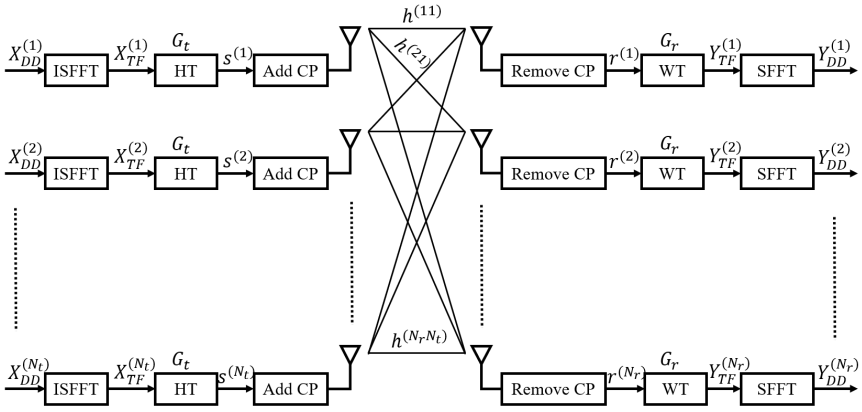


Figure 1: MIMO-OTFS block diagram.

$$\boldsymbol{X}_{TF}^{(p)}[n,m] = \frac{1}{\sqrt{MN}} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} \boldsymbol{X}_{DD}^{(p)}[k,l] \exp\left(j \frac{2\pi}{NM} \left(nkM - mlN \right) \right), \tag{1}$$

where $n = 0 : N-1$ and $m = 0 : M-1$. Next, the time-frequency domain signal is converted to a time domain continuous signal $s(t)$ using the Heisenberg transform (HT) as follows:

$$s^{(p)}(t) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \boldsymbol{X}_{TF}^{(p)}[n,m] \exp\left(j2\pi m\Delta f(t-nT) \right) g_t(t-nT), \tag{2}$$

where g_t is the pulse shaping filter. $s^{(p)}(t)$ can be written in a vector form as follows:

$$\boldsymbol{s}^{(p)} = \left(\boldsymbol{F}_N^H \otimes \boldsymbol{G}_t \right) \boldsymbol{x}_{DD}^{(p)}, \tag{3}$$

where $\boldsymbol{s}^{(p)} = vec(\boldsymbol{S}^{(p)})$ with $\boldsymbol{S}^{(p)} = \boldsymbol{G}_t \boldsymbol{F}_M^H (\boldsymbol{F}_M \boldsymbol{X}_{DD}^{(p)} \boldsymbol{F}_N^H) = \boldsymbol{G}_t \boldsymbol{X}_{DD}^{(p)} \boldsymbol{F}_N^H$, $\boldsymbol{x}_{DD}^{(p)} = vec(\boldsymbol{X}_{DD}^{(p)})$ and $\otimes$ is the Kronecker product. $\boldsymbol{F}_K$ and $\boldsymbol{F}_K^H$ denote, the K-point direct and inverse discrete Fourier transform matrices. $\boldsymbol{G}_t = \text{diag}[g_t(0), g_t(T/M), ..., g_t((M-1)T/M)]$. One CP is added to $s^{(p)}(t)$ before its transmission.

The MIMO multi-paths time-varying wireless channel is sparse in the DD domain involving only a few parameters. The channel between the p^{th} transmitting antenna and the q^{th} receiving antenna has P taps. Thus, the baseband channel impulse response can be expressed as

$$h_{qp}(\tau,\nu) = \sum_{i=1}^{P} h_i^{(qp)} \delta(\tau - \tau_i^{(qp)}) \delta(\nu - \nu_i^{(qp)}), \tag{4}$$

where $h_i^{(qp)}$, $\tau_i^{(qp)}$, and $\nu_i^{(qp)}$ denote the complex gain, delay shift, and Doppler shift of the i^{th} path, respectively. The i^{th} delay and Doppler taps $\left(l_i^{(qp)}, k_i^{(qp)} \right)$ can be written as $l_i^{(qp)} = \tau_i^{(qp)} M\Delta f$ and $k_i^{(qp)} = \nu_i^{(qp)} NT$.

After transmitting through a multi-path time-varying channel $h_{qp}(\tau,\nu)$, the received signal $\boldsymbol{r}^{(q)}$ is defined by the following linear system for $q = 1 : N_r$:

$$\boldsymbol{r}^{(q)} = \sum_{i=1}^{N_t} \boldsymbol{H}_{qp}\boldsymbol{s}^{(p)} + \boldsymbol{n}^{(q)}, \tag{5}$$

where $\boldsymbol{n}^{(q)} \sim \mathcal{CN}(0, \sigma_n^2 \boldsymbol{I}_{MN})$ is an additive complex Gaussian noise variable, $\boldsymbol{H}_{qp} = \sum_{i=1}^{P} h_i^{(qp)} \boldsymbol{\Pi}^{l_i^{(qp)}} \boldsymbol{\Delta}^{k_i^{(qp)}} \in \mathbb{C}^{MN \times MN}$ denotes the channel matrix between the p^{th} Tx antenna and the q^{th} Rx antenna. $\boldsymbol{\Pi}$ denotes the permutation matrix and $\boldsymbol{\Delta} = \text{diag}[\exp(j2\pi(0)/MN), \exp(j2\pi(1)/MN), ..., \exp(j2\pi(MN-1)/MN)]$. Letting $\boldsymbol{r} = [(\boldsymbol{r}^{(1)})^T, (\boldsymbol{r}^{(2)})^T, ..., (\boldsymbol{r}^{(N_r)})^T]^T \in \mathbb{C}^{MNN_r}$ and $\boldsymbol{s} = [(\boldsymbol{s}^{(1)})^T, (\boldsymbol{s}^{(2)})^T, ..., (\boldsymbol{s}^{(N_t)})^T]^T \in \mathbb{C}^{MNN_t}$, we get

$$\boldsymbol{r} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{n}, \tag{6}$$

where $\boldsymbol{n} = [(\boldsymbol{n}^{(1)})^T, (\boldsymbol{n}^{(2)})^T, ..., (\boldsymbol{n}^{(N_r)})^T]^T \in \mathbb{C}^{MNN_r}$ and

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{H}_{11} & \cdots & \boldsymbol{H}_{1N_t} \\ \vdots & \ddots & \vdots \\ \boldsymbol{H}_{N_r1} & \cdots & \boldsymbol{H}_{N_rN_t} \end{pmatrix} \in \mathbb{C}^{MN \times MN}. \tag{7}$$

At the receiver end, the signal $\boldsymbol{r}^{(q)}$ first undergoes an adapted filter bank, which provides the cross-ambiguity function between $g_r(t)$ and $r^{(q)}(t)$. Then, the output $\boldsymbol{Y}_{TF}^{(q)}(t,f)$ is sampled leading to $\boldsymbol{Y}_{TF}^{(q)}[n,m] = \boldsymbol{Y}_{TF}^{(q)}(t,f)|_{(t=nT, f=m\Delta f)}$. This procedure is called the Wigner transform (WT). After that, the DD domain signal $\boldsymbol{Y}_{DD}^{(q)}$ is obtained by applying the symplectic finite Fourier transform (SFFT) to $\boldsymbol{Y}_{TF}^{(q)}$ as follows:

$$\boldsymbol{Y}_{DD}^{(q)}[k,l] = \frac{1}{\sqrt{MN}} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \boldsymbol{Y}_{TF}^{(q)}[n,m] \exp\left(-j\frac{2\pi}{NM}(nkM - mlN)\right). \tag{8}$$

The DD input/output relationship between $\boldsymbol{Y}_{DD}^{(q)}$ and $\boldsymbol{X}_{DD}^{(p)}$ can be expressed as follows:

$$\boldsymbol{Y}_{DD}^{(q)}[k,l] = \sum_{p=1}^{N_t} \sum_{k'=0}^{N-1} \sum_{l'=0}^{M-1} \boldsymbol{X}_{DD}^{(p)}[k',l'] h_w[[k-k']_N, [l-l']_M] + \tilde{\boldsymbol{n}}^{(q)}[k,l], \tag{9}$$

where $h_w[k,l]$ is the DD effective channel and $\tilde{\boldsymbol{n}}^{(q)}[k,l]$ is the DD noise sample. $[.]_K$ stands for modulo K operator. Finally, based on (3), (5) and the fact that $\boldsymbol{y}_{DD}^{(q)} = (\boldsymbol{F}_M \otimes \boldsymbol{G}_r)\boldsymbol{r}^{(q)}$, where $\boldsymbol{y}_{DD}^{(q)} = vec(\boldsymbol{Y}_{DD}^{(q)})$, letting $\boldsymbol{y} = [(\boldsymbol{y}_{DD}^{(1)})^T, (\boldsymbol{y}_{DD}^{(2)})^T, ..., (\boldsymbol{y}_{DD}^{(N_r)})^T]^T$, $\boldsymbol{x} = [(\boldsymbol{x}_{DD}^{(1)})^T, (\boldsymbol{x}_{DD}^{(2)})^T, ..., (\boldsymbol{x}_{DD}^{(N_t)})^T]^T$, and $\tilde{\boldsymbol{n}} = [(\boldsymbol{F}_N \otimes \boldsymbol{G}_r)\boldsymbol{n}^{(1)})^T, (\boldsymbol{F}_N \otimes \boldsymbol{G}_r)\boldsymbol{n}^{(2)})^T, ..., (\boldsymbol{F}_N \otimes \boldsymbol{G}_r)\boldsymbol{n}^{(N_r)})^T]^T$, the received signal in the DD domain can be expressed in a vector form as follows:

$$\boldsymbol{y} = \boldsymbol{\mathcal{H}}\boldsymbol{x} + \tilde{\boldsymbol{n}}, \tag{10}$$

where $\boldsymbol{\mathcal{H}}_{qp} = (\boldsymbol{F}_N \otimes \boldsymbol{G}_r)\boldsymbol{H}_{qp}(\boldsymbol{F}_N^H \otimes \boldsymbol{G}_t)$ and

$$\boldsymbol{\mathcal{H}} = \begin{pmatrix} \boldsymbol{\mathcal{H}}_{11} & \cdots & \boldsymbol{\mathcal{H}}_{1N_t} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\mathcal{H}}_{N_r1} & \cdots & \boldsymbol{\mathcal{H}}_{N_rN_t} \end{pmatrix} \in \mathbb{C}^{MN \times MN}. \tag{11}$$

4. Proposed scheme

In this section, we describe the proposed algorithm for CE, starting with the suggested pilot pattern. Then, we provide details of the proposed scheme.

4.1. Proposed pilot pattern

Channel estimation in this context requires the use of pilots. Two approaches are commonly used. The first one, called CPA, consists in separating transmission of the pilots from that of the information symbols., i.e.: two frames will be sent. The first only contains pilots; whereas, the second carries information symbols [14]. The second approach, which is the most used, consists in putting pilots and data symbols in the same frame while separating them by guard intervals to avoid data/pilot interferences [3]. The pilot placement in the DD domain for the latter approach is expressed as follows:

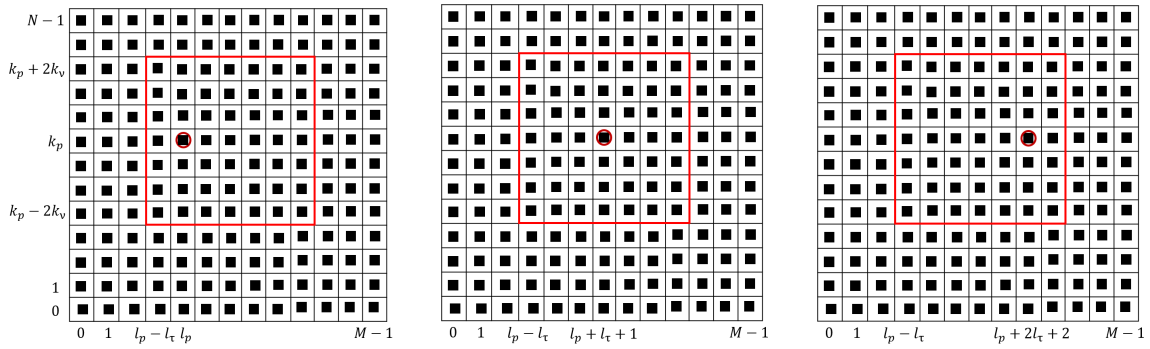
$$\boldsymbol{X}_{DD}^{(p)}[k,l] = \begin{cases} x_p, & k = k_p,\ l = l_p + (p-1)(l_\tau + 1), \\ 0, & k_p - 2k_\nu \leq k \leq k_p + 2k_\nu \text{ and } \ l_p - l_\tau \leq l \leq l_p + N_t(l_\tau + 1) - 1, \\ \boldsymbol{X}_d^{(p)}[k,l] & \text{otherwise}, \end{cases} \tag{12}$$

where $\boldsymbol{X}_d^{(p)}[k,l]$ denotes the information symbols at location $[k,l]$ of the p^{th} Tx antenna and l_τ, k_ν denote the maximum delay tap and the maximum Doppler tap. We can notice that $N_G = (4k_\nu + 1)(N_t l_\tau + N_t + l_\tau) - 1$ guard symbols are used in each frame in this scheme. The SE can be defined as [34]:

$$SE = (1 - \eta)\log|\mathcal{A}|, \tag{13}$$

where $\mathcal{A}$ denotes the constellation set to which the data symbols belong and η is the pilot overhead. Hence, the pilot overhead relative to the scheme (12) is given by $\eta = (N_G + 1)/NM$.

Clearly, N_G and η increase with N_t and the SE decreases as N_t increase. Furthermore, the expression of N_G depends also on k_ν, i.e., in high mobility environments (high Doppler shifts) N_G should be large to avoid data/pilot interference and this also causes a degradation of the SE. Therefore, in a massive MIMO system



(a) Antenna 1 (b) Antenna 2 (c) Antenna 3

Figure 2: Tx pilot and data symbols for MIMO-OTFS system (■: data, ○: pilot, the red square area represents the space allocated for guard intervals in classical approaches).

or in a high mobility environment the DD frame size should be very large, increasing thus the computational complexity. In addition, in the new wireless communication systems the frame duration is limited. In such cases, the SE is severely affected. Therefore, in this paper, we propose a CE algorithm for MIMO-OTFS systems, which uses the entire OTFS frame for data symbol transmission. The proposed pilot pattern in the DD domain is illustrated in Fig 2 (example of 3×3 MIMO-OTFS), where the area bounded by the red square denotes the guard space adopted in the classic OTFS CE scheme. As shown in Fig 2, we suggest the following superimposed arrangement:

$$\boldsymbol{X}_{DD}^{(p)}[k,l] = \begin{cases} x_p + \boldsymbol{X}_d^{(p)}[k,l], & k = k_p,\ l = l_p + (p-1)(l_\tau+1), \\ \boldsymbol{X}_d^{(p)}[k,l] & \text{otherwise}, \end{cases} \tag{14}$$

We can observe that this proposed pilot pattern does not use guard intervals. This will increase the SE of the system but will introduce interference between the pilot and the data symbols. Hence, we need an interference cancellation scheme and an efficient CE algorithm.

The Rx symbol pattern at each antenna is illustrated in Fig 3a. At the q^{th} Rx antenna, the received symbols $\boldsymbol{Y}_{DD}^{(q)}[k,l]$ for $k \in [k_p - k_\nu, k_p + k_\nu]$, $l \in [l_p + (p-1)(l_\tau+1), l_p + N_t(l_\tau+1) - 1]$ are used for CE to the p^{th} Tx antenna. These symbols are affected by the data symbols, pilot signal of the p^{th} Tx antenna, and the channel between the p^{th} Tx and q^{th} Rx antennas, as shown in Fig 3a. The received symbols related to the pilot can be rewritten as follows:

$$\boldsymbol{Y}_{DD}^{(q)}[k,l] = x_p h^{(qp)}[[k-k_p]_N, [l-l_p]_M] + \sum_{p=1}^{N_t} I_{k,l}^{(qp)} + \tilde{\boldsymbol{n}}^{(q)}[k,l], \tag{15}$$

where $I_{k,l}^{(qp)}$ denotes the interference introduced by the data symbols, given by

$$I_{k,l}^{(qp)} = \sum_{k'=k-k_\nu}^{k+k_\nu} \sum_{l'=l-l_\tau}^{l+l_\tau} \boldsymbol{X}_{DD}^{(p)}[k,l] h^{(qp)}[[k-k_p]_N, [l-l_p]_M] = \sum_{i \in \mathcal{S}_{k,l}} h_i^{(qp)} \boldsymbol{X}_{DD}^{(p)}[[k-k_i]_N, [l-l_i]_M] \alpha_{k,l}, \quad (16)$$

where $\mathcal{S}_{k,l}$ is the set containing indices of all data symbols contributing to $\boldsymbol{Y}_{DD}^{(q)}[k,l]$ (the pilot is interfered by P data symbols, therefore $|\mathcal{S}_{k,l}| = P$), and

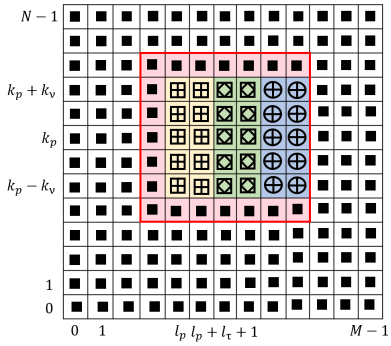
$$\alpha_{k,l} = \begin{cases} \exp\left(j2\pi \frac{k_i[l-l_i]_M}{MN}\right), & l_i \leq l \leq M, \\ \exp\left(-j2\pi \frac{l_i}{N}\right) \exp\left(j2\pi \frac{k_i[l-l_i]_M}{MN}\right) & 0 \leq l \leq l_i. \end{cases} \quad (17)$$

Because of the interference between the pilot and data symbols, the CE cannot be done simply from (15) using the threshold $\rho = 3\sigma$ as suggested in [3]. Therefore, we need to define a new decision threshold and an algorithm to manage these interferences. This is the objective of the next section.

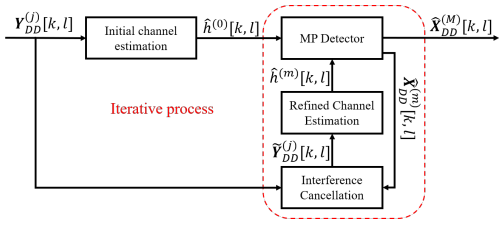
4.2. Proposed ICEDD algorithm

The proposed iterative channel estimation and data detection (ICEDD) scheme is shown in Fig 3b in the form of a block diagram. The first step of this algorithm is to find an initial estimate of the channel $\{(\boldsymbol{h}^{(qp)})^{(0)}\}$ for $p = 1 : N_t$ and for $q = 1 : N_r$ from (15), followed by data symbol detection using the estimated channel coefficients and applying the MP algorithm. Then, an interference cancellation scheme is applied to remove the interference from the received signal $\boldsymbol{Y}_{DD}^{(q)}[k,l]$, leading to the signal $\tilde{\boldsymbol{Y}}_{DD}^{(q)}[k,l]$ (this scheme is fed by the received signal $\boldsymbol{Y}_{DD}^{(q)}[k,l]$ and decided symbols). Afterwards, a refined channel estimate is obtained using the signal $\tilde{\boldsymbol{Y}}_{DD}^{(q)}[k,l]$. This new estimated channel will then be used for the detection of the data symbols. The process containing interference cancellation, refined CE, and data detection is done in an iterative way, as shown in Fig 3b.

The initial channel estimate is obtained by comparing the received signal $\boldsymbol{Y}_{DD}^{(q)}[k,l]$ given by equation (15) to a predetermined threshold. This threshold, unlike the one used in the classical method, takes into account the interference $\sum_{p=1}^{N_t} I_{k,l}^{(qp)}$ and the noise effect. Threshold determination requires knowledge of the noise-plus-interference energy. From (16), provided data symbols are independent from the channel coefficients, we can compute the interference energy as follows:



(a) Rx symbol pattern at one antenna of MIMO-OTFS system (■: data detection, symbols in the yellow, green, and blue boxes: CE for Tx antenna 1, 2, and 3, respectively).



(b) Block diagram for the proposed CE and data detection scheme for MIMO-OTFS systems.

Figure 3: (a): Rx symbol pattern at one antenna of MIMO-OTFS system. (b): Block diagram for the proposed CE and data detection scheme for MIMO-OTFS systems.

$$\mathbb{E}\left\{|\sum_{p=1}^{N_t} I_{k,l}^{(qp)}|^2\right\} = \sum_{p=1}^{N_t}\sum_{i\in\mathcal{S}_{k,l}} \mathbb{E}\{|h_i^{(qp)}|^2\}\mathbb{E}\{|\boldsymbol{X}_{DD}^{(p)}[[k-k_i]_N,[l-l_i]_M]|^2\} + \sum_{p=1}^{N_t}\sum_{\substack{i'\in\mathcal{S}_{k,l}\\ i\neq i'}} \mathbb{E}\{h_i^{(qp)}(h_i^{(qp)})^\star\}$$
$$\times\mathbb{E}\{\boldsymbol{X}_{DD}^{(p)}[[k-k_i]_N,[l-l_i]_M](\boldsymbol{X}_{DD}^{(p)})^\star[[k-k_{i'}]_N,[l-l_{i'}]_M]\}\exp\left(-j2\pi\frac{k_il_i-k_{i'}l_{i'}}{MN}\right). \tag{18}$$

If we assume that the channel coefficients are independent for the different paths, we get $\mathbb{E}\{h_i^{(qp)}(h_{i'}^{(qp)})^\star\} = 0$ for $i \neq i'$. Then, (18) can be simplified as $\mathbb{E}\left\{|\sum_{p=1}^{N_t} I_{k,l}^{(qp)}|^2\right\} = \sigma_d^2\sum_{p=1}^{N_t}\sum_{i\in\mathcal{S}_{k,l}}\mathbb{E}\{|h_i^{(qp)}|^2\}$, where $\sigma_d^2 = \mathbb{E}\{|\boldsymbol{X}_{DD}^{(p)}[k,l]|^2\}$. If we further assume that the channel coefficients follow an exponential power delay profile (PDP) (the Extended Vehicular A model [37] of the 3GPP, the standardization body for 5G cellular communications is adopted) and their powers are normalized [38–40], yielding $\sum_{i\in\mathcal{S}_{k,l}}\mathbb{E}\{|h_i^{(qp)}|^2\} = 1$, we get $\mathbb{E}\{|\sum_{p=1}^{N_t} I_{k,l}^{(qp)}|^2\} = N_t\sigma_d^2$. As a result, the threshold expression will be given as follows:

$$\rho = 3\sqrt{N_t\sigma_d^2+\sigma^2}. \tag{19}$$

It is shown in [3] that the BER performance improves as the threshold increases, because at small threshold values, the path false alarm probability becomes higher (detection of non-existent paths becomes favourable) and the BER performance degrades. However, increasing the threshold above a certain value may cause the non-detection of some low power paths. Hence, there is an optimal threshold to balance the false alarm and miss detection probabilities. That's why the choice of 3 standard deviations for the threshold is optimal. Subsequently, the estimation of channel coefficients is performed as follows:

$$\hat{\boldsymbol{h}}^{(qp)}[[k-k_p]_N,[l-l_p-(p-1)(l_\tau+1)]_M]=\frac{\boldsymbol{Y}_{DD}^{(q)}[k,l]}{x_p}\mathbf{1}_{|\boldsymbol{Y}_{DD}^{(q)}[k,l]|\geq\rho}$$
$$k_p-k_\nu\leq k\leq k_p+k_\nu,\ l_p+(p-1)(l_\tau+1)\leq l\leq l_p+pl_\tau+p-1, \tag{20}$$

Equations (20) allow to estimate the effective channel $\hat{\boldsymbol{h}}^{(qp)}$ and its corresponding delay and Doppler taps $\{\hat{l}_i,\hat{k}_i\}$. By exploiting the 2D convolution property of the OTFS system, and the fact that $l_i=\tau_i M\Delta f$, and $k_i=\nu_i NT$, we can easily deduce the channel parameters $\{\hat{h}_i,\hat{\tau}_i,\hat{\nu}_i\}$. Then an initial estimate of the channel matrix $\boldsymbol{\mathcal{H}}$ is obtained, namely $\hat{\boldsymbol{\mathcal{H}}}^{(0)}$. Once the channel is estimated, an initial detection of data symbols is obtained using an adapted version of the MP algorithm proposed in [13, 14]. The aim here is to estimate the vector of data symbols $\boldsymbol{x}_d$ from the following model using $\hat{\boldsymbol{\mathcal{H}}}^{(0)}$ instead of $\boldsymbol{\mathcal{H}}$:

$$\boldsymbol{y}_d=\boldsymbol{y}-\hat{\boldsymbol{\mathcal{H}}}^{(0)}\boldsymbol{x}_p=\boldsymbol{\mathcal{H}}\boldsymbol{x}_d+\tilde{\boldsymbol{w}}, \tag{21}$$

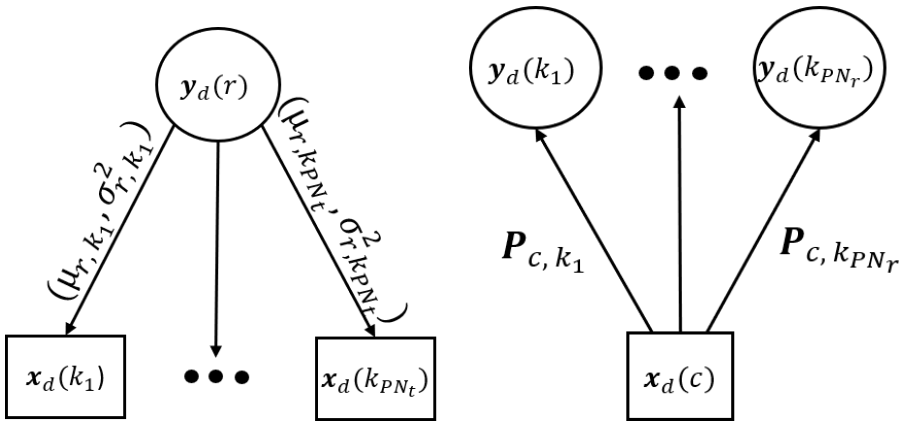
where $\boldsymbol{x}_p$ is an $MNN_t\times 1$ vector containing N_t non-zero elements corresponding to the pilot, $\boldsymbol{x}_d$ is a data symbols vector, and $\tilde{\boldsymbol{w}}=\tilde{\boldsymbol{n}}+(\boldsymbol{\mathcal{H}}-\hat{\boldsymbol{\mathcal{H}}}^{(0)})\boldsymbol{x}_p$ is a vector formed by the additive noise and the CE error $\boldsymbol{\mathcal{H}}-\hat{\boldsymbol{\mathcal{H}}}^{(0)}$.

Let us note that $\boldsymbol{\mathcal{H}}$ is sparse: each row and each column of $\boldsymbol{\mathcal{H}}$ contains only PN_t and PN_r non-zero elements, respectively. Therefore, the system is modelled as a sparsely connected factor graph with MNN_t variable nodes and MNN_r observation nodes corresponding to $\boldsymbol{x}_d$ and $\boldsymbol{y}_d$, respectively. The factor graph of the MP algorithm, which consists of observation nodes and variable nodes, is shown in Fig 4. Thanks to the channel sparsity in the DD domain, each variable node $\boldsymbol{x}_d(c)$ is relied to only PN_r observation nodes $\{\boldsymbol{y}_d(k_i),k_i\in\mathcal{I}_c\}$, and each observation node $\boldsymbol{y}_d(r)$ is relied to only PN_t variable nodes $\{\boldsymbol{x}_d(k_i),k_i\in\mathcal{I}_r\}$, where $\mathcal{I}_c$ and $\mathcal{I}_r$ denote the sets of non-zero indices in the c^{th} column and the r^{th} row of $\boldsymbol{\mathcal{H}}$. The application of the maximum a posteriori (MAP) criterion on (21) yields:

$$\hat{\boldsymbol{x}}_d=\underset{\boldsymbol{x}_d\in\mathcal{A}^{NMN_r}}{\arg\max}\ p(\boldsymbol{x}_d|\boldsymbol{y}_d,\hat{\boldsymbol{\mathcal{H}}}^{(0)}). \tag{22}$$

Problem (22) is a demanding combinatorial problem. Thus, the following symbol by symbol MAP criterion is used:

$$\begin{aligned}\hat{\boldsymbol{x}}_d(r)&=\underset{a_j\in\mathcal{A}}{\arg\max}\ p(\boldsymbol{x}_d(r)=a_j|\boldsymbol{y}_d,\hat{\boldsymbol{\mathcal{H}}}^{(0)}),\\&=\underset{a_j\in\mathcal{A}}{\arg\max}\ \frac{1}{|\mathcal{A}|}p(\boldsymbol{y}_d|\boldsymbol{x}_d(r)=a_j,\hat{\boldsymbol{\mathcal{H}}}^{(0)}),\\&\approx\underset{a_j\in\mathcal{A}}{\arg\max}\ \prod_{k\in\mathcal{I}_r}p(\boldsymbol{y}_d(k)|\boldsymbol{x}_d(r)=a_j,\hat{\boldsymbol{\mathcal{H}}}^{(0)}),\end{aligned} \tag{23}$$



(a) Observation node messages (b) Variable node messages

Figure 4: Messages in factor graph.

for $r = 0 : NMN_t - 1$ and $k = 0 : NMN_r - 1$. It is important to keep in mind that the entries of the vector $\boldsymbol{y}_d$ are independent for a given $\boldsymbol{x}_d(k)$ due to the sparsity of the matrix $\boldsymbol{\mathcal{H}}$. It is assumed that all entries in the vector $\boldsymbol{x}_d$ are equally likely. The MP algorithm, which is described below, can be used to solve problem (23). The message that passes to $\boldsymbol{y}_d(c)$, for $c \in \mathcal{I}_r$, from $\boldsymbol{x}_d(r)$, for $r = 0 : NMN_t - 1$ is the probability mass function (pmf) given as:

$$\boldsymbol{p}_{rc} = \{p_{rc}(a_i) | a_i \in \mathcal{A}\}. \tag{24}$$

The r^{th} observation node $\boldsymbol{y}_d(r)$ is related to the c^{th} variable node $\boldsymbol{y}_d(c)$ by the following formula:

$$\boldsymbol{y}_d(r) = \boldsymbol{x}_d(c)\hat{\boldsymbol{\mathcal{H}}}^{(0)}(r,c) + \psi_{rc}, \tag{25}$$

where $\hat{\boldsymbol{\mathcal{H}}}^{(0)}(r,c)$ is the $(r,c)^{th}$ element of $\hat{\boldsymbol{\mathcal{H}}}^{(0)}$ and ψ_{rc} denotes the interference-plus-noise term whose expression is given as follows:

$$\psi_{rc} = \sum_{k \in \mathcal{I}_r, k \neq c} \boldsymbol{x}_d(k)\hat{\boldsymbol{\mathcal{H}}}^{(0)}(r,k) + \tilde{\boldsymbol{w}}(r). \tag{26}$$

The term ψ_{rc} is approximated by a Gaussian random variable because $\tilde{\boldsymbol{w}}(r)$ follows a Gaussian distribution and, by using the central-limit theorem, $\sum_{k \in \mathcal{I}_r, k \neq c} \boldsymbol{x}_d(k)\hat{\boldsymbol{\mathcal{H}}}^{(0)}(r,k)$ is also approximated by a Gaussian distribution. In the m^{th} iteration of MP, messages that goes from the observation node $\boldsymbol{y}_d(r)$ to the variable node $\boldsymbol{x}_d(c)$ contain the mean $\mu_{rc}^{(m)}$ and the variance $(\sigma_{r,c}^{(m)})^2$ of the interference term (26), where $c \in \mathcal{I}_r$.

The major difference between the versions of the MP algorithm suggested in [13, 14] and the version proposed here lies in the calculation of $\mu_{rc}^{(m)}$ and $(\sigma_{r,c}^{(m)})^2$. Unlike the versions in [13, 14] where the interference

term contains only additive noise, in our case, this term is a function of additive noise and the mixture containing the symbols of pilots and the error of CE. This mixture is not present in earlier versions of the MP algorithm due to the use of guard intervals between pilot and data symbols. The second difference is that these earlier versions of the MP algorithm assume perfect knowledge of CSI, i.e., the MIMO matrix $\boldsymbol{\mathcal{H}}$ is perfectly known. However, in our case, we do not know $\boldsymbol{\mathcal{H}}$ but only an estimate $\hat{\boldsymbol{\mathcal{H}}}$.

The expressions for the mean $\mu_{rc}^{(m)}$ and variance $(\sigma_{r,c}^{(m)})^2$ of the interference term are given as follows:

$$\mu_{r,c}^{(t)} = \sum_{k \in \mathcal{I}_r, k \neq c} \sum_{i=1}^{|\mathcal{A}|} a_i p_{r,k}^{(m-1)}(a_i) \hat{\boldsymbol{\mathcal{H}}}(r,k) + m_{\tilde{\boldsymbol{w}}(k)}, \tag{27}$$

$$(\sigma_{r,c}^{(t)})^2 = \sum_{k \in \mathcal{I}_r, k \neq c} \sum_{i=1}^{|\mathcal{A}|} |a_i|^2 p_{r,k}^{(m-1)}(a_i) |\hat{\boldsymbol{\mathcal{H}}}(r,k)|^2 - |\mu_{r,c}^{(m)}|^2 + var\{\tilde{\boldsymbol{w}}(k)\}, \tag{28}$$

where $m_{\tilde{\boldsymbol{w}}(k)}$ and $var\{\tilde{\boldsymbol{w}}(k)\}$ are the mean and the variance of $\tilde{\boldsymbol{w}}(k)$ whose expressions are derived below. We have $\tilde{\boldsymbol{w}} = \tilde{\boldsymbol{n}} + (\boldsymbol{\mathcal{H}} - \hat{\boldsymbol{\mathcal{H}}})\boldsymbol{x}_p$ and we know that $\boldsymbol{x}_p$ is a vector containing only N_t non-zero elements whose value of each of these elements is equal to x_p. Therefore, the k^{th} element of $\tilde{\boldsymbol{w}}$ can take the following form:

$$\tilde{\boldsymbol{w}}(k) = \tilde{\boldsymbol{n}}(k) + \sum_{i \in \mathcal{I}_l} (\boldsymbol{\mathcal{H}}(k,i) - \hat{\boldsymbol{\mathcal{H}}}(k,i)), \tag{29}$$

where $\mathcal{I}_l$ is the set in the DD domain the pilots contribute to. Since the elements of $\tilde{\boldsymbol{n}}$ are centred and have a variance σ^2 and since x_p have zero mean and variance σ_p^2, we can establish the values of $m_{\tilde{\boldsymbol{w}}(k)}$ and $var\{\tilde{\boldsymbol{w}}(k)\}$ as follows:

$$\begin{aligned} m_{\tilde{\boldsymbol{w}}(k)} &= 0, \\ var\{\tilde{\boldsymbol{w}}(k)\} &= \sigma^2 + \sigma_p^2 \sum_{i \in \mathcal{I}_l} |\boldsymbol{\mathcal{H}}(k,i) - \hat{\boldsymbol{\mathcal{H}}}(k,i)|^2. \end{aligned} \tag{30}$$

After that, the entries of the pmf vector $\boldsymbol{P}_{r,c}^{(m)}$ are passed from the variable node $\boldsymbol{x}_d(c)$ to the observation node $\boldsymbol{y}_d(r)$. The elements of this pmf vector are calculated using a damping method proposed in [35]:

$$p_{c,r}^{(m)}(a_i) = \Delta p_{c,r}^{(m-1)}(a_i) + (1-\Delta) p_{c,r}^{(m-2)}(a_i), \tag{31}$$

where $\Delta \in (0,1]$ is the damping factor allowing the improvement of the convergence of MP algorithm, and

$$p_{c,r}^{(m)} \propto \prod_{k \in \mathcal{I}_c, k \neq c} \exp\left(- \frac{|\boldsymbol{y}_d(k) + \mu_{k,c}^{(m)} + \hat{\boldsymbol{\mathcal{H}}}(k,c) a_i|}{(\sigma_{r,c}^{(m)})^2} \right). \tag{32}$$

The MP algorithm stops if the maximum number of iterations m_{iter} is reached or if $|p_{c,r}^{(m)} - p_{c,r}^{(m-1)}| < \epsilon$

(for all (r, c) pairs), where ϵ is a small value. Finally, the decision on the detected data symbols is given by

$$\hat{\boldsymbol{x}}_d(c) = \underset{a_i \in \mathcal{A}}{\arg\max}\, p_{r,c}^{(m)}(a_i),\ c = 0 : NMN_t - 1. \tag{33}$$

The detected data symbols $\hat{\boldsymbol{X}}_{DD}^{(0)}[k,l]$ and the initial channel estimate $\{(\hat{\boldsymbol{h}}^{(qp)})\}^{(0)}$ for $p = 1 : N_t$ and for $q = 1 : N_r$ will then be used to cancel the received signal interference caused by the data symbols. The residual term after interference cancellation takes the following form:

$$\tilde{\boldsymbol{Y}}_{DD}^{(q)}[k,l] = \hat{\boldsymbol{Y}}_{DD}^{(q)}[k,l] - \sum_{p=1}^{N_t} \sum_{k'} \sum_{l'} \hat{\boldsymbol{X}}_d^{(p)}[k',l'] \hat{\boldsymbol{h}}^{(qp)}[[k-k']_N, [l-l']_M]. \tag{34}$$

In the case where the interferences are perfectly eliminated, we can easily get a good estimate of the DD channel by the threshold method using the classical threshold $\rho = 3\sigma$, because the residual term in this case contains only the pilot and the additive noise. In our case, due to the noise influence, the interference cancellation is imperfect. In this case, the iterative scheme of Fig 3b is proposed to perform CE and data detection in an iterative way. With this scheme, a good CE allows for a good detection of data symbols and an efficient interference cancellation for performing the threshold-based CE.

In this algorithm, the data symbols at the m^{th} iteration $\hat{\boldsymbol{X}}_d^{(m)}[k,l]$ are detected using the MP algorithm by feeding it with the DD channel estimated at iteration $m-1$, i.e., $(\hat{\boldsymbol{h}}^{(ji)})^{(m)}[k,l]$. Then, $\hat{\boldsymbol{X}}_d^{(m)}[k,l]$ and $(\hat{\boldsymbol{h}}^{(ji)})^{(m)}[k,l]$ will be used for interference cancellation (34). Finally, a threshold-based refined CE technique will be applied. Specifically, after the application of the interference cancellation scheme, the threshold ρ used in the CE process need to be updated. The number of channel paths is unknown at the receiver. As a compromise, we consider a relatively large number P' in order to avoid channel paths miss detection with small gains. Assuming that we can estimate Q paths instead of P' paths at the $(m-1)^{th}$ iteration. In this case, after the application of interference cancellation scheme, there are still other interferences related to the error $P' - Q$ about the number of estimated paths. The energy of these residual interferences can be approximated by $((P'-Q)/P')N_t\sigma_d^2$. Therefore, the new threshold for the threshold-based refined CE at the m^{th} iteration is given as follows:

$$\rho^{(m)} = 3\sqrt{\left(\frac{P'-Q}{P'}\right) N_t\sigma_d^2 + \sigma^2}. \tag{35}$$

The iterative algorithm ends if the number of iterations reaches the maximum number M_{max} or if the performance gain of CE by running more iterations is marginal. The proposed algorithm is summarized in Algorithm 1.

Algorithm 1 PROPOSED ALGORITHM FOR CHANNEL ESTIMATION AND DATA DETECTION.

Input: x_p, $\{\boldsymbol{Y}_{DD}^{(j)}\}_{j=1:N_r}$, M_{max}, N_t, N_r, M, N, m_{iter}, e

Calculate the decision threshold: $\rho^{(0)} = 3\sqrt{N_t\sigma_d^2 + \sigma^2}$

Calculate an initial channel estimate: $(\hat{\boldsymbol{h}}^{(ji)})^{(0)}$:

if $|\boldsymbol{Y}_{DD}^{(j)}[k,l]| > \rho^{(0)}$ **then**

$(\hat{\boldsymbol{h}}^{(ji)})^{(0)}[[k-k_p]_N, [l-l_p-(i-1)(l_\tau+1)]_M] = \boldsymbol{Y}_{DD}^{(j)}[k,l]/x_p$

end if

while $m < M_{max}$ & $\|\hat{\boldsymbol{h}}^{(m)} - \hat{\boldsymbol{h}}^{(m-1)}\| < e$ **do**

Use the MP algorithm (22)-(33) for symbol detection:

$\left(\{\hat{\boldsymbol{x}}_d^{(i)}\}_{i=1:N_t}\right)^{(m)} = MP\left((\hat{\boldsymbol{h}}^{(ji)})^{(m-1)}, \{\boldsymbol{Y}_{DD}^{(j)}\}_{j=1:N_r}, m_{\tilde{\boldsymbol{w}}(k)}, var_{\tilde{\boldsymbol{w}}(k)}, m_{iter}\right)$

Apply the interference cancellation scheme:

$\tilde{\boldsymbol{Y}}_{DD}^{(j)}[k,l] = \hat{\boldsymbol{Y}}_{DD}^{(j)}[k,l] - \sum_{i=1}^{N_t}\sum_{k'}\sum_{l'}\left(\hat{\boldsymbol{X}}_d^{(i)}\right)^{(m)}[k',l'](\hat{\boldsymbol{h}}^{(ji)})^{(m-1)}[[k-k']_N, [l-l']_M]$

Update the threshold: $\rho^{(m)} = 3\sqrt{((P'-Q)/P')N_t\sigma_d^2 + \sigma^2}$

Calculate the refined channel estimate:

if $|\tilde{\boldsymbol{Y}}_{DD}^{(j)}[k,l]| > \rho^{(m)}$ **then**

$(\hat{\boldsymbol{h}}^{(ji)})^{(m)}[[k-k_p]_N, [l-l_p-(i-1)(l_\tau+1)]_M] = \tilde{\boldsymbol{Y}}_{DD}^{(j)}[k,l]/x_p$

end if

end while

Output: $\{\hat{\boldsymbol{h}}^{(ji)}\}_{i=1:N_t,j=1:N_r}$, $\{\hat{\boldsymbol{x}}_d^{(i)}\}_{i=1:N_t}$.

5. Computational complexity and numerical results

In this section, we evaluate the computational complexity and the performance of ICEDD in terms of CE-NMSE, BER, and SE. Then, we compare obtained results against five state-of-the-art methods.

The parameters of the simulations are given in Table 2. The DD channel is modelled as follows: the delay is a 5-tap model whose parameters are given in [35]. Each delay tap corresponds to a single Doppler shift in the form $\nu_i = \nu_{max}cos(\theta_i)$, where $\theta_i \in [0, \pi]$ and ν_{max} is the maximum Doppler shift of the channel. The maximum delay tap $l_\tau = 4$ and the maximum Doppler tap $k_\nu = 2$ correspond to a high mobility scenario with a maximum relative speed $v_{max} = 500\ km/h$.

Table 2: Simulation parameters.

Parameter	Value	Parameter	Value
$(N,\ M)$	(16, 16)	$(t_c(\mu s),\ T(ms))$	(541, 1)
$(N_t,\ N_r)$	(2, 2)	$v_{max}(km/h)$	500
$(N_p,\ M_p)$	(4, 5)	P	5
$(f_c(GHz),\ \Delta f(kHz))$	(4, 15)	Modulation	BPSK
$(k_\nu,\ l_\tau)$	(2, 4)	Pulse shaping	Rectangular

5.1. Computational complexity

ICEDD iterates between two steps: a CE step and a data detection step. Thus, the complexity of the proposed algorithm is written as follows: $C = n_{iter}(C_{CE} + C_{DD})$, where C_{CE} and C_{DD} denote the overall

computational complexity of the CE step and the data detection step, respectively. For the sake of simplicity, we assume that $N_t = N_r = N_a$.

The complexity of one iteration of the detection algorithm requires the computation of (27), (28), (31), and (33). Each of (27), (28), and (31) has a complexity of $\mathcal{O}(MNN_aP|\mathcal{A}|)$. In addition, the complexity of (33) is $\mathcal{O}(MNN_a|\mathcal{A}|)$. Therefore, the overall complexity of the data detection algorithm is dominated by $C_{DD} = O(m_{iter}MNN_aP|\mathcal{A}|)$, where m_{iter} is the number of iterations required to the convergence of the data detection algorithm. For the CE step, the computational complexity is dominated by $C_{CE} = \mathcal{O}(l_\tau N_a N)$. The overall computational complexity of the proposed algorithm for CE and data detection is given by $C = n_{iter}(\mathcal{O}(l_\tau N_a N) + \mathcal{O}(m_{iter}MNN_aP|\mathcal{A}|))$.

The computational complexity of each of the state-of-the-art methods is presented in the Table 3. Letting $K = (2k_\nu+2Q+1)(l_\tau+1)$, $G = (M_\tau+1)(G_\nu+1)$ and noting that C_m represents the computational complexity of the method m.

Table 3: Computational complexities of the proposed and state-of-the-art CE schemes.

Method	**Computational complexity**
EP [3]	$\mathcal{O}(l_\tau N_a N)$
BSBL-BR [16]	$\mathcal{O}(N_{iter}N_a^3LK^2)$
RG-OMP [15]	$\mathcal{O}(M^3N_p^3)$
RG-BL [15]	$\mathcal{O}(G^3N_a^3)$
ICEDD (CE)	$\mathcal{O}(n_{iter}l_\tau N_a N)$

It can be seen that the proposed algorithm and the EP method are the least complex because their complexities vary linearly with N. The EP scheme is the least complex because it is executed once while ICEDD requires n_{iter} iterations to converge. It can be also seen that RG-OMP is the most complex because its complexity is proportional to M^3.

In practice, $l_\tau < M_\tau < M$, $k_\nu < N_\nu < N$, $N_\nu << G_\nu$, and $N_{iter} = 10$. Therefore, $N_{iter}LK^2 < G^3$, leading to $C_{BSBL-BR} < C_{RG-BL}$. Since $n_{iter} \approx 4$, we can easily check $n_{iter}l_\tau N < N_{iter}LK^2N_a^2$, thus, $C_{ICEDD} < C_{BSBL-BR}$. Finally, we can classify the complexities of the five methods as follows: $C_{EP} \approx C_{ICEDD} < C_{BSBL-BR} < C_{RG-BL} < C_{RG-OMP}$.

5.2. Numerical results

Table 4 shows the expressions and values of the proposed scheme pilot overhead as well as those of the state-of-the-art methods CPA [14], EP [3], RG-BL and RG-OMP [15], BSBL-BR [16], OMP [31], and SoBaP [32]. SE refers to the rate of information that can be transmitted over a given bandwidth in a specific communication system. By definition, its expression can be given as follows:

$$SE = \frac{D}{B} \text{ (bits/s/Hz)}, \tag{36}$$

where D and B denote the bit rate and bandwidth, respectively. For an OTFS system, D and B are expressed as follows:

$$D = \frac{N_b}{T} \text{ and } B = M\Delta f, \tag{37}$$

where N_b is the number of information bits sent on an OTFS frame. For M-ary modulation (QAM or MPSK), $N_b = N_s \log |\mathcal{A}|$, where N_s is the number of information symbols sent on an OTFS frame.

When N_p symbols are reserved for pilots and N_g symbols for guard intervals, $N_s = MN - N_p - N_g$. Thus, the SE expression can be given as follows:

$$SE = \frac{(MN - N_p - N_g) \log |\mathcal{A}|}{M\Delta f NT}. \tag{38}$$

Since $T\Delta f = 1$, the SE of a scheme s is finally expressed as follows:

$$SE = (1 - \eta_s) \log |\mathcal{A}|, \tag{39}$$

where $\eta_s = (Np + N_g)/NM$ is the pilot overhead of scheme s. Therefore, for a given constellation (BPSK is used in our case), the scheme with a small pilot overhead is the scheme with the best SE. Consequently, from Table 4, we can see that the designed scheme is the best compared to the state-of-the-art methods in terms of SE.

Table 4: Pilot overhead of the proposed and state-of-the-art schemes.

Scheme (s)	**Pilot overhead (η_s)**	**Numerical value**
CPA [14]	$MN/2MN$	50 %
EP [3]	$(N_G + 1)/MN$	48 %
RG-OMP + RG-BL [15]	$N_p/(N_p + N)$	20 %
BSBL-BR [16]	L/MN	9 %
OMP [31] + SobaP [32]	$N_t(l_\tau + M_p)/MN$	7 %
ICEDD (CE)	$0/MN$	0 %

The signal-to-noise ratio and the pilot signal-to-noise ratio are defined as follows:

$$SNR = \frac{\sigma_d^2}{\sigma^2} \text{ and } SNR_p = \frac{\sigma_p^2}{\sigma^2}. \tag{40}$$

Fig 5 shows the BER performance of the proposed ICEDD algorithm versus the number of iterations with $SNR = 5$, 10 dB, $\sigma_p^2/\sigma_d^2 = 20$ dB, and BPSK modulation. From Fig 5, it can be seen that, for both

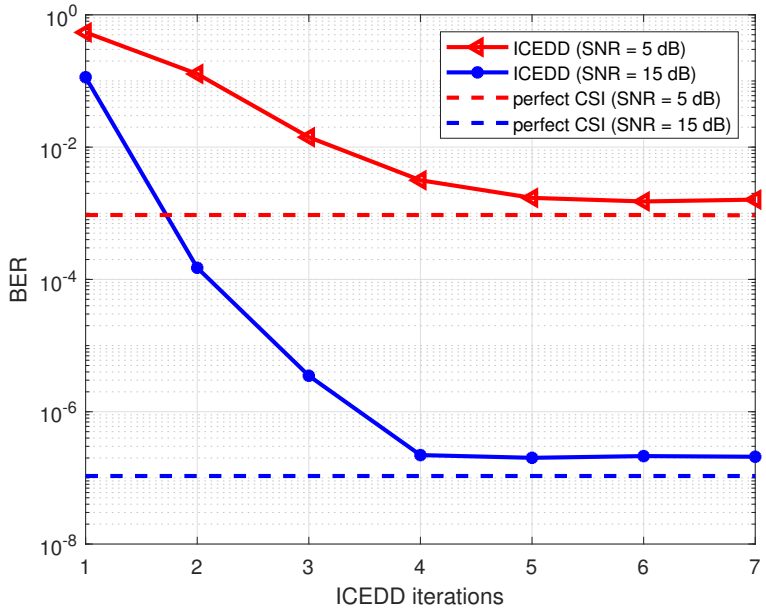
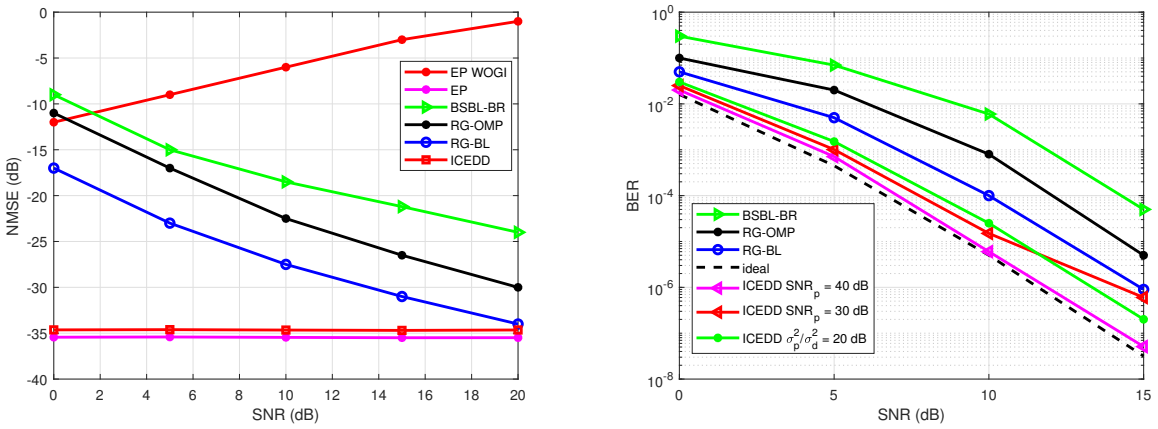


Figure 5: BER performance of the ICEDD algorithm versus the number of iterations.

SNR values, the BER decreases with increasing number of iterations, and it saturates after about 4 or 5 iterations. The convergence and the effectiveness of the proposed ICEDD algorithm are confirmed by the steady states reached by the two curves of the BER after a few iterations. It can be also observed that, as expected, the scenario with higher SNR exhibits superior convergence and detection accuracy.

We now investigate the variation of NMSE as a function of SNR: $NMSE = 1 - (|\boldsymbol{h}^H \hat{\boldsymbol{h}}|/\|\boldsymbol{h}\|_2 \|\hat{\boldsymbol{h}}\|_2)^2$. We assume that $SNR_p = 40$ dB. Fig 6a shows a comparison between ICEDD and EP [3], BSBSL-BR [16], RG-OMP and RG-BL [15] methods in terms of NMSE. We observe that ICEDD has the lowest NMSE value compared to BSBSL-BR [16], RG-OMP, and RG-BL [15] methods. We also observe that the ICEDD and the EP method has a similar NMSE performance. However, the EP method uses guard intervals and the proposed scheme improves the SE (see Table 4). This improvement becomes more important in a massive MIMO-OTFS and in MU-MIMO-OTFS systems, where more guard symbols are required for avoiding inter-user interference. To highlight the influence of neglecting the interferences on the performance in terms of NMSE of the EP scheme, we consider the EP scheme but without the use of guard intervals (EP-WOGI) and threshold $\rho = 3\sigma$. We see that the NMSE performance of the EP-WOGI scheme is weaker than that of ICEDD and other state-of-the-art methods. Its performance degrades with the increase of SNR because the interference suffered by the pilot symbol becomes more severe. Therefore, even though EP-WOGI has the same SE as the proposed scheme, its NMSE performance is degraded. Consequently, neglecting the interference severely degrades the CE performance.



(a) NMSE versus SNR

(b) BER versus SNR

Figure 6: NMSE and BER performances of the proposed and state-of-the-art methods.

We now investigate the variation of BER as a function of SNR. We consider two distinct values for SNR_p, 30 and 40 dB. Fig 6b shows a comparison between ICEDD, BSBSL-BR [16], RG-OMP, and RG-BL [15]. When $SNR_p = 40$ dB, the performance of ICEDD is close to that of the oracle (perfectly known CSI). When $SNR_p = 30$ dB, the performance of the proposed algorithm degrades at high SNRs. This degradation is due to the CE error caused by a high interference energy to the fixed pilot symbol energy. To avoid this problem, we opt for an adaptive pilot symbol energy. We plot the BER performance of the proposed algorithm by using $\sigma_p^2/\sigma_d^2 = 20$ dB. In this scenario, SNR_p also rises with the increase in SNR, improving the BER performance as well as the accuracy of CE.

In Fig 7, we investigate the BER performance of the ICEDD algorithm for different user speeds with $\sigma_p^2/\sigma_d^2 = 20$ dB, $l_\tau = 4$, and BPSK modulation. Consider user speeds of 250, 500, and 1000 km/h corresponding to maximum Doppler tap $k_\nu = 1$, 2, and 4, respectively. From Fig 7, it is easy to observe that the proposed ICEDD algorithm exhibits highly similar performance under different mobilities. This is due to the fact that OTFS is insensitive to Doppler frequency shifts. This property allows OTFS to distinguish similar scattering paths with different velocities. Therefore, a similar BER performance of the ICEDD algorithm is observed for different user speeds.

6. Conclusion

We have proposed an algorithm for MIMO-OTFS channel estimation and data detection for a system with superimposed pilot symbols. The proposed algorithm iterates between data-aided channel estimation and message-passing-aided data detection. The channel estimation step is based on a simple threshold method and the data detection step is based on a message-passing algorithm. We have derived an optimal threshold

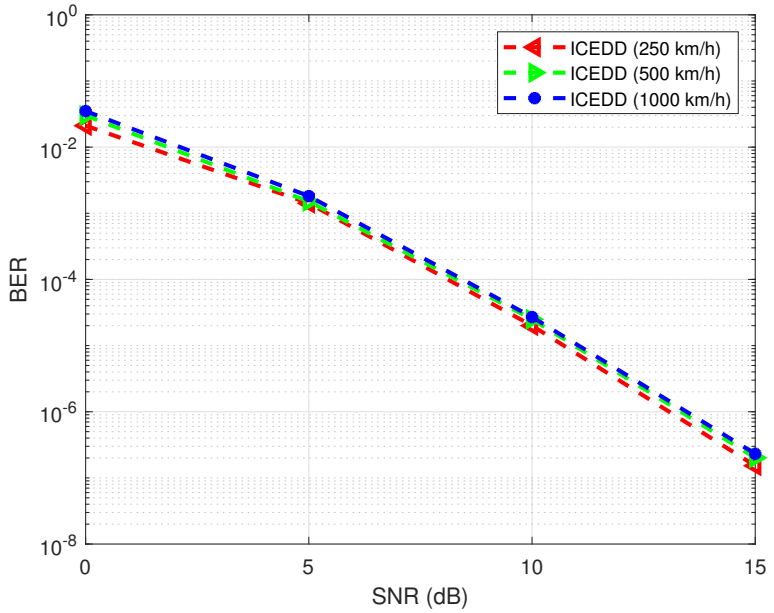


Figure 7: BER performance of the ICEDD algorithm with different mobilities.

expression for the channel estimation method proposed in this work. Unlike classical methods [3], that involve guard interval and fixed threshold for the detection of channel paths, for the determination of the threshold we take into account the additive noise and the interference caused by data symbols. To improve the channel estimation efficiency, we have proposed an interference cancellation scheme which is executed at each iteration of the algorithm. The comparison conducted in high-mobility environment against state-of-the-art methods has shown that the proposed scheme achieves a good compromise between computational complexity and performance in terms of channel estimation NMSE, BER and spectral efficiency.

References

[1] Y. Zhang, Q. Zhang, C. He, Y. Zhou, L. Jing, Delay-doppler domain decision feedback turbo equalization for OTFS modulation, Physical Communication 52 (2022) 101699.

[2] R. Melki, H. N. Noura, M. M. Mansour, A. Chehab, A survey on OFDM physical layer security, Physical Communication 32 (2019) 1–30.

[3] P. Raviteja, K. T. Phan, Y. Hong, Embedded pilot-aided channel estimation for OTFS in delay–doppler channels, IEEE transactions on vehicular technology 68 (5) (2019) 4906–4917.

[4] A. Monk, R. Hadani, M. Tsatsanis, S. Rakib, OTFS-orthogonal time frequency space, arXiv preprint arXiv:1608.02993.

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

[6] S. M. Pishvaei, B. M. Tazehkand, J. Pourrostam, Design and performance evaluation of FBMC-based orthogonal time–frequency space (OTFS) system, Physical Communication 53 (2022) 101723.

[7] H. Bitra, P. Ponnusamy, S. Chintagunta, S. Pragadeshwaran, Nonlinear companding transforms for reducing the PAPR of OTFS signal, Physical Communication 53 (2022) 101729.

[8] G. Surabhi, R. M. Augustine, A. Chockalingam, On the diversity of uncoded OTFS modulation in doubly-dispersive channels, IEEE transactions on wireless communications 18 (6) (2019) 3049–3063.

[9] Z. Ding, R. Schober, P. Fan, H. V. Poor, OTFS-NOMA: An efficient approach for exploiting heterogenous user mobility profiles, IEEE Transactions on Communications 67 (11) (2019) 7950–7965.

[10] R. M. Augustine, G. Surabhi, A. Chockalingam, Space-time coded OTFS modulation in high-doppler channels, in: 2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring), IEEE, 2019, pp. 1–6.

[11] G. Surabhi, A. Chockalingam, Low-complexity linear equalization for 2× 2 MIMO-OTFS signals, in: 2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), IEEE, 2020, pp. 1–5.

[12] M. Li, S. Zhang, F. Gao, P. Fan, O. A. Dobre, A new path division multiple access for the massive MIMO-OTFS networks, IEEE journal on selected areas in communications 39 (4) (2020) 903–918.

[13] P. Raviteja, K. T. Phan, Y. Hong, E. Viterbo, Interference cancellation and iterative detection for orthogonal time frequency space modulation, IEEE transactions on wireless communications 17 (10) (2018) 6501–6515.

[14] M. K. Ramachandran, A. Chockalingam, MIMO-OTFS in high-doppler fading channels: Signal detection and channel estimation, in: 2018 IEEE Global Communications Conference (GLOBECOM), IEEE, 2018, pp. 206–212.

[15] S. Srivastava, R. K. Singh, A. K. Jagannatham, L. Hanzo, Bayesian learning aided simultaneous row and group sparse channel estimation in orthogonal time frequency space modulated MIMO systems, IEEE Transactions on Communications 70 (1) (2021) 635–648.

[16] L. Zhao, J. Yang, Y. Liu, W. Guo, Block sparse bayesian learning-based channel estimation for MIMO-OTFS systems, IEEE Communications Letters 26 (4) (2022) 892–896.

[17] I. A. Khan, S. K. Mohammed, Low complexity channel estimation for OTFS modulation with fractional delay and doppler, arXiv preprint arXiv:2111.06009.

[18] M. Zhang, F. Wang, X. Yuan, L. Chen, 2D structured turbo compressed sensing for channel estimation in OTFS systems, in: 2018 IEEE International Conference on Communication Systems (ICCS), IEEE, 2018, pp. 45–49.

[19] O. K. Rasheed, G. Surabhi, A. Chockalingam, Sparse delay-doppler channel estimation in rapidly time-varying channels for multiuser OTFS on the uplink, in: 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), IEEE, 2020, pp. 1–5.

[20] F. Gómez-Cuba, Compressed sensing channel estimation for OTFS modulation in non-integer delay-doppler domain, arXiv preprint arXiv:2111.12382.

[21] K. Murali, A. Chockalingam, On OTFS modulation for high-doppler fading channels, in: 2018 Information Theory and Applications Workshop (ITA), IEEE, 2018, pp. 1–10.

[22] W. Shen, L. Dai, J. An, P. Fan, R. W. Heath, Channel estimation for orthogonal time frequency space (OTFS) massive MIMO, IEEE Transactions on Signal Processing 67 (16) (2019) 4204–4217.

[23] Y. Liu, S. Zhang, F. Gao, J. Ma, X. Wang, Uplink-aided high mobility downlink channel estimation over massive MIMO-OTFS system, IEEE Journal on Selected Areas in Communications 38 (9) (2020) 1994–2009.

[24] X. Wu, S. Ma, X. Yang, Tensor-based low-complexity channel estimation for mmwave massive MIMO-OTFS systems, Journal of Communications and Information Networks 5 (3) (2020) 324–334.

[25] D. Shi, W. Wang, L. You, X. Song, Y. Hong, X. Gao, G. Fettweis, Deterministic pilot design and channel estimation for downlink massive MIMO-OTFS systems in presence of the fractional doppler, IEEE Transactions on Wireless Communications 20 (11) (2021) 7151–7165.

[26] T. Li, C. Han, R. Yao, Y. Fan, X. Zuo, Low pilot overhead channel estimation for CP-OFDM-based massive MIMO OTFS system, IET Communications 16 (10) (2022) 1071–1082.

[27] Y. Liu, Y. L. Guan, D. González, Near-optimal BEM OTFS receiver with low pilot overhead for high-mobility communications, IEEE Transactions on Communications 70 (5) (2022) 3392–3406.

[28] F. Zhang, W. Ji, L. Qiu, Channel estimation for massive MIMO-OTFS systems via sparse bayesian learning with 2-D local beta process, in: 2022 IEEE Wireless Communications and Networking Conference (WCNC), IEEE, 2022, pp. 1383–1388.

[29] R. Muzavazi, O. O. Oyerinde, Channel estimation and data detection schemes for orthogonal time frequency space massive MIMO systems, Computers and Electrical Engineering 102 (2022) 108215.

[30] R. Mallaiah, K. Vejandla, V. Mani, A. Kumar, M. Sellathurai, An embedded pilot power based channel estimation and low-complexity feedback equalization scheme for OTFS system, Physical Communication 55 (2022) 101875.

[31] S. Wang, J. Guo, X. Wang, W. Yuan, Z. Fei, Pilot design and optimization for OTFS modulation, IEEE Wireless Communications Letters 10 (8) (2021) 1742–1746.

[32] R. Ouchikh, A. Aïssa-El-Bey, T. Chonavel, M. Djeddou, Sparse channel estimation algorithms for OTFS system, IET Communications 16 (18) (2022) 2158–2170.

[33] M. Li, S. Zhang, Y. Ge, F. Gao, P. Fan, Joint channel estimation and data detection for hybrid RIS aided millimeter wave OTFS systems, IEEE Transactions on Communications 70 (10) (2022) 6832–6848.

[34] W. Yuan, S. Li, Z. Wei, J. Yuan, D. W. K. Ng, Data-aided channel estimation for OTFS systems with a superimposed pilot and data transmission scheme, IEEE wireless communications letters 10 (9) (2021) 1954–1958.

[35] H. B. Mishra, P. Singh, A. K. Prasad, R. Budhiraja, OTFS channel estimation and data detection designs with superimposed pilots, IEEE transactions on wireless communications 21 (4) (2021) 2258–2274.

[36] R. Ouchikh, A. Aïssa-El-Bey, T. Chonavel, M. Djeddou, Iterative channel estimation and data detection algorithm for OTFS modulation, in: 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2022, pp. 5263–5267.

[37] S. Ahmadi, LTE-Advanced: a practical systems approach to understanding 3GPP LTE releases 10 and 11 radio access technologies, Academic Press, 2013.

[38] T. Thaj, E. Viterbo, Low complexity iterative rake decision feedback equalizer for zero-padded OTFS systems, IEEE transactions on vehicular technology 69 (12) (2020) 15606–15622.

[39] T. Thaj, E. Viterbo, Low complexity iterative rake detector for orthogonal time frequency space modulation, in: 2020 IEEE Wireless Communications and Networking Conference (WCNC), IEEE, 2020, pp. 1–6.

[40] Y. Hong, T. Thaj, E. Viterbo, Delay-doppler communications: Principles and applications.