I/O Relation Estimation and Detection in MIMO-Zak-OTFS with Embedded Pilots

Abhishek Bairwa and Ananthanarayanan Chockalingam
Department of ECE, Indian Institute of Science, Bangalore

***Abstract*—In this paper, we consider the problem of input-output (I/O) relation estimation and signal detection in multiple-input multiple-output Zak-OTFS (MIMO-Zak-OTFS) systems. Towards this, we consider an embedded pilot framework, where pilot symbols meant for I/O relation estimation and data symbols co-exist in a given frame. We present an embedded pilot frame structure for a 2×2 MIMO-Zak-OTFS system and propose an I/O relation estimation scheme, considering different delay-Doppler (DD) pulse shaping filters, including sinc, Gaussian, and Gaussian-sinc filters. We evaluate the mean square error (MSE) performance of the proposed estimation scheme in Vehicular-A channel model with fractional delays and Dopplers. Our simulation results show that, due to its highly localized nature in the DD domain and hence low inter-symbol interference levels, Gaussian filter gives better MSE performance compared to the other two filters. Using the estimated I/O relation in a given frame, we detect the data symbols in the same frame using a local search based detection scheme. Our results show that, because of its good balance between main lobe and sidelobe characteristics, the Gaussian-sinc filter outperforms the other two filters in terms of bit error rate performance with estimated I/O relation.**

***Index Terms*—Delay-Doppler domain, MIMO-Zak-OTFS modulation, I/O relation estimation, signal detection, embedded pilots, delay-Doppler pulse shaping.**

I. Introduction

Orthogonal time frequency space (OTFS) is a promising modulation waveform for highly time-selective channels, which are expected in thenext generation mobile communication systems [1]-[3]. In multicarrier OTFS (MC-OTFS) introduced in [1], information symbols mounted in the delay-Doppler (DD) domain are converted to time domain (TD) in two steps, namely, DD domain to time-frequency (TF) domain conversion followed by TF domain to time domain conversion. In Zak transform based OTFS (Zak-OTFS) [4]-[6], transformation to a TD signal is carried out in a single step using inverse Zak transform. Zak-OTFS has been shown to be more robust to large delay and Doppler spreads compared to MC-OTFS [5]. The basic information carrier in Zak-OTFS is a quasi-periodic pulse in the DD domain. In order to limit the bandwidth and time duration of transmission, a DD domain pulse shaping filter is employed at the transmitter. A matching DD filter is used at the receiver. The choice of these filters influence the receiver performance. Sinc, Gaussian, root raised cosine and Gaussian-sinc pulse shaping filters have been considered in the Zak-OTFS literature [5], [7], [8].

DD channel estimation and signal detection are crucial tasks in OTFS receivers. Pilot symbols are sent in OTFS frames for the purpose of channel estimation. In MC-OTFS, frames with exclusive pilot, embedded pilot, and superimposed pilots have been widely considered [9], [10], [11]. Similar pilot frames have been used in Zak-OTFS also [5], [7], [12]. In Zak-OTFS, the end to end effective channel comprises of the DD spreads due to the transmit and receive DD filters in addition to the spreads caused by the physical channel. Estimation of this end to end effective channel in Zak-OTFS is termed as I/O relation estimation, which is needed for detection. Zak-OTFS allows the estimation of this relation without explicitly estimating the parameters of the physical channel, termed as 'model-free' approach, which is made possible because of the predictability attribute of Zak-OTFS [5].

I/O relation estimation schemes using model-free approach in Zak-OTFS have been reported in [5], [7], [8]. These works consider different types of pilot frames and DD filters, but mainly in single-input single-output (SISO) settings. I/O relation estimation and detection in Zak-OTFS in multiple-input multiple-output (MIMO) settings remains largely unexplored, which forms the main focus in this paper. Our new contribution in this regard is that we investigate the problem of model-free I/O relation estimation in a MIMO setting under the embedded pilot framework, which has not been reported before. Our contributions in this paper can be summarized as follows.

- We present an embedded pilot frame structure for 2×2 MIMO-Zak-OTFS, where the frames sent by each transmit antenna consists of a pilot symbol and data symbols with carefully crafted guard space in between in such a way to limit the inter-antenna and inter-symbol interference to achieve good estimation performance.
- We propose an I/O relation estimation scheme for the above embedded frame and evaluate its mean square error (MSE) performance in the vehicular-A (Veh-A) channel with fractional delays and Dopplers [14], for sinc, Gaussian, and Gaussian-sinc DD filters. Our simulation results show that, due to its highly localized nature in the DD domain and hence low inter-symbol interference levels, Gaussian filter gives better MSE performance compared to the other two filters.
- Using the estimated I/O relation in a given frame, we detect the data symbols in the same frame using a local search based detection algorithm [13]. Our results show that, because of its good balance between main lobe and sidelobe characteristics, the Gaussian-sinc filter outperforms the other two filters in terms of bit error rate (BER) performance with estimated I/O relation.

This work was supported in part by the J. C. Bose National Fellowship, Department of Science and Technology, Government of India.

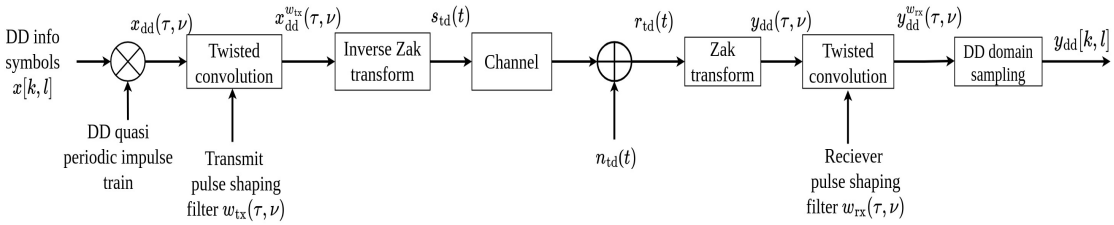

Fig. 1: Block diagram of Zak-OTFS transceiver.

II. MIMO-Zak-OTFS system model

This section introduces SISO-Zak-OTFS system model and extends it to a MIMO setting. The basic information carrier in Zak-OTFS is a quasi-periodic DD domain pulse localized in a fundamental DD period $\mathcal{D}_0$, defined as $\mathcal{D}_0 = \{(\tau,\nu) \mid 0 \le \tau < \tau_p, 0 \le \nu < \nu_p\}$, where τ and ν denote delay and Doppler variables, respectively, and $\tau_{\rm p}$ and $\nu_{\rm p}$ are delay and Doppler periods, respectively, such that $\tau_{\rm p}\nu_{\rm p} = 1$. The time domain representation of a quasi-periodic DD domain pulse is a pulsone, which is a time domain pulse train modulated by a frequency tone. The delay period $\tau_{\rm p}$ is sliced into M delay bins and the Doppler period $\nu_{\rm p}$ is sliced into N Doppler bins, and MN information symbols are mounted on MN DD pulses located in these MN DD bins. M and N are chosen such that $MN = BT$, where T and B are the time duration and bandwidth of transmission of a frame. That is, the delay resolution $\Delta\tau$ is $\Delta\tau = \frac{1}{B} = \frac{\tau_{\rm p}}{M}$ and the Doppler resolution $\Delta\nu$ is $\Delta\nu = \frac{1}{T} = \frac{\nu_{\rm p}}{N}$. To limit the time duration and bandwidth to T and B, respectively, a DD domain pulse shaping filter is used at the transmitter. Figure 1 shows the block diagram of the Zak-OTFS transceiver.

The MN information symbols, denoted by $x[k,l]$s, $k = 0,\cdots,M-1$, $l = 0,\cdots,N-1$, are drawn from a modulation alphabet $\mathbb{A}$. The DD grid on which these MN symbols are multiplexed is given by $\Lambda_{\rm dd} \triangleq \left\{\left(k\frac{\tau_p}{M}, l\frac{\nu_p}{N}\right) \mid k = 0,\cdots,M-1, l = 0,\cdots,N-1\right\}$, The $x[k,l]$s on $\Lambda_{\rm dd}$ are encoded as quasi-periodic discrete DD information as $x_{\rm dd}[k+nM, l+mN] = x[k,l]e^{j2\pi n\frac{l}{N}}, n,m \in \mathbb{Z}$, which is converted into a continuous DD signal by mounting on a continuous DD domain quasi-periodic impulse train, resulting in the continuous DD domain information bearing signal $x_{\rm dd}(\tau,\nu)$ as

$$x_{\rm dd}(\tau,\nu) = \sum_{k,l\in\mathbb{Z}} x_{\rm dd}[k,l]\delta(\tau - k\Delta\tau)\,\delta(\nu - l\Delta\nu), \quad (1)$$

where $\delta(\cdot)$ denotes Kronecker delta function. The signal $x_{\rm dd}(\tau,\nu)$ is also a quasi-periodic, i.e., $x_{\rm dd}(\tau+n\tau_p, \nu+m\nu_p) = e^{j2\pi n\nu\tau_p}\, x_{\rm dd}(\tau,\nu),\ \forall\, n,m \in \mathbb{Z}$. The $x_{\rm dd}(\tau,\nu)$ signal is bandwidth and time limited by filtering through the Tx DD domain filter $w_{\rm tx}(\tau,\nu)$, to obtain $x_{\rm dd}^{w_{\rm tx}}(\tau,\nu) = w_{\rm tx}(\tau,\nu) *_\sigma x_{\rm dd}(\tau,\nu)$, where $*_\sigma$ denotes twisted convolution[1] operation. The time domain signal for transmission is obtained using inverse Zak transform as

$$s_{\rm td}(t) = \mathcal{Z}_t^{-1}\left(x_{\rm dd}^{w_{\rm tx}}(\tau,\nu)\right) = \sqrt{\tau_p}\int_0^{\nu_p} x_{\rm dd}^{w_{\rm tx}}(t,\nu)d\nu. \quad (2)$$

The transmit signal $s_{\rm td}(t)$ passes through the channel whose impulse response in the DD domain is given by $h(\tau,\nu) = \sum_{i=1}^{P} h_i\delta(\tau-\tau_i)\delta(\nu-\nu_i)$, where h_i, τ_i, and ν_i are the channel gain, delay, and Doppler of the ith path, respectively, and P is the number of paths. The received time domain signal is

$$r_{\rm td}(t) = \int\int h(\tau,\nu)s_{\rm td}(t-\tau)e^{j2\pi\nu(t-\tau)}d\tau d\nu + n(t), \quad (3)$$

where $n(t)$ is the additive white Gaussian noise. Zak transform is used to convert the received time domain signal to DD domain as $y_{\rm dd}(\tau,\nu) = \mathcal{Z}_t(r_{\rm td}(t))$, i.e.,

$$y_{\rm dd}(\tau,\nu) = \sqrt{\tau_p}\sum_{k\in\mathbb{Z}} r_{\rm td}(\tau + k\tau_p)e^{-j2\pi\nu k\tau_p} + n_{\rm dd}(\tau,\nu), \quad (4)$$

where $n_{\rm dd}(\tau,\nu) = \mathcal{Z}_t(n(t))$ is the noise in DD domain. The DD signal $y_{\rm dd}(\tau,\nu)$ is filtered through the Rx filter $w_{\rm rx}(\tau,\nu)$, which is matched to the Tx filter, i.e., $w_{\rm rx}(\tau,\nu) = w_{\rm tx}^*(-\tau,-\nu)e^{j2\pi\tau\nu}$. The output of the Rx filter is given by $y_{\rm dd}^{w_{\rm rx}}(\tau,\nu) = w_{\rm rx}(\tau,\nu) *_\sigma y_{\rm dd}(\tau,\nu)$, i.e.,

$$y_{\rm dd}^{w_{\rm rx}}(\tau,\nu) = h_{\rm eff}(\tau,\nu) *_\sigma x_{\rm dd}(\tau,\nu) + n_{\rm dd}^{w_{\rm rx}}(\tau,\nu), \quad (5)$$

where $h_{\rm eff}(\tau,\nu)$ is the effective channel consisting of the cascade of the Tx filter, physical channel, and Rx filter, given by $h_{\rm eff}(\tau,\nu) = w_{\rm rx}(\tau,\nu) *_\sigma h(\tau,\nu) *_\sigma w_{\rm tx}(\tau,\nu)$, and $n_{\rm dd}^{w_{\rm rx}}(\tau,\nu) = w_{\rm rx}(\tau,\nu) *_\sigma n_{\rm dd}(\tau,\nu)$ is the filtered noise.

The DD signal $y_{\rm dd}^{w_{\rm rx}}(\tau,\nu)$ is sampled on the information lattice, resulting in the discrete quasi-periodic DD domain received signal $y_{\rm dd}[k,l]$ as

$$y_{\rm dd}[k,l] = y_{\rm dd}^{w_{\rm rx}}\Big(\tau = \frac{k\tau_{\rm p}}{M}, \nu = \frac{l\nu_{\rm p}}{N}\Big), \quad k,l \in \mathbb{Z}, \quad (6)$$

which is given by

$$y_{\rm dd}[k,l] = h_{\rm eff}[k,l] *_{\sigma{\rm d}} x_{\rm dd}[k,l] + n_{\rm dd}[k,l], \quad (7)$$

where $*_{\sigma{\rm d}}$ is twisted convolution in discrete DD domain, i.e.,

$$h_{\rm eff}[k,l] *_{\sigma{\rm d}} x_{\rm dd}[k,l] = \sum_{k',l'\in\mathbb{Z}} h_{\rm eff}[k-k', l-l']x_{\rm dd}[k',l']e^{\frac{j2\pi k'(l-l')}{MN}}, \quad (8)$$

where the discrete effective channel filter $h_{\rm eff}[k,l]$ and filtered noise samples $n_{\rm dd}[k,l]$ are given by $h_{\rm eff}[k,l] = h_{\rm eff}\Big(\tau = \frac{k\tau_p}{M}, \nu = \frac{l\nu_p}{N}\Big)$, $n_{\rm dd}[k,l] = n_{dd}^{w_{\rm rx}}\Big(\tau = \frac{k\tau_p}{M}, \nu = \frac{l\nu_p}{N}\Big)$. Due to the quasi-periodicity in the DD domain, it is sufficient to

[1]Twisted convolution operation between two DD functions $u(\tau,\nu)$ and $v(\tau,\nu)$ is defined as $u(\tau,\nu) *_\sigma v(\tau,\nu) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(\tau',\nu')v(\tau-\tau',\nu-\nu')e^{j2\pi\nu'(\tau-\tau')}d\tau'd\nu'$. Twisted convolution operation is associative and it preserves quasi-periodicity [3].

consider the received samples $y_{\rm dd}[k,l]$ within the fundamental period $\mathcal{D}_0$. The $y_{\rm dd}[k,l]$s are written in a vector form and the end-to-end DD domain I/O relation is written in matrix-vector form as

$$\mathbf{y} = \mathbf{H}_{\rm eff}\mathbf{x} + \mathbf{n}, \tag{9}$$

where $\mathbf{x}, \mathbf{y}, \mathbf{n} \in \mathbb{C}^{MN\times 1}$, such that their $(kN+l+1)^{\rm th}$ entries are given by $x_{kN+l+1} = x_{\rm dd}[k,l]$, $y_{kN+l+1} = y_{\rm dd}[k,l]$, $n_{kN+l+1} = n_{\rm dd}[k,l]$, and $\mathbf{H}_{\rm eff} \in \mathbb{C}^{MN\times MN}$ is the effective channel matrix such that

$$\mathbf{H}_{\rm eff}[k'N+l'+1, kN+l+1] = \sum_{m,n\in\mathbb{Z}} h_{\rm eff}[k'-k-nM, l'-l-mN] e^{j2\pi nl/N} e^{j2\pi \frac{(l'-l-mN)(k+nM)}{MN}}, \tag{10}$$

where $k', k = 0, \cdots, M-1$ and $l', l = 0, \cdots, N-1$. The above system model is extended to MIMO setting in the following subsection.

A. MIMO-Zak-OTFS system model

Consider a spatially multiplexed MIMO-Zak-OTFS system with n_t transmit antennas and n_r receive antennas, $n_r \geq n_t$. Let $h_{qp}(\tau,\nu)$ denote the impulse response of the channel in DD domain between pth transmit antenna and qth receive antenna, and let P denote the number of paths between each pair of transmit and receive antennas. Then $h_{qp}(\tau,\nu)$ can be written as $h_{qp}(\tau,\nu) = \sum_{i=1}^{P} h_{qpi}\delta(\tau-\tau_{qpi})\delta(\nu-\nu_{qpi})$, where the ith path between pth transmit antenna and qth receive antenna has delay τ_{qpi}, Doppler ν_{qpi}, and gain h_{qpi}, where $p = 1, \cdots, n_t$, $q = 1, \cdots, n_r$, and $i = 1, \cdots, P$. The effective channel between pth transmit antenna and qth receive antenna is given by the cascaded twisted convolution as $h_{{\rm eff},qp}(\tau,\nu) = w_{\rm rx}(\tau,\nu) *_\sigma h_{qp}(\tau,\nu) *_\sigma w_{\rm tx}(\tau,\nu)$. Using (7), the discrete DD received signal at qth receive antenna is

$$y_{{\rm dd},q}[k,l] = \sum_{p=1}^{n_t} h_{{\rm eff},qp}[k,l] *_{\sigma{\rm d}} x_{{\rm dd},p}[k,l] + n_{{\rm dd},q}[k,l]. \tag{11}$$

Vectorizing the above relation, we can write

$$\mathbf{y}_q = \sum_{p=1}^{n_t} \mathbf{H}_{qp}\, \mathbf{x}_p + \mathbf{n}_q, \tag{12}$$

where $\mathbf{y}_q$, $\mathbf{x}_p$ and $\mathbf{n}_q$ are $MN\times 1$ vectors and $\mathbf{H}_{qp}$ is a $MN\times MN$ matrix. Concatenating these received signal vectors gives

$$\underbrace{\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\\ \vdots\\ \mathbf{y}_{n_r}\end{bmatrix}}_{\mathbf{y}_{\rm mimo}} = \underbrace{\begin{bmatrix}\mathbf{H}_{11} & \mathbf{H}_{12} & \dots & \mathbf{H}_{1n_t}\\ \mathbf{H}_{21} & \mathbf{H}_{22} & \dots & \mathbf{H}_{2n_t}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{H}_{n_r1} & \mathbf{H}_{n_r2} & \dots & \mathbf{H}_{n_rn_t}\end{bmatrix}}_{\mathbf{H}_{\rm eff,mimo}} \underbrace{\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\\ \vdots\\ \mathbf{x}_{n_t}\end{bmatrix}}_{\mathbf{x}_{\rm mimo}} + \underbrace{\begin{bmatrix}\mathbf{n}_1\\ \mathbf{n}_2\\ \vdots\\ \mathbf{n}_{n_r}\end{bmatrix}}_{\mathbf{n}_{\rm mimo}}, \tag{13}$$

resulting in a vectorized I/O relation for MIMO-Zak-OTFS as

$$\mathbf{y}_{\rm mimo} = \mathbf{H}_{\text{eff, mimo}}\, \mathbf{x}_{\rm mimo} + \mathbf{n}_{\rm mimo}, \tag{14}$$

where $\mathbf{y}_{\rm mimo}$, $\mathbf{n}_{\rm mimo} \in \mathbb{C}^{n_rMN\times 1}$, $\mathbf{x}_{\rm mimo} \in \mathbb{C}^{n_tMN\times 1}$ and $\mathbf{H}_{\text{eff, mimo}} \in \mathbb{C}^{n_rMN\times n_tMN}$. With sufficiently large spacing between receive antennas, the noise vectors at the receive antennas are independent of each other, i.e.,

$$\mathbf{C}_{\mathbf{n}_{\rm mimo}} = \mathbb{E}\left[\mathbf{n}_{\rm mimo}\mathbf{n}_{\rm mimo}^H\right] = \begin{bmatrix}\mathbf{C}_{n_1} & \mathbf{0} & \dots & \mathbf{0}\\ \mathbf{0} & \mathbf{C}_{n_2} & \dots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{C}_{n_r}\end{bmatrix}, \tag{15}$$

where $\mathbf{C}_{n_q} = \mathbb{E}\left[\mathbf{n}_q\mathbf{n}_q^H\right]$ for $q = 1, \cdots, n_r$.

III. MIMO I/O RELATION ESTIMATION AND DETECTION

In this section, we present the proposed embedded pilot frame structure and the model-free I/O relation estimation and detection schemes for MIMO-Zak-OTFS.

A. Embedded pilot frame structure

Here, we present the proposed embedded pilot frame structure for MIMO-Zak-OTFS with $n_t = n_r = 2$. Consider a fundamental region $\mathcal{F} = \{(k,l) : k = 0,1,\cdots,M-1, l = 0,1,\cdots,N-1\}$. This $\mathcal{F}$ region is divided into pilot regions ($\mathcal{P}_1$ and $\mathcal{P}_2$), guard regions ($\mathcal{G}_1$ and $\mathcal{G}_2$) and information region ($\mathcal{I}$). These regions are defined as follows:

$$\mathcal{P}_1 \triangleq \left\{(k,l)\in\mathcal{F} \mid \left|k-\frac{M}{2}\right| + \left|l-\frac{N}{2}\right| \leq a_1\right\}, \tag{16}$$

$$\mathcal{G}_1 \triangleq \left\{(k,l)\in\mathcal{F} \mid a_1 < \left|k-\frac{M}{2}\right| + \left|l-\frac{N}{2}\right| \leq a_2\right\}, \tag{17}$$

$$\mathcal{P}_2 \triangleq \left\{(k,l)\in\mathcal{F} \;\middle|\; \begin{aligned} k+l &\leq a_1 \vee\\ -(k-M)+l &\leq a_1 \vee\\ k-(l-N) &\leq a_1 \vee\\ -(k-M)-(l-N) &\leq a_1\end{aligned}\right\}, \tag{18}$$

$$\mathcal{G}_2 \triangleq \left\{(k,l)\in\mathcal{F} \;\middle|\; \begin{aligned} a_1 < k+l &\leq a_2 \vee\\ a_1 < -(k-M)+l &\leq a_2 \vee\\ a_1 < k-(l-N) &\leq a_2 \vee\\ a_1 < -(k-M)-(l-N) &\leq a_2\end{aligned}\right\}, \tag{19}$$

$$\mathcal{I} \triangleq \mathcal{F} - (\mathcal{P}_1 \cup \mathcal{G}_1 \cup \mathcal{P}_2 \cup \mathcal{G}_2) \tag{20}$$

where a_1 and a_2 are parameters affecting the size of these regions and their values can be chosen according to the system parameters.

Figure 2 shows the regions described above and the pilot placement for 2×2 MIMO. $\mathcal{P}_1$ and $\mathcal{P}_2$ are represented by green and red shaded regions, respectively. $\mathcal{G}_1$ and $\mathcal{G}_2$ are shown by yellow and grey shaded regions, respectively. Data region $\mathcal{I}$ is represented by blue shaded region. Pilots are placed at $(\frac{M}{2}, \frac{N}{2})$ and at $(0,0)$ for transmitter antenna 1 (Tx_1) and transmitter antenna 2 (Tx_2), respectively. The symbol $x_1[k,l]$ in the frame at Tx_1 is given by

$$x_1[k,l] = \begin{cases}\sqrt{\frac{E_{\rm p}}{2}} & (k,l) = \left(\frac{M}{2}, \frac{N}{2}\right)\\ \sqrt{\frac{E_{\rm d}}{2|\mathcal{I}|}}\, x_{{\rm d},1}[k,l] & (k,l)\in\mathcal{I}\\ 0, & \text{otherwise,}\end{cases} \tag{21}$$

where $x_{{\rm d},1}[k,l]$ is information symbol taken from unit energy modulation alphabet. Likewise, the symbol $x_2[k,l]$ in the frame at Tx_2 is given by

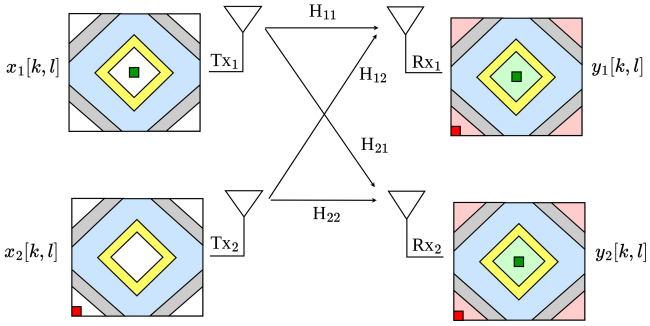

Fig. 2: Embedded frame structure and pilots placement in a 2×2 MIMO-Zak-OTFS system.

$$x_2[k,l] = \begin{cases} \sqrt{\frac{E_\mathrm{p}}{2}} & (k,l) = (0,0) \\ \sqrt{\frac{E_\mathrm{d}}{2\,|\mathcal{I}|}}\, x_{\mathrm{d},2}[k,l] & (k,l) \in \mathcal{I} \\ 0, & \text{otherwise,} \end{cases} \tag{22}$$

where $x_{\mathrm{d},2}[k,l]$ is information symbol taken from unit energy modulation alphabet. Normalizing the channel gains to unity and using unit energy DD filters, we get the received average energy at both receiver antennas (Rx$_1$ and Rx$_2$) as $E_\mathrm{p} + E_\mathrm{d}$. The pilot-to-data ratio (PDR) is defined as $\frac{E_\mathrm{p}}{E_\mathrm{d}}$. This proposed frame structure with the above guard, pilot, and data regions is found to achieve good estimation and detection performance.

B. MIMO I/O relation estimation

We use the above defined structure for embedded pilot frames to estimate the MIMO I/O relation. Embedded pilot frames $x_1[k,l]$ and $x_2[k,l]$ are transmitted from Tx$_1$ and Tx$_2$, respectively, as shown in Fig. 2. We obtain the received frames $y_1[k,l]$ and $y_2[k,l]$ at Rx$_1$ and Rx$_2$, respectively. Reading off the green shaded region in the received frame $y_1[k,l]$, denoted by $y_1^{(g)}[k,l]$, the effective I/O relation between Tx$_1$ and Rx$_1$ is estimated as

$$\hat{h}_{\text{eff},11}[k,l] = \begin{cases} y_1^{(g)}\left[k+\frac{M}{2}, l+\frac{N}{2}\right] e^{-j\pi\frac{l}{N}}, & -\frac{M}{2} \le k < \frac{M}{2}, \\ & -\frac{N}{2} \le l < \frac{N}{2} \\ 0, & \text{otherwise,} \end{cases} \tag{23}$$

where

$$y_1^{(g)}[k,l] = \begin{cases} \sqrt{\frac{2}{E_\mathrm{p}}}\, y_1[k,l], & (k,l) \in \mathcal{P}_1, \\ 0, & \text{otherwise.} \end{cases} \tag{24}$$

Likewise, reading off the red shaded region in the received frame $y_1[k,l]$, denoted by $y_1^{(r)}[k,l]$, the effective I/O relation $\hat{h}_{\text{eff},12}[k,l]$ between Tx$_2$ and Rx$_1$ is estimated as

$$\hat{h}_{\text{eff},12}[k,l] = \begin{cases} y_1^{(r)}[k,l], & 0 \le k < \frac{M}{2}, \\ & 0 \le l < \frac{N}{2} \\ y_1^{(r)}[k+M,l]\, e^{-j2\pi\frac{l}{N}}, & -\frac{M}{2} \le k < 0, \\ & 0 \le l < \frac{N}{2} \\ y_1^{(r)}[k,l+N], & 0 \le k < \frac{M}{2}, \\ & -\frac{N}{2} \le l < 0 \\ y_1^{(r)}[k+M,l+N]\, e^{-j2\pi\frac{l}{N}}, & -\frac{M}{2} \le k < 0, \\ & -\frac{N}{2} \le l < 0 \\ 0, & \text{otherwise,} \end{cases} \tag{25}$$

where

$$y_1^{(r)}[k,l] = \begin{cases} \sqrt{\frac{2}{E_\mathrm{p}}}\, y_1[k,l], & (k,l) \in \mathcal{P}_2 \\ 0, & \text{otherwise} \end{cases} \tag{26}$$

In a similar way, the estimates $\hat{h}_{\text{eff},21}[k,l]$ and $\hat{h}_{\text{eff},22}[k,l]$ can be obtained by reading off $y_2^{(g)}[k,l]$ and $y_2^{(r)}[k,l]$, respectively. Using these DD effective channel estimates, the matrices $\hat{\mathbf{H}}_{11}$, $\hat{\mathbf{H}}_{12}$, $\hat{\mathbf{H}}_{21}$, and $\hat{\mathbf{H}}_{22}$ are constructed using (10). The $\hat{\mathbf{H}}_{\text{eff, mimo}}$ is obtained by concatenating these matrices as

$$\hat{\mathbf{H}}_{\text{eff, mimo}} = \begin{bmatrix} \hat{\mathbf{H}}_{11} & \hat{\mathbf{H}}_{12} \\ \hat{\mathbf{H}}_{21} & \hat{\mathbf{H}}_{22} \end{bmatrix}. \tag{27}$$

C. MIMO detection

The received symbols in pilot regions ($\mathcal{P}_1$ and $\mathcal{P}_2$) are not used for detecting the information symbols present in the information region ($\mathcal{I}$). In order to do that, a particular sub-matrix of the $\hat{\mathbf{H}}_{\text{eff, mimo}}$ matrix ($\hat{\mathbf{H}}_{\text{eq}}$) is used for signal detection. This sub-matrix is obtained as follows. The set of rows $\{k_1 N + l_1 + 1 \mid (k_1,l_1) \in \mathcal{I} \cup \mathcal{G}_1 \cup \mathcal{G}_2\}$ and the set of columns $\{k_2 N + l_2 + 1 \mid (k_2,l_2) \in \mathcal{I}\}$ are retained in each of $\hat{\mathbf{H}}_{11}$, $\hat{\mathbf{H}}_{12}$, $\hat{\mathbf{H}}_{21}$, and $\hat{\mathbf{H}}_{22}$. The remaining rows and columns are not considered for equalization. These irrelevant rows and columns are removed from $\hat{\mathbf{H}}_{\text{eff, mimo}}$ to get $\hat{\mathbf{H}}_{\text{eq}}$. In addition, the received symbols at locations $\{(k,l) \in \mathcal{I} \cup \mathcal{G}_1 \cup \mathcal{G}_2\}$ in $y_1[k,l]$ and $y_2[k,l]$ are considered for equalization. These symbols are arranged in a vector $\mathbf{y}_{\text{eq}}$. We aim to obtain an estimate of transmitted symbols at locations $\{(k,l) \in \mathcal{I}\}$ in $x_1[k,l]$ and $x_2[k,l]$ arranged in a vector $\mathbf{x}_\mathrm{d}$. As widely considered in the Zak-OTFS literature, we first consider minimum mean square error (MMSE) detection. We also consider a local search based algorithm, namely, likelihood ascent search (LAS) algorithm [13], to achieve improved BER performance. The matrix that minimizes the mean square error $\mathbb{E}\left[\|\mathbf{x}_\mathrm{d} - \mathbf{G}\mathbf{y}_{\text{eq}}\|^2\right]$ is $\mathbf{G}_{\text{mmse}} = \mathbf{H}_{\text{eq}}^H \left(\mathbf{H}_{\text{eq}}\mathbf{H}_{\text{eq}}^H + \mathbf{C}_{\mathbf{n}_{\text{eq}}}\mathbf{I}\right)^{-1}$, where $\mathbf{C}_{\mathbf{n}_{\text{eq}}}$ is obtained by retaining the rows and columns $\{k_1 N + l_1 + 1 \mid (k_1,l_1) \in \mathcal{I} \cup \mathcal{G}_1 \cup \mathcal{G}_2\}$ in $\mathbf{C}_{n_1}$ and $\mathbf{C}_{n_2}$ (see Eq. (15)). The MMSE detector output vector $\hat{\mathbf{x}}_\mathrm{d}$ is obtained as $\hat{\mathbf{x}}_\mathrm{d} = f(\mathbf{G}_{\text{mmse}}\mathbf{y}_{\text{eq}})$, where $f(.)$ maps each entry of $\mathbf{G}_{\text{mmse}}\mathbf{y}_{\text{eq}}$ to a symbol in the modulation alphabet $\mathbb{A}$ based on minimum Euclidean distance.

We improve upon the MMSE solution by performing a local neighborhood search using the LAS algorithm in [13] with MMSE detector output vector as the initial solution vector. We call this detector as the MMSE-LAS detector. The neighborhood is defined as the set of vectors which differ from the current solution vector in one coordinate. If the best neighbor of the current solution vector is better than the current solution vector in terms of maximum-likelihood (ML) cost $\|\mathbf{y}_{\text{eq}} - \mathbf{H}_{\text{eq}}\mathbf{x}_\mathrm{d}\|_2^2$, then that neighbor is taken as the current solution and the algorithm proceeds to the next iteration. The process continues till a local minima is reached, and the local minima vector is declared as the detected vector.

IV. Results and Discussions

Simulations are carried out for the considered 2×2 MIMO-Zak-OTFS system with the following parameters:

TABLE I: Power-delay profile of Veh-A channel model

Path number (i)	1	2	3	4	5	6
τ_i (μs)	0	0.31	0.71	1.09	1.73	2.51
Relative power (p_i) dB	0	-1	-9	-10	-15	-20

$\nu_{\rm p} = 15$ kHz, $\tau_{\rm p} = \frac{1}{\nu_{\rm p}} = 66.66$ μs, $M = 12$, $N = 14$, $T = N\tau_{\rm p} = 0.93$ ms, $B = M\nu_{\rm p} = 180$ kHz, and BPSK. Vehicular-A channel model [14] with fractional DDs, $P = 6$ paths, maximum Doppler $\nu_{\max} = 815$ Hz, maximum delay $\tau_{\max} = 2.51\mu$s, and power delay profile given in Table I, is considered. The Doppler associated with the ith path is modeled as $\nu_i = \nu_{\max}\cos(\theta_i)$ where θ_is are independent and uniformly distributed in $[0, 2\pi)$. Sinc, Gaussian, and Gaussian-sinc pulse shaping filters are considered. The sinc filter is given by $w_{\rm tx}(\tau,\nu) = \sqrt{B}{\rm sinc}(B\tau)\sqrt{T}{\rm sinc}(T\nu)$, Gaussian filter is given by $w_{\rm tx}(\tau,\nu) = \left(\frac{2\alpha_\tau B^2}{\pi}\right)^{\frac{1}{4}} e^{-\alpha_\tau B^2\tau^2}\left(\frac{2\alpha_\nu T^2}{\pi}\right)^{\frac{1}{4}} e^{-\alpha_\nu T^2\nu^2}$, and Gaussian-sinc filter is given by $w_{\rm tx}(\tau,\nu) = \Omega_\tau\Omega_\nu\sqrt{BT}{\rm sinc}(B\tau){\rm sinc}(T\nu)e^{-\alpha_\tau B^2\tau^2}e^{-\alpha_\nu T^2\nu^2}$. In order to contain 99% of the energy to be within B and T, $\alpha_\tau = \alpha_\nu = 1.584$ is used for Gaussian filter. Similarly, for Gaussian-sinc filter, $\alpha_\tau = \alpha_\nu = 0.044$ and $\Omega_\tau = \Omega_\nu = 1.0278$ are used. Embedded pilot frame parameters a_1 and a_2 are chosen to be 3 and 4 respectively. These values are chosen to make pilot region and guard region sufficiently large to get good estimation performance and less interference between pilot and data symbols.

Figure 3 shows the MSE vs data SNR performance of the considered 2×2 MIMO-Zak-OTFS system for different filters at 0 dB PDR. The MSE is defined as $\frac{\|\mathbf{H}_{\text{eff, mimo}} - \hat{\mathbf{H}}_{\text{eff,mimo}}\|_{\mathbf{F}}^2}{\|\mathbf{H}_{\text{eff, mimo}}\|_{\mathbf{F}}^2}$. From Fig. 3, it is seen that the sinc filter gives relatively poor MSE performance compared to Gaussian and Gaussian-sinc filters. This is because of the high sidelobes in the sinc filter, which causes interference between pilot and data symbols and interference due to quasi-periodic aliases. Gaussian filter achieves the best MSE performance, which is attributed to its very low sidelobes and hence very good localization of the DD pulse. The Gaussian-sinc filter has lower sidelobes compared to the sinc filter but higher sidelobes compared to Gaussian filter [8]. So, the Gaussian-sinc filter performance is in between those of sinc and Gaussian filters.

Figure 4 shows the BER performance achieved by the MMSE and MMSE-LAS detectors for sinc, Gaussian, and Gaussian-sinc filters in the considered 2×2 MIMO-Zak-OTFS system with 0 dB PDR. We observe that sinc filter gives better BER performance compared to Gaussian filter. This is because sinc filter has nulls at DD sampling points, and hence there is no inter-symbol interference at the sampling points. Whereas, Gaussian filter does not have nulls at the sampling points, and hence is susceptible to inter-symbol interference. Like sinc filter, Gaussian-sinc filter also has nulls at DD sampling points. It also has low sidelobes resulting in a good estimate of the I/O relation. This makes its BER performance better than sinc and Gaussian filters. Also, MMSE-LAS detector is found to give a significant improvement in BER performance compared to MMSE detector due to the neighborhood search.

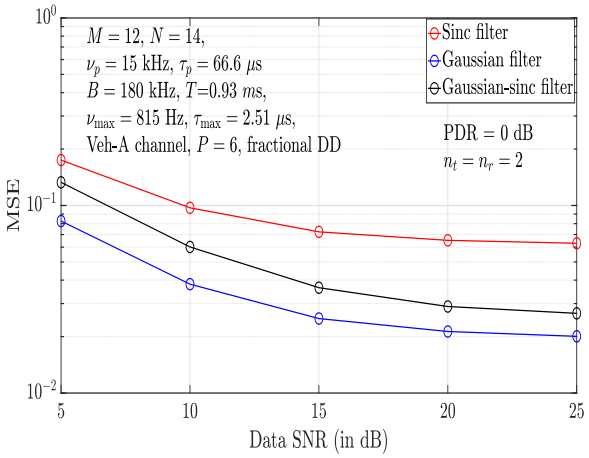

Fig. 3: MSE vs data SNR performance of 2×2 MIMO-Zak-OTFS with embedded pilot frame at 0 dB PDR.

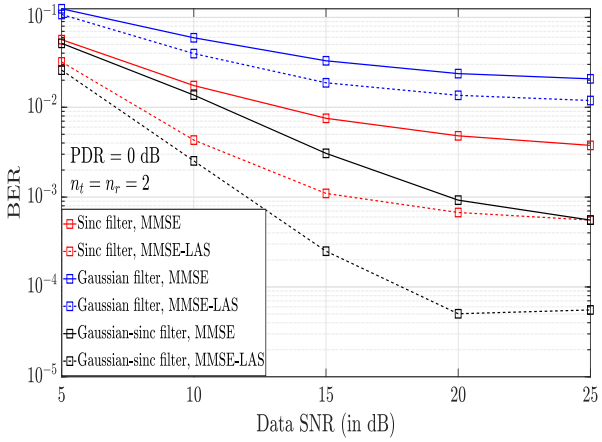

Fig. 4: BER vs data SNR performance of 2×2 MIMO-Zak-OTFS with embedded pilot frame at 0 dB PDR.

Next, Figs. 5 and 6 present a comparison of MSE and BER performance between SISO and 2×2 MIMO Zak-OTFS systems. Figure 5 presents the MSE performance comparison as a function of PDR at a data SNR of 15 dB. From Fig. 5, it can be seen that, as the PDR increases, the pilot power increases and hence the MSE performance improves. At low PDRs, the pilot power is less, and, therefore, the effect of noise in the read-off region dominates. As MIMO-Zak-OTFS has a smaller read-off region for estimation, the noise effect is less and the MSE performance is slightly better than that of SISO-Zak-OTFS at low PDRs. For high PDRs, effect of pilot power dominates over noise while estimating the I/O relation from the read-off region. As SISO-Zak-OTFS has a larger read-off region for estimation, it captures more received pilot energy for estimation. Thus, at high PDRs, SISO-Zak-OTFS has better MSE performance.

Figure 6 shows the BER vs data SNR performance comparison at a PDR of 0 dB. It is observed that, with MMSE detection, the BER performance of the 2×2 MIMO system is worse than that of the SISO system, because of the spatial and DD domain interferences and the sub-optimality of

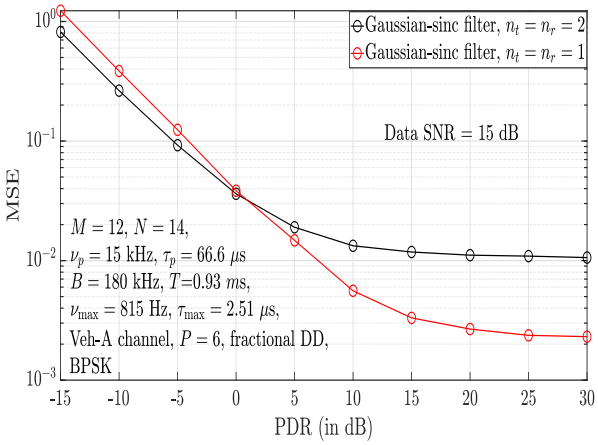

Fig. 5: MSE vs PDR performance of 2×2 MIMO-Zak-OTFS as compared with that of SISO-Zak-OTFS at 15 dB data SNR.

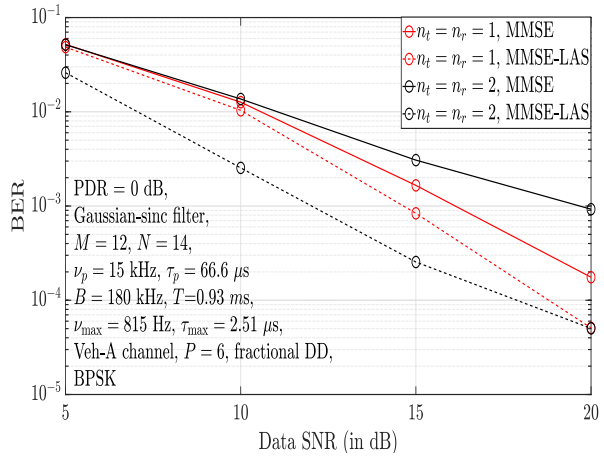

Fig. 6: BER vs data SNR performance of 2×2 MIMO-Zak-OTFS as compared with SISO-Zak-OTFS at 0 dB PDR.

MMSE detection. However, with MMSE-LAS detection, the MIMO system performance is better than the SISO system performance. For example, at a BER of 10^{-3}, there is about 2 to 3 dB performance improvement in favor of MIMO. This is because MIMO performance with optimum detection is better than SISO performance, and MMSE-LAS detection achieves near-optimal performance in large dimensions [13]. This illustrates the effectiveness of MMSE-LAS detection in MIMO-Zak-OTFS. Similar performance behavior is observed for 8-QAM modulation with rate-1/3 coding (see Fig. 7).

V. Conclusions

We investigated the problem of model-free Zak-OTFS I/O relation estimation and signal detection with embedded pilots in a MIMO setting, which has not been reported before. We proposed an embedded pilot frame structure for a 2×2 MIMO-Zak-OTFS system and an I/O relation estimation scheme based on a simple read-off operation in the pilot region of the frame. Our simulation results showed that the proposed schemes achieved good MSE and BER performance. Among the considered DD filters (sinc, Gaussian, Gaussian-sinc filter), Gaussian-sinc filter was found to achieve better BER performance due to its favorable main lobe and sidelobe

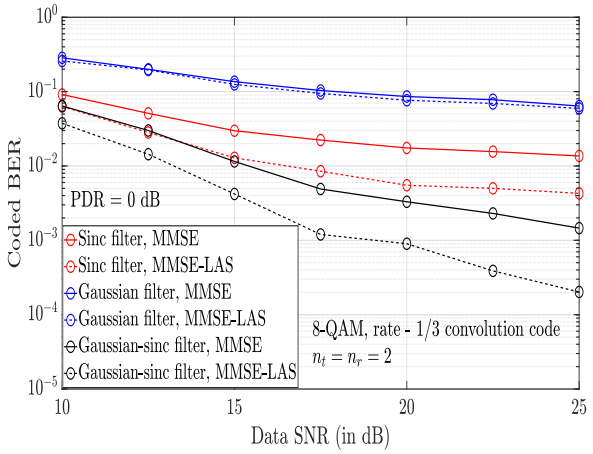

Fig. 7: Coded BER vs data SNR performance of 2×2 MIMO-Zak-OTFS for 8-QAM with rate 1/3 coding at 0 dB PDR.

characteristics leading to a balanced estimation and detection performance. Superimposed pilot frames can alleviate the throughput loss due to pilot and guard regions at the cost of more sophisticated signal processing to process the overlapping pilot and data symbols. Superimposed pilot frames for MIMO-Zak-OTFS can be considered for future work.

References

[1] R. Hadani *et al.*, "Orthogonal time frequency space modulation," *Proc. IEEE WCNC'2017*, pp. 1-6, Mar. 2017.

[2] Best Readings in Orthogonal Time Frequency Space (OTFS) and Delay Doppler Signal Processing, https://www.comsoc.org/publications/best-readings/orthogonal-time-frequency-space-otfs-and-delay-doppler-signal-processing.

[3] S. K. Mohammed, R. Hadani, and A. Chockalingam, *OTFS Modulation - Theory and Applications*, IEEE-Wiley, 2024.

[4] S. K. Mohammed, R. Hadani, A. Chockalingam, and R. Calderbank, "OTFS — a mathematical foundation for communication and radar sensing in the delay-Doppler domain," *IEEE BITS The Inform. Theory Mag.*, vol. 2, no. 2, pp. 36-55, 1 Nov. 2022.

[5] S. K. Mohammed, R. Hadani, A. Chockalingam, and R. Calderbank, "OTFS — predictability in the delay-Doppler domain and its value to communication and radar sensing," *IEEE BITS The Inform. Theory Mag.*, vol. 3, no. 2, pp. 7-31, Jun. 2023.

[6] S. Gopalam, I. B. Collings, S. V. Hanly, H. Inaltekin, S. R. B. Pillai, P. Whiting, "Zak-OTFS implementation via time and frequency windowing," *IEEE Trans. Commun.*, vol. 72, no. 7, pp. 3873-3889, Jul. 2024.

[7] J. Jayachandran, R. K. Jaiswal, S. K. Mohammed, R. Hadani, A. Chockalingam, and R. Calderbank, "Zak-OTFS: pulse shaping and the tradeoff between time/bandwidth expansion and predictability," available online: arxiv:2405.02718v1 [eess.SP] 4 May 2024.

[8] A. Das, F. Jesbin, and A. Chockalingam, "A Gaussian-sinc pulse shaping filter for Zak-OTFS," online arXiv:2502.03904v1 [cs.IT] 6 Feb 2025.

[9] M. K. Ramachandran and A. Chockalingam, "MIMO-OTFS in high-Doppler fading channels: signal detection and channel estimation," *Proc. IEEE GLOBECOM'2018*, pp. 206-212, Dec. 2018.

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

[11] H. B. Mishra, P. Singh, A. K. Prasad, and R. Budhiraja, "OTFS channel estimation and data detection designs with superimposed pilots," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2258-2274, Apr. 2022.

[12] M. Ubadah, S. K. Mohammed, R. Hadani, S. Kons, A. Chockalingam, and R. Calderbank, "Zak-OTFS for integration of sensing and communication," available online: arxiv:2404.04182v1 [eess.SP] 5 Apr 2024.

[13] A. Chockalingam and B. Sundar Rajan, *Large MIMO Systems*, Cambridge University Press, 2014.

[14] ITU-R M.1225,"Guidelines for evaluation of radio transmission technologies for IMT-2000," *International Telecommunication Union Radio communication*, 1997.