Improved Twisted Convolution Precoding for Zak-OTFS

Naveed Bin Nazir and Ananthanarayanan Chockalingam
Department of ECE, Indian Institute of Science, Bangalore

***Abstract*—Recently, a novel twisted convolution (TC) based precoding scheme that enabled single-tap equalization at the Zak-OTFS receiver has been proposed. This precoder maximized the per-carrier signal-to-interference plus noise ratio (SINR) at the receiver. In this paper, we extend this work in two useful directions. First, we propose a TC precoder that maximizes the correlation between the actual precoded channel response vector and a desired target precoded channel response vector. The proposed correlation-maximizing precoder achieves a better localized response at the receiver, and hence achieves improved symbol error rate (SER) performance. Second, we propose a deep neural network (DNN)-based 'denoiser' at the transmitter that enhances the accuracy of the read-off based model-free channel state information at the transmitter (CSIT) estimate, which further improves the SER performance. Heatmaps of the precoded channel response and SER simulation results using Vehicular-A channel model with fractional delays and Dopplers demonstrate the effectiveness of the proposed precoder for different delay-Doppler pulse shaping filters.**

***Index Terms*—Zak-OTFS modulation, delay-Doppler domain, twisted convolution, precoding, single-tap equalization, DNN-enhanced CSIT estimation.**

I. Introduction

Orthogonal time frequency space modulation based on a Zak transform [1] framework (Zak-OTFS), which allows a direct conversion between the delay-Doppler (DD) domain and the time domain, has recently been proposed in [2],[3]. Zak-OTFS achieves robust performance over a large range of delay and Doppler spreads. It has a non-fading and predictable input-output (I/O) relation which allows the time-varying channel to be efficiently acquired (using a simple model-free read-off operation) and equalized. Due to these attractive features, Zak-OTFS has received increased research interest. Recent works have considered several aspects of Zak-OTFS including optimum receiver, I/O relation estimation techniques, DD pilot architectures, near-optimal signal detection, DD pulse shaping filter design, differential communication, MIMO operation, learning techniques in receiver design, etc. [4]-[10].

The topic of precoding for Zak-OTFS is open for research. While Zak-OTFS without precoding requires a complex equalizer at the receiver (e.g., MMSE equalization) for reliable data detection, a carefully designed precoder at the transmitter can concentrate the highly dispersive channel energy towards the desired DD channel taps, making a simple single-tap equalizer possible at the receiver. An early investigation of Zak-OTFS precoder design that achieves the above possibility has recently been reported in [11]. This precoder design exploits channel reciprocity (e.g., as in a TDD system), where a pilot frame (exclusive or embedded pilot frame) transmitted by the user equipment (UE) on the uplink to the base station (BS) can be used by the BS to obtain an estimate of the channel state information at the transmitter (CSIT). This CSIT can then be used to obtain the precoder for downlink transmission from the BS to the UE. A key novelty of the precoder design is that the precoding operation is a twisted-convolution (TC) filtering operation, which gels naturally with the effective channel, where the effective channel itself is a TC cascade of the transmit DD pulse shaping filter, the physical channel, and the receive filter. This TC cascade of the precoder filter with the effective channel allows the derivation of the optimum precoder in closed-form. In particular, it obtains the optimum precoder that maximizes the per-carrier (i.e., per DD bin) signal-to-interference plus noise ratio (SINR) at the UE receiver on the downlink. The precoder offers the following advantages on the UE side: enabling single-tap equalization lowering receiver complexity, improving spectral efficiency by reducing pilot overhead on the downlink frames, and achieving good downlink symbol error rate (SER) performance. Inspired by these advantages, in this paper, we extend the work in [11] in two novel and useful ways (which are new contributions in this paper), as summarized below.

- First, we propose a TC precoder that maximizes the correlation between the actual precoded channel response vector and a desired target precoded channel response vector. The proposed correlation-maximizing precoder achieves a better localized response at the receiver, and therefore achieves improved SER performance compared to that of the SINR-maximizing precoder in [11].
- Second, recognizing the need/potential to improve the accuracy of the estimated CSIT, we propose a deep neural network (DNN)-based 'denoiser' that enhances the accuracy of the read-off based model-free CSIT estimate, which further improves the SER performance.
- We evaluate the SER performance achieved by the proposed precoder in Vehicular-A channels [12] with fractional DDs. Through heatmaps of the precoded channel response and simulated SER results, we demonstrate the effectiveness of the proposed precoder for different DD pulse shaping filters including sinc, Gaussian, and Gaussian-sinc filters.

This work was supported in part by the J. C. Bose National Fellowship and the NSF-DST project (International Cooperation Division), Department of Science and Technology, Government of India.

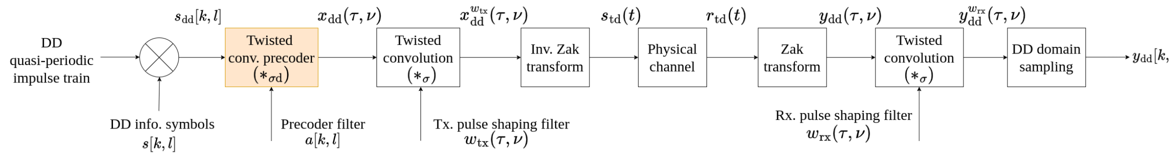


Fig. 1: Block diagram of the TC precoded Zak-OTFS transceiver.

II. TC precoded Zak-OTFS system Model

The block diagram of the TC precoded Zak-OTFS transceiver is shown in Fig. 1. It depicts the signal flow in the overall end-to-end chain, including the precoder operation through discrete twisted convolution. In Zak-OTFS, a pulse in the DD domain is the basic information carrier. A DD pulse is a quasi-periodic localized function defined by a delay period $\tau_{\rm p}$ and a Doppler period $\nu_{\rm p} = \frac{1}{\tau_{\rm p}}$. The fundamental period in the DD domain is defined as $\mathcal{D}_0 = \{(\tau,\nu) : 0 \leq \tau < \tau_{\rm p}, 0 \leq \nu < \nu_{\rm p}\}$, where τ and ν represent the delay and Doppler variables, respectively. The fundamental period is discretized into M bins on the delay axis and N bins on the Doppler axis, as $\left\{(k\frac{\tau_{\rm p}}{M}, l\frac{\nu_{\rm p}}{N}) | k = 0,\ldots,M-1, l = 0,\ldots,N-1\right\}$. The Zak-OTFS frame is limited to a time duration $T = N\tau_{\rm p}$ and a bandwidth $B = M\nu_{\rm p}$. In each frame, MN information symbols drawn from a modulation alphabet $\mathbb{A}$, $s[k,l] \in \mathbb{A}$, $k = 0,\ldots,M-1$, $l = 0,\ldots,N-1$, are multiplexed in the DD domain. The $s[k,l]$s are encoded as per the following equation to obtain a quasi-periodic extension of the signal in the discrete DD domain:

$$s_{\rm dd}[k+nM,\, l+mN] = s[k,l]\, e^{j2\pi n \frac{l}{N}}, \quad n,m \in \mathbb{Z}. \tag{1}$$

A transmit-side precoding filter $a[k,l]$ is then applied via discrete twisted convolution to obtain the precoded signal $x_{\rm dd}[k,l]$ given by

$$x_{\rm dd}[k,l] = a[k,l] *_{\sigma\rm d} s_{\rm dd}[k,l], \tag{2}$$

where $*_{\sigma\rm d}$ denotes the discrete twisted convolution[1], which preserves quasi-periodicity. The precoded discrete DD signal $x_{\rm dd}[k,l]$ in (2) is lifted to continuous DD domain through 2D impulse modulation as

$$x_{\rm dd}(\tau,\nu) = \sum_{k,l\in\mathbb{Z}} x_{\rm dd}[k,l]\, \delta(\tau - k\Delta\tau)\, \delta(\nu - l\Delta\nu), \tag{3}$$

which is quasi-periodic, i.e., $x_{\rm dd}(\tau + n\tau_p,\, \nu + m\nu_p) = e^{j2\pi n\nu\tau_p} x_{\rm dd}(\tau,\nu)$, $n,m \in \mathbb{Z}$. To confine the transmit signal within the desired time-frequency support (frame duration T and bandwidth B), a DD domain transmit pulse shaping filter $w_{\rm tx}(\tau,\nu)$ is applied as $x_{\rm dd}^{w_{\rm tx}}(\tau,\nu) = w_{\rm tx}(\tau,\nu) *_\sigma x_{\rm dd}(\tau,\nu)$, where $*_\sigma$ denotes the continuous twisted convolution operation[2] The time domain signal corresponding to $x_{\rm dd}^{w_{\rm tx}}(\tau,\nu)$ is obtained via inverse Zak transform as $s_{\rm td}(t) = \mathcal{Z}_t^{-1}(x_{\rm dd}^{w_{\rm tx}}(\tau,\nu)) = \sqrt{\tau_p}\int_0^{\nu_p} x_{\rm dd}^{w_{\rm tx}}(t,\nu)\, d\nu$. The signal $s_{\rm td}(t)$ propagates through a doubly-dispersive channel with P paths, modeled as $h(\tau,\nu) = \sum_{i=1}^{P} h_i\, \delta(\tau-\tau_i)\, \delta(\nu-\nu_i)$, where h_i, τ_i, and ν_i represent the complex gain, delay, and Doppler shift of the ith path, respectively. The received time domain signal is given by $r_{\rm td}(t) = \iint h(\tau,\nu)\, s_{\rm td}(t-\tau)\, e^{j2\pi\nu(t-\tau)} d\tau d\nu + n(t)$, where $n(t)$ is additive white Gaussian noise (AWGN).

At the receiver, the Zak transform[3] converts the received time domain signal to its DD domain form as $y_{\rm dd}(\tau,\nu) = \mathcal{Z}_t(r_{\rm td}(t)) + n_{\rm dd}(\tau,\nu)$, where $n_{\rm dd}(\tau,\nu)$ is the DD domain noise. The received signal is then filtered by a receive filter matched to the transmit filter, $w_{\rm rx}(\tau,\nu) = w_{\rm tx}^*(-\tau,-\nu)\, e^{j2\pi\tau\nu}$, resulting in

$$y_{\rm dd}^{w_{\rm rx}}(\tau,\nu) = h_{\rm eff}(\tau,\nu) *_\sigma x_{\rm dd}(\tau,\nu) + n_{\rm dd}^{w_{\rm rx}}(\tau,\nu), \tag{4}$$

where $h_{\rm eff}(\tau,\nu) = w_{\rm rx}(\tau,\nu) *_\sigma h(\tau,\nu) *_\sigma w_{\rm tx}(\tau,\nu)$ is the effective DD domain channel response (a.k.a. the 'effective channel'). Sampling (4) on the information grid points $\tau = \frac{k\tau_{\rm p}}{M}, \nu = \frac{l\nu_{\rm p}}{N}$ yields the discrete model

$$y_{\rm dd}[k,l] = h_{\rm eff}[k,l] *_{\sigma\rm d} x_{\rm dd}[k,l] + n_{\rm dd}[k,l]. \tag{5}$$

Substituting the precoded signal from (2) into (5), we obtain

$$\begin{aligned} y_{\rm dd}[k,l] &= h_{\rm eff}[k,l] *_{\sigma\rm d} a[k,l] *_{\sigma\rm d} s_{\rm dd}[k,l] + n_{\rm dd}[k,l] \\ &= h_{\rm a}[k,l] *_{\sigma\rm d} s_{\rm dd}[k,l] + n_{\rm dd}[k,l], \end{aligned} \tag{6}$$

where $h_{\rm a}[k,l] \triangleq w_{\rm rx}[k,l] *_{\sigma\rm d} h[k,l] *_{\sigma\rm d} w_{\rm tx}[k,l] *_{\sigma\rm d} a[k,l]$, i.e., $h_{\rm a}[k,l] = h_{\rm eff}[k,l] *_{\sigma\rm d} a[k,l]$ denotes the effective precoded DD domain channel response (a.k.a 'precoded effective channel'). Equation (6) thus represents the discrete input-output (I/O) relation of the TC precoded Zak-OTFS system.

Support sets of $h_{\rm eff}[k,l]$, $h_a[k,l]$, $a[k,l]$: Let the support set of $h_{\rm eff}[k,l]$, denoted by $\mathcal{S}_{h_{\rm eff}}$, be $\mathcal{S}_{h_{\rm eff}} \triangleq \{(k,l) \,|\, k_{\rm min} \leq k \leq k_{\rm max}\ ,\ l_{\rm min} \leq l \leq l_{\rm max}\}$, where $k_{\rm min} \leq 0 \leq k_{\rm max}$, $(k_{\rm max} - k_{\rm min}) < M$, and $0 \leq l_{\rm max} < N/2$. The support set of the precoder filter $a[k,l]$, denoted by $\mathcal{S}_a$, is chosen in such a way that the support set of $h_{\rm a}[k,l]$, denoted by $\mathcal{S}_{h_{\rm a}}$, is $\mathcal{S}_{h_{\rm a}} \triangleq \left\{(k,l) \middle| \frac{-M}{2} \leq k \leq \frac{M}{2} - 1,\ \frac{-N}{2} \leq l \leq \frac{N}{2} - 1\right\}$. Accordingly, $\mathcal{S}_a$ is chosen as

$$\mathcal{S}_a \triangleq \left\{ (k,l) \;\middle|\; \begin{array}{l} -\dfrac{M}{2} - k_{\rm min} \leq k \leq \dfrac{M}{2} - k_{\rm max} - 1, \\ -\dfrac{N}{2} + l_{\rm max} \leq l \leq \dfrac{N}{2} - l_{\rm max} - 1 \end{array} \right\}. \tag{7}$$

I/O relation in matrix-vector-form: We write the end-to-end DD I/O relation in (6) in a matrix-vector form as

$$\mathbf{y} = \mathbf{H}_{\rm eff}\mathbf{s} + \mathbf{n}, \tag{8}$$

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

[1] $(u *_{\sigma\rm d} v)[k,l] = \sum_{k',l'\in\mathbb{Z}} u[k-k',\, l-l']\, v[k',l']\, e^{j2\pi\frac{k'(l-l')}{MN}}$.

[2] $(u *_\sigma v)(\tau,\nu) = \iint u(\tau',\nu')\, v(\tau-\tau',\nu-\nu')\, e^{j2\pi\nu'(\tau-\tau')} d\tau' d\nu'$. Twisted convolution operation is associative and preserves quasi-periodicity.

[3] $\mathcal{Z}_t(x(t)) = \sqrt{\tau_p}\sum_{k\in\mathbb{Z}} x(\tau + k\tau_p)\, e^{-j2\pi\nu k\tau_p}$.

$$
\mathbf{H}_{\text{eff}}[k'N+l'+1,\ kN+l+1] = \begin{cases} h_{\text{a}}[k'-k,\ l'-l]\, e^{j2\pi \frac{(l'-l)k}{MN}}, \\ \quad (k'-k,\ l'-l) \in \mathcal{S}_{h_{\text{a}}}, \\ 0, \quad \text{otherwise}, \end{cases} \tag{9}
$$

where $k', k = 0, \ldots, M-1$ and $l', l = 0, \ldots, N-1$. Assuming i.i.d. zero-mean symbols $s[k,l]$ with energy E, i.e., $\mathbb{E}\big[|h_{\text{a}}[k,l] *_{\sigma \text{d}} s_{\text{dd}}[k,l]|^2\big] = E$, unit-energy transmit and matched receive filters, and a normalized precoding filter satisfying $\sum_{(k,l)\in\mathcal{S}_a} |a[k,l]|^2 = 1$, the average received energy per DD bin is E and the DD domain noise variance is N_0, giving the per-bin signal-to-noise ratio (SNR) $\rho \triangleq E/N_0$.

Precoder vector: Arranging the MN taps of $h_{\text{a}}[k,l]$ into a vector $\mathbf{h}_{\text{a}}$, we write $\mathbf{h}_{\text{a}} = \big(h_{\text{a},0}\,, h_{\text{a},1}\,, \cdots, h_{\text{a},(MN-1)}\big)^T$, where $h_{\text{a}, k+\frac{M}{2}+\left(l+\frac{N}{2}\right)M} = h_{\text{a}}[k,l]$, $(k,l) \in \mathcal{S}_{h_{\text{a}}}$. From discrete twisted convolution definition, it follows that

$$
\mathbf{h}_{\text{a}} = \mathbf{H}'\,\mathbf{a}, \tag{10}
$$

where $\mathbf{a} \in \mathbb{C}^{K\times 1}$, $K = (k_{\min} - k_{\max} + M)(N - 2l_{\max})$, is the precoder vector consisting of the precoding filter taps given by $\mathbf{a} = (a_0\,, a_1\,, \cdots, a_{K-1})^T$, where $a_{k+\frac{M}{2}+k_{\min}+\left(l+\frac{N}{2}-l_{\max}\right)(M-k_{\max}+k_{\min})} \triangleq a[k,l]$ with $(k,l) \in \mathcal{S}_a$. The matrix $\mathbf{H}' \in \mathbb{C}^{MN\times K}$ has entries in its $r(k,l)$th row and $c(k',l')$th column with $(k,l) \in \mathcal{S}_{h_{\text{a}}}$, $(k',l') \in \mathcal{S}_a$, $r(k,l) \triangleq k + \frac{M}{2} + (l + \frac{N}{2})M$, $c(k',l') \triangleq k' + \frac{M}{2} + k_{\min} + (l' + \frac{N}{2} - l_{\max})(M - k_{\max} + k_{\min})$, as

$$
H'_{r(k,l),\, c(k',l')} \triangleq \begin{cases} h_{\text{eff}}[k-k', l-l']\, e^{j2\pi k' \frac{(l-l')}{MN}}, \\ \qquad (k-k', l-l') \in \mathcal{S}_{h_{\text{eff}}}, \\ 0, \quad \text{otherwise}. \end{cases} \tag{11}
$$

Single-tap channel estimator/equalizer at the Rx: If the system operates deep within the crystalline regime [3], i.e., $(k_{\max} - k_{\min}) \ll M, \quad 2l_{\max} \ll N$, the delay and Doppler periods are much larger than the respective channel spreads, the I/O relation becomes predictable, i.e., the entire channel can be characterized from the estimate at a single pilot position. Since the precoder concentrates the channel energy towards the origin, the downlink receiver needs only to estimate $h_{\text{a}}[0,0]$. So, a single pilot symbol without any guard region around it in the downlink frame is sufficient, yielding high spectral efficiency. The pilot symbol is located at $(k_{\text{p}}, l_{\text{p}}) = (M/2, N/2)$ and all other bins are loaded with data symbols, with $s[k_{\text{p}}, l_{\text{p}}] = \sqrt{\eta(MN-1)E}$, where η is the pilot-to-data power ratio (PDR), and the pilot and data energies are $\eta(MN-1)E$ and $(MN-1)E$, respectively. The estimate of $h_{\text{a}}[0,0]$ from the received pilot $y_{\text{dd}}[k_{\text{p}}, l_{\text{p}}]$ is obtained as

$$
\widehat{h}_{\text{a}}[0,0] \triangleq \frac{y_{\text{dd}[k_{\text{p}}, l_{\text{p}}]}}{\sqrt{\eta(MN-1)E}\left(1 + \frac{1}{\rho\eta(MN-1)}\right)}. \tag{12}
$$

Since the precoder pre-compensates the channel spread, the interference between precoded channel taps becomes very small, making a single-tap equalization sufficient. So, the equalized (k,l)th information symbol is obtained as

$$
\widehat{s}[k,l] \;=\; \frac{\widehat{h}^*_{\text{a}}[0,0]\, y_{\text{dd}}[k,l]}{\left(|\widehat{h}_{\text{a}}[0,0]|^2 + \frac{1}{\rho}\right)}. \tag{13}
$$

III. Proposed correlation-maximizing TC precoder

The work in [7] derives the TC precoder vector $\mathbf{a}$ by maximizing the SINR at the downlink receiver. In our current work here, we propose an alternative TC precoder design based on maximization of the correlation between the actual precoded channel response and a desired target precoded channel response. The proposed precoder achieves superior energy localization of $h_{\text{eff}}[k,l]$ around a single dominant DD tap, thus offering improved single-tap estimation/equalization performance compared to the SINR-maximizing (SINR-max) precoder. To formalize the proposed correlation-maximizing (Corr-max) precoder, we define the desired target response $h_{\text{target}}[k,l]$ as

$$
h_{\text{target}}[k,l] \triangleq \begin{cases} \sum\limits_{(k,l)\in\mathcal{S}_{h_{\text{eff}}}} |h_{\text{eff}}[k,l]|^2, & (k,l) = (0,0) \\ 0, & \text{otherwise}, \end{cases} \tag{14}
$$

i.e., all energy from $h_{\text{eff}}[k,l]$ is focused at the origin and zero energy in all other bins. The precoder design problem then aims to find a normalized precoder vector $\mathbf{a}$ such that the resulting effective channel $\mathbf{h}_{\text{a}} = \mathbf{H}'\mathbf{a}$ aligns as closely as possible with the desired target response $\mathbf{h}_{\text{target}}$. Accordingly, we formulate the optimization problem that maximizes the normalized correlation between $\mathbf{h}_{\text{a}}$ and $\mathbf{h}_{\text{target}}$ as

$$
\mathbf{a}_{\text{opt}} = \arg\max_{\|\mathbf{a}\|=1} \underbrace{\frac{|\langle \mathbf{H}'\mathbf{a}, \mathbf{h}_{\text{target}}\rangle|^2}{\|\mathbf{H}'\mathbf{a}\|_2^2}}_{\mathcal{J}(\mathbf{a})}. \tag{15}
$$

To solve (15), define $\mathbf{g} \triangleq \mathbf{H}'^H \mathbf{h}_{\text{target}}, \quad \mathbf{R} \triangleq \mathbf{H}'^H\mathbf{H}'$, so that the objective in (15) becomes

$$
\mathcal{J}(\mathbf{a}) \;=\; \frac{|\mathbf{a}^H\mathbf{g}|^2}{\mathbf{a}^H\mathbf{R}\mathbf{a}} \;=\; \frac{\mathbf{a}^H(\mathbf{g}\mathbf{g}^H)\mathbf{a}}{\mathbf{a}^H\mathbf{R}\mathbf{a}}. \tag{16}
$$

Let $\mathbf{G} \triangleq \mathbf{g}\mathbf{g}^H$ ($\mathbf{G}$ is Hermitian). The maximization of (16) is therefore the generalized Rayleigh quotient problem

$$
\max_{\|\mathbf{a}\|=1} \frac{\mathbf{a}^H\mathbf{G}\mathbf{a}}{\mathbf{a}^H\mathbf{R}\mathbf{a}}. \tag{17}
$$

Assume[4] $\mathbf{R} \succ 0$ (positive definite). Let $\mathbf{b} \triangleq \mathbf{R}^{1/2}\mathbf{a}$, so that $\mathbf{a} = \mathbf{R}^{-1/2}\mathbf{b}$ and $\mathbf{a}^H\mathbf{R}\mathbf{a} = \mathbf{b}^H\mathbf{b}$. Substituting into (16) gives

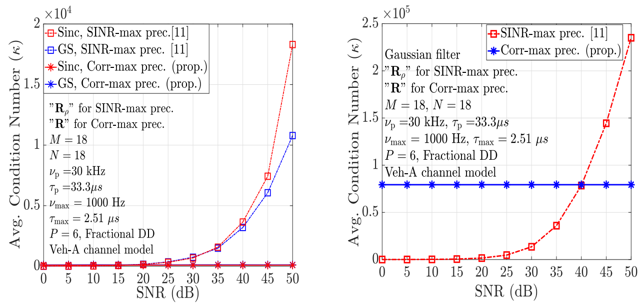
$$
\frac{\mathbf{a}^H\mathbf{G}\mathbf{a}}{\mathbf{a}^H\mathbf{R}\mathbf{a}} = \frac{\mathbf{b}^H\big(\mathbf{R}^{-1/2}\mathbf{G}\mathbf{R}^{-1/2}\big)\mathbf{b}}{\mathbf{b}^H\mathbf{b}} \triangleq \frac{\mathbf{b}^H\mathbf{M}\mathbf{b}}{\mathbf{b}^H\mathbf{b}}, \tag{18}
$$

where $\mathbf{M} \triangleq \mathbf{R}^{-1/2}\mathbf{G}\mathbf{R}^{-1/2}$ is Hermitian. By the Rayleigh-Ritz theorem, the maximum of the quadratic form above (over nonzero $\mathbf{b}$) equals the largest eigenvalue $\lambda_{\max}(\mathbf{M})$, attained when $\mathbf{b}$ is in the direction of the principal eigenvector of $\mathbf{M}$. Using the rank-one structure of $\mathbf{G}$, we obtain

$$
\mathbf{M} \;=\; \mathbf{R}^{-1/2}\mathbf{g}\mathbf{g}^H\mathbf{R}^{-1/2} = \big(\mathbf{R}^{-1/2}\mathbf{g}\big)\big(\mathbf{R}^{-1/2}\mathbf{g}\big)^H. \tag{19}
$$

Thus, $\mathbf{M}$ is an outer product $\mathbf{u}\mathbf{u}^H$ with $\mathbf{u} = \mathbf{R}^{-1/2}\mathbf{g}$, which has a single non-zero eigenvalue $\lambda_{\max}(\mathbf{M}) = \|\mathbf{u}\|^2 = \|\mathbf{R}^{-1/2}\mathbf{g}\|^2$ and principal eigenvector proportional to $\mathbf{u}$. Hence, $\mathbf{b}_{\text{opt}} \propto \mathbf{R}^{-1/2}\mathbf{g}$ and $\mathbf{a}_{\text{opt}} \propto \mathbf{R}^{-1}\mathbf{g} = \mathbf{R}^{-1}\mathbf{H}'^H\mathbf{h}_{\text{target}}$.

[4]If $\mathbf{R}$ is only positive semidefinite (PSD), replace $\mathbf{R}^{-1}$ by the Moore–Penrose pseudoinverse $\mathbf{R}^{\dagger}$.



(a) Sinc and Gaussian-sinc pulses

(b) Gaussian pulse.

Fig. 2: Condition number of $\mathbf{R}_\rho$ and $\mathbf{R}$ vs. SNR for sinc, Gaussian-sinc, and Gaussian pulse shaping filters.

Normalizing to unit norm gives the optimum precoder in closed-form as

$$\mathbf{a}_{\text{opt}} = \frac{\mathbf{R}^{-1}\mathbf{H}'^H\mathbf{h}_{\text{target}}}{\|\mathbf{R}^{-1}\mathbf{H}'^H\mathbf{h}_{\text{target}}\|} = \frac{(\mathbf{H}'^H\mathbf{H}')^{-1}\mathbf{H}'^H\mathbf{h}_{\text{target}}}{\|(\mathbf{H}'^H\mathbf{H}')^{-1}\mathbf{H}'^H\mathbf{h}_{\text{target}}\|}. \tag{20}$$

On the structure and condition number of* $\mathbf{R}$*: Let $p = \frac{M}{2}(N+1)$ and denote the rows of $\mathbf{H}'$ by $\{\mathbf{h}'_i\}_{i=0}^{MN-1}$, with $\mathbf{h}'_p \triangleq \mathbf{H}'^H\mathbf{e}_p$ representing the response at the pivot DD bin. The SINR-max [11] and the proposed Corr-max precoders can be expressed, respectively, as

$$\mathbf{a}_{\text{SINR}}(\rho) = \frac{\mathbf{R}_\rho^{-1}\mathbf{h}'_p}{\|\mathbf{R}_\rho^{-1}\mathbf{h}'_p\|}, \qquad \mathbf{R}_\rho = \frac{\mathbf{I}}{\rho} + \mathbf{R}_\infty, \tag{21}$$

$$\mathbf{a}_{\text{Corr}} = \frac{\mathbf{R}^{-1}\mathbf{h}'_p}{\|\mathbf{R}^{-1}\mathbf{h}'_p\|}, \qquad \mathbf{R} = \mathbf{R}_\infty + \mathbf{h}'_p\mathbf{h}'^H_p, \tag{22}$$

where the common term $\mathbf{R}_\infty \triangleq \sum_{i\neq p}\mathbf{h}'_i\mathbf{h}'^H_i$ captures the contributions from all non-pivot DD domain components, which are typically weak. As a result, $\mathbf{R}_\infty$ is often highly ill-conditioned; for instance, simulations with a sinc pulse yield an average condition number on the order of 10^6. The key distinction between the two precoders lies in the matrices being inverted; $\mathbf{R}_\rho$ for $\mathbf{a}_{\text{SINR}}(\rho)$ and $\mathbf{R}$ for $\mathbf{a}_{\text{Corr}}$. To compare their conditioning, we note that the regularization term $\mathbf{I}/\rho$ in $\mathbf{R}_\rho$ improves numerical stability at low-to-moderate SNRs by preventing the eigenvalues from approaching zero. However, as ρ increases, this term decreases and $\mathbf{R}_\rho \to \mathbf{R}_\infty$, leading to instability in the inversion of $\mathbf{R}_\rho$ in the high-SNR regime. In contrast, the proposed Corr-max precoder inherently includes the rank-one term $\mathbf{h}'_p\mathbf{h}'^H_p$, which, according to *Weyl's inequality*, reduces the eigenvalue spread and therefore the condition number. Thus, $\mathbf{R}$ is inherently better conditioned than $\mathbf{R}_\infty$, ensuring that $\mathbf{a}_{\text{Corr}}$ remains numerically stable even at high SNRs. For the Gaussian pulse, we observe a similar trend with overall higher condition numbers. Specifically, $\mathbf{R}_\infty$ exhibits an average condition number of approximately 4×10^5, while the inclusion of the rank-one term $\mathbf{h}'_p\mathbf{h}'^H_p$ in $\mathbf{R}$ moderately improves conditioning, yielding values around 8×10^4. Furthermore, the regularized matrix $\mathbf{R}_\rho$ remains well-conditioned at low SNRs but eventually surpasses $\mathbf{R}$ as ρ increases. The above points and trends are illustrated in Fig. 2, which shows the variation of the condition number of $\mathbf{R}_\rho$ and $\mathbf{R}$ as a function of SNR for sinc, Gaussian, and Gaussian-sinc (GS) pulse shaping filters.

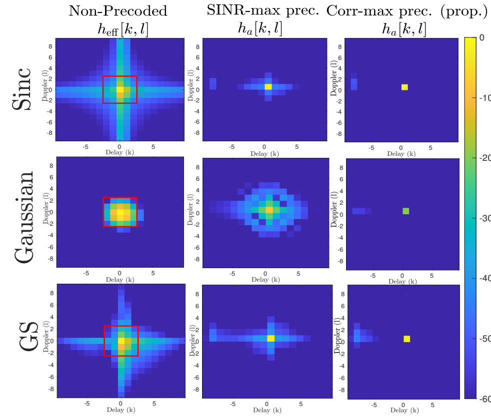


Fig. 3: Heatmaps of non-precoded effective channel and precoded effective channels of SINR-max and proposed Corr-max precoders for different pulse shaping filters.

Comparison of heatmaps for* $\mathbf{a}_{\text{SINR}}$ *and* $\mathbf{a}_{\text{Corr}}$ *precoders: Figure 3 compares the energy localization and interference behavior of the proposed Corr-max precoder relative to that of the SINR-max precoder. It shows the heatmaps of the non-precoded effective channel and the precoded effective channels of SINR-max and Corr-max precoders for sinc, Gaussian, and GS pulse shaping filters. The following system parameters are used to generate these heatmaps: Vehicular A (Veh-A) channel model [12] with $P = 6$ paths, maximum delay $\tau_{\max} = 2.51\mu$s, maximum Doppler $\nu_{\max} = 1$ kHz, $\nu_p = 30$ kHz, $\tau_p = 1/\nu_p = 33.33$ μs, $M = N = 18$, $B = M\nu_p = 540$ kHz, $T = N\tau_p = 0.6$ ms, and $\rho = 15$ dB. The parameters $k_{\max}$, $k_{\min}$, $l_{\max}$, $l_{\min}$ that define the support sets $\mathcal{S}_{h_{\text{eff}}}$, $\mathcal{S}_a$ are chosen such that the crystallization condition[5] is satisfied, which is given by

$$k_{\max} \triangleq \left\lceil \frac{M\,\tau_{\max}}{\tau_p} \right\rceil < M, \quad l_{\max} \triangleq \left\lceil \frac{2N\,\nu_{\max}}{\nu_p} \right\rceil < N, \tag{23}$$

with $k_{\min} = -k_{\max}$ and $l_{\min} = -l_{\max}$.

From the energy distribution of the effective channels before and after precoding in Fig. 3, we can observe the following. 1) Among the considered pulse shaping filters, the sinc pulse achieves the strongest precoded channel energy concentration, and the concentration achieved by Gaussian pulse is the weakest. This behavior arises from two key factors. First, within the support region $\mathcal{S}_{h_{\text{eff}}}$ (indicated by the red bounding box in the effective channel heatmaps), the Gaussian pulse spreads its energy more uniformly, as indicated by the broader yellow regions in the heatmap. In contrast, the sinc and GS pulses exhibit sharper concentration with distinct peaks and weaker sidelobes. Second, the Gaussian pulse does not satisfy the Nyquist criterion on the information grid, since it exhibits non-zero values at the sampling points. The sinc and GS pulses, by contrast, have zeros at these Nyquist points, resulting in better interference suppression. Between the two, the sinc pulse provides a slightly better energy confinement than the GS pulse. 2) The proposed Corr-max precoder further

[5] Crystallization condition refers to the condition where the delay period τ_{p} and the Doppler period ν_{p} are larger than the maximum delay spread $\tau_{\max}$ and the maximum Doppler spread $\nu_{\max}$ of the channel, respectively [3].

enhances energy localization by concentrating the effective channel energy at a single dominant DD point. Although minor leakage energy persists, it remains significantly lower than that of the SINR-max precoder, highlighting the strong focusing capability of the proposed precoder. The residual leakage energy, expressed as a fraction of the total effective channel energy, is computed as

$$\eta_{\text{res}} = \frac{\sum_{(k,l)\in\mathcal{F}\setminus\mathcal{C}} |h_{\text{a}}[k,l]|^2}{\sum_{(k,l)\in\mathcal{F}} |h_{\text{a}}[k,l]|^2}, \tag{24}$$

where $\mathcal{C}$ denotes the dominant DD tap (center point) and $\mathcal{F}$ represents the region of the fundamental period. Table I summarizes the residual leakage energy η_{res} for different pulse shaping filters, comparing the SINR-max and Corr-max precoders. These results demonstrate that the sinc pulse

TABLE I: Residual leakage energy η_{res} (% of total energy) for different pulse shaping filters at $\rho = 15$ dB.

Precoder	Sinc	Gaussian	GS
SINR-max	4.03%	20.38%	4.48%
Corr-max (prop.)	2.58%	3.70%	2.89%

yields the most favorable energy concentration in TC precoded systems, and the proposed Corr-max precoder achieves better precoded energy concentration than the SINR-max precoder.

A. *DNN-aided CSIT estimation*

Knowledge of CSIT is required for precoding. In TDD systems where channel reciprocity holds, this can be obtained through pilot transmission from UE on the uplink and channel estimation at the BS, and the estimated CSIT is used for TC precoding. Here, we propose such a CSIT estimation scheme aided by a DNN. A Zak-OTFS frame (exclusive or embedded pilot frame) is sent from UE to BS, which estimates the channel using a model-free read-off operation. The noisy estimates from the read-off operation are enhanced through 'denoising' by a DNN (see Fig. 4). The proposed DNN architecture for this purpose is described below.

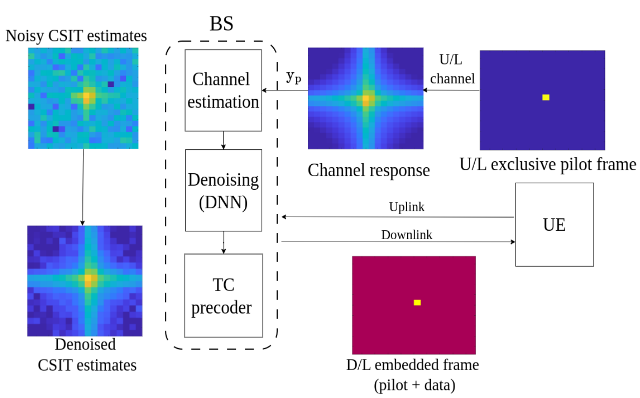


Fig. 4: Uplink pilot-based CSIT estimation with DNN-aided denoising.

Training dataset: Let the UE transmit a pilot frame $\mathbf{x}_{\text{p}} \in \mathbb{C}^{MN\times 1}$ with a pilot at $(k_{\text{p}}, l_{\text{p}}) = (M/2, N/2)$ with some pilot SNR ρ_{p}. The received pilot at the BS is

$$\mathbf{y}_{\text{p}} = \mathbf{H}_{\text{eff}}\mathbf{x}_{\text{p}} + \mathbf{n}_{\text{p}}, \tag{25}$$

from which the BS obtains a noisy estimate of the effective channel $\hat{h}_{\text{eff}}[k,l]$ through a model-free read-off [3]. Vectorizing the estimated effective channel coefficients over the support region $\mathcal{S}_{h_{\text{eff}}}$, we obtain $\hat{\mathbf{h}}_{\mathcal{S}_h}$. The real and imaginary components of $\hat{\mathbf{h}}_{\mathcal{S}_h}^{(i)}$ are treated as independent real-valued inputs to the DNN, denoted by

$$\mathbf{x}_{\text{train}}^{(i)} \in \{\Re(\hat{\mathbf{h}}_{\mathcal{S}_h}^{(i)}), \Im(\hat{\mathbf{h}}_{\mathcal{S}_h}^{(i)})\}, \tag{26}$$

with corresponding target outputs

$$\mathbf{y}_{\text{train}}^{(i)} \in \{\Re(\mathbf{h}_{\mathcal{S}_h}^{(i)}), \Im(\mathbf{h}_{\mathcal{S}_h}^{(i)})\}. \tag{27}$$

Each component is treated as a separate training sample, allowing the DNN to learn real-valued mappings independently for real and imaginary parts.

DNN architecture: The proposed DNN has three hidden layers with decreasing widths, as follows:
Input $\rightarrow$ *80* $\rightarrow$ *SiLU* $\rightarrow$ *LayerNorm* $\rightarrow$ *60* $\rightarrow$ *SiLU* $\rightarrow$ *LayerNorm* $\rightarrow$ *40* $\rightarrow$ *SiLU* $\rightarrow$ *LayerNorm* $\rightarrow$ *Output.*
The detailed configuration and training parameters of the proposed DNN are summarized in Table II.

Complexity: The total number of trainable parameters is $\sum_{l=1}^{L}(n_{l-1}n_l + 3n_l)$, where L is the number of layers, n_{l-1} and n_l are the number of input and output neurons of the lth layer, including weights, biases, and LayerNorm parameters. For the proposed DNN, the number of trainable parameters is 10765, indicating a compact and computationally efficient architecture.

TABLE II: Proposed DNN and training parameters.

Parameters	Proposed DNN
Input / Output neurons	25 / 25
Hidden layers	3 (SiLU activation)
Output activation	Linear
Training examples	20,000
Epochs / Batch size	60 / 128
Optimizer / Loss	Adam / RMSE

Loss function: We define the RMSE loss in terms of real and imaginary error norms:

$$\mathcal{L}_{\text{RMSE}} = \sqrt{\frac{1}{N_{\text{train}}}\sum_{i=1}^{N_{\text{train}}}\left(E_{\text{Re}}^{(i)} + E_{\text{Im}}^{(i)}\right)}, \tag{28}$$

where N_{train} is the number of training samples, and the per-sample errors are

$$E_{\text{Re}}^{(i)} = \|\Re(\tilde{\mathbf{h}}_{\mathcal{S}_h}^{(i)}) - \Re(\mathbf{h}_{\mathcal{S}_h}^{(i)})\|_2^2, \quad E_{\text{Im}}^{(i)} = \|\Im(\tilde{\mathbf{h}}_{\mathcal{S}_h}^{(i)}) - \Im(\mathbf{h}_{\mathcal{S}_h}^{(i)})\|_2^2, \tag{29}$$

where $\tilde{\mathbf{h}}_{\mathcal{S}_h}^{(i)} = g_{\Theta}(\hat{\mathbf{h}}_{\mathcal{S}_h}^{(i)})$ denotes the denoised channel estimate produced by the DNN for the ith input $\hat{\mathbf{h}}_{\mathcal{S}_h}^{(i)}$, and Θ represents all learnable network parameters.

IV. Results and Discussion

In this section, we present the simulated SER performance of the proposed Corr-max precoder in comparison with that of the SINR-max precoder in [11]. The system parameters used are the same as the ones used for Fig. 3. Figure 5 shows the SER performance of the SINR-max and the proposed Corr-max precoders for sinc, Gaussian, and GS pulses, assuming

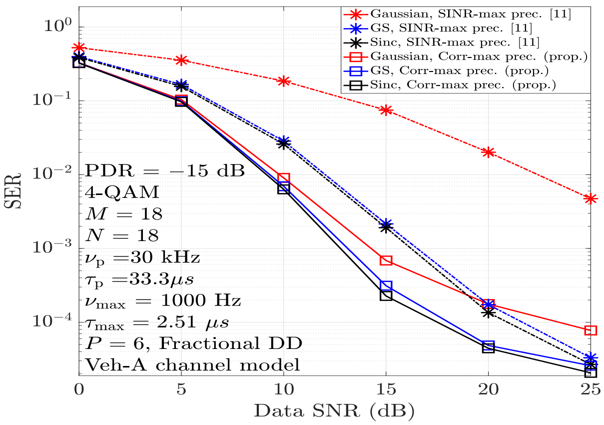


Fig. 5: SER performance of SINR-max and proposed Corr-max precoders for different pulse shaping filters. Perfect CSIT.

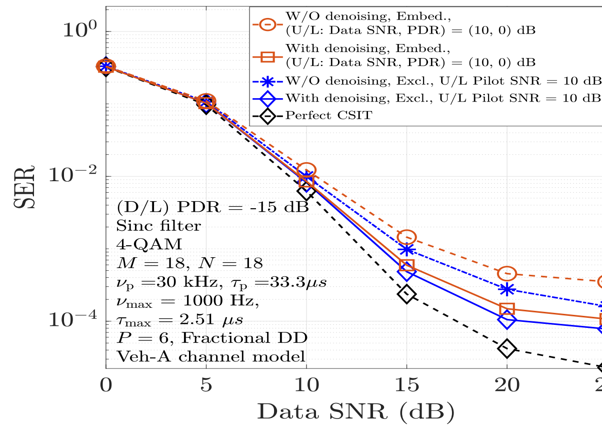


Fig. 6: SER performance of proposed Corr-max precoder for sinc pulse with estimated CSIT (with/without DNN denoising).

perfect CSIT (Fig. 6 shows the performance with estimated CSIT). The PDR is fixed at -15 dB. It can be seen that the proposed Corr-max precoder gives substantial performance gains over the SINR-max precoder for all the pulses considered. For example, at an SER of 10^{-3}, the proposed Corr-max precoder achieves an SNR gain of about 4 dB, >12 dB, and 4 dB for sinc, Gaussian, and GS pulses, respectively, compared to the SINR-max precoder. The gain is essentially due to the superior concentration of the precoded channel energy achieved by the proposed Corr-max precoder (as observed in the heatmaps in Fig. 3). Also, the tendency of the SER to floor at high SNRs can be attributed to the increased interference between pilot and data symbols for increasing SNR. Note that, i) for a fixed PDR, increasing SNR increases pilot power, ii) there is no guard space between the pilot and data symbols, and iii) a small fraction of leakage energy remains despite the energy concentration achieved by the precoders (see Table I).

Figure 6 shows the impact of estimated CSIT on the SER performance of the proposed Corr-max precoder with sinc pulse, comparing the cases with and without DNN denoising. Perfect CSIT is also shown as a benchmark. CSIT estimation is performed using either exclusive or embedded pilot frames, both featuring a pilot symbol at $(M/2, N/2)$. In the exclusive frame, the pilot is transmitted with an SNR of 10 dB. In the embedded frame, the uplink efficiency is 61.1%, with dedicated pilot and guard regions and the remaining region used for data at a data SNR of 10 dB and PDR of 0 dB. The results demonstrate that the proposed DNN-based denoising significantly improves performance, and that the SER for both exclusive and embedded frames approaches the perfect-CSIT benchmark, highlighting the effectiveness of the proposed Corr-max precoder and the DNN-based denoising. Also, due to interference from the data in the embedded frame, its performance is slightly worse than that of the exclusive frame.

V. Conclusion

We presented a twisted convolution (TC) based precoder for Zak-OTFS and its SER performance. While a recent work on this new topic obtained the optimum precoder that maximized the per-carrier (i.e., per DD bin) SINR, we obtained the optimum precoder that maximized the correlation between the actual precoded channel response vector and a desired target precoded channel response vector. The proposed Corr-max precoder was shown to achieve superior energy concentration in the DD domain at the receiver, and hence superior SER performance compared to the SINR-max precoder. We also proposed a DNN-based denoiser that enhanced the CSIT estimate for precoding and hence improved the SER. Future work could include TC precoders for multiuser downlink scenarios and learning-based precoder design for Zak-OTFS.

References

[1] A. J. E. M. Janssen, "The Zak transform: a signal transform for sampled time-continuous signals," *Philips J. Res.*, 43, pp. 23-69, 1988.

[2] 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.

[3] 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.

[4] S. Gopalam, H. Inaltekin, I. B. Collings, and S. V. Hanly, "Optimal Zak-OTFS receiver and its relation to the radar matched filter," *IEEE Open J. of the Commun. Soc.*, vol. 5, pp. 4462-4482, 2024.

[5] J. Jayachandran, et al., "Zak-OTFS with interleaved pilots to extend the region of predictable operation," *IEEE Trans. Veh. Tech.*, doi: 10.1109/TVT.2025.3579394, 2025.

[6] F. Jesbin and A. Chockalingam, "Near-optimal detection of Zak-OTFS signals," *IEEE ICC'2024*, pp. 4476-4481, Jun. 2024.

[7] A. Das, F. Jesbin, and A. Chockalingam, "A Gaussian-sinc pulse shaping filter for Zak-OTFS," *IEEE Trans. Veh. Tech.*, 2025. doi: 10.1109/TVT.2025.3609781.

[8] S. R. Mattu, N. Mehrotra, and R. Calderbank, "Differential communication in channels with mobility and delay spread using Zak-OTFS," *IEEE Wireless Commun. Lett.*, doi: 10.1109/LWC.2025.3600660.

[9] A. Bairwa and A. Chockalingam, "MIMO-Zak-OTFS in doubly-selective channels: I/O relation estimation and detection," *IEEE VTC'2025-Spring*, pp. 1-6, Jun. 2025.

[10] N. B. Nazir, A. Chockalingam, "DNN-pased joint I/O relation estimation and detection in Zak-OTFS," *IEEE VTC'2025-Spring*, pp. 1-6, Jun. 2025.

[11] S. K. Mohammed, et al., "Precoded Zak-OTFS for per-carrier equalization," arXiv:2507.09894, 18 Jul 2025.

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