# Predictive Beamforming for OTFS-Enabled URLLC in High-Mobility Vehicular Networks

Jianzhe Xue, *Graduate Student Member, IEEE*, Tiankai Jiang, *Graduate Student Member, IEEE*, Zhanxi Ma, *Graduate Student Member, IEEE*, Yunting Xu, *Member, IEEE*, Haibo Zhou, *Senior Member, IEEE*, and Xuemin Shen, *Fellow, IEEE*

***Abstract*—Ultra reliable low latency communication (URLLC) in vehicular networks is pivotal for meeting the stringent requirements of transportation safety. However, achieving it is very challenging due to the high mobility of vehicles and the complex propagation environment. Orthogonal time frequency space (OTFS) modulation addresses these issues by modulating symbols into the delay-Doppler (DD) domain, which can leverage full time-frequency diversity by spreading each DD domain symbol across the entire time-frequency plane. In this paper, we propose a novel OTFS-enabled ultra reliable low latency vehicular network architecture for downlink transmission. To achieve low latency, we adopt frequency division duplex (FDD) mode to transmit data frames as soon as they arrive to minimize scheduling delays. Furthermore, to enhance the received signal strength at the receiver, beamforming is applied at the transmitter. Due to the channel state information (CSI) feedback delay in FDD systems, we design a deep learning algorithm, the DD-domain Convolutional Transformer (DDCT), for predictive beamforming based on historical DD-domain CSI. In DDCT, a convolutional neural network extracts spatial features from the DD domain, and a transformer captures their temporal correlations. Extensive simulation results demonstrate the effectiveness of the proposed vehicular network architecture and the superiority of the deep learning algorithm for predictive beamforming.**

***Index Terms*—URLLC, vehicular network, OTFS, deep learning, beamforming.**

## I. Introduction

ULTRA reliable low latency communications (URLLC) vehicular networks are envisioned as a cornerstone for enabling safer and smarter transportation systems [1]. URLLC has garnered significant attention as a key technology in reliable-sensitive applications, and the international telecommunication union has defined hyper reliable low latency communication (HRLLC) as one of the six usage scenarios of international mobile telecommunications (IMT)-2030 as its extension [2]. However, achieving both reliability and latency requirements is particularly challenging in high-mobility environments like vehicular networks due to severe Doppler shifts and rapidly changing channel conditions [3]. Therefore, designing suitable wireless transceivers that account for the characteristics of high-mobility channels is critical for ensuring URLLC in vehicular networks.

The stringent latency requirements of vehicular networks demand that data frames be transmitted immediately upon arrival at the base station (BS), thereby minimizing scheduling delays [4]. In traditional time division duplexing (TDD) systems, time is partitioned into alternating uplink and downlink slots, which forces data frames to wait for the designated downlink slot before transmission. For example, the switching interval in LTE systems can be as long as 5 or 10 ms, which is intolerable for vehicular applications [5]. In contrast, frequency division duplexing (FDD) allocates separate frequency bands for uplink and downlink operations, enabling continuous and simultaneous transmission and reception of signals. This simultaneous operation helps maintain low latency by eliminating the need for slot switching, allowing vehicular data frames to be transmitted immediately upon arrival at the BS via the downlink channel. Consequently, FDD offers enhanced flexibility and is better suited for next-generation URLLC transceivers in vehicular networks, effectively meeting their stringent latency requirements [6], [7].

To enhance the reliability of vehicular networks, the orthogonal time frequency space (OTFS) modulation is proposed to address the complex distortions of high-mobility vehicular channels [8]. Unlike quasi-static channels that are only frequency-selective fading due to multi-path effects, vehicular channels are also time-selective fading due to Doppler effects induced by relative motion among the transmitter, scatterers, and receiver. Conventional orthogonal frequency division multiplexing (OFDM) modulation operating in the time-frequency (TF) domain is mainly designed for frequency-selective fading channels, and suffers from severe inter-carrier interference when applied to such doubly-selective fading channels [9]. OTFS maps data symbols into the delay-Doppler (DD) domain, where each DD domain symbol is spread across

Received 19 December 2024; revised 18 April 2025; accepted 29 June 2025. Date of publication 8 July 2025; date of current version 29 December 2025. This work is supported in part by the National Natural Science Foundation of China under Grant 623B2052 and 62271244. The associate editor coordinating the review of this article and approving it for publication was T. H. Luan. *(Corresponding author: Haibo Zhou.)*

Jianzhe Xue, Tiankai Jiang, Zhanxi Ma, and Haibo Zhou are with the School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China (e-mail: jianzhexue@smail.nju.edu.cn; tiankaijiang@smail.nju.edu.cn; zx_ma@smail.nju.edu.cn; haibozhou@nju.edu.cn).

Yunting Xu is with the School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China, and also with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: yunting.xu@ntu.edu.sg).

Xuemin Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: sshen@uwaterloo.ca).

Digital Object Identifier 10.1109/TCCN.2025.3587126

the entire TF domain via the inverse symplectic finite Fourier transform, exploiting full time-frequency channel diversity and enabling the high-mobility channel to have a nearly uniform effect on each data symbol [10]. In other words, OTFS transforms the TF domain channel of OFDM, which undergoes both frequency and time fading, into a DD domain channel where the channel conditions are approximately constant. This transformation positions OTFS as a promising approach for advancing physical layer modulation in vehicular networks, ensuring robust and reliable communication in high-mobility scenarios [11].

Moreover, enhancing the received signal strength (RSS) at the receiver is essential for improving reliability in vehicular networks. Beamforming techniques, which use antenna arrays to transmit directional signals, can significantly improve RSS [12]. However, acquiring real-time channel state information (CSI) for beamforming poses a major challenge in FDD-based vehicular networks due to inherent feedback delays of FDD systems and the short coherence time of high-mobility vehicular channels. Firstly, the frequency separation between uplink and downlink channels implies they may exhibit different characteristics, preventing the direct use of uplink CSI for downlink beamforming [13]. Secondly, the CSI feedback loop in FDD systems, which includes channel estimation at the receiver, feedback transmission to the transmitter, and processing at the transmitter, introduces delays that can render the CSI outdated, especially in rapidly changing vehicular channels. Moreover, conventional reactive beamforming, which adjusts the beamforming matrix after packet arrival, may not respond quickly enough for vehicular networks requiring timely communication. Consequently, predictive beamforming, which leverages historical CSI to compute the beamforming matrix for upcoming data transmissions, emerges as a promising solution for URLLC in FDD-based vehicular networks. By forecasting future channel states and adjusting the beamforming matrix accordingly, predictive beamforming ensures consistent signal quality and enhances reliability [14].

Achieving predictive beamforming depends on effective capturing features within historical CSI, which entails understanding channel conditions across historical time slots and modeling their temporal variations [15], [16]. Deep learning (DL) is regarded as a promising solution to build a connection between historical CSI and predictive beamforming matrix, due to its powerful learning ability [17], [18]. The CSI of the OTFS modulation system is represented in the DD domain by a set of complex numbers arranged in a two-dimensional matrix. Since the DD domain CSI can be interpreted as grid data, convolutional neural networks (CNNs) can efficiently extract the channel property by learning the spatial features of each CSI. Moreover, modeling the channel variation is important for predictive beamforming. By sorting the historical CSI in time, we can obtain a sequence of data that contains the characteristics of channel variations. Neural networks such as the transformer can effectively model the channel variation by learning the temporal features of sequential historical CSI [19]. Consequently, the data characteristics of historical DD domain CSI and the learning ability of DL make DL an excellent tool for predictive beamforming for the OTFS-enabled vehicular networks [20].

In this paper, we propose an OTFS-enabled ultra reliable low latency vehicular network architecture in the FDD mode, and a corresponding DL algorithm for predictive beamforming. To reduce the latency caused by the time slot scheduling in traditional TDD mode and meet the stringent latency requirement of vehicular networks, the FDD mode is used to offer more time flexibility. Considering the high mobility of the vehicular channel, we adopt the OTFS modulation to improve transmission reliability. Moreover, beamforming is implemented at the transmitter to improve the RSS at the receiver. Given the restriction of FDD and latency requirement, there is a necessity for predictive beamforming using historical CSI. To this end, a DL algorithm is designed for predictive beamforming based on the historical DD domain CSI, where the convolutional neural network extracts spatial features in each CSI and the transformer learns their temporal correlations. Furthermore, we evaluate the performance of the proposed vehicular network architecture integrated with the DL-based predictive beamforming algorithm through comprehensive simulations. The main contributions of the paper are summarized as follows:

- We propose a ultra reliable low latency vehicular network architecture in FDD mode. The vehicular data packet can be transmitted as soon as it arrives at the transmitter without waiting for the time slot scheduling in the traditional TDD mode, which further improves the delay performance.
- We leverage the OTFS modulation for the transceiver design in vehicular networks. By modulating symbols into the DD domain, the influence of the Doppler shifts and delay spreads of high-mobility wireless channels can be reduced, which improves the reliability of the physical layer.
- We design a DL algorithm named DD-domain Convolutional Transformer (DDCT) for predictive beamforming by leveraging historical DD domain CSI. The CNN is used for learning the spatial feature of each two-dimensional CSI, while the transformer is used for learning the temporal features of channel variations.

The structure of this paper is as follows. Section II reviews the related works on OTFS and vehicular networks. Section III outlines the OTFS modulation and DD domain channel in vehicular network. Section IV presents the design framework and key components of the transceiver. Section V presents the proposed DL-based predictive beamforming algorithm. Section VI discusses comprehensive simulation results, followed by the conclusions of this work in Section VII.

## II. Related Work

Vehicular network is one of the main use cases expected in 6G wireless networks, as it enables vehicular safety-related applications that require URLLC. To meet its stringent performance requirements, advanced transceivers are needed, which can provide effective signal transmission and reception in a challenging vehicular environment. Due to the high

mobility of vehicles and the complex multi-path propagation, the vehicular wireless channel exhibits time-frequency doubly-selectivity, which poses significant challenges for conventional modulation schemes. OTFS is a promising technique that transforms the doubly-selective fading channels into nearly flat fading channels in the DD domain and thus establishes reliable communication for high-mobility devices [21]. An OTFS-enabled transceiver is developed with both analog beamforming and hybrid beamforming, where a two-stage approach is designed for transmit and receiver to maximize the directional beamforming gains [22]. To obtain the CSI, a channel estimation scheme based on the embedded pilot is designed by arranging data, guard, and pilot symbols in the DD domain to mitigate the interference between data and pilot symbols [23]. For the OTFS receiver, a linear complexity detector that extracts and coherently combines received components of transmitted symbols using maximal ratio combining is studied to improve the detection accuracy [24].

The application of the OTFS modulation technique in vehicular networks has received significant attention. An OTFS system with spatial-index modulation is designed to improve the reliability of vehicular communications, in which a three-dimensional index modulation approach is adopted by utilizing the transmit antenna, Doppler, and delay indexes [25]. The minimum mean square error and zero-forcing receivers with low complexity are designed and analyzed for OTFS-enabled vehicular networks, considering both perfect and imperfect CSI [26]. To minimize the future data frame error rate, a DL-based symbol precoder is designed utilizing the estimated historical CSI [27]. Noted that the symbol precoding here is to adjust the symbols in the DD domain and can be applied to the single-input single-output system, rather than using multiple antennas for beamforming. To maximize the spectral efficiency, an embedded OTFS pilot pattern for vehicular networks is designed to minimize the air time of the transceiver and conserve the resources, in which the position of the pilot is optimized with respect to speed [28]. Most of the existing works about OTFS for vehicular networks focus on physical layer signal processing, and less on transceiver design.

Recently, DL techniques have been exploited for 6G communication systems, developing efficient beamforming strategies for its ability of learning features from data without prior knowledge [29], [30], [31], [32], [33]. In 5G FDD systems that need huge amounts of the measured downlink CSI feedback, the uplink CSI is used to infer the downlink CSI by an attention-based DL approach to eliminate the feedback overhead [34]. A joint deep learning framework incorporating power optimization and channel prediction is proposed for beamforming prediction using only historical channel data but not current channel information [20]. Predictive beamforming and beam tracking are explored with the aid of radar-communication signals, in which the locations of vehicles are estimated from the reflected echo signals [35]. The predictive beamforming for supporting URLLC in vehicular networks is envisioned with the potential of combining sensing data to infer the vehicle states [36], [37]. Beamforming for OTFS systems based on codebook matrices has not been investigated yet. The challenge of DL-based beamforming is to establish an accurate mapping from known input information to the beamforming matrix. Specifically, the historical DD domain CSI has high dimensions and complex spatial-temporal correlations, which poses challenges to feature extraction in deep learning.

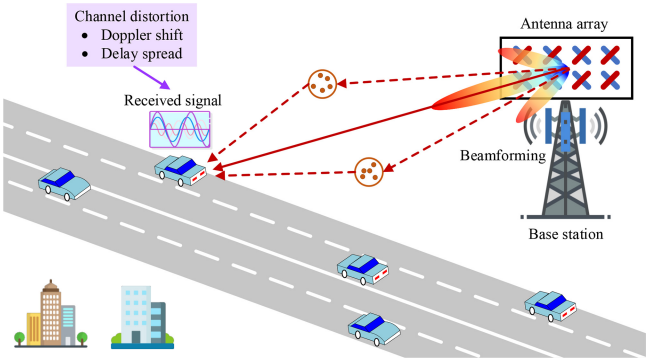



Fig. 1. An illustration of vehicular network downlink scenario.

TABLE I
NOTATIONS

| Parameter | Notation |
|---|---|
| Transmitted DD domain symbol matrix | $\mathbf{X_{TF}}$ |
| Received DD domain symbol matrix | $\mathbf{Y_{DD}}$ |
| Transmitted time domain signal | $s(t)$ |
| Received time domain signal | $r(t)$ |
| Delay spread | $\tau_p$ |
| Doppler shift | $\nu_p$ |
| Channel response coefficient | $h_p$ |
| Historical DD domain channel | $\mathbf{H}_{\mathbf{DD}}^{t}$ |
| The number of delay bins | $M$ |
| The number of Doppler bins | $N$ |

Different from the existing studies, in the follows, we focus on designing an OTFS-enabled ultra reliable low latency vehicular network architecture operating in FDD mode, and leveraging DL to build the connection between historical DD domain CSI and predictive beamforming matrix in codebook.

## III. System Model

In this section, we commence by presenting a vehicular network scenario under consideration. We elucidate the fundamental principles and mathematical frameworks underpinning OTFS modulation and channel model, as well as detailing the input-output relationship of signals within the DD domain. For convenience, Table I summarizes the major notations.

### A. Vehicular Network Scenario

The scenario depicted in Fig. 1 involves an OTFS-enabled vehicular network architecture that is configured as a multiple-input single-output (MISO) system. This system comprises a BS equipped with $N_{\mathrm{tx}}$ transmit antennas, serving as the transmitter, and multiple vehicles, each possess a single receive antenna $N_{\mathrm{rx}} = 1$, serving as the receivers. Given the dynamic nature of vehicular channels, characterized by high mobility and the presence of scatterers, the received signal is subject to alterations in the form of Doppler shift and delay spread. OTFS

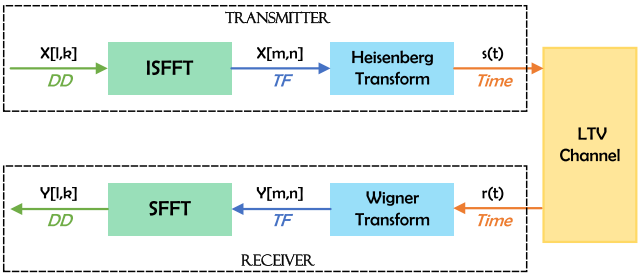



Fig. 2. The transmitter and receiver for OTFS modulation.

modulation is applied to address these challenges. Moreover, to fulfill the stringent reliability requirement of vehicular networks, it becomes imperative to employ the BS's antenna array for the implementation of predictive beamforming.

### B. OTFS Modulation and Demodulation

We delineate the architecture of an OTFS transceiver system, as illustrated in Fig. 2. We designate $M$ as the number of delay bins or subcarriers and $N$ as the number of Doppler bins or time slots, with $T$ representing the symbol duration and $\triangle f = 1/T$ denoting the subcarrier spacing. A constellation of $M \times N$ quadrature amplitude modulation (QAM) symbols, $\mathbf{X_{DD}}[l,k] \in \mathbb{C}^{M\times N}$, $l = 1,\ldots,M$ and $k = 1,\ldots,N$, is systematically arranged on the two-dimensional DD grid.

Initially, the DD domain symbol matrix, $\mathbf{X_{DD}}$, undergoes a transformation into the TF domain, $\mathbf{X_{TF}}[m,n] \in \mathbb{C}^{M\times N}$, $m = 1,\ldots,M$ and $n = 1,\ldots,N$, utilizing the inverse symplectic finite Fourier transform (ISFFT), given as

$$\mathbf{X_{TF}}[m,n] = \frac{1}{\sqrt{MN}} \sum_{k=1}^{N}\sum_{l=1}^{M} \mathbf{X_{DD}}[l,k] e^{j2\pi\left(\frac{nk}{N}-\frac{ml}{M}\right)}. \quad (1)$$

The ISFFT amounts to an $M$-point FFT of the columns and an $N$-point IFFT of the rows of $\mathbf{X}$, which can be expressed in matrix form as $\mathbf{X_{TF}} = \mathbf{F}_M \mathbf{X_{DD}} \mathbf{F}_N^H$, with superscript $H$ representing the Hermitian transpose and $\mathbf{F}_N$ denoting the discrete Fourier transform (DFT) matrix of size $N \times N$. Subsequently, the Heisenberg transform, typically realized through OFDM modulation or the inverse fast Fourier transform (IFFT), is applied to $\mathbf{X_{TF}}$ based on the pulse waveform $g_{\rm tx}(t)$ to generate the time domain signal, *s(t)*, given by

$$s(t) = \sum_{n=1}^{N}\sum_{m=1}^{M} \mathbf{X_{TF}}[m,n] g_{\rm tx}(t-nT) e^{j2\pi m\Delta f(t-nT)}. \quad (2)$$

The corresponding matrix formulation, using an $M$-point IFFT along with the pulse waveform, is given by $\mathbf{S} = \mathbf{G}_{\rm tx}\mathbf{F}_M^H \mathbf{X_{TF}} = \mathbf{G}_{\rm tx}\mathbf{X_{DD}}\mathbf{F}_N^H$, where $\mathbf{G}_{\rm tx} = \mathrm{diag}\{g_{\rm tx}(\frac{iT}{M})\}_{i=1}^{M}$. The $MN$-samples of the transmitted signal can be derived by vectorizing $\mathbf{S}$, given as

$$\mathbf{s} = \mathrm{vec}(\mathbf{S}) = \left(\mathbf{F}_N^H \otimes \mathbf{G}_{\rm tx}\right)\mathbf{x} \in \mathbb{C}^{MN\times 1}, \quad (3)$$

where $\otimes$ denotes the Kronecker product and $\mathbf{x} = \mathrm{vec}(\mathbf{X_{DD}}) \in \mathbb{C}^{MN\times 1}$ is the vectorization of transmitted DD domain symbols.

At the receiver, time-domain signal *r(t)* is converted back to the TF domain symbol matrix $\mathbf{Y_{TF}}[m,n] \in \mathbb{C}^{M\times N}$ through the Wigner transform, employing the matched pulse waveform $g_{\rm rx}(t)$, given as

$$\mathbf{Y_{TF}}[m,n] = \int g_{\rm rx}^*(t-nT) r(t) e^{-j2\pi m\Delta f(t-nT)} dt, \quad (4)$$

where superscript $*$ denotes the conjugation. Mirroring the transmitter's approach, this process can be implemented by OFDM demodulation or fast Fourier transform (FFT). Therefore, (4) can be represented in matrix form, using an $M$-point FFT along with the pulse waveform, as $\mathbf{Y_{TF}} = \mathbf{F}_M \mathbf{G}_{\rm rx} \mathbf{R}$, where $\mathbf{G}_{\rm rx} = \mathrm{diag}\{g_{\rm rx}^*(\frac{iT}{M})\}_{i=1}^{M}$. The received DD domain symbol matrix, $\mathbf{Y_{DD}}[l,k] \in \mathbb{C}^{M\times N}$, is then derived using the symplectic finite Fourier transform (SFFT), given as

$$\mathbf{Y_{DD}}[l,k] = \frac{1}{\sqrt{MN}} \sum_{n=1}^{N}\sum_{m=1}^{M} \mathbf{Y_{TF}}[m,n] e^{-j2\pi\left(\frac{nk}{N}-\frac{ml}{M}\right)}. \quad (5)$$

Similar to the transmitter side, the matrix representation of SFFT amounts to an $M$-point IFFT of the columns and an $N$-point FFT of the rows of $\mathbf{X}$, which can be expressed as $\mathbf{Y_{DD}} = \mathbf{F}_M^H \mathbf{Y_{TF}} \mathbf{F}_N = \mathbf{G}_{\rm rx}\mathbf{R}\mathbf{F}_N$. The corresponding $MN$-received DD domain symbols can be subsequently obtained by vectorizing $\mathbf{Y_{DD}}$ as

$$\mathbf{y} = \mathrm{vec}(\mathbf{Y_{DD}}) = (\mathbf{F}_N \otimes \mathbf{G}_{\rm rx})\mathbf{r} \in \mathbb{C}^{MN\times 1}, \quad (6)$$

where $\mathbf{r} = \mathrm{vec}(\mathbf{R}) \in \mathbb{C}^{MN\times 1}$ is the vectorization of received time domain samples.

### C. DD Domain Channel Model

The vehicular channel encompassing $P$ resolvable paths can be comprehensively characterized within the DD domain by parameters $(\tau_p, \nu_p, h_p)$ for $p = 1,\ldots,P$. Here, $\tau_p$, $\nu_p$, and $h_p$ signify the delay spread, Doppler shift, and channel response coefficient for the $p$-th path, respectively. Thereby, the DD domain channel impulse response can be conveniently expressed as

$$h(\tau,\nu) = \sum_{p=1}^{P} h_p \delta(\tau-\tau_p)\delta(\nu-\nu_p), \quad (7)$$

where $\delta(\cdot)$ denotes the Dirac delta function. For the MISO system depicted in Fig. 1, the fading coefficient in (7) can be replaced by $\mathbf{h}_p \in \mathbb{C}^{1\times \mathrm{N_{tx}}}$, which represents the complex fading matrix between each transmit-receive antenna pair for the $p$-th path. Consequently, the received time domain signal, *r(t)*, resulting from *s(t)* transmitted over the vehicular channel, is formulated as

$$\begin{aligned} r(t) &= \iint h(\tau,\nu)\mathbf{w}s(t-\tau)e^{j2\pi\nu(t-\tau)}\mathrm{d}\tau\mathrm{d}\nu + n(t) \\ &= \sum_{p=1}^{P} \mathbf{h_p}\mathbf{w}s(t-\tau_p)e^{j2\pi\nu_p(t-\tau_p)} + n(t) \end{aligned} \quad (8)$$

where $\mathbf{w} \in \mathbb{C}^{N_{\rm tx}\times 1}$ is the beamforming matrix designed at the transmitter and *n(t)* is the additive white Gaussian

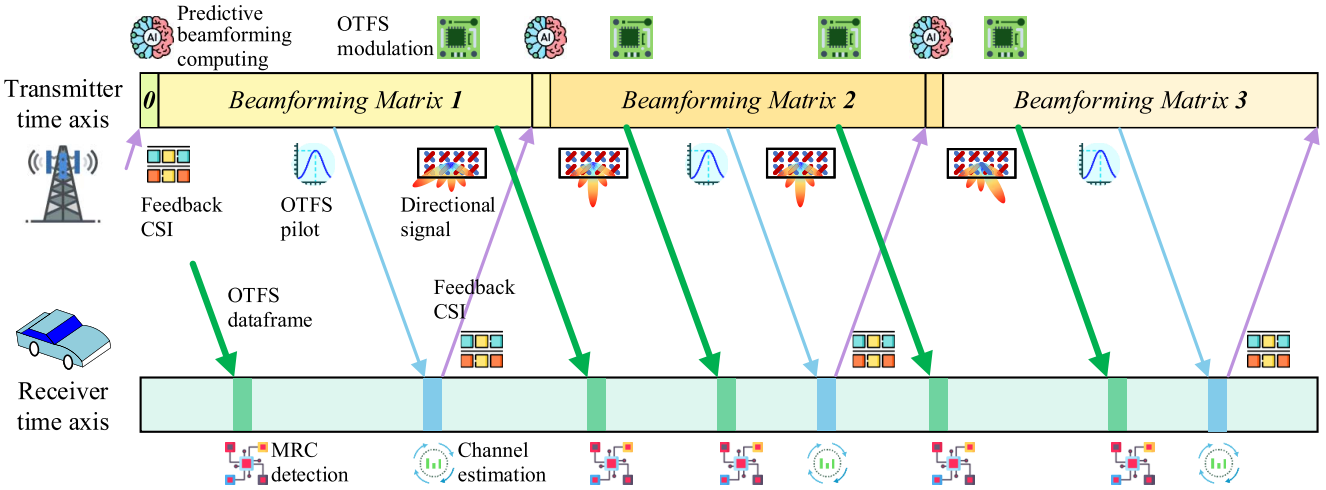



Fig. 3. Transceiver framework in vehicular networks.

noise (AWGN). The delay and Doppler shift for each path is quantified by

$$\tau_p = \frac{l_p}{M\Delta f}, \nu_p = \frac{k_p}{NT}. \tag{9}$$

For simplification, these shifts are considered as integer multiples of $1/M\Delta f$ and $1/NT$, implying that $l_p$ and $k_p$ are integers. The maximum delay and Doppler shifts are denoted by $\tau_{\max} = l_{\max}/M\Delta f$ and $\nu_{\max} = k_{\max}/NT$, where $l_{\max}$ and $k_{\max}$ are the maximum delay shift and Doppler shift of the channel, respectively. We assume an under-spread channel condition, where $l_p \leq l_{\max} < M$, $-N/2 < -k_{\max} \leq k_p \leq k_{\max} < N/2$, and $\tau_{\max}\nu_{\max} \ll 1$. The $q$-th received signal sample $r(q) = \{r(t)|_{t=\frac{q}{M\Delta f}}\}_{q=1}^{MN}$ can be obtained by substituting (9) in (8) as

$$r(q) = \sum_{p=1}^{P} \mathbf{h_p w} s\left(\left[q - l_p\right]_{MN}\right) e^{j2\pi \frac{k_p(q-l_p)}{MN}} + n(q), \tag{10}$$

where $[\cdot]_n$ represents the mod-$n$ operation.

The permutation matrix $\mathbf{\Pi} \in \mathbb{R}^{MN\times MN}$ that symbolizes delay shifts, and the diagonal matrix $\mathbf{\Delta} \in \mathbb{C}^{MN\times MN}$ that characterizes Doppler shifts, are defined as

$$\mathbf{\Pi} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 1 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}, \tag{11}$$

$$\mathbf{\Delta} = \text{diag}\left\{z^0, z^1, \ldots, z^{MN-1}\right\}, \tag{12}$$

where $z = e^{\frac{j2\pi}{MN}}$. Each path introduces an $l_p$-step cyclic shift of the transmitted signal vector $\mathbf{s}$, modeled by $\mathbf{\Pi}^{l_p}$, and modulates it with a carrier at frequency $k_p$, modeled by $\mathbf{\Delta}^{k_p}$. Accordingly, the time domain channel matrix $\mathbf{H_T} \in \mathbb{C}^{MN\times MN}$ can be constructed as

$$\mathbf{H_T} = \sum_{p=1}^{P} \mathbf{h_p w} \mathbf{\Delta}^{k_p} \mathbf{\Pi}^{l_p}. \tag{13}$$

This leads to the vectorized form of (10), given as

$$\mathbf{r} = \mathbf{H_T s} + \mathbf{n}. \tag{14}$$

Substituting (3) and (6) in (14), the input-output relationship of the MISO system in the DD domain is expressed as

$$\begin{aligned} \mathbf{y} &= (\mathbf{F}_N \otimes \mathbf{G}_{\text{rx}})\mathbf{H_T}\left(\mathbf{F}_N^H \otimes \mathbf{G}_{\text{tx}}\right)\mathbf{x} + (\mathbf{F}_N \otimes \mathbf{G}_{\text{rx}})\mathbf{n} \\ &= \mathbf{H_{DD}}\mathbf{x} + \mathbf{n_{DD}}, \end{aligned} \tag{15}$$

where $\mathbf{H_{DD}} = (\mathbf{F}_N \otimes \mathbf{G}_{\text{rx}})\mathbf{H_T}(\mathbf{F}_N^H \otimes \mathbf{G}_{\text{tx}}) \in \mathbb{C}^{MN\times MN}$ denotes the effective DD domain channel matrix and $\mathbf{n_{DD}} = (\mathbf{F}_N \otimes \mathbf{G}_{\text{rx}})\mathbf{n} \in \mathbb{C}^{MN\times 1}$ is the DD domain noise vector. Equation (15) illustrates the comprehensive transformation of the transmitted DD domain symbols through the vehicular channel and their reception within the OTFS transceiver system.

## IV. Design of Vehicular Network Architecture

In this section, we delineate the architecture of ultra reliable low latency vehicular network, including the overall framework, the deployment of predictive beamforming at the transmitter, and methodologies for channel estimation and signal detection at the receiver.

### A. Transceiver Framework in Vehicular Networks

The transceiver framework for ultra reliable low latency vehicular networks is shown in Fig. 3. The BS segments time into uniform intervals, with each interval ending in a predictive beamforming design stage where both newly acquired and historical CSI are employed to compute the beamforming matrix for the upcoming interval. Within each interval, the BS rapidly transmits OTFS data frames to vehicles using the beamforming matrix computed in the previous interval and intermittently sends pilot signals to gather up-to-date CSI feedback from the vehicles. Notably, the BS maintains continuous data transmission even while processing the predictive beamforming design. At the vehicle side, the receiver initially identifies the type of signal received from the BS: if a pilot signal is detected, the vehicle performs DD domain channel estimation and transmits the CSI back to the BS, whereas

upon receiving an OTFS data frame, the vehicle carries out DD domain channel estimation followed by signal detection.

### B. Predictive Beamforming With Type I Codebook

The BS's reliance on instantaneous CSI for beamforming matrix design is inherently impractical due to the unavailability of instant channel estimation at the beginning of OTFS data frames, particularly in rapidly time-varying vehicular networks. It is worth emphasizing that the utilization of historical CSI to predict the beamforming matrix for the following time segment holds great promise, which bypasses the acquisition of instantaneous CSI and ensures reliable subsequent data transmission. Specifically, we leverage a series of past estimated CSI matrices, $\mathbf{H_{DD}}^{t-1}, \mathbf{H_{DD}}^{t-2}, \cdots, \mathbf{H_{DD}}^{t-\epsilon}$, $\forall t > \epsilon$, where $t$ denotes the current time segment and $\epsilon$ is the number of historical observations. Based on the historical DD domain CSI, we exploit the temporal dependency of successive channels, allowing signals to be beamformed and transmitted without delay.

In 3GPP NR, codebook-based precoding is specified for uplink and downlink beam management to support beamforming [38]. The Type I codebook is constructed from a two-dimensional DFT matrix with spatial oversampling, where the precoding matrix indicator simply identifies the column of the DFT matrix that corresponds to one single beam pointing to the desired spatial direction. The DFT matrix size is determined by the number of antenna ports $(A_1, A_2)$ in the horizontal and vertical dimensions for each polarization of the dual-polarized antenna array. Additional oversampling factors $(O_1, O_2)$ are specified to improve beamforming resolution. Horizontal beam vector $\mathbf{v}_{\zeta_1} \in \mathbb{C}^{\mathbf{A}_1 \times 1}$ and vertical beam vector $\mathbf{v}_{\zeta_2} \in \mathbb{C}^{\mathbf{A}_2 \times 1}$ are respectively defined as

$$\mathbf{v}_{\zeta_1} = \left[1 \; e^{j2\pi \frac{\zeta_1}{O_1 A_1}} \; \cdots \; e^{j2\pi \frac{\zeta_1 (A_1-1)}{O_1 A_1}}\right], \tag{16}$$

$$\mathbf{v}_{\zeta_2} = \left[1 \; e^{j2\pi \frac{\zeta_2}{O_2 A_2}} \; \cdots \; e^{j2\pi \frac{\zeta_2 (A_2-1)}{O_2 A_2}}\right], \tag{17}$$

where $\zeta_1 = n_1 O_1 + m_1$, $m_1 = 0, 1, \ldots, O_1 - 1$, $n_1 = 0, 1, \ldots, A_1 - 1$, and $\zeta_2 = n_2 O_2 + m_2$, $m_2 = 0, 1, \ldots, O_2 - 1$, $n_2 = 0, 1, \ldots, A_2 - 1$. The combined beam candidate set $\mathcal{C}$ is the Kronecker product of horizontal and vertical beam vectors, given as

$$\mathcal{C} = \left\{\tilde{\mathbf{w}}_{\zeta_1,\zeta_2}\right\} = \left\{\mathbf{v}_{\zeta_1} \otimes \mathbf{v}_{\zeta_2}\right\}, \tag{18}$$

which contains a total of $A_1 O_1 A_2 O_2$ beam candidates $\tilde{\mathbf{w}}_{\zeta_1,\zeta_2} \in \mathbb{C}^{A_1 A_2 \times 1}$. Once $\tilde{\mathbf{w}}_{\zeta_1,\zeta_2}$ is selected for transmission, the optimal beamforming matrix $\mathbf{w} \in \mathbb{C}^{2A_1 A_2 \times 1}$ takes the form as

$$\mathbf{w} = \begin{bmatrix} \tilde{\mathbf{w}}_{\zeta_1,\zeta_2} \\ \varphi_s \tilde{\mathbf{w}}_{\zeta_1,\zeta_2} \end{bmatrix}, \tag{19}$$

where $\varphi_s = e^{j\pi s/2}$, $s \in \{0, 1, 2, 3\}$ is the phase difference between polarizations.

Consider the BS-vehicle downlink transceiver system where the BS is equipped with $N_{\text{tx}} = 2A_1 A_2$ antenna ports. The adopted Type I codebook allows the BS and vehicles to agree upon a common codebook configuration for beamforming in FDD mode, which is achieved with minimal overhead by selecting beams directly from the codebook.

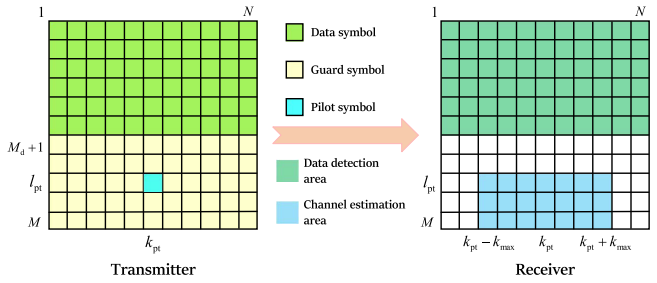



Fig. 4. OTFS channel estimation.

### C. DD Domain Channel Estimation

In the pursuit of optimizing both predictive beamforming and signal detection, CSI acquisition becomes paramount [23]. We present a channel estimation strategy that uses the DD domain pilot within the OTFS data frame, as delineated in Fig. 4. The DD grid consists of $M_{\text{d}}$ rows of data symbols $x_{\text{d}}$, and $M - M_{\text{d}}$ rows of zero-padding. The length of the guard interval should be at least twice the maximum delay shift $l_{\max}$, since both the data symbol and pilot symbol require a guard interval of length $l_{\max}$ in the delay dimension to avoid interference. The pilot symbol $x_{\text{pt}}[l_{\text{pt}}, k_{\text{pt}}]$ is placed at the exact center of the guard symbols to preclude interference across both delay and Doppler dimensions. Based on the above requirements, we choose $M_{\text{d}} = M - 2l_{\max} - 1$ and $l_{\text{pt}} = M - l_{\max}$ and $k_{\text{pt}} = \lfloor N/2 \rfloor$ to achieve a balanced trade-off between estimation precision and data transmission efficiency. The OTFS data frame can thus be expressed as a function of the DD grid coordinates, given as

$$\mathbf{X_{DD}}[l,k] = \begin{cases} x_{\text{pt}} & k = k_{\text{pt}}, l = l_{\text{pt}}, \\ 0 & 1 \le k \le N, \\ & M_{\text{d}} + 1 \le l \le M, \\ x_{\text{d}} & \text{otherwise.} \end{cases} \tag{20}$$

In the case of rectangular pulse waveforms, the input-output relation in (15) exhibits a two-dimensional circular convolution between $\mathbf{X_{DD}}$ and $\mathbf{H_{DD}}$, with each path contributing to a circular shift of the transmitted symbols by its delay and Doppler indices, given as

$$\mathbf{Y_{DD}}[l,k] = \sum_{k_p=-k_{\max}}^{k_{\max}} \sum_{l_p=0}^{l_{\max}} b[l_p,k_p] h[l_p,k_p] \beta[l,k] \cdot \mathbf{X_{DD}}\left([l-l_p]_M, [k-k_p]_N\right) + n[l,k]. \tag{21}$$

In (21), $\beta[l,k]$ is the known phase shift given by

$$\begin{cases} e^{j2\pi\left(\frac{l-l_p}{M}\right)\frac{k_p}{N}} & l_p + 1 \le l \le M \\ \frac{N-1}{N} e^{j2\pi\left(\frac{l-l_p}{M}\right)\frac{k_p}{N}} e^{-j2\pi\left(\frac{[k-k_p]_N}{N}\right)} & 1 \le l \le l_p \end{cases}, \tag{22}$$

and $b[l_p, k_p] \in \{0, 1\}$ is defined as a binary indicator. Specifically, $b[l_p, k_p] = 1$ signifies the presence of a path characterized by delay shift $l_p$, Doppler shift $k_p$, and corresponding fading magnitude $h[l_p, k_p]$. Conversely, $b[l_p, k_p] = 0$ implies the nonexistence of a path with such delay and Doppler

combination. Following this, the received symbols pertinent to channel estimation can be expressed as

$$\mathbf{Y_{DD}}[l,k] = b\big[l-l_{\rm pt}, k-k_{\rm pt}\big]h\big[l-l_{\rm pt}, k-k_{\rm pt}\big] \cdot \beta[k,l]x_{\rm pt} + n[l,k], \tag{23}$$

for $l \in [l_{\rm pt}, M]$ and $k \in [k_{\rm pt}-k_{\max}, k_{\rm pt}+k_{\max}]$.

Consequently, the receiver can extract the DD domain channel impact based on the received pilot symbols within the channel estimation region.In the DD domain, the presence of each path manifests itself as a scaling and shifting of the pilot symbols, which allows the channel estimation algorithm to infer the DD domain CSI based on the magnitude of the received symbols with respect to a predefined positive threshold $\Gamma$. Note that this positive threshold serves to minimize the impact of AWGN. Specifically, if the magnitude of a received polit symbols exceeds $\Gamma$, we set $b[l-l_{\rm pt}, k-k_{\rm pt}] = 1$ and $h[l-l_{\rm pt}, k-k_{\rm pt}] = \mathbf{Y_{DD}}[l,k]/\beta[k,l]x_{\rm pt}$. Otherwise, $b[l-l_{\rm pt}, k-k_{\rm pt}]$ and $h[l-l_{\rm pt}, k-k_{\rm pt}]$ are set to zero. Symbols outside the estimation area, i.e., $l \in [1, M_{\rm d}]$, $k \in [1, N]$, are allocated for signal detection.

### D. MRC Signal Detection

This subsection elucidates the methodology for delay-time domain based maximal ratio combining (MRC) signal detection aimed at reconstructing the originally transmitted signal [24]. We define $\tilde{\mathbf{X}}$ as the matrix consisting of delay-time sample vectors, which correlate with the DD domain symbols via the following transformation given as

$$\tilde{\mathbf{X}}^{\rm T} = [\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_M] = \mathbf{F}_N^H[\mathbf{x}_1, \ldots, \mathbf{x}_M] = \mathbf{F}_N^H \cdot \mathbf{X_{DD}}^{\rm T}, \tag{24}$$

where the superscript T represents transpose. Analogously, the received DD domain symbols and delay-time sample vectors are related as

$$\mathbf{Y_{DD}}^{\rm T} = [\mathbf{y}_1, \ldots, \mathbf{y}_M] = \mathbf{F}_N[\tilde{\mathbf{y}}_1, \ldots, \tilde{\mathbf{y}}_M] = \mathbf{F}_N \cdot \tilde{\mathbf{Y}}^{\rm T}. \tag{25}$$

The convolution of the transmitted and received DD domain symbol vectors can be simplified as

$$\mathbf{y}_l = \sum_{l_p \in \mathcal{L}} \mathbf{K}_{l,l_p} \cdot \mathbf{x}_{l-l_p}. \tag{26}$$

The set $\mathcal{L} = \{l_p\} \in \mathbb{Z}$ encompasses distinct integer delay shifts among the $P$ paths identified through DD domain channel estimation, and $\mathcal{K}_{l_p} = \{k_p \mid l = l_p\} \in \mathbb{Z}$ comprises the corresponding distinct integer Doppler shifts. The circulant matrices $\mathbf{K}_{l,l_p} \in \mathbb{C}^{N\times N}$ in (26) embodies the vehicular channel between the receiver delay index $l$ and the transmitter delay index $l-l_p$, given as

$$\mathbf{K}_{l,l_p} = \mathrm{circ}\Big[\boldsymbol{\nu}_{l,l_p}(1), \ldots, \boldsymbol{\nu}_{l,l_p}(N)\Big], \tag{27}$$

$$\boldsymbol{\nu}_{l,l_p}(k_p) = b\big[l_p, k_p\big]h\big[l_p, k_p\big]z^{k_p(l-l_p)}, \tag{28}$$

where $\boldsymbol{\nu}_{l,l_p}(k_p)$ is the Doppler spread vector in the $l_p$-th delay shift experienced by all the symbols in the $(l-l_p)$-th row of the OTFS data frame. To attain the input-output relationship within the delay-time domain, we split time domain sample index $q = 1, \ldots, MN$ into delay and Doppler indices as $q = l + kM$. This decomposition allows us to further express (26) in terms of delay-time symbol vectors as

$$\tilde{\mathbf{y}}_l(k) = \sum_{l_p \in \mathcal{L}} \tilde{\boldsymbol{\nu}}_{l,l_p}(k)\tilde{\mathbf{x}}_{l-l_p}(k), \tag{29}$$

$$\tilde{\boldsymbol{\nu}}_{l,l_p}(k) = \sum_{k_p \in \mathcal{K}_{l_p}} b\big[l_p, k_p\big]h\big[l_p, k_p\big]z^{k_p(l-l_p)}e^{\frac{j2\pi k_p k}{N}}. \tag{30}$$

The iterative MRC process introduces a residual noise plus interference (RNPI) term in the $i$-th iteration, given as

$$\Delta\mathbf{y}_l^{(i)} = \mathbf{y}_l - \sum_{l_p \in \mathcal{L}} \mathbf{K}_{l,l_p} \cdot \hat{\mathbf{x}}_{l-l_p}^{(i)}. \tag{31}$$

The computational complexity of this process is optimized by storing and updating the RNPI vectors as

$$\Delta\mathbf{y}_{l+l_p}^{(i)} \leftarrow \Delta\mathbf{y}_{l+l_p}^{(i)} - \mathbf{K}_{l+l_p,l_p} \cdot \Big(\mathbf{x}_l^{(i)} - \mathbf{x}_l^{(i-1)}\Big). \tag{32}$$

After the above definition, the MRC output of each iteration is formulated as

$$\mathbf{c}_l^{(i)} = \hat{\mathbf{x}}_l^{(i-1)} + \left(\sum_{l_p \in \mathcal{L}} \mathbf{K}_{l+l_p,l_p}^H \cdot \mathbf{K}_{l+l_p,l_p}\right)^{-1} \cdot \Delta\mathbf{g}_l^{(i)}, \tag{33}$$

$$\Delta\mathbf{g}_l^{(i)} = \sum_{l_p \in \mathcal{L}} \mathbf{K}_{l+l_p,l_p}^H \cdot \Delta\mathbf{y}_{l+l_p}^{(i)}, \tag{34}$$

where $\Delta\mathbf{g}_l^{(i)}$ is the RNPI's maximal ratio combining across all delay branches containing the symbol vector $\mathbf{x}_l$.

According to (26) and (29), the iterative MRC procedure of (33) and (34) in the DD domain can be further streamlined by converting matrix-vector products into element-wise vector products within the delay-time domain, given as

$$\tilde{\mathbf{c}}_l^{(i)} = \tilde{\mathbf{x}}_l^{(i-1)} + \Delta\tilde{\mathbf{g}}_l^{(i)} \oslash \sum_{l_p \in \mathcal{L}} \tilde{\boldsymbol{\nu}}_{l+l_p,l_p}^H \circ \tilde{\boldsymbol{\nu}}_{l+l_p,l_p}, \tag{35}$$

$$\Delta\tilde{\mathbf{g}}_l^{(i)} = \sum_{l_p \in \mathcal{L}} \tilde{\boldsymbol{\nu}}_{l+l_p,l_p}^* \circ \Delta\tilde{\mathbf{y}}_{l+l_p}^{(i)}, \tag{36}$$

Likewise, (32) can be expressed in the delay-time domain as

$$\Delta\tilde{\mathbf{y}}_{l+l_p}^{(i)} \leftarrow \Delta\tilde{\mathbf{y}}_{l+l_p}^{(i)} - \tilde{\boldsymbol{\nu}}_{l+l_p,l_p} \circ \Big(\tilde{\mathbf{x}}_l^{(i)} - \tilde{\mathbf{x}}_l^{(i-1)}\Big), \tag{37}$$

where the superscript, $\sim$, denotes the $N$-IFFT operation, $\oslash$ denotes the Hadamard division, i.e., the element-wise division, and $\circ$ denotes the Hadamard product, i.e., the element-wise multiplication. The refined approach leverages the intrinsic properties of the delay-time domain representation to minimize computational overhead, effectively enhancing the efficiency of the iterative MRC signal detection.

## V. Deep Learning-Based Predictive Beamforming

In this section, we first analyze the data characteristics of historical DD domain CSI, and then propose the DL neural network named DDCT for predictive beamforming.

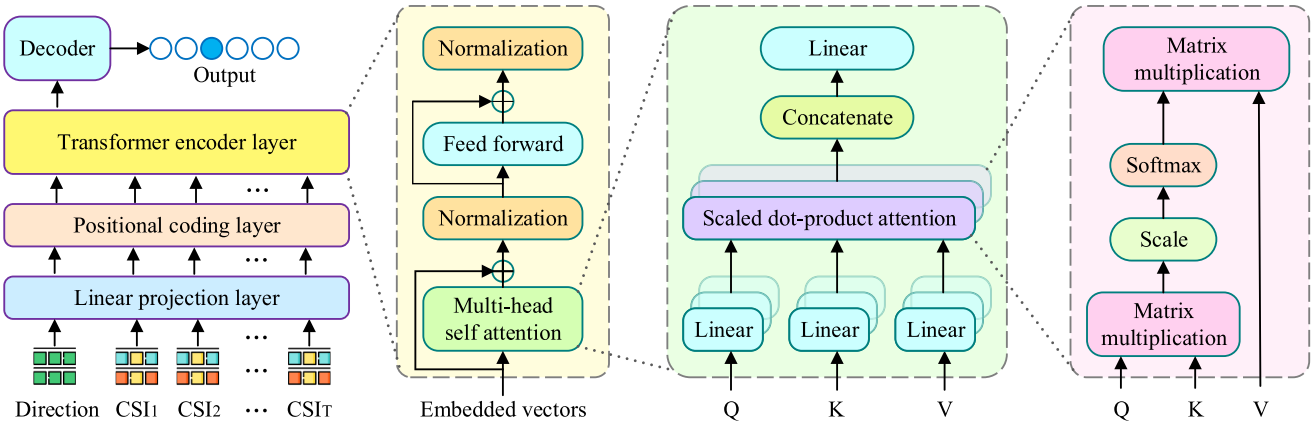



Fig. 5. The neural network structure of DDCT.

### A. Features in Historical DD Domain CSI

Predictive beamforming relies on capturing the dependencies and variations of the historical CSI in the DD domain. In OTFS transceiver systems, the CSI fed back by the receiver is represented in the DD domain. It can be visualized as a two-dimensional grid-like data, where one dimension represents the time delay of the channel response component and the other dimension represents the Doppler information. By analyzing spatial features, such as the intensity and location of the response components, the delay and Doppler of each channel path can be obtained. Moreover, by considering multiple DD domain historical CSI over a pastime interval, we can obtain a dataset that can be viewed as a sequence of grid-like data with temporal features. Hence, the DL algorithm can learn to select the predictive beamforming matrix by extracting the spatial and temporal features of the historical DD domain CSI.

### B. DDCT Algorithm

The neural network structure of the proposed DDCT algorithm for predictive beamforming is shown in Fig. 5. It leverages the multi-head attention mechanism within the transformer encoder and enables nuanced learning of temporal correlations and spatial features in sequential CSI data, enhancing the capabilities of predictive beamforming in vehicular networks.

The historical DD domain CSI data needs to be preprocessed to meet the input requirements of the neural network. The DD domain CSI at the time segment $t$ can be represented as a two-dimensional complex matrix, $\mathbf{H}_{\mathrm{DD}}^{t} \in \mathbb{C}^{(l_{\max}+1)\times(2k_{\max}+1)}$. However, given that neural networks require real number inputs, this complex matrix must be transformed into a real number representation. A complex number is analogous to a vector in a two-dimensional plane, which can be decomposed into two real numbers either through rectangular coordinates into real and imaginary parts, or through polar coordinates into magnitude and phase. Consequently, the real-number representation of the DD domain CSI becomes a three-dimensional real matrix $\mathbf{H}_{\mathrm{R}}^{t} \in \mathbb{R}^{2\times(l_{\max}+1)\times(2k_{\max}+1)}$, suitable as input for the neural network. For predictive beamforming at time $t$, the deep learning algorithm uses the historical DD domain CSI from the previous $\epsilon$ time segments. The sequential input is denoted as $[\mathbf{H}_{\mathrm{R}}^{t-\epsilon}, \mathbf{H}_{\mathrm{R}}^{t-(\epsilon-1)}, \ldots, \mathbf{H}_{\mathrm{R}}^{t-1}]$.

The linear projection layer using CNN is employed to extract features in each three-dimensional real matrix CSI, as well as to flat it to vector. Matrix $\mathbf{H}_{\mathrm{R}}^{t}$ can be analogized to an image with 2 channels, where $(l_{\max}+1)$ and $(2k_{\max}+1)$ correspond to the image's height and width, respectively. Employing the CNN with a kernel size of $[l_{\max}+1, 2k_{\max}+1]$, matrix $\mathbf{H}_{\mathrm{R}}^{t}$ undergoes linear projection to a new vector, $E_t' \in \mathbb{R}^{d_{\mathrm{e}}}$, given by

$$E_t' = \sum_{c=1}^{2} \sum_{l=1}^{l_{\max}+1} \sum_{k=1}^{2k_{\max}+1} (W_{\mathrm{P}})_{d,c,l,k} \cdot \left(\mathbf{H}_{\mathrm{R}}^{t}\right)_{c,l,k} + Z, \quad (38)$$

where $W_{\mathrm{P}} \in \mathbb{R}^{d_{\mathrm{e}}\times 2\times(l_{\max}+1)\times(2k_{\max}+1)}$ and $Z \in \mathbb{R}^{d_{\mathrm{e}}}$ represent the parameters of the neural network. Vector $E_t'$ is then processed through an activation function to obtain the normalized embedding vector as

$$E_t = sigmoid\left(E_t'\right), \quad (39)$$

where $sigmoid(\cdot)$ is an activation function that maps vector values to a range between 0 and 1, defined as

$$sigmoid(a) = \frac{1}{1+e^{-a}}. \quad (40)$$

This linear projection operation is uniformly applied to all DD domain CSI matrices using consistent neural network parameters. After the linear projection layer, the historical DD domain CSI data from the previous $T$ time segments are transformed into a sequence of vectors $[E_{t-T}, E_{t-(T-1)}, \ldots, E_{t-1}]$ for further processing. Since the vehicle driving direction has a great influence on the Doppler effect, we include the vehicle direction as part of the inputs. To maintain the consistency of the data format, the direction information is also expanded into a format similar to the CSI and then passed through a specialized linear projection to obtain the embedded vector $D \in \mathbb{R}^{d_{\mathrm{e}}}$.

The transformer encoder is employed to extract the channel variation features hidden among temporal correlation of the sequential vectors. Given the transformer encoder's ability to inherently discern the sequential order of input data, positional

encoding becomes indispensable. To infuse the transformer with positional awareness, we employ sine and cosine functions of varying frequencies as the basis for our encoding scheme. The predetermined positional encoding matrix, $\mathscr{P} \in \mathbb{R}^{(T+1)\times d_{\mathrm{e}}}$, is given by

$$\mathscr{P}(u,i) = \begin{cases} sin\left(\frac{u}{50^{\frac{i}{d_{\mathrm{e}}}}}\right), & \text{if i is even,} \\ cos\left(\frac{u}{50^{\frac{i-1}{d_{\mathrm{e}}}}}\right), & \text{if i is odd,} \end{cases} \tag{41}$$

where $u$ represents the position within the sequence, $d_{\mathrm{e}}$ is the dimension of each vector and $i$ indexes the dimension. By augmenting our sequence of vectors with the positional encoding matrix, we obtain the encoded inputs, denoted as

$$E = \left[D; E_{t-T}; E_{t-(T-1)}; \cdots; E_{t-1}\right] + \mathscr{P}. \tag{42}$$

The multi-head attention mechanism, a cornerstone of the self-attention framework within the transformer, excels in extracting features related to temporal correlations. The inputs to the self-attention module consist of three matrices: query $Q$, key $K$, and value $V$. These matrices are all in dimension of $\mathbb{R}^{(T+1)\times d_{\mathrm{k}}}$, and they are formulated as

$$\begin{cases} Q = EW_q, \\ K = EW_k, \\ V = EW_v, \end{cases} \tag{43}$$

where $W_q$, $W_k$, and $W_v$ are trainable parameters of the transformer. Then they are used for the dot-product attention $\mathcal{A}(\cdot)$, given by

$$\mathcal{A}(Q,K,V) = softmax\left(\frac{QK^{\mathrm{T}}}{\sqrt[2]{d_{\mathrm{e}}}}\right)V. \tag{44}$$

In (44), $softmax(\cdot)$ is an activation function that normalizes the vector elements to ensure their sum equals 1, defined as

$$softmax(a_i) = \frac{e^{a_i}}{\sum_j e^{a_j}}. \tag{45}$$

The multi-head attention mechanism significantly augments the algorithm's ability to concurrently process information from diverse perspectives within the input sequence. This enhancement is achieved by projecting the original input vectors into multiple segments, known as "heads," each independently calculating distinct attention weights. It is achieved through the linear projection of the queries, keys, and values $h$ times, utilizing learned and distinct linear projections for each projection, thereby enabling the algorithm to assimilate and synthesize information across various representational subspaces. The function of the multi-head attention mechanism, denoted as $\mathcal{M}(\cdot)$, is expressed as

$$\mathcal{M}(Q,K,V) = [h_1, h_2, \ldots, h_i]\, W_o, \tag{46}$$

where $W_o$ represents the neural network parameters. In (46), $h_i$ signifies the outputs from an individual attention head and is calculated by

$$h_i = \mathcal{A}\left(QW_i^Q, KW_i^K, VW_i^V\right), \tag{47}$$

with $W_i^Q$, $W_i^K$ and $W_i^V$ being the respective head-specific weight matrices for queries, keys, and values. The outputs from the multi-head attention module are synthesized by concatenating the outputs of all heads and subsequently applying a linear transformation. Subsequently, the algorithm integrates two "add & norm" layers with an intermediary "feed-forward" layer, which together give the outputs of the transformer encoder layer with the same structural format as its inputs. Then, the first vector from the output sequence is selected and subjected to a sigmoid activation function, followed by a linear decoding layer. This layer contains a number of neurons equivalent to the total number of beams in the codebook, with each neuron representing a potential beam direction. The outputs are further normalized using the softmax function. Finally, we select the beam matrix whose representing neuron has the highest value as the predictive beamforming matrix for the upcoming time segment, $t$.

## VI. Simulation Results and Analyses

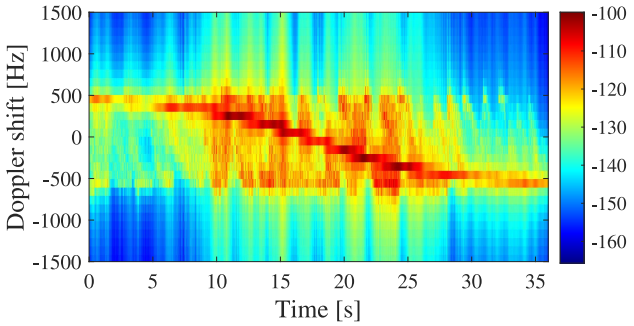
In this section, the experiment scenario and simulation settings are first presented. Then, we analyze the Doppler shift property of the vehicular channel and the benefit of beamforming. Next, we evaluate the predictive beamforming performance of different DL approaches, showing the advantage of our proposed DDCT. Moreover, based on the proposed DDCT, we show the effectiveness of predictive beamforming for the OTFS-enabled vehicular networks.

### A. Experiment Setting

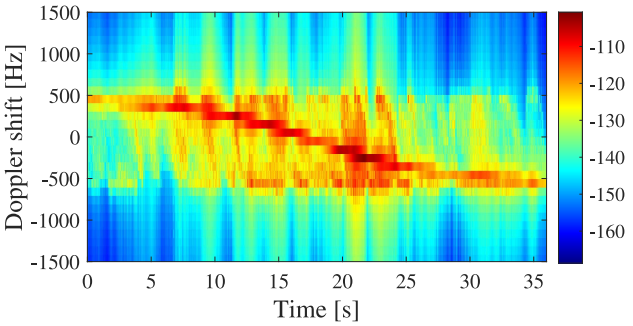
We simulate an east-west highway scenario, with a BS located 300 meters away from road center. The highway has 3 lanes in both directions, each spaced 5 meters apart. We use quasi deterministic radio channel generator (Quadriga), whose channel model follows a geometry-based stochastic channel modeling approach, to simulate wireless channels of vehicular networks. We simulate both line-of-sight (LOS) and non-line-of-sight (NLOS) channels. We take 2 cars from each direction for training, and 2 of the remaining cars with different directions are used for testing. The cross-entropy is used as the training loss. Other important simulation parameters are given in Table II. For DDCT's parameters, its CNN module uses a hidden dimension of 16 with a kernel size of (5, 9), while the transformer module employs a single-layer transformer encoder with 2 heads and a hidden dimension of 16.

To evaluate the predictive beamforming performance of the DDCT, we choose 3 classic algorithms that can be used for sequential grid-like data processing as benchmarks:

- DNN: a 4-layer simple neural network with sigmoid as the activation function. The high-dimensional matrix is flattened as its inputs. The dimensions of the hidden layer include 256 and 512.
- ConvLSTM: a neural network architecture that combines the spatial feature extraction ability of CNN with the temporal sequence modeling ability of the LSTM network. The hidden dimension is 16 and the kernel size is (3, 3).
- 3DCNN: an extension of the traditional CNN, designed to process data with three spatial dimensions by applying

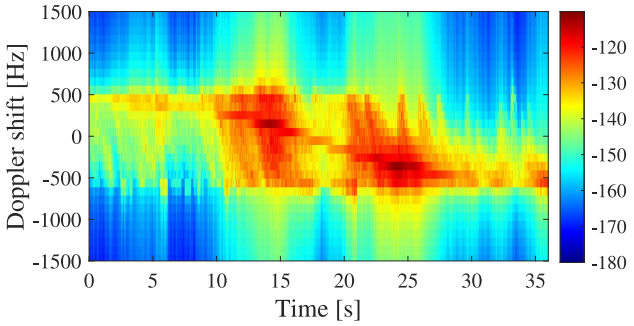



(a) Doppler spectrum of vehicle facing east.

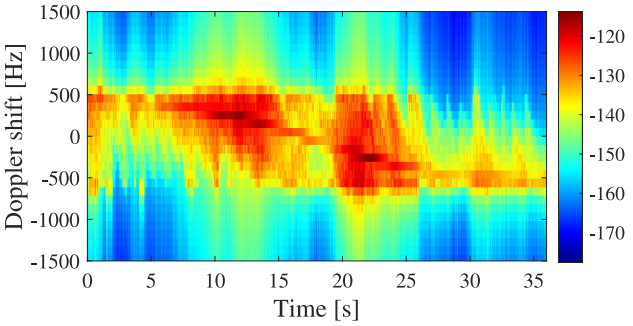



(b) Doppler spectrum of vehicle facing west.

Fig. 6. The Doppler spectrum of LOS channel.

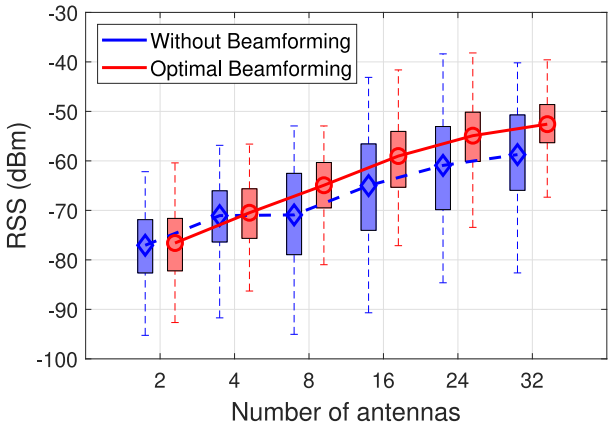



(a) Doppler spectrum of vehicle facing east.

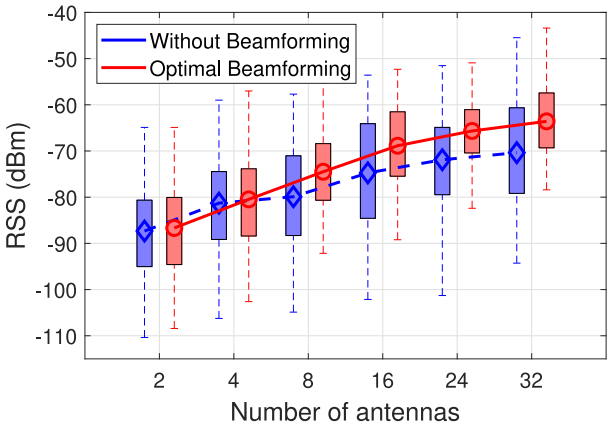



(b) Doppler spectrum of vehicle facing west.

Fig. 7. The Doppler spectrum of NLOS channel.




(a) Beamforming benefit of LOS channel.




(b) Beamforming benefit of NLOS channel.

Fig. 8. The benefits of beamforming.

3D convolution to the input data. The 3D convolution is applied 2 times and their kernel size are (1, 3, 3).

### *B. Analysis of Channel and Beamforming*

The impact of the Doppler effect on vehicular channels, engendered by the relative motion between the transmitter and receiver, is elucidated through the analysis of the Doppler shift spectrum. Fig. 6 and Fig. 7 expound upon the Doppler shift for LOS and NLOS channels, respectively. The Doppler shift, ranging from +500 Hz to -500 Hz, dynamically evolves over time as vehicles initially approach and subsequently recede from the BS. As a vehicle approaches the base station, it experiences a positive shift that reduces to zero at the closest point, then turns negative as it moves away, creating the diagonal pattern in the spectrum. The spectral concentration in the LOS scenario signifies strong signal presence from the primary propagation path, whereas the dispersed Doppler shifts in the NLOS scenario show attenuated signal strength from the main path, indicative of more multi-path propagation effects. This dispersion of the Doppler spectrum is a testament to the complex interaction of the transmitted signal with the environment, thus enabling the design and optimization of predictive beamforming.

The efficacy of beamforming is demonstrated in Fig. 8, which compares the RSS with varying numbers of transmit

TABLE II
SIMULATION PARAMETERS

| Parameter | Value |
|---|---|
| Road length | 1000 m |
| Vehicular antenna height | 1.5 m |
| BS antenna height | 25 m |
| Vehicular speed | 100 km/h |
| Carrier frequency | 5.9 GHz |
| Bandwidth of each subcarrier | 1 MHz |
| Transmit power of BS | 20 dBm |
| Noise spectral density | -100 dBm |
| OTFS data frame size | $50 * 50$ |
| Guard interval length | 10 |
| Pilot position | (45,25) |
| Historical CSI window size | 5 |
| Time interval for beamforming update | 10 ms |

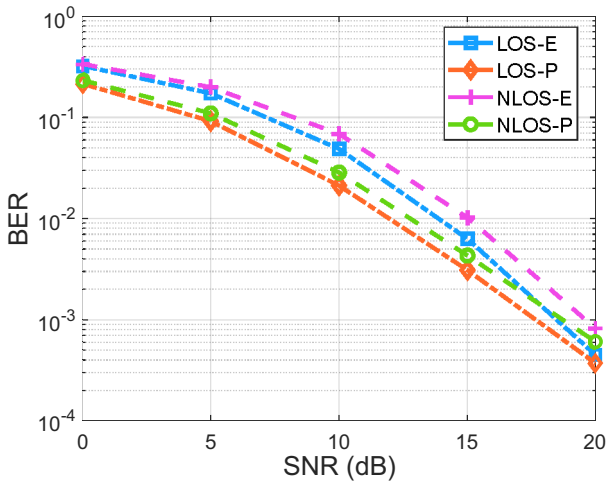
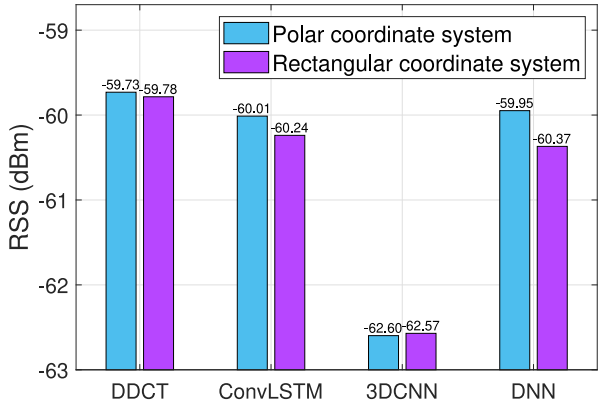
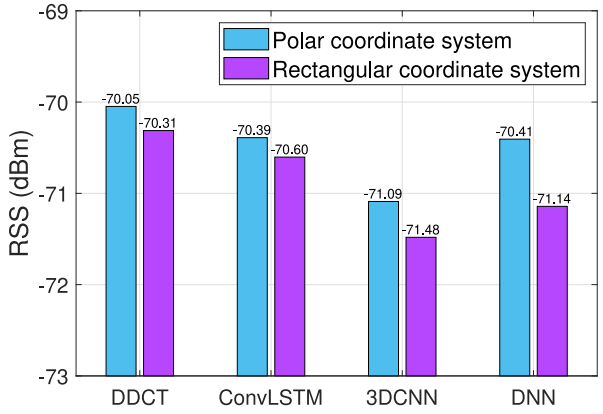



Fig. 9. BER of OTFS in different SNR.




(a) Average RSS of LOS channel.




(b) Average RSS of NLOS channel.

Fig. 10. RSS comparison of different DL approaches.

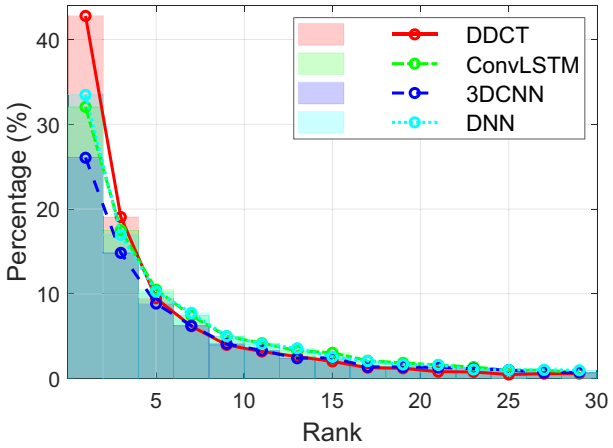
antennas under optimal beamforming and none beamforming conditions. The RSS is observed to improve with an increase in the number of antennas at the BS, attributable to the amplification afforded by the collective transmission power of the antenna array. Notably, the disparity in RSS between scenarios employing optimal beamforming and those devoid of it escalates with the number of antennas, substantiating the enhanced efficacy of beamforming in larger arrays. This trend is consistent across both LOS and NLOS channels. Additionally, the LOS channels manifest an RSS that is approximately 10 dBm higher than the RSS of NLOS channels under analogous conditions, underlining the impact of environmental factors on beamforming performance.

Fig. 9 shows the BER performance of the OTFS system employing 16 antennas, evaluated across various signal-to-noise ratio (SNR) levels for both LOS and NLOS channels. The depicted curves are derived by averaging the BER across all channel sampling points of entire vehicular journey. Note that the legend distinguishes between receiver decoding processes utilizing perfect CSI, denoted as "P", and those employing estimated CSI, indicated as "E". The results reveal a discernible downward trajectory of BER corresponding to incremental enhancements in SNR. It can be found that the BER obtained using the adopted DD domain channel estimation method has a comparable performance to that of perfect CSI, and progressively converges to that of perfect CSI as SNR increases. This parallel in the BER performance substantiates the efficacy of the adopted DD domain channel estimation for OTFS transceiver systems.

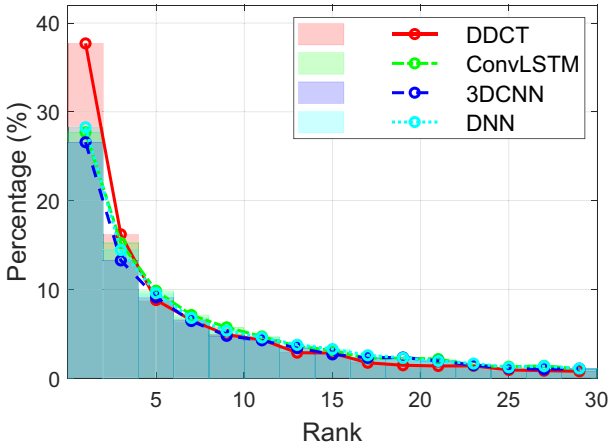
### C. Performance Evaluation of DL Approaches

To evaluate the performance of our proposed DDCT algorithm, the following comparisons are conducted in the 16-antenna case with both LOS and NLOS channels.

Fig. 10 shows the performance of various DL algorithms in executing predictive beamforming, depicted through the averages of the RSS. Note that polar and rectangular coordinate systems in the legend represent two different methods of converting complex numbers to real numbers. Notably, DL algorithms employing the polar coordinate system for preprocessing the CSI generally surpass their rectangular system counterparts in performance metrics. Furthermore, the DDCT algorithm achieves superior RSS on average relative to alternative DL algorithms, substantiating its efficacy across both coordinate systems. This enhancement is consistent and observable within the results obtained from both LOS and NLOS channels, indicating the robustness of the DDCT algorithm in diverse propagation environments.

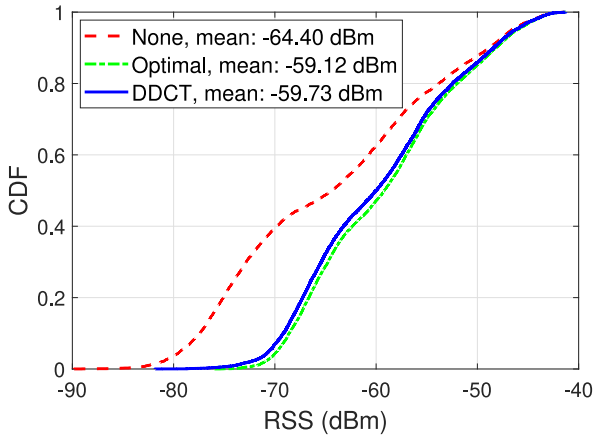



(a) Rank performance of LOS channel.

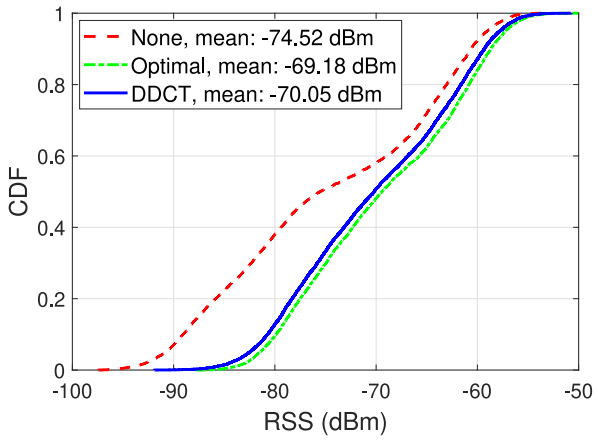



(b) Rank performance of NLOS channel

Fig. 11. Rank comparison of different DL approaches.




(a) CDF for RSS of LOS channel.




(b) CDF for RSS of NLOS channel

Fig. 12. The CDF of RSS improvement.

Fig. 11 shows the rank distribution of beamforming matrices as determined by various DL algorithms. The evaluation process entails iterating over all potential beamforming matrices within the codebook, each corresponding to a distinct directional beam, and computing the RSS subsequent to beamforming for every channel sampling point. Subsequently, beamforming matrices are ordered based on the RSS for each sampling point, from highest to lowest, to establish a ranking sequence where the top-ranked matrix represents the optimal beamforming solution for that specific sampling point. The frequency with which matrices chosen by the DL algorithm appear in each rank position accumulates to establish a rank distribution. The results indicate a notable propensity for the proposed DDCT algorithm to select beamforming matrices that rank within the top two positions at a frequency of approximately 40% and within the third or fourth positions at about 20% across the sorted orders. The powerful spatial-temporal correlation extraction capability of DDCT enables it to better understand the changes in the vehicular channel, and thus its performance is significantly better than similar algorithms. These results shows DDCT's superior predictive capabilities for beamforming in vehicular networks.

### *D. Effectiveness of Predictive Beamforming*

To show the effectiveness of predictive beamforming with our proposed DDCT algorithm, we present the comparison of none beamforming, optimal beamforming, and the DDCT-based predictive beamforming in the 16 antennas case with both LOS and NLOS channels. The polar coordinate system is used for preprocessing the inputs of DDCT.

The cumulative distribution function (CDF) curves in Fig. 12 show the RSS distribution for different beamforming strategies in both LOS and NLOS channels, along with the respective mean RSS values. In these scenario, the mean RSS without beamforming is significantly weaker than both the optimal beamforming and the DDCT-based predictive beamforming. The close alignment of the DDCT's performance to the optimal beamforming suggests that the algorithm effectively approximates the ideal strategy, even though it uses historical channel data rather than real-time CSI. Notably, the DDCT algorithm significantly improves the overall RSS distribution, particularly by elevating the lower bound of signal strength, which helps mitigate the risks of experiencing weak signals and ensures more reliable communication. This enhancement is consistent across various channel conditions, showcasing the robustness of the DDCT

algorithm in maintaining high signal strength and avoiding performance drops. Overall, the DDCT algorithm enhances both signal strength and consistency, making it a valuable tool for improving reliability in dynamic vehicular networks where real-time CSI may not be available.

## VII. Conclusion

In this paper, we have proposed an OTFS-enabled ultra reliable low latency vehicular network architecture operating in the FDD mode, and a corresponding DL algorithm for predictive beamforming. In our design, the FDD mode and OTFS modulation are used to enhance the latency and reliability performance, respectively. Moreover, we have designed a DL algorithm, named DDCT, for predictive beamforming based on historical DD domain CSI. Extensive simulations have verified the effectiveness of the proposed OTFS-based transceiver with predictive beamforming algorithm. For the future work, we will explore resource allocation to further enhance the performance of OTFS-enabled vehicular networks.

## References


[1] C.-X. Wang et al., "On the road to 6G: Visions, requirements, key technologies, and testbeds," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 905–974, 2nd Quart., 2023.

[2] *Framework and Overall Objectives of the Future Development of IMT for 2030 and Beyond*, ITU-Rec. D. SG2, Int. Telecommun. Union, Geneva, Switzerland, 2023.

[3] J. Xue, K. Yu, T. Zhang, H. Zhou, L. Zhao, and X. Shen, "Cooperative deep reinforcement learning enabled power allocation for packet duplication URLLC in multi-connectivity vehicular networks," *IEEE Trans. Mobile Comput.*, vol. 23, no. 8, pp. 8143–8157, Aug. 2024.

[4] T. Tao, Y. Wang, D. Li, Y. Wan, P. Baracca, and A. Wang, "6G hyper reliable and low-latency communication: Requirement analysis and proof of concept," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, 2023, pp. 1–5.

[5] W. Kim, H. Ji, and B. Shim, "Channel aware sparse transmission for ultra low-latency communications in TDD systems," *IEEE Trans. Commun.*, vol. 68, no. 2, pp. 1175–1186, Feb. 2020.

[6] Y. Zhao and W. Xie, "Physical layer latency analysis for 5G NR," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB)*, 2023, pp. 1–6.

[7] Y. Xu, B. Qian, K. Yu, T. Ma, L. Zhao, and H. Zhou, "Federated learning over fully-decoupled RAN architecture for two-tier computing acceleration," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 3, pp. 789–801, Mar. 2023.

[8] R. Hadani et al., "Orthogonal time frequency space modulation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2017, pp. 1–6.

[9] J. Wu and P. Fan, "A survey on high mobility wireless communications: Challenges, opportunities and solutions," *IEEE Access*, vol. 4, pp. 450–476, 2016.

[10] W. Yuan et al., "New delay doppler communication paradigm in 6G era: A survey of orthogonal time frequency space (OTFS)," *China Commun.*, vol. 20, no. 6, pp. 1–25, 2023.

[11] L. Gaudio, G. Colavolpe, and G. Caire, "OTFS vs. OFDM in the presence of sparsity: A fair comparison," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4410–4423, Jun. 2022.

[12] X. Fu et al., "A tutorial on downlink Precoder selection strategies for 3GPP MIMO codebooks," *IEEE Access*, vol. 11, pp. 138897–138922, 2023.

[13] J. Chen, X. Liang, J. Xue, Y. Sun, H. Zhou, and X. Shen, "Evolution of RAN architectures towards 6G: Motivation, development, and enabling technologies," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 3, pp. 1950–1988, 3rd Quart., 2024.

[14] Z. Du et al., "Integrated sensing and communications for V2I networks: Dynamic predictive beamforming for extended vehicle targets," *IEEE Trans. Wireless Commun.*, vol. 22, no. 6, pp. 3612–3627, Jun. 2023.

[15] J. Xue et al., "Large AI model for delay-doppler domain channel prediction in 6G OTFS-based vehicular networks," *Sci. China Inf. Sciences*, vol. 68, no. 7, pp. 184–198, 2025.

[16] C. Liu et al., "Learning-based predictive Beamforming for integrated sensing and communication in vehicular networks," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 8, pp. 2317–2334, Aug. 2022.

[17] J. Xue, T. Jiang, H. Huang, Z. Zhang, H. Zhou, and X. Shen, "Enabling data-driven OTFS modulatioof for 6G hyper reliable high-mobility communications," *IEEE Wireless Commun.*, vol. 32, no. 3, pp. 196–203, Jun. 2025.

[18] H. A. Kassir, Z. D. Zaharis, P. I. Lazaridis, N. V. Kantartzis, T. V. Yioultsis, and T. D. Xenos, "A review of the state of the art and future challenges of deep learning-based beamforming," *IEEE Access*, vol. 10, pp. 80869–80882, 2022.

[19] J. Xue et al., "Sparse mobile crowdsensing for cost-effective traffic state estimation with spatio–temporal transformer graph neural network," *IEEE Internet Things J.*, vol. 11, no. 9, pp. 16227–16242, May 2024.

[20] J. Zhang, G. Zheng, Y. Zhang, I. Krikidis, and K.-K. Wong, "Deep learning based predictive beamforming design," *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 8122–8127, Jun. 2023.

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

[22] S. Srivastava, R. K. Singh, A. K. Jagannatham, A. Chockalingam, and L. Hanzo, "OTFS transceiver design and sparse doubly-selective CSI estimation in analog and hybrid Beamforming aided mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, pp. 10902–10917, Dec. 2022.

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

[24] T. Thaj and E. Viterbo, "Low complexity iterative rake decision feedback equalizer for zero-padded OTFS systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 15606–15622, Dec. 2020.

[25] Y. Yang, Z. Bai, K. Pang, S. Guo, H. Zhang, and K. S. Kwak, "Spatial-index modulation based orthogonal time frequency space system in vehicular networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 6, pp. 6165–6177, Jun. 2023.

[26] P. Singh, A. Gupta, H. B. Mishra, and R. Budhiraja, "Low-complexity ZF/MMSE MIMO-OTFS receivers for high-speed vehicular communication," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 209–227, 2022.

[27] C. Liu, S. Li, W. Yuan, X. Liu, and D. W. K. Ng, "Predictive precoder design for OTFS-enabled URLLC: A deep learning approach," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 7, pp. 2245–2260, Jul. 2023.

[28] C. S. Reddy, P. Priya, D. Sen, and C. Singhal, "Spectral efficient modem design with OTFS modulation for vehicular-IoT system," *IEEE Internet Things J.*, vol. 10, no. 3, pp. 2444–2458, Feb. 2023.

[29] X. Shen, J. Gao, W. Wu, M. Li, C. Zhou, and W. Zhuang, "Holistic network virtualization and pervasive network intelligence for 6G," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 1–30, 1st Quart., 2022.

[30] Z. Qin and H. Yin, "A review of Codebooks for CSI feedback in 5G new radio and beyond," 2023, *arXiv:2302.09222*.

[31] Z. Tong, J. Wang, X. Hou, C. Jiang, and J. Liu, "UAV-assisted covert federated learning over mmWave massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 11785–11798, Sep. 2024.

[32] B. Guo, S. He, M. Shi, K. Yu, J. Chen, and X. Shen, "Semantic-DARTS: Elevating semantic learning for mobile differentiable architecture search," *IEEE Internet Things J.*, vol. 12, no. 2, pp. 1673–1687, Jan. 2025.

[33] B. Guo et al., "Latency-aware neural architecture performance predictor with query-to-tier technique," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 34, no. 7, pp. 5868–5883, Jul. 2024.

[34] Y. Jiang, W. Lin, W. Zhao, and C. Wang, "AcsiNet: Attention-based deep learning network for CSI prediction in FDD MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 12, no. 3, pp. 471–475, Mar. 2023.

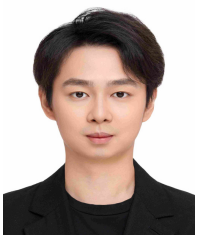
[35] F. Liu and C. Masouros, "A tutorial on joint radar and communication transmission for vehicular networks—Part-III: Predictive Beamforming without state models," *IEEE Commun. Lett.*, vol. 25, no. 2, pp. 332–336, Feb. 2021.

[36] W. Yuan, J. Zou, Y. Cui, X. Li, J. Mu, and K. Han, "Orthogonal time frequency space and predictive beamforming-enabled URLLC in vehicular networks," *IEEE Wireless Commun.*, vol. 30, no. 2, pp. 56–62, Apr. 2023.

[37] C. Liu, X. Liu, S. Li, W. Yuan, and D. W. K. Ng, "Deep CLSTM for predictive beamforming in integrated sensing and communication-enabled vehicular networks," *J. Commun. Inf. Netw.*, vol. 7, no. 3, pp. 269–277, Sep. 2022.
[38] "NR; physical layer procedures for data; (Release 17), Version 17.1.0," 3GPP, Sophia Antipolis, France, Rep. TS 38.214, May 2022.

**Jianzhe Xue** (Graduate Student Member, IEEE) received the B.S. degree in communication engineering from Xidian University, Xi'an, China, in 2021. He is currently pursuing the Ph.D. degree with the School of Electronic Science and Engineering, Nanjing University, China. His current research interests include Internet of Vehicles, orthogonal time frequency space modulation, and machine learning for wireless communications.

**Tiankai Jiang** (Graduate Student Member, IEEE) received the B.S. degree in communication engineering from Nanjing University, Nanjing, China, in 2024, where he is currently pursuing the M.S. degree with the School of Electronic Science and Engineering. His current research interests include intelligent transportation system and Internet of Vehicles.



**Zhanxi Ma** (Graduate Student Member, IEEE) received the B.S. degree in communication engineering from Xiamen University, Xiamen, China, in 2023. He is currently pursuing the Ph.D. degree with the School of Electronic Science and Engineering, Nanjing University, China. His current research interests include Internet of Vehicles, channel generation, and space-terrestrial integrated network.

**Yunting Xu** (Member, IEEE) received the Ph.D. degree from Nanjing University in 2024. He is currently a Research Fellow with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. He mainly focuses on the artificial intelligence and networking optimization in the field of emerging wireless networks.

**Haibo Zhou** (Senior Member, IEEE) received the Ph.D. degree in information and communication engineering from Shanghai Jiao Tong University, Shanghai, China, in 2014. From 2014 to 2017, he was a Postdoctoral Fellow with the Broadband Communications Research Group, Department of Electrical and Computer Engineering, University of Waterloo. He is currently a Full Professor with the School of Electronic Science and Engineering, Nanjing University, Nanjing, China. His research interests include resource management and protocol design in B5G/6G networks, vehicular ad hoc networks, and space-air-ground integrated networks. He was a recipient of the 2019 IEEE ComSoc Asia–Pacific Outstanding Young Researcher Award, the IEEE ComSoc Distinguished Lecturer from 2023 to 2024, the IEEE VTS Distinguished Lecturer from 2023 to 2025, and the 2023 IEEE ComSoc WTC Young Researcher Award. He was a Highly Cited Researcher by Clarivate Analytics in 2022 and 2020, respectively. He served as a Track/Symposium CoChair for IEEE/CIC ICCC 2019, IEEE VTC-Fall 2020, IEEE VTC-Fall 2021, WCSP 2022, IEEE GLOBECOM 2022, and IEEE ICC 2024. He is currently an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE INTERNET OF THINGS JOURNAL, IEEE NETWORK MAGAZINE, and *Journal of Communications and Information Networks*. He was elected as an IET Fellow in 2022.

**Xuemin Shen** (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, USA, in 1990. He is a University Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research focuses on network resource management, wireless network security, Internet of Things, 5G and beyond, and vehicular ad hoc and sensor networks. He received the Canadian Award for Telecommunications Research from the Canadian Society of Information Theory in 2021, the R.A. Fessenden Award in 2019 from IEEE, Canada, the Award of Merit from the Federation of Chinese Canadian Professionals (Ontario) in 2019, the James Evans Avant Garde Award in 2018 from the IEEE Vehicular Technology Society, the Joseph LoCicero Award in 2015, the Education Award in 2017 from the IEEE Communications Society, the Technical Recognition Award from Wireless Communications Technical Committee in 2019, and the AHSN Technical Committee in 2013. He has also received the Excellent Graduate Supervision Award in 2006 from the University of Waterloo and the Premiers Research Excellence Award in 2003 from the Province of Ontario, Canada. He served as the Technical Program Committee Chair/CoChair for IEEE Globecom 16, IEEE Infocom14, IEEE VTC10 Fall, IEEE Globecom07, and the Chair for the IEEE Communications Society Technical Committee on Wireless Communications. He is the President of the IEEE Communications Society. He was a Vice President for Technical & Educational Activities, a Vice President for Publications, a Member-at-Large on the Board of Governors, a Chair of the Distinguished Lecturer Selection Committee, a Member of IEEE Fellow Selection Committee of the ComSoc. He is a registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, a Chinese Academy of Engineering Foreign Member, and a Distinguished Lecturer of the IEEE Vehicular Technology Society and Communications Society. He served as an Editor-in-Chief of IEEE INTERNET OF THINGS JOURNAL, IEEE NETWORK, and *IET Communications*.