

Sparse Signal Recovery Neural Network with Application to High-Mobility Massive MIMO-OTFS Communication Systems

Mostafa Payami, and Steven D. Blostein, *Senior Member, IEEE*

Abstract—A deep learning-based sparse signal recovery network SSRnet is designed. This network is built on the proposed neural network PositionNet+, which takes the received signal as input and obtains the support of the desired sparse matrix without requiring a sensing matrix. Using PositionNet+, SSRnet is able to recover the sparse signal precisely, outperforming conventional methods, including least-squares (LS) estimation with perfectly known support, by virtue of its denoising behavior, while offering substantially reduced computation. The network is then utilized to perform channel estimation of high-mobility massive multiple-input multiple-output orthogonal time frequency space (MIMO-OTFS) wireless systems which is cast as a sparse signal recovery problem. In OTFS, data is modulated in the delay-Doppler domain to transform a fast time-varying and frequency-selective fading channel into a quasi-static and sparse channel. To maximize performance, OTFS systems require accurate channel estimation and low pilot signaling which are provided by SSRnet. Simulation and computational comparisons demonstrate that the proposed approach enhances performance in terms of bit error rate (BER) and normalized mean squared error (NMSE), reduces pilot symbol overhead, as well as lowers computational complexity.

Index Terms—Sparse signal recovery, orthogonal time frequency space (OTFS), deep learning (DL), convolutional neural network (CNN), massive multiple-input multiple-output (MIMO), channel estimation.

I. INTRODUCTION

ULTRA high data rate communications in high-mobility scenarios is an envisioned requirement in sixth-generation (6G) wireless cellular networks. Currently, orthogonal frequency division multiplexing (OFDM) is the dominant modulation technique for 5G communication systems. However, in high-mobility applications, OFDM might experience considerable inter-carrier interference (ICI), a consequence of the Doppler effect associated with frequency-selective time-varying fading channels. In addition to the stochastic complex gain, the fading channel also imposes a different transmission delay and Doppler frequency shift on each transmission path. While OFDM is capable of mitigating the inter-symbol interference (ISI) by appending a cyclic prefix to the transmitted frames, the remaining ICI caused by the Doppler spread may lead to severe performance degradation [1]–[3].

Manuscript received 22 May 2024; revised 24 December 2024; accepted 4 March 2025. This work was supported by the Natural Research Council of Canada Discovery Grant RGPIN-2019-06237.

Mostafa Payami and Steven D. Blostein are with the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON, Canada, e-mail: mostafa.payami@queensu.ca and steven.blostein@queensu.ca.

Sparse signal recovery arises in a wide range of applications from wireless communications to signal processing and image compression, where the main goal is the extraction of a sparse signal from a potentially under-sampled or compressed set of measurements. Several preeminent algorithms have been developed to solve the sparse recovery problems. Orthogonal matching pursuit (OMP) has been recognized as a robust method, offering computational efficiency, simplicity, and versatility across diverse applications. Sparse Bayesian learning (SBL) is another powerful signal recovery method known to be more sparsity-inducing than other methods that incorporate Laplacian priors. SBL is also better able to handle ill-conditioned measurement matrices, thereby improving signal recovery performance. However, SBL has the drawback of prohibitively high computational complexity of matrix inversion in the expectation-maximization (EM) iterations. Employing approximate message passing (AMP), the developed inverse-free AMP-SBL algorithm can improve the computational efficiency. Nonetheless, a large number of iterations and careful hand-tuning are still required for convergence [4]–[7].

Channel estimation of high-mobility orthogonal time frequency space (OTFS)-based communication networks has recently been formulated as a sparse signal recovery problem. OTFS modulates information-bearing symbols on the delay-Doppler (DD) plane using the inverse symplectic finite Fourier transform (ISFFT), spreading them over the time-frequency grid. By applying the symplectic finite Fourier transform (SFFT), the delay-Doppler domain channel representation is obtained [8], [9]. Excellent performance has been demonstrated by OTFS compared to traditional OFDM in high-mobility environments. The improvements offered by OTFS are explained by fact that the channel impulse response (CIR) of doubly-dispersive environments in the delay-Doppler domain is quasi-time-invariant and sparse. During these quasi-stationary time periods of the doubly-dispersive channel, the received OTFS signal becomes a 2D convolution of the data symbols with the effective channel response. As a result, the OTFS can leverage time and frequency diversities inherent in the channel to improve communications reliability [10], [11].

Massive multiple-input multiple-output (MIMO) can be combined with OTFS to improve the spectral and energy efficiency, diminish adverse effects of the propagation environment, and suppress multiuser interference. To fully capitalize on the benefits of massive MIMO-OTFS systems, obtaining accurate channel state information (CSI) of the fading environment, which is cast as a sparse signal recovery problem, is

a prerequisite and becomes more complicated as the number of transmit or receive antennas increases [12], [13].

A. Related Work

To obtain accurate CSI at the transmitter, a compact angular-delay-Doppler (ADD)-domain representation of MIMO fading channels is formulated in [14] for OTFS-based communication systems. An OMP algorithm is then developed to acquire the 4D-sparse CSI of this MIMO-OTFS system. A notable feature of the derived channel model is its efficacy in extracting the 4D-sparsity in the ADD domain occurring due to a few dominant scatterers, with a low pilot overhead ratio. The CSI acquisition problem of MIMO-OTFS systems is analyzed in [15]. Based on delay-Doppler domain characteristics, this problem is cast as a block-sparse signal recovery problem, which is then solved by the introduced block-sparse Bayesian learning with block reorganization (BSBL-BR) algorithm, where the size of non-sparse blocks is updated iteratively to achieve higher estimation accuracy.

A channel estimation method for downlink massive MIMO-OTFS is proposed in [16] which considers environments with fractional Doppler and employs deterministic pilot designs. Since the fading channel geometry changes slowly relative to the communication signaling timescale, a modified sensing matrix-based channel estimation (MSMCE) algorithm is designed to obtain the CSI. In [17], downlink channel estimation of massive MIMO-OTFS is formulated as a sparse signal recovery problem, and a 3D-structured OMP (3D-SOMP) algorithm is proposed to estimate the MIMO channel matrix. The 3D-SOMP method can extract the structured sparsity in MIMO-OTFS channels, i.e., normal, block, and burst sparsity along the delay, Doppler and angular dimensions, respectively.

There has also been emerging interest in communication systems that feature wide-band millimeter-wave (mmWave) frequencies [18] as well as in integrated sensing and communications (ISAC), where the objective is to estimate positions, angles and velocities of mobile users during communication [19]. Both [18] and [19] consider the estimation of parameters of a model consisting of known numbers of target and multipaths in a MIMO-OFDM system using tensor decomposition-based frameworks. Subspace estimation for structured matrices are employed with promising results. In general, however, numbers of targets and propagation paths are not known a priori and may be difficult to both detect and estimate. This motivates the transformation of the model from time-frequency to delay-Doppler representation such that the resulting signal representation becomes sparse. As a result, parameters can be both detected by the presence of energy in delay-Doppler bins as well as explicitly estimated simply by their indices rather than requiring signal array processing algorithms based on an assumed problem structure.

Deep learning-aided approaches to MIMO-OTFS channel estimation have also been investigated [4], [20]–[27]. A two-step OMP method for beamspace channel estimation of mmWave massive MIMO systems is developed in [4], which is based on deep learning and compressed sensing. First, a composite convolution kernel function is introduced

for coarsely estimating the angles of departure (AoDs) and angles of arrival (AoAs) from the correlation matrix. In the second step, a Squeeze-and-Excitation residual network (SE-ResNet) with Noise-to-Void learning strategy is devised to denoise the correlation matrix and refine the estimation of AoAs and AoDs. The approach in [4] does not require labeled data and is robust in the low signal-to-noise ratio (SNR) regime. A denoising convolutional neural network (DnCNN)-assisted channel estimation framework for OTFS-based high-speed vehicular communications is proposed in [20] by taking advantage of hybrid dilated convolution and incorporating residual paths into a convolutional neural network (CNN). A coarsely estimated channel matrix, $\hat{\mathbf{H}}_{\text{OMP}}$ is first obtained by the OMP and is then fed into the DnCNN as input, while the true channel matrix, \mathbf{H} , is used for labeled training. Through two noise cancellation modules, the DnCNN learns the difference between $\hat{\mathbf{H}}_{\text{OMP}}$ and \mathbf{H} to extract the noise matrix $\hat{\mathbf{H}}_n$. This scheme can mitigate the noise and interference and thus provides a higher accuracy.

Deep neural network (DNN)-based channel estimation for OTFS modulation in air-to-ground communication scenarios with fractional and highly dynamic Doppler environments are considered in [21]. The proposed DNN is composed of four fully-connected layers with rectified linear unit $\text{ReLU}(x) = \max(0, x)$ activation function. The received pilot symbols in the delay-Doppler domain are processed by the neural network to estimate the channel parameters that are used in the maximal ratio combining stage to demodulate the received signal. The DNN in [21] can achieve similar performance compared to conventional methods with a pilot energy reduction of 16 dB. In [22], sparse channel estimation is first performed by OMP and then a DNN consisting of five fully-connected layers is used to enhance the accuracy. Since the introduced DL-based scheme in [22] is a model-driven paradigm, it has the advantages of a small set of training data and a short training time. In [23], a data-driven framework is proposed for joint measurement matrix design and support recovery of complex sparse signals. The designed network consists of two modules: an autoencoder and a hard thresholding module. The introduced approach can extract the sparsity patterns and perform sparse support recovery with a low computational complexity. This scheme is then applied to massive machine type communications (mMTC) for device activity detection in grant-free massive access, leading to an improved performance compared to prevailing techniques.

In [24], by unfolding the iterative soft thresholding algorithm (ISTA), a trainable ISTA (TISTA) is proposed and applied to the sparse signal recovery problem. The TISTA algorithm includes a linear estimation component and a minimum mean squared error (MMSE)-based shrinkage component. TISTA includes adjustable parameters that can control step size and error variance for the MMSE shrinkage function, and moreover, the number of trainable variables is nearly equal to the number of layers in the unfolded network, and therefore results in a highly stable and fast training process. Joint activity detection and channel estimation in massive random access is investigated in [25]. When the receiver is equipped with multiple antennas, the problem is one of joint

sparse recovery with multiple measurement vectors. In [25], a modified MMSE shrinkage function is applied to the TISTA. This shrinkage function is then learned through a model-driven deep network. The required preamble length is dramatically decreased compared to that of TISTA.

A frequency-selective wideband mmWave system is examined in [26] and two networks based on deep learning with compressive sensing (DL-CS) are proposed for channel estimation. The networks learn useful a priori information from the training dataset to achieve accurate channel estimates with a small number of training symbols. The initial DL-CS scheme estimates the support of the channel matrix in the frequency domain, which is subsequently utilized to reconstruct the fading channel. The second scheme uses the obtained support and implements a low-complexity fine-tuning algorithm to improve estimation accuracy. The DL-based schemes surpass the conventional OMP method in terms of normalized mean squared error (NMSE) and spectral efficiency, especially in the low SNR regime. In [27], reconfigurable intelligent surface (RIS)-aided multiuser mmWave massive MIMO systems are investigated over cascaded frequency-selective and flat fading channels, and compressive sensing based data-driven neural networks for CSI estimation with low training overhead are designed. The methods proposed in [27] extract sparsity among different subcarriers as well as extract sparsity of the angular cascaded channel matrices. A denoising deep network is employed to precisely recover the support of the channel, and as a result, improved NMSE performance with reduced pilot overhead and computational complexity can be achieved.

While conventional methods provide channel estimates with acceptable accuracy, they have high computational complexity due to hundreds of iterations, where each iteration contains the computationally intensive Moore-Penrose pseudo-inverse. The pilot overhead also tends to be very high, rendering these methods costly or infeasible for practical implementation. On the other hand, although the DL-based solutions described above can yield improved results, they are still based on conventional computationally intensive OMP-based methods, followed by several DNN layers that only refine the estimation.

B. Contributions

Motivated by the aforementioned considerations, DNN-based methods are investigated to address the sparse signal recovery problem that avoid certain drawbacks of conventional signal-based recovery methods. Deep learning is motivated by its potential to improve estimation accuracy, reduce computational complexity, reduce pilot overhead, as well as extract sparsity patterns in the data [28], [29]. To realize neural networks' great potential, innovative design is required. In this paper, a unique design is proposed to capture the sparsity-related patterns of the signal, resulting in significantly improved performance.

Specifically, the design utilizes CNN-based networks, which are well-suited for extracting spatial features of multi-dimensional data [30], [31]. Recently, PositionNet introduced in [32], extracts spatial features by taking the correlation vector $|\Phi^H \mathbf{y}_{DD}|$ as input, where \mathbf{y}_{DD} is the delay-Doppler

domain received signal and Φ is the sensing matrix, and finds the indices of nonzero elements (support) of the sparse channel matrix. Here, we build upon that idea and propose a novel CNN-based network PositionNet+, which only requires the received signal \mathbf{y}_{DD} as input, and processes its spatial features to obtain the support of the desired sparse matrix. The proposed PositionNet+ has two main advantages over PositionNet: (i) It can find the support of the sparse matrix more accurately, and (ii) it does not require knowledge of the sensing matrix. In fact, PositionNet+ can be used standalone in applications where the goal is only to find the locations of nonzero elements, e.g., gaining environment-related information in integrated sensing and communications (ISAC), or it can alternatively be incorporated into other neural networks for a variety of purposes such as sparse signal recovery.

The PositionNet+ support extractor is next incorporated into the proposed design of SSRnet, a neural network being able to precisely recover the sparse signal. That is, SSRnet solves the problem of channel estimation in high-mobility massive MIMO-OTFS cellular networks. This problem is formulated as a sparse signal recovery problem, given the limited number of scatterers or reflectors in the delay-Doppler domain. The training data is obtained from simulating the 3GPP spatial channel model for realistic urban macrocellular scenarios [34], to establish DL-based schemes that are practically applicable. Finally, the NMSE and bit error rate (BER) performance are evaluated, with perfectly known support followed by LS estimation of CSI as the benchmark, and the results are compared to existing methods.

The rest of this paper is organized as follows. In Section II, fundamentals of OTFS and massive MIMO-OTFS system model are presented, and the channel estimation problem is formulated. The architecture of the proposed PositionNet+ and SSRnet are described in Section III and the training procedure is presented. Simulation results are illustrated in Section IV, and finally, concluding remarks are drawn in Section V.

II. SYSTEM MODEL AND PROBLEM FORMULATION

In this section, the delay-Doppler representation of OTFS signals is first presented. Following that, OTFS modulation and demodulation are briefly reviewed. The massive MIMO-OTFS system model is then derived and the channel estimation problem is formulated.

A. Preliminaries of OTFS

Let M and N denote the numbers of information symbols placed along the delay and Doppler axes, respectively. The 2D delay-Doppler domain lattice includes M points along the delay dimension with spacing $\Delta\tau$, and N points along the Doppler dimension with spacing $\Delta\nu$, establishing a set of MN points inside the DD-domain rectangle expressed by

$$\mathcal{F} = \{(\ell\Delta\tau, k\Delta\nu); \ell = 0, 1, \dots, M-1, k = 0, 1, \dots, N-1\}, \quad (1)$$

in which $\Delta\tau = \frac{\tau_r}{M}$ and $\Delta\nu = \frac{\nu_r}{N}$, where τ_r and ν_r are the delay period and Doppler period, respectively, chosen to satisfy the condition $\tau_r\nu_r = 1$. A doubly-selective fading

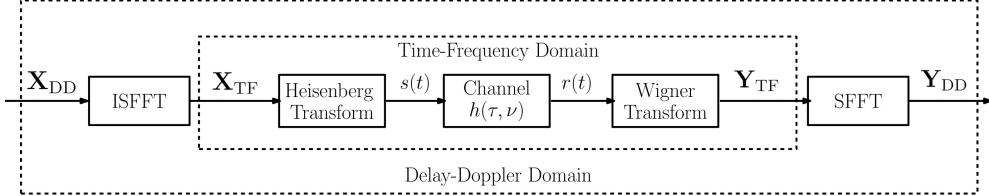


Fig. 1. OTFS modulation and demodulation [35].

channel in the delay-Doppler domain is delimited by the finite support $[0, \tau_{\max}]$ along the delay dimension and by $[-\nu_{\max}, \nu_{\max}]$ along the Doppler dimension, where τ_{\max} and ν_{\max} represent the maximum delay spread and Doppler spread of the time-varying frequency-selective channel and in general the underspread condition $\tau_{\max}\nu_{\max} \ll 1$ is satisfied. The delay period τ_r and Doppler period ν_r are selected so that they guarantee $\nu_r \gg \nu_{\max}$ and $\tau_r \gg \tau_{\max}$ which is feasible since $\tau_r\nu_r = 1$. For instance, in a typical urban environment $\tau_{\max} = 5 \mu\text{s}$, and for a mobile user moving at a velocity of 250 km/h and receiving service from a BS operating at a carrier frequency of $f_c = 4.25 \text{ GHz}$, $\nu_{\max} = 984 \text{ Hz}$ and thus, $\tau_{\max}\nu_{\max} = 4.92 \times 10^{-3} \ll 1$. Consequently, the delay period and Doppler period can be chosen as $\tau_r = 100 \mu\text{s}$ and $\nu_r = 10 \text{ KHz}$, respectively.

The reciprocal time-frequency domain lattice consists of M points along the frequency dimension with spacing $\Delta f = \frac{1}{\tau_r}$, and N points along the time dimension with spacing $\Delta t = \frac{1}{\nu_r}$ expressed as

$$\mathcal{G} = \{(m\Delta f, n\Delta t); m = 0, 1, \dots, M-1, n = 0, 1, \dots, N-1\}. \quad (2)$$

The CIR in the delay-Doppler domain manifests quasi-static and sparse characteristics. By contrast, it exhibits irregular and non-sparse characteristics with rapid fluctuations in the time-frequency domain. The sparsity is beneficial for channel estimation, CSI feedback compression or hybrid analog and digital beamforming.

B. OTFS Modulation and Demodulation

The general OTFS scheme is depicted in Fig. 1. First, the MN data symbols \mathbf{X}_{DD} are positioned on the delay-Doppler domain lattice \mathcal{F} . By employing the ISFFT, the symbols are transformed into the time-frequency domain as

$$\mathbf{X}_{\text{TF}}[m, n] = \frac{1}{\sqrt{MN}} \sum_{\ell=0}^{M-1} \sum_{k=0}^{N-1} \mathbf{X}_{\text{DD}}[\ell, k] e^{j2\pi(\frac{nk}{N} - \frac{m\ell}{M})}, \quad (3)$$

then the Heisenberg transform is applied to transform the signal into the time domain. After adding a cyclic prefix in the time domain, the signal transmitted through the doubly-dispersive channel is expressed as

$$s(t) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mathbf{X}_{\text{TF}}[m, n] g_{\text{tx}}(t - n\Delta t) e^{j2\pi m\Delta f(t - n\Delta t)}, \quad (4)$$

where $g_{\text{tx}}(\cdot)$ represents the transmit pulse shaping function.

The received signal impacted by the time-varying fading channel is a superimposition of several dominant paths of the transmitted signal. Each path experiences a distinct delay and Doppler spread and is filtered by a delay-Doppler CIR, represented by $h_d(\tau, \nu)$. Upon the removal of the cyclic prefix, the received signal can be expressed as

$$r(t) = \int_{-\nu_{\max}}^{\nu_{\max}} \int_0^{\tau_{\max}} s(t - \tau) h_d(\tau, \nu) e^{j2\pi\nu(t - \tau)} d\tau d\nu + z(t), \quad (5)$$

where $z(t)$ models the continuous-time additive white Gaussian noise (AWGN). The output of the receiver matched filter is written as

$$\mathbf{Y}_{\text{TF}}[m, n] = \int r(t) g_{\text{rx}}^*(t - n\Delta t) e^{j2\pi m\Delta f(t - n\Delta t)} dt, \quad (6)$$

in which $g_{\text{rx}}(\cdot)$ is the receive pulse shaping function. The above operation that transforms the received signal from time domain to time-frequency domain is known as the Wigner transform. Finally, the SFFT is applied to transform the time-frequency domain waveform into the delay-Doppler domain signal which is given by

$$\mathbf{Y}_{\text{DD}}[\ell, k] = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mathbf{Y}_{\text{TF}}[m, n] e^{-j2\pi(\frac{nk}{N} - \frac{m\ell}{M})}. \quad (7)$$

C. System Model

We consider a massive MIMO-OTFS system equipped with N_t antennas at the BS as a large-scale antenna array providing service to N_u single-antenna mobile users. To mitigate the inter-user interference, downlink precoding, e.g., Tomlinson-Harashima, is performed based on the CSI acquired through uplink feedback channel in frequency-division duplexing (FDD) communication links [11]. Since the channel estimation procedure is identical for all the users in the downlink, without loss of generality the problem is analyzed and formulated for an arbitrary user.

As mentioned, a set of MN symbols \mathbf{X}_{DD} comprising information-bearing, guard and pilot symbols are multiplexed in the delay-Doppler domain. It is assumed that there are N_p dominant propagation paths in the downlink time-varying fading channel and each path consists of N_s sub-paths. Therefore, the delay-Doppler channel response $h_d(\tau, \nu)$ is given by [16]

$$h_d(\tau, \nu) = \sum_{p=1}^{N_p} \sum_{s=1}^{N_s} \alpha_s \delta(\tau - \tau_p) \delta(\nu - \nu_s), \quad (8)$$

where α_s and ν_s signify the complex path gain and Doppler shift associated with the s^{th} sub-path of p^{th} dominant path, respectively. All signals that pass through any sub-path of the p^{th} dominant path can be considered to arrive with the same delay τ_p [36]. Assuming $g_{\text{rx}}(\cdot)$ and $g_{\text{tx}}(\cdot)$ are rectangular waveforms, the time-varying channel of the $(r+1)^{\text{th}}$ transmit antenna, where $r = 0, 1, \dots, N_t - 1$, on the delay tap ℓ at time kT_s can be written as [37]

$$h_t[\ell, k, r] = \sum_{p=1}^{N_p} \sum_{s=1}^{N_s} \alpha_s e^{j2\pi k T_s \nu_s} \delta(\ell T_s - \tau_p) e^{-j2\pi \theta_s r}, \quad (9)$$

in which $T_s = \frac{1}{M \Delta f}$ indicates the receiver sampling interval and $\theta_s = \frac{d}{\lambda} \sin \phi_s$ is the spatial angle associated with the AoD of s^{th} sub-path, ϕ_s , considering a uniform linear antenna array at the BS, where d and λ are the antenna spacing and carrier wavelength, respectively. The spatial-delay-Doppler (SDD)-domain channel response $\mathcal{H}_{\text{SDD}} \in \mathbb{C}^{M \times N \times N_t}$ of the massive MIMO-OTFS system can then be expressed as [17]

$$\begin{aligned} \mathcal{H}_{\text{SDD}}[\ell, k, r] &= \sum_{n=0}^{N-1} h_t[\ell, n(M + N_{\text{cp}}) + 1, r] e^{-j2\pi \frac{n k}{N}} \\ &= \sum_{p=1}^{N_p} \sum_{s=1}^{N_s} \rho_s \Lambda_N(\nu_s N(M + N_{\text{cp}}) T_s - k) \\ &\quad \times \delta(\ell T_s - \tau_p) e^{-j2\pi \theta_s r}, \end{aligned} \quad (10)$$

where $\rho_s = \alpha_s e^{j2\pi \nu_s T_s}$, N_{cp} is the length of cyclic prefix and $\Lambda_N(x) = \sum_{n=0}^{N-1} e^{j2\pi \frac{n x}{N}} = \frac{\sin(\pi x)}{\sin(\pi \frac{x}{N})} \exp\left(j\pi \frac{x(N-1)}{N}\right)$.

The ADD-domain channel response $\mathcal{H}_{\text{ADD}} \in \mathbb{C}^{M \times N \times N_t}$ is obtained by applying the inverse discrete Fourier transform (DFT) to \mathcal{H}_{SDD} along the spatial dimension:

$$\mathcal{H}_{\text{ADD}}[\ell, k, b] = \sum_{r=0}^{N_t-1} \mathcal{H}_{\text{SDD}}[\ell, k, r] e^{j2\pi \frac{rb}{N_t}}, \quad (11)$$

where $b = -\frac{N_t}{2}, \dots, 0, \dots, \frac{N_t}{2} - 1$ denotes the angular (beam) index. By substituting (11) into (10), the ADD-domain channel \mathcal{H}_{ADD} can be written as

$$\begin{aligned} \mathcal{H}_{\text{ADD}}[\ell, k, b] &= \sum_{p=1}^{N_p} \sum_{s=1}^{N_s} \rho_s \Lambda_N(\nu_s N(M + N_{\text{cp}}) T_s - k) \\ &\quad \times \delta(\ell T_s - \tau_p) \Lambda_{N_t}(b - \theta_s N_t). \end{aligned} \quad (12)$$

A key characteristic of \mathcal{H}_{ADD} is its 3D sparsity, i.e., normal sparsity, block sparsity and burst sparsity along the delay, Doppler and angular dimensions, respectively. This feature can be exploited to acquire the CSI efficiently with a small pilot overhead ratio.

D. Channel Estimation Problem Formulation

The structure of an OTFS frame in the delay-Doppler domain is illustrated in Fig. 2. It consists of data and pilot symbols as well as guard intervals to avoid interference between data and pilot symbols. The length of pilots along the delay and Doppler dimensions are M_τ and N_ν , respectively. The pilots are transmitted with the same power as data

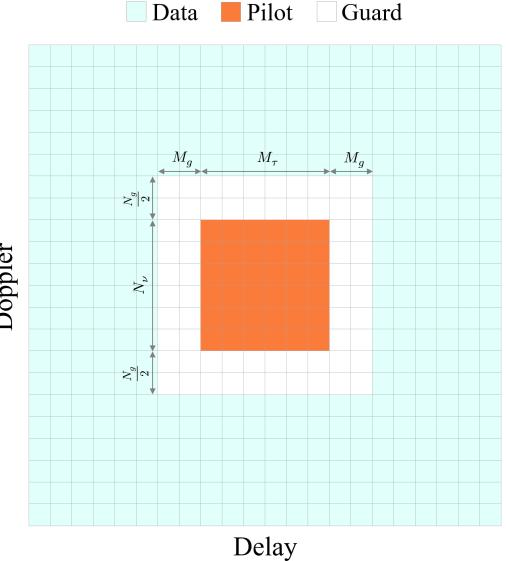


Fig. 2. The structure of an OTFS frame in the delay-Doppler domain.

symbols. The length of guard intervals is chosen so as to avoid interference between pilots and data.

The sparse channel tensor \mathcal{H}_{ADD} in massive MIMO-OTFS systems has finite support along both delay and Doppler dimensions. Therefore, within each OTFS frame, $M_\tau \times N_\nu$ pilot symbols are mounted along the delay and Doppler axes, selected in a sufficiently large number to recover the nonzero part of \mathcal{H}_{ADD} . The pilot-containing OTFS frames are then modulated and transmitted simultaneously over the N_t transmit antennas. At the receiver, the pilot symbols can be expressed as [17]

$$\begin{aligned} \mathbf{Y}_{\text{DD}}^{\text{P}}[\ell, k] &= \sum_{r=0}^{N_t-1} \sum_{\tilde{\ell}=0}^{M_\tau-1} \sum_{\tilde{k}=-\frac{N_\nu}{2}}^{\frac{N_\nu}{2}-1} \psi[\ell - \tilde{\ell}, \tilde{k}] \mathbf{H}_{\text{SDD}}[\tilde{\ell}, \tilde{k}, r] \\ &\quad \times \mathbf{P}_{\text{SDD}}[\ell - \tilde{\ell}, k - \tilde{k}, r] + \mathbf{Z}_{\text{DD}}^{\text{P}}[\ell, k], \end{aligned} \quad (13)$$

where $\mathbf{Y}_{\text{DD}}^{\text{P}} \in \mathbb{C}^{M_\tau \times N_\nu}$, $\psi_{\ell, k} = \exp\left(\frac{j2\pi \ell k}{N(M + N_{\text{cp}})}\right)$ is the compensation phase, $\mathbf{H}_{\text{SDD}} \in \mathbb{C}^{M_\tau \times N_\nu \times N_t}$ is a truncated replica of \mathcal{H}_{SDD} along delay and Doppler dimensions which encompasses the nonzero part of \mathcal{H}_{SDD} , $\mathbf{P}_{\text{SDD}} \in \mathbb{C}^{M_\tau \times N_\nu \times N_t}$ is the SDD-domain pilot tensor chosen from complex Gaussian random sequences, and $\mathbf{Z}_{\text{DD}}^{\text{P}} \in \mathbb{C}^{M_\tau \times N_\nu}$ is the discrete-time delay-Doppler domain AWGN impacting the pilot signals. Note that here $\ell = 0, 1, \dots, M_\tau - 1$ and $k = -\frac{N_\nu}{2}, \dots, 0, \dots, \frac{N_\nu}{2} - 1$. The channel response \mathbf{H}_{ADD} is expressed as the DFT of the corresponding ADD-domain sparse channel $\mathbf{H}_{\text{ADD}} \in \mathbb{C}^{M_\tau \times N_\nu \times N_t}$, which contains the nonzero part of \mathcal{H}_{ADD} , along the angular dimension:

$$\mathbf{H}_{\text{SDD}}[\ell, k, r] = \sum_{b=-\frac{N_t}{2}}^{\frac{N_t}{2}-1} \mathbf{H}_{\text{ADD}}[\ell, k, b] e^{-j2\pi \frac{br}{N_t}}, \quad (14)$$

by substituting (14) into (13) and defining $\mathbf{P}_{\text{ADD}} = \sum_{r=0}^{N_t-1} \mathbf{P}_{\text{SDD}} e^{j2\pi \frac{rb}{N_t}}$,

$$\begin{aligned} \mathbf{Y}_{\text{DD}}^{\text{P}}[\ell, k] = & \sum_{b=-\frac{N_t}{2}}^{\frac{N_t}{2}-1} \sum_{\tilde{\ell}=0}^{M_\tau-1} \sum_{\tilde{k}=-\frac{N_\nu}{2}}^{\frac{N_\nu}{2}-1} \psi[\ell - \tilde{\ell}, \tilde{k}] \mathbf{H}_{\text{ADD}}[\tilde{\ell}, \tilde{k}, b] \\ & \times \mathbf{P}_{\text{ADD}}[\ell - \tilde{\ell}, k - \tilde{k}, b] + \mathbf{z}_{\text{DD}}^{\text{P}}[\ell, k], \end{aligned} \quad (15)$$

to streamline the model, $\mathbf{Y}_{\text{DD}}^{\text{P}}$ is vectorized to form $\mathbf{y}_{\text{DD}} \in \mathbb{C}^{M_\tau N_\nu \times 1}$ and \mathbf{H}_{ADD} is rearranged into N_t vectors $\mathbf{h}_b \in \mathbb{C}^{M_\tau N_\nu \times 1}$. Consequently, the above model can be equivalently expressed as

$$\mathbf{y}_{\text{DD}} = \sum_{b=-\frac{N_t}{2}}^{\frac{N_t}{2}-1} (\Psi \odot \mathbf{P}_b) \mathbf{h}_b + \mathbf{z}_{\text{DD}}, \quad (16)$$

where $\mathbf{P}_b \in \mathbb{C}^{M_\tau N_\nu \times M_\tau N_\nu}$ is a quasi-Toeplitz convolution matrix whose $(\ell N_\nu + k + \frac{N_\nu}{2} + 1, \tilde{\ell} N_\nu + \tilde{k} + \frac{N_\nu}{2} + 1)$ -th element is equal to $\mathbf{P}_{\text{ADD}}[\ell - \tilde{\ell}, k - \tilde{k}, b]$, and $\Psi \in \mathbb{C}^{M_\tau N_\nu \times M_\tau N_\nu}$ is a matrix whose $(\ell N_\nu + k + \frac{N_\nu}{2} + 1, \tilde{\ell} N_\nu + \tilde{k} + \frac{N_\nu}{2} + 1)$ -th element is equal to $\psi[\ell - \tilde{\ell}, \tilde{k}]$. We denote $\mathbf{h}_{\text{ADD}} = [\mathbf{h}_{-\frac{N_t}{2}}^T, \dots, \mathbf{h}_0^T, \dots, \mathbf{h}_{\frac{N_t}{2}-1}^T]^T$ and $\Phi = [\Psi \odot \mathbf{P}_{-\frac{N_t}{2}}, \dots, \Psi \odot \mathbf{P}_0, \dots, \Psi \odot \mathbf{P}_{\frac{N_t}{2}-1}]$, in which \odot indicates the Hadamard product. Therefore, the end-to-end massive MIMO-OTFS system model is finally formulated as the following sparse signal recovery problem

$$\mathbf{y}_{\text{DD}} = \Phi \mathbf{h}_{\text{ADD}} + \mathbf{z}_{\text{DD}}, \quad (17)$$

where \mathbf{y}_{DD} is the observation vector, $\Phi \in \mathbb{C}^{M_\tau N_\nu \times M_\tau N_\nu N_t}$ is the sensing matrix depending on the pilot symbols, $\mathbf{h}_{\text{ADD}} \in \mathbb{C}^{M_\tau N_\nu N_t \times 1}$ is a sparse vector that is rearranged into a $M_\tau \times N_\nu \times N_t$ tensor to obtain the ADD-domain sparse channel tensor \mathbf{H}_{ADD} , and $\mathbf{z}_{\text{DD}} \in \mathbb{C}^{M_\tau N_\nu \times 1}$ is the AWGN vector. The reader is referred to [17] for further details on the properties of high-mobility fading channel and the structure of the sensing matrix.

While conventional methods, e.g., Algorithm 1 in [14] and Algorithm 1 in [17], require computationally intensive tasks which require many iterations that each compute the Moore-Penrose pseudo-inverse of a matrix, we instead propose a CNN-based network to solve this sparse signal recovery problem and obtain the estimated channel $\hat{\mathbf{H}}_{\text{ADD}}$ with improved performance compared to previously proposed algorithms. To ensure real-valued computation in the neural network architecture, the complex-valued signal model is represented by its equivalent real-valued counterpart with double the dimensionality, i.e., Eq. (17) is expressed as

$$\begin{bmatrix} \text{Re}(\mathbf{y}_{\text{DD}}) \\ \text{Im}(\mathbf{y}_{\text{DD}}) \end{bmatrix} = \begin{bmatrix} \text{Re}(\Phi) & -\text{Im}(\Phi) \\ \text{Im}(\Phi) & \text{Re}(\Phi) \end{bmatrix} \begin{bmatrix} \text{Re}(\mathbf{h}_{\text{ADD}}) \\ \text{Im}(\mathbf{h}_{\text{ADD}}) \end{bmatrix} + \begin{bmatrix} \text{Re}(\mathbf{z}_{\text{DD}}) \\ \text{Im}(\mathbf{z}_{\text{DD}}) \end{bmatrix}. \quad (18)$$

III. PROPOSED PositionNet+ AND SSRnet

In this section, the architectures of PositionNet+ and SSRnet are each described along with their design rationale.

A. PositionNet+

The first step towards designing the SSRnet is to create a network that can obtain the support of the desired sparse channel matrix. In [32], the neural network PositionNet is proposed that takes the correlation-based proxy $|\Phi^H \mathbf{y}_{\text{DD}}|$ as input and produces a sparse *zero-one* tensor $\hat{\mathcal{I}}$ which is an estimation of the true support $\mathcal{I} = \text{sign}(|\mathbf{H}_{\text{ADD}}|)$, where the *ones* indicate the existence of nonzero elements. In the following, the approach is improved to a CNN-based network design, termed PositionNet+, that instead only requires the received signal \mathbf{y}_{DD} as input and finds the support of the sparse matrix without any sensing matrix. The proposed PositionNet+ provides the following principal advantages over PositionNet: (i) knowledge of the sensing matrix is not required, and (ii) improved accuracy in support recovery of sparse tensors compared to state-of-the-art methods.

The architecture of PositionNet+ is illustrated in the block diagram Fig. 3. The design is an alternative to the OMP algorithm, and the reader is referred to Algorithm 1 of [14] for a typical OMP implementation. To process the spatial features, the received signal is passed through three 2D Convolutional layers, each with $N_{\text{fm}} = 32$ filters with kernel size 3×3 in order to generate 32 feature maps, followed by a Leaky ReLU ($x = \max(\alpha x, x)$) activation function for nonlinearity with $\alpha = 0.1$ and a Layer Normalization (LN) layer. This is followed by another 2D Convolutional layer with one filter with kernel size 3×3 . Next, a complex Dense (fully-connected) layer, a Dropout layer with rate 0.25 and a Squared Absolute Value layer are used. This stage is analogous to $|\Phi^H \mathbf{y}_{\text{DD}}|$ in PositionNet [32], but with coefficients learned during network training rather than requiring a sensing matrix, and is responsible for obtaining features to determine the existence of nonzero elements. It should be noted that this complex layer was implemented in Keras by separating real and imaginary parts, and here we use the complex Dense layer for simplicity in illustration.

The processing continues with a Reshape layer and three 3D Convolutional layers with $N_{\text{fm}} = 32$ filters with kernel size $3 \times 3 \times 3$ to produce 32 feature maps, followed by a Leaky ReLU layer with $\alpha = 0.1$, an LN layer, a Batch Normalization (BN) layer and a Rescaling layer. Then, inspired by the argmax function in conventional methods, $\text{Softmax}(\mathbf{z}) = \frac{\exp(z_i)}{\sum_i \exp(z_i)}$ is used which is a key layer in the design. It is worth noting that Softmax's function is similar to the Attention mechanism in deep learning [38], and in fact Softmax is its core function. As simulations reveal, the Softmax layer, applied in conjunction with the Convolutional layers, results in enhanced performance compared to that of a standard Attention architecture. Finally, a Dense layer is used for nonlinear mapping whose activation function is the sigmoid $\sigma(x) = \frac{1}{1+e^{-x}}$ ensuring output values $\hat{\mathcal{I}}$ are within $[0, 1]$, accomplishing a binary classification task. Note that a simple hard thresholding binary classifier can next be applied to the output tensor to ensure that all entries of $\hat{\mathcal{I}}$ are exactly *zero* or *one*. It is worth noting that the DNN with $N_{\text{fm}} = 16$ would still yield a satisfactory performance.

To place the above process in context, OMP-based sparse CSI estimation methods reported in [14] and [17] obtain the

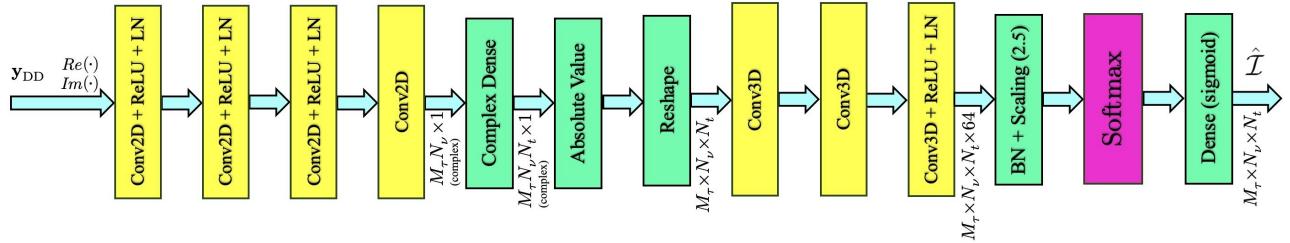


Fig. 3. The architecture of PositionNet+.

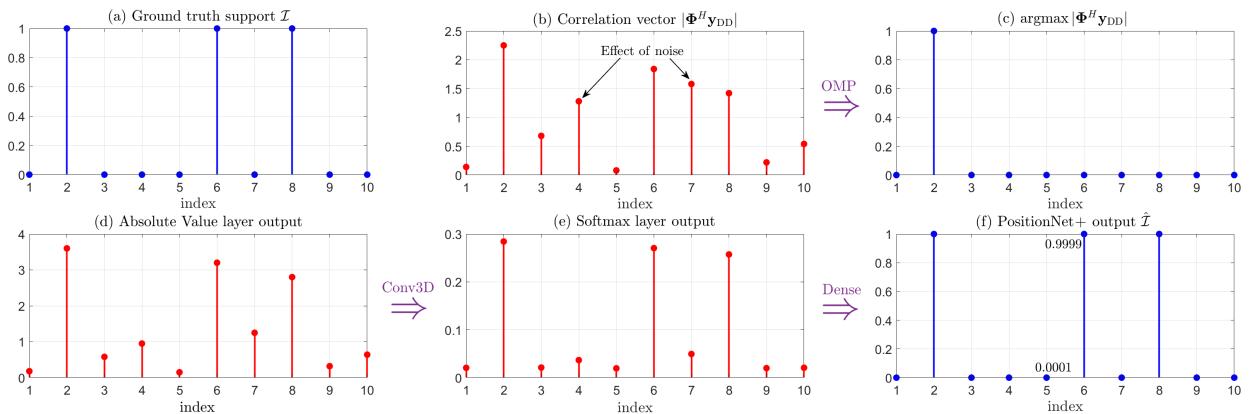
nonzero channel indices (support) by minimizing the residual between the sensing pilots and the observation signal. These are greedy methods that iteratively determine the maximum correlation match, and remove the corresponding component from the residual signal. These algorithms can be used in a wide range of applications but have two major drawbacks: (i) high computational complexity, and (ii) severe performance degradation in noisy environments. The source of computational complexity of these methods arises from the iterative process of obtaining the sparse channel support index-by-index, where each iteration requires expensive computation of the Moore-Penrose pseudo-inverse of a matrix, which increases in proportion to the cardinality of support. When the observation vector y_{DD} is noisy, OMP-based algorithm's nonzero indices may be incorrectly detected from the noisy correlation-based proxy $|\Phi^H y_{DD}|$.

In contrast, the proposed PositionNet+ makes soft decisions based on nonlinear mappings to detect multiple nonzero channel indices simultaneously. This is achieved by the Softmax layer, which processes the optimized correlation vector produced by preceding Convolutional and Complex Dense layers. PositionNet+, by virtue of supervised learning, is able to detect the support with high accuracy in the presence of severe noise, exhibiting **denoising behavior**. This can be seen from Fig. 7, which displays accuracy vs. SNRs for $N_t = 16$, $\eta = 26\%$. The results demonstrate denoising behavior: the accuracy of

PositionNet+ in the low SNR regime is close to that of the high SNR regime. While the proposed approach requires training data, which comes at some cost, the DNN has the benefit of capturing nonlinear relationships and extracting channel support with faster runtime and higher accuracy.

To illustrate this process, a conceptual example for the support recovery of a one-dimensional sparse signal is depicted in Fig. 4, which compares the recovery process of PositionNet+ versus the iterative OMP class of algorithms. Fig. 4(b) shows one iteration of OMP where the correlation vector $|\Phi^H y_{DD}|$ is fed to the argmax function, with output in Fig. 4(c). By contrast, Fig. 4(d) depicts the output of the proposed Absolute Value layer, a DNN optimized correlation vector, determined by the preceding Convolutional layer with input y_{DD} . The Softmax layer output is shown in Fig. 4(e), and PositionNet+ output \hat{I} which is an estimation of ground-truth support $I = \text{sign}(|\mathbf{H}_{ADD}|)$, is shown in Fig. 4(f), presenting the result of the simultaneous support recovery process.

PositionNet+ has $2M_r^2N_v^2N_t + 98,600$ parameters, in which the term $M_r^2N_v^2N_t$ is the number of complex-valued entries of the sensing matrix Φ , which instead of relying on a sensing matrix, it can learn the optimized coefficients in the training process. The remaining 98,600 parameters come from the other neural network layers. To train this network, two approaches can be adopted:

Fig. 4. Conceptual comparison of the support recovery process between PositionNet+ and OMP-based algorithms for a one-dimensional sparse signal: (a) Ground truth support $I = \text{sign}(|\mathbf{H}_{ADD}|)$, (b) OMP correlation vector $|\Phi^H y_{DD}|$, (c) argmax output of one OMP iteration, (d) Output of optimized correlation vector at Absolute Value layer based on Convolutional layers with input y_{DD} , (e) Output of Softmax layer, and (f) Output of PositionNet+, the estimated support \hat{I} .

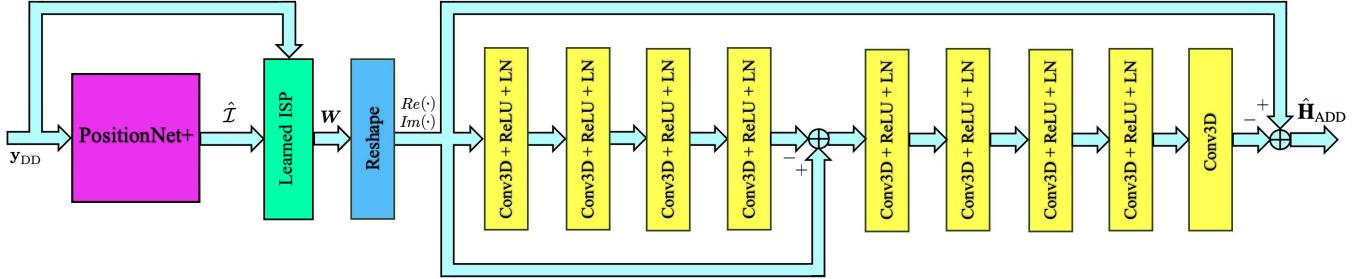


Fig. 5. The architecture of SSRnet.

- 1) choosing negated cosine similarity as the loss function:

$$\mathcal{L}_P = -\frac{\langle \mathcal{I}, \hat{\mathcal{I}} \rangle_F}{\|\mathcal{I}\|_F \|\hat{\mathcal{I}}\|_F}, \quad (19)$$

where $\langle \cdot, \cdot \rangle_F$ and $\|\cdot\|_F$ denote the Frobenius inner product and Frobenius tensor norm, respectively. Since \mathcal{I} and $\hat{\mathcal{I}}$ are both non-negative, $\mathcal{L}_P \in [-1, 0]$ and it reaches its minimum when the two tensors are fully aligned. Here, output of the sigmoid, i.e. the estimated support $\hat{\mathcal{I}}$, will be highly sharp but may miss a small percentage of nonzero indices. This method may be useful for standalone applications such as ISAC, which can be combined with conventional methods, such as the 3D-SOMP or Subspace Pursuit (SP) [39] methods, to obtain the few remaining nonzero indices.

- 2) choosing binary cross-entropy as an alternative loss function for binary classification:

$$\mathcal{L}_P = -[\mathcal{I} \log(\hat{\mathcal{I}}) + (1 - \mathcal{I}) \log(1 - \hat{\mathcal{I}})]. \quad (20)$$

Here, $\mathcal{L}_P \in [0, \infty)$ and it reaches its minimum when the corresponding entries in \mathcal{I} and $\hat{\mathcal{I}}$ are exactly the same. Given that the signal is sparse and consequently the data, i.e., the number of zero's and one's in the support \mathcal{I} , is not balanced, a customized weighted binary cross-entropy layer to compensate the data imbalance would be preferred. This way, the output $\hat{\mathcal{I}}$ will not be sharp, but would yield a small percentage more nonzero indices whose corresponding values can be suppressed by methods such as the SSRnet, which is described next.

B. SSRnet

We next incorporate PositionNet+ into the SSRnet to recover the sparse signal. The architecture of SSRnet is depicted in Fig. 5. To obtain the values of nonzero elements of 3D sparse channel matrix, the subnetwork Learned ISP is designed based on the deep unfolding Iterative Sparsification-Projection (ISP) method [40]. The ISP method solves problems of the following form:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{Ax}\|_2 \leq \epsilon, \quad (21)$$

where $\|\cdot\|_0$ is the ℓ_0 (pseudo) norm, defined as the number of nonzero entries. Since the ℓ_0 norm function is discontinuous and non-differentiable, by approximating the ℓ_0

norm with a smooth function such as $\|\mathbf{x}\|_0 = \lim_{\sigma \rightarrow 0} \|1 - \exp(-|\mathbf{x}|^2/\sigma^2)\|_1$ [41], the ISP iteratively finds the sparsest LS solution. The Improved ISP algorithm [42], which includes a momentum, is presented in Algorithm 1, where $\mathcal{S}_\sigma(\cdot)$ and $\mathcal{P}_{A_\epsilon}(\cdot)$ indicate the sparsification function and the projection step, respectively, defined by

$$\mathcal{S}_\sigma(\mathbf{x}) = \mathbf{x} \odot \left(1 - \exp\left(-\frac{|\mathbf{x}|^2}{\sigma^2}\right)\right), \quad (22)$$

$$\mathcal{P}_{A_\epsilon}(\mathbf{x}) = \mathbf{x} - \mathbf{A}^\dagger(\mathbf{Ax} - \mathbf{y}), \quad (23)$$

where \odot denotes the Hadamard (element-wise) product and the exponential function is applied entrywise. The reader is referred to [40] and [42] for details of the ISP algorithm.

Based on the ISP principle, Learned ISP, a deep unfolded subnetwork is developed which employs a momentum to achieve accelerated convergence. The schematic diagram of one layer of the Learned ISP is depicted in Fig. 6. Here, we have $\mathbf{x} = \mathbf{h}_{ADD}$, $\mathbf{A} = \Phi$, and $\mathbf{y} = \mathbf{y}_{DD}$. The initial vector \mathbf{x}_0 in the ISP algorithm is usually chosen as the regular LS solution, or a zero vector that is iteratively forced to become sparse, thereby requiring hundreds of iterations. Since PositionNet+ yields an estimated support $\hat{\mathcal{I}}$, $\hat{\mathbf{h}}_{ADD}^{(0)} = (\Phi^\dagger \mathbf{y}_{DD}) \odot \hat{\mathcal{I}}$ is fed to the Learned ISP as the initial sparse vector. In the design, the Exponential sparsification function is employed [43]

$$\mathcal{S}_\theta(\mathbf{x}) = \theta_1 \mathbf{x} - \theta_2 \mathbf{x} \odot \exp\left(-\frac{|\mathbf{x}|^2}{\theta_3^2}\right), \quad (24)$$

where learnable parameters $\boldsymbol{\theta} = \{\theta_1, \theta_2, \theta_3\}$ as well as the acceleration parameter w are obtained during training. The

Algorithm 1 Improved ISP Method

```

Inputs:  $\mathbf{A}, \mathbf{y}, \mathbf{x}_0, \sigma, w, c \in (0, 1)$ 
Initialization:  $k = 0, \mathbf{x}_{-1} = \mathbf{0}$ 
while  $\sigma > \sigma_{\min}$  do
    while  $\|\mathbf{x}_k - \mathbf{x}_{k-1}\|_2 > \epsilon$  do
         $\mathbf{z}_k = \mathbf{x}_k + w(\mathbf{x}_k - \mathbf{x}_{k-1})$ 
         $\mathbf{s}_k = \mathcal{S}_\sigma(\mathbf{z}_k)$ 
         $\mathbf{x}_{k+1} = \mathcal{P}_{A_\epsilon}(\mathbf{s}_k)$  ▷ Sparsification
    end while
     $\sigma = c \sigma$  ▷ Projection
end while
Output:  $\mathbf{x}$ 

```

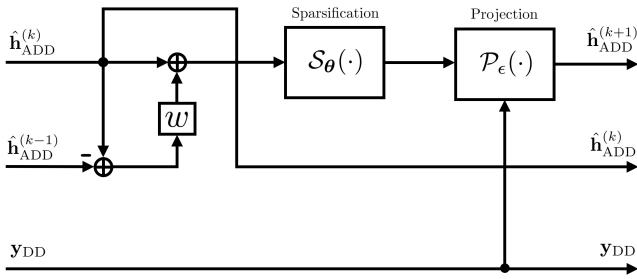


Fig. 6. Schematic diagram of one layer of Learned ISP.

estimated support $\hat{\mathcal{I}}$ is used within S_θ as $\theta_3 = \theta'_3(1 - \alpha\hat{\mathcal{I}})$, in which $\alpha \in [0, 1]$ is a learnable parameter, so that nonzero values at already known indices are not shrunk. Given that the initial vector $\hat{\mathbf{h}}_{ADD}^{(0)}$ is typically close to the true channel vector, the estimated support is used in the sparsification step, and the optimized values for learnable variables θ , α and w are found during training, the designed Learned ISP subnetwork consists of only $N_L = 12$ layers. It is worth noting that high computational complexity of OMP-based methods arises from the iterative process of obtaining the sparse channel support index-by-index, where each iteration requires expensive computation of the Moore-Penrose pseudo-inverse of a unique sub-matrix $\Phi_{\mathcal{I}}$, which is formed by collecting different columns of the sensing matrix Φ according to detected support. Conversely, the Learned ISP requires only the pseudo-inverse of Φ , which is a one-time computation that does not need to be recomputed unless the pilot symbols are changed. Computational reduction is significant, since, in communication systems, a set of pilot symbols are typically in use for a longer time period than the data transmission period. Therefore, the most computationally intensive task of the proposed approach is matrix-vector multiplication, as discussed in detail in Section III.C. A Reshape layer then rearranges the obtained sparse vector into a $M_\tau \times N_\nu \times N_t$ tensor.

An alternative to the Learned ISP block can be a computational non-trainable layer that constructs $\Phi_{\mathcal{I}} = \Phi(:, \hat{\mathcal{I}})$ by collecting columns of the sensing matrix according to the obtained support, computing its Moore-Penrose pseudo-inverse $\Phi_{\mathcal{I}}^\dagger$ followed by $\Phi_{\mathcal{I}}^\dagger \mathbf{y}_{DD}$ using TensorFlow's linear algebra module to obtain an LS solution, and then scatters the values of nonzero elements obtained as a non-sparse vector into a sparse tensor of size $M_\tau \times N_\nu \times N_t$ at indices specified by the obtained support $\hat{\mathcal{I}}$.

Next, real and imaginary parts of the signal are separated and fed into eight 3D Convolutional layers with residual (skip) connections, each with $N_{fm} = 32$ filters with kernel size $3 \times 3 \times 3$ to yield 32 feature maps, followed by a Leaky ReLU layer with $\alpha = 0.1$ and an LN layer. Finally, one 3D Convolutional layer with one filter with kernel size $3 \times 3 \times 3$ and ℓ_1 Activity Regularizer for sparsity promotion is used to produce the estimated signal. SSRnet, including the PositionNet+, has $2M_\tau^2N_\nu^2N_t + 490,700$ parameters.

The Learned ISP subnetwork after PositionNet+ yields a

coarse and noisy LS estimation of the values of the nonzero indices detected by PositionNet+. However, three potential problems remain: 1) there may be mistakenly detected indices (false positives) or missed indices (false negatives), 2) values of nonzero indices are noisy, and 3) the sparse recovery problem in Eq. (17) is an approximation to the underlying nonlinear relationship between \mathbf{y}_{DD} and \mathbf{h}_{ADD} , whose accuracy degrades in scenarios such as low pilot overhead ratios. To mitigate the effects of the above-mentioned issues, SSRnet is designed to benefit from the fast and accurate support recovery capability of PositionNet+ and the subsequent 3D Convolutional layers that serve a twofold purpose: (i) suppressing values at mistakenly detected indices and reproducing values at rarely missed indices, and (ii) implicitly denoising the obtained 3D sparse channel matrix. This implies that the overall accuracy of SSRnet in finding the support of a sparse matrix can potentially approach 100%.

To train this neural network, a customized Keras layer is implemented that computes the logarithm of NMSE of the sparse channel matrix \mathbf{H}_{ADD} as the loss function

$$\mathcal{L} = \log(\text{NMSE}) = \log \left(\frac{\|\mathbf{H}_{ADD} - \hat{\mathbf{H}}_{ADD}\|_F^2}{\|\mathbf{H}_{ADD}\|_F^2} \right). \quad (25)$$

The rationale for using $\log(\text{NMSE})$ is to emphasize better performance in the high-SNR regime.

C. Computational Complexity

The overall computational complexity of PositionNet+ and SSRnet are determined by the complexity orders of Convolutional and Dense layers, as well as the Learned ISP layers obtaining an LS solution. These are presented in Table I in comparison to complexity orders of the conventional OMP and SBL algorithms. In Table I, K_τ , whose maximum value is equal to N_p in the investigated system, indicates the sparsity along the delay dimension.

From the comparisons in Table I, it is observed that the greatest complexity order of SSRnet is $N_L \mathcal{O}(M_\tau^2 N_\nu^2 N_t)$, arising from the matrix-vector multiplication within Learned ISP layers, whereas for OMP it is $\mathcal{O}(K_\tau^3 M_\tau N_\nu^4 N_t^3) + \mathcal{O}(K_\tau M_\tau^2 N_\nu^3 N_t^2)$. It is clear that OMP's computation order is significantly larger as N_t drastically increases in massive MIMO systems. The complexity order of SBL is $\mathcal{O}(M_\tau^3 N_\nu^3 N_t^3)$, also considerably higher than that of SSRnet. Therefore, the proposed SSRnet has a lower complexity order compared with OMP and SBL algorithms. Note that if the computational layer which computes the LS solution is used as an alternative to the Learned ISP block, the computational complexity associated with the one-time pseudo-inverse calculation will become $\mathcal{O}(K_\tau^3 N_\nu^3 N_t^3)$. The computational cost associated with the fixed 490,700 parameters of SSRnet is of order $\mathcal{O}(M_\tau N_\nu N_t)$. The number of parameters is determined by the kernel size, number of filters, and number of input channels of the Conv2D and Conv3D layers, and does not depend on the spatial dimensions of the input tensor, whereas the computational load during training and inference depends on the spatial dimensions as well.

TABLE I
COMPUTATIONAL COMPLEXITY OF DIFFERENT SCHEMES

Scheme	Complexity order
PositionNet+	$\mathcal{C}_{\text{PN}} = \mathcal{O}(M_\tau^2 N_\nu^2 N_t)$
SSRnet	$\mathcal{C}_{\text{SSR}} = N_L \mathcal{O}(M_\tau^2 N_\nu^2 N_t) + \mathcal{C}_{\text{PN}}$
OMP	$\mathcal{C}_{\text{OMP}} = \mathcal{O}(K_\tau^3 M_\tau N_\nu^4 N_t^3) + \mathcal{O}(K_\tau M_\tau^2 N_\nu^3 N_t^2)$
SBL	$\mathcal{C}_{\text{SBL}} = N_{\text{iter}} \mathcal{O}(M_\tau^3 N_\nu^3 N_t^3)$

The high complexity of OMP-based methods is attributed to iterative computation of the pseudo-inverse, which increases in proportion to the cardinality of support of the desired channel matrix, which in practical massive MIMO systems, typically represents tens to hundreds of iterations. The complexity order of SBL increases drastically with the dimensions of the channel matrix, arising due to the matrix inversion step which requires tens to hundreds of iterations to achieve a convergence. In massive MIMO-OTFS systems of very large dimension, the required computational capability for running the SBL can therefore be prohibitively high. In contrast, SSRnet benefits from PositionNet+ by obtaining all nonzero indices of the fading channel matrix simultaneously, and computing the nonzero values by adopting a layered Learned ISP approach based on matrix-vector multiplication rather than Moore-Penrose pseudo-inverse computation. This lowers SSRnet's computational complexity compared to those of the existing OMP and SBL methods.

D. Implementation Platform

The computing platform used was an Intel Core i7-12700 CPU with NVIDIA RTX A2000 GPU. Experiments were conducted using TensorFlow framework and Keras deep learning library on a PyCharm integrated development environment (IDE). To train the models, a dataset of 10^5 fading channel samples was generated using the 3GPP spatial channel model in MATLAB, 75% of which was used for training and 25% for validation. A separate dataset of 10^3 channel realizations was generated for performance evaluation and comparison to the 3D-SOMP and OMP methods. The training was performed by stochastic gradient descent, the optimizer AdamW was used given its computational efficiency and a small memory requirement, and a scheduled learning rate of $(10^{-3}, 10^{-4}, 10^{-5})$. The training of both PositionNet+ and SSRnet is completed in only few epochs, requiring a few minutes. Simulation codes for the experimental results are available online at the repository [45].

IV. SIMULATION RESULTS

The channel estimation performance of massive MIMO-OTFS system is investigated in the sub-6 GHz frequency band using the 3GPP spatial channel model for urban macrocellular scenarios. A summary of system parameters is presented in Table II. The size of the Doppler dimension N used for the simulations was limited by available computational power. The pilot overhead ratio η is defined as the ratio of resources

TABLE II
SYSTEM PARAMETERS FOR MASSIVE MIMO-OTFS SIMULATION

System Parameters	Values
Carrier frequency f_c	2.5 GHz
Mobile user's maximum velocity	$30 \sim 580$ km/h
OTFS frame size (M, N)	$(600, 10)$
Subcarrier spacing Δf	15 KHz
Cyclic prefix duration	$16.67 \mu\text{s}$
FFT size of base OFDM modem	1024
Duplexing method	FDD
Number of transmit antennas N_t	$1 \sim 64$
Number of mobile user antennas N_r	1
Number of dominant channel paths N_p	4, 6, 8
Number of sub-paths per dominant path N_s	20

allocated for pilot transmission to the total available resources in the delay-Doppler domain, expressed as a percentage. Given a total number of $M_\tau N_\nu$ pilots, we set $N_\nu = N$ due to the limited impact of Doppler dimension sparsity when N is not significantly large.

The accuracy of PositionNet+ in obtaining the support, and NMSE and BER performance of SSRnet are evaluated across various scenarios, including SNR, number of transmit antennas, mobile user velocity and pilot overhead ratio. Performance is compared to conventional impulse-based channel estimation [44], OMP algorithm [14], and 3D-SOMP scheme [17]. The 3D-SOMP method identifies the support of the high-mobility channel tensor along the delay, Doppler and angular dimensions in an iterative consecutive manner for each dominant path, as presented in Algorithm 1 of [17].

Fig. 7 presents the accuracy of PositionNet+, refined by Learned ISP, in correctly finding the support of the sparse channel matrix when $N_t = 16$ and $\eta = 26\%$ versus SNRs of $\{0, 8, 15\}$ dB. The network was trained using $\text{SNR} = \{0, 10, 20, 30\}$ dB and for a random but fixed pilot set. Accuracy is defined as the ratio of correctly detected indices to all nonzero indices. As can be observed, the proposed PositionNet+ is more accurate in finding the support of the desired sparse channel matrix than that of OMP, 3D-SOMP [17] as well as the recently proposed PositionNet [32]. It can be observed that PositionNet+ can detect the support with high accuracy even in the presence of noise. Accuracy performance in the low SNR regime is observed to be close to that of the high SNR regime, revealing **denoising behavior**, which is mainly attributed to effectively extracting the nonlinear mapping between the observation vector and the support of the 3D sparse tensor from the data.

Fig. 8 presents the accuracy of PositionNet+ in obtaining the support of the sparse channel matrix when $N_t = 16$, $\text{SNR} = 8$ dB and $\eta = 26\%$ for different numbers of propagation paths N_p . The presented accuracy is obtained from the neural network which is trained only with $N_p = 6$ paths. Nonetheless as observed, PositionNet+ can still recover the support accurately for other numbers of paths that correspond

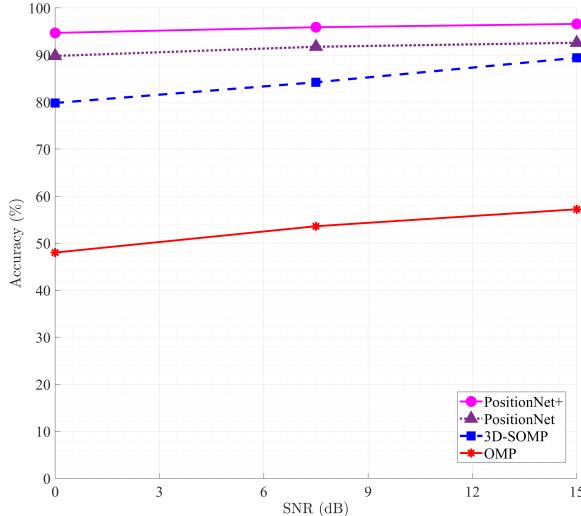


Fig. 7. The accuracy of PositionNet+ in finding the support against SNR for $N_t = 16$, $\eta = 26\%$.

to different sparsity levels. In other words, PositionNet+ is a sparsity-blind algorithm [33], i.e., it can recover the support of a sparse signal without requiring prior knowledge of sparsity level. This also makes PositionNet+ adaptable, flexible and robust to signals with varying degrees of sparsity, unlike traditional OMP-based sparse recovery algorithms, that often rely on knowledge of sparsity level for sparse signal recovery. This is a considerable advantage of PositionNet+ since in many real-world applications, the sparsity level is imprecisely unknown, may vary, and is difficult to estimate.

Fig. 9 illustrates the NMSE performance of the proposed SSRnet versus SNR for $N_t = 32$ with 28% pilot overhead when the mobile user is moving at a velocity of 360 km/h, compared with that of OMP, 3D-SOMP, LS estimation with perfectly known support, and the conventional impulse-based technique, where the base station transmits impulses in the

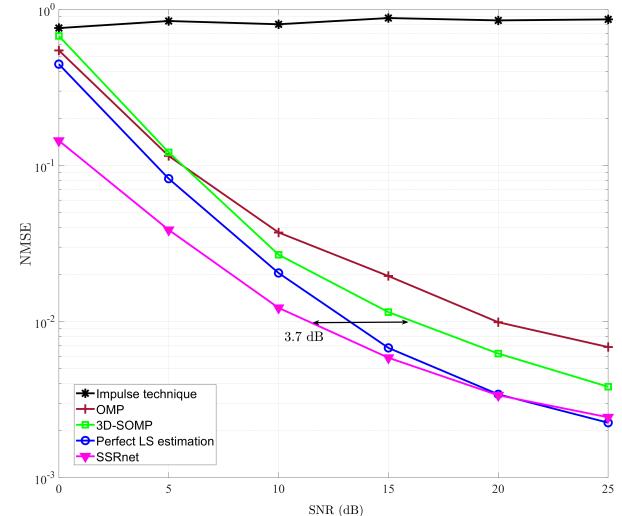


Fig. 9. The NMSE performance against the SNR for $N_t = 32$ and $\eta = 28\%$.

delay-Doppler domain as pilot signals. As can be seen, SSRnet yields 3.7 dB improvement in SNR compared to 3D-SOMP at $\text{NMSE} = 10^{-2}$ and even outperforms LS estimation with perfectly known support, due to its denoising behavior.

The NMSE performance of SSRnet versus number of transmit antennas when $\text{SNR} = 10$ dB, $\eta = 25\%$ overhead and user velocity 420 km/h, is compared with that of OMP, 3D-SOMP, perfect LS estimation, and the impulse-based technique is depicted in Fig. 10. The NMSE performance versus user velocity for $\text{SNR} = 15$ dB and $\eta = 20\%$ overhead is illustrated in Fig. 11 where the network was trained for user velocities of $\nu = \{50, 100, 150\}$ m/s. This also shows some degree of robust performance to training mismatch. As can be seen, SSRnet yields improved performance compared to other methods.

Fig. 12 shows the NMSE performance for different values of pilot overhead ratio η for $N_t = 16$, $\text{SNR} = 12$ dB and

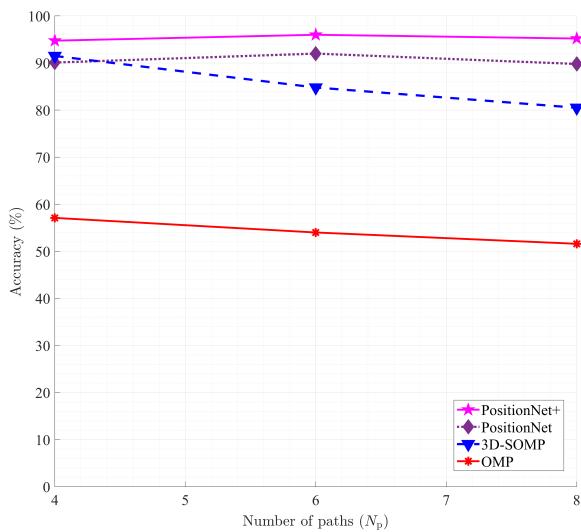


Fig. 8. The accuracy of PositionNet+ in finding the support against number of paths N_p for $\text{SNR} = 8$ dB, $N_t = 16$ and $\eta = 26\%$.

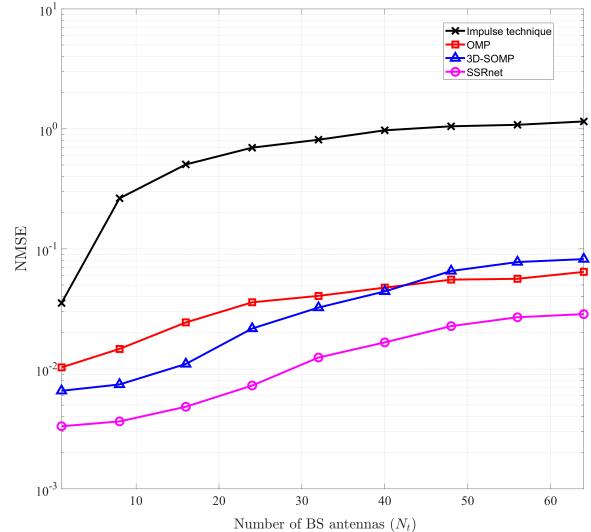


Fig. 10. The NMSE performance against the number of transmit antennas for $\eta = 25\%$ and $\text{SNR} = 10$ dB.

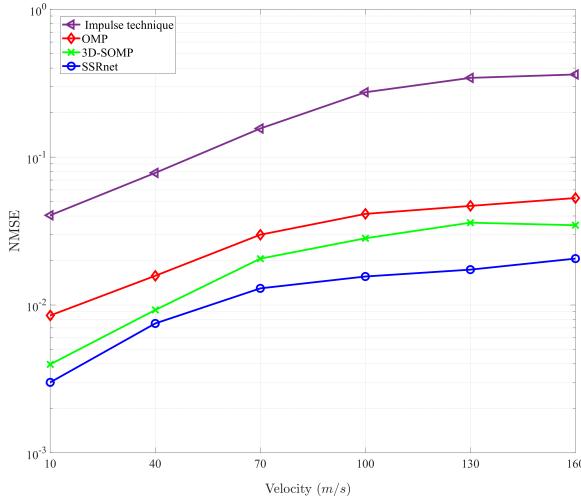


Fig. 11. The NMSE performance versus the mobile user's velocity for $\eta = 20\%$ and $\text{SNR} = 15 \text{ dB}$.

velocity 280 km/h. SSRnet provides a significant reduction of 50% in the required pilot overhead to achieve the same NMSE performance compared to 3D-SOMP in the low-pilot overhead ratio regime. A major disadvantage of 3D-SOMP is its poor performance when available resources for training symbols are limited. However, in practical communication systems, it is desirable that only a small portion of resources be dedicated to pilot transmission and CSI acquisition purposes.

Uncoded BER performance with matched filter detection is depicted in Fig. 13 where binary phase shift keying (BPSK) modulation is employed over the same 3GPP urban macrocellular channel, $\eta = 30\%$ and the user velocity is 380 km/h. The proposed SSRnet is compared with 3D-SOMP, the impulse-based technique, OFDM modulator-demodulator (modem) and perfect-CSI as the benchmark. It is observed that the performance of OTFS with SSRnet is close to that of LS estimation of perfectly obtained support and even surpasses it in the low-SNR regime by benefiting from the denoising

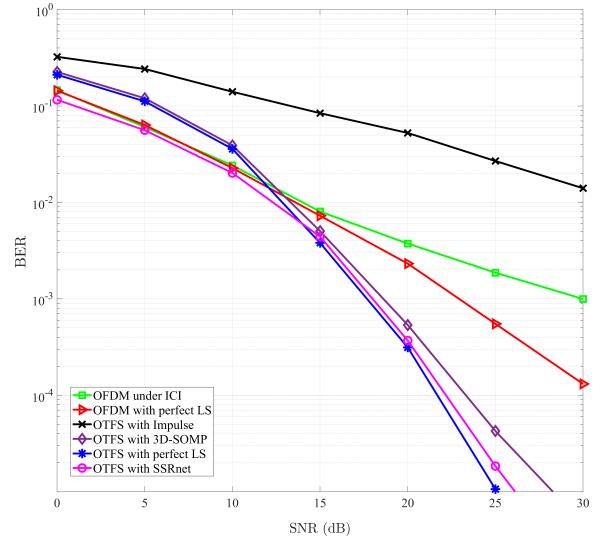


Fig. 13. Uncoded BER performance against SNR. The user velocity is 380 km/h.

behavior originating from the Convolutional layers.

Finally, Fig. 14 depicts the computation time needed to perform the estimation of one 3D fading channel tensor against different numbers of antennas. It is observed that a significant computation time reduction of 76.8% is achieved compared to that of the 3D-SOMP algorithm. This reduction is expected since a large proportion of complexity in conventional methods is spent calculating the Moore-Penrose pseudo-inverse in each iteration to find the support. On the other hand, PositionNet+ obtains all the support at once without any need for the Moore-Penrose pseudo-inverse computation.

Overall, these simulations highlight several advantages of the proposed approach over traditional channel estimation techniques: (i) effectiveness of the neural network in learning the complex relationship between the received signal and the true channel, (ii) improvement in channel estimation accuracy

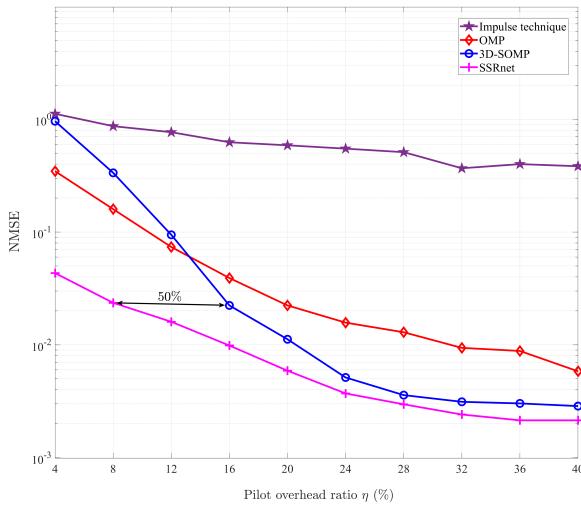


Fig. 12. The NMSE performance against pilot symbols ratio η for $N_t = 16$ and $\text{SNR} = 12 \text{ dB}$.

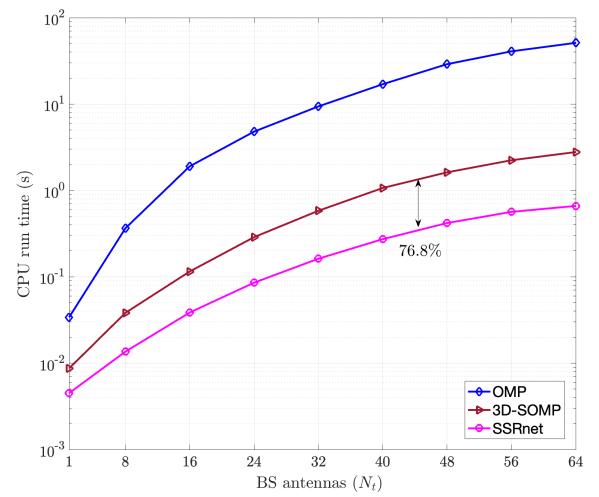


Fig. 14. The computation time for performing the channel estimation.

with low pilot overhead ratio, even in challenging channel conditions, (iii) greater computational efficiency compared to traditional techniques, and (iv) denoising behavior.

V. CONCLUSION

A deep learning-based framework for sparse signal recovery is proposed and applied to the problem of channel estimation of high-mobility massive MIMO-OTFS communication systems in the sub-6 GHz frequency band. PositionNet+ finds the support of the desired sparse matrix without requiring knowledge of the sensing matrix. Next, SSRnet, is able to accurately recover the sparse matrix values. Training of the designed neural networks was performed using a dataset of 10^5 doubly-selective fading channel samples. The proposed framework, experimentally verified by simulation, is shown to significantly improve NMSE performance by 3.7 dB in SNR gain, lower BER, lower pilot symbol overhead by 50%, as well as reduce computation time by 76.8% compared to state-of-the-art algorithms.

In future work, improving the performance of high mobility systems may be realized by (i) training the system based on complex-valued neural network processing by employing complex-valued convolutional layers, and (ii) investigating DL-based dynamic pilot transmission schemes to decrease pilot signaling overhead and more fully capture doubly-dispersive channel characteristics.

REFERENCES

- [1] H. Liu, Y. Liu, M. Yang and Q. Zhang, "On the Characterizations of OTFS Modulation Over Multipath Rapid Fading Channel," *IEEE Trans. Wireless Comm.*, vol. 22, no. 3, pp. 2008-2021, Mar. 2023.
- [2] Z. Yin *et al.*, "Deep CSI Compression for Massive MIMO: A Self-Information Model-Driven Neural Network," *IEEE Trans. Wireless Comm.*, vol. 21, no. 10, pp. 8872-8886, Oct. 2022.
- [3] S. Li *et al.*, "A Novel ISAC Transmission Framework Based on Spatially-Spread Orthogonal Time Frequency Space Modulation," *IEEE J. Selec. Areas Comm.*, vol. 40, no. 6, pp. 1854-1872, Jun. 2022.
- [4] W. Tong *et al.*, "Deep Learning Compressed Sensing-based Beamspace Channel Estimation in mmWave Massive MIMO Systems," *IEEE Wireless Comm. Lett.*, vol. 11, no. 9, pp. 1935-1939, Sep. 2022.
- [5] P. Zhang *et al.*, "Hybrid Beamforming Based on an Unsupervised Deep Learning Network for Downlink Channels With Imperfect CSI," *IEEE Wireless Comm. Lett.*, vol. 11, no. 7, pp. 1543-1547, Jul. 2022.
- [6] C. K. Thomas *et al.*, "Energy Efficient Sparse Bayesian Learning using Learned Approximate Message Passing," 2021 IEEE 22nd Int. Workshop Signal Process. Adv. Wireless Comm. (SPAWC), Lucca, Italy, 2021, pp. 271-275.
- [7] Y. Cui, S. Li and W. Zhang, "Jointly Sparse Signal Recovery and Support Recovery via Deep Learning With Applications in MIMO-Based Grant-Free Random Access," *IEEE J. Selec. Areas Comm.*, vol. 39, no. 3, pp. 788-803, Mar. 2021.
- [8] Y. Liu, S. Zhang, F. Gao, J. Ma and X. Wang, "Uplink-Aided High Mobility Downlink Channel Estimation Over Massive MIMO-OTFS System," *IEEE J. Selec. Areas Comm.*, vol. 38, no. 9, pp. 1994-2009, Sep. 2020.
- [9] H. Lin and J. Yuan, "Orthogonal Delay-Doppler Division Multiplexing Modulation," *IEEE Trans. Wireless Comm.*, vol. 21, no. 12, pp. 11024-11037, Dec. 2022.
- [10] R. Hadani *et al.*, "Orthogonal Time Frequency Space Modulation," 2017 IEEE Wireless Commun. Netw. Conf. (WCNC), San Francisco, USA, Mar. 2017, pp. 1-6.
- [11] R. Hadani and A. Monk, "OTFS: A New Generation of Modulation Addressing the Challenges of 5G," 2018, arXiv:1802.02623.
- [12] H. Qu, G. Liu, M. A. Imran, S. Wen and L. Zhang, "Efficient Channel Equalization and Symbol Detection for MIMO OTFS Systems," *IEEE Trans. Wireless Comm.*, vol. 21, no. 8, pp. 6672-6686, Aug. 2022.
- [13] Y. Shan and F. Wang, "Low-Complexity and Low-Overhead Receiver for OTFS via Large-Scale Antenna Array," *IEEE Trans. Veh. Tech.*, vol. 70, no. 6, pp. 5703-5718, Jun. 2021.
- [14] S. Srivastava, R. K. Singh, A. K. Jagannatham and L. Hanzo, "Delay-Doppler and Angular Domain 4D-Sparse CSI Estimation in OTFS Aided MIMO Systems," *IEEE Trans. Veh. Tech.*, vol. 71, no. 12, pp. 13447-13452, Dec. 2022.
- [15] L. Zhao, J. Yang, Y. Liu and W. Guo, "Block Sparse Bayesian Learning-Based Channel Estimation for MIMO-OTFS Systems," *IEEE Comm. Lett.*, vol. 26, no. 4, pp. 892-896, Apr. 2022.
- [16] D. Shi *et al.*, "Deterministic Pilot Design and Channel Estimation for Downlink Massive MIMO-OTFS Systems in Presence of the Fractional Doppler," *IEEE Trans. Wireless Comm.*, vol. 20, no. 11, pp. 7151-7165, Nov. 2021.
- [17] W. Shen, L. Dai, J. An, P. Fan and R. W. Heath, "Channel Estimation for Orthogonal Time Frequency Space (OTFS) Massive MIMO," *IEEE Trans. Signal Process.*, vol. 67, no. 16, pp. 4204-4217, 15 Aug. 2019.
- [18] R. Zhang *et al.*, "Tensor Decomposition-Based Channel Estimation for Hybrid mmWave Massive MIMO in High-Mobility Scenarios," *IEEE Trans. Comm.*, vol. 70, no. 9, pp. 6325-6340, Sep. 2022.
- [19] R. Zhang *et al.*, "Integrated Sensing and Communication With Massive MIMO: A Unified Tensor Approach for Channel and Target Parameter Estimation," *IEEE Trans. Wireless Comm.*, vol. 23, no. 8, pp. 8571-8587, Aug. 2024.
- [20] B. He *et al.*, "Denoising CNN Based Channel Estimation for Vehicular OTFS Communication System," 25th Int. Conf. Advanced Comm. Tech. (ICACT), Korea, Republic of, 2023, pp. 54-58.
- [21] L. Guo, P. Gu, J. Zou, G. Liu and F. Shu, "DNN-Based Fractional Doppler Channel Estimation for OTFS Modulation," *IEEE Trans. Veh. Tech.*, vol. 72, no. 11, pp. 15062-15067, Nov. 2023.
- [22] Q. Li, Y. Gong, F. Meng, L. Han and Z. Xu, "A Novel Channel Estimation Method based on Deep Neural Network for OTFS System," 2022 15th Int. Cong. Image and Signal Process., BioMedical Eng. and Informatics (CISP-BMEI), China, 2022, pp. 1-6.
- [23] S. Li *et al.*, "Joint Design of Measurement Matrix and Sparse Support Recovery Method via Deep Auto-Encoder," *IEEE Signal Process. Lett.*, vol. 26, no. 12, pp. 1778-1782, Dec. 2019.
- [24] D. Ito, S. Takabe and T. Wadayama, "Trainable ISTA for Sparse Signal Recovery," *IEEE Trans. Signal Process.*, vol. 67, no. 12, pp. 3113-3125, Jun. 2019.
- [25] U. K. S. Shiv, S. Bhashyam, C. R. Srivatsa and C. R. Murthy, "Learning-Based Sparse Recovery for Joint Activity Detection and Channel Estimation in Massive Random Access Systems," *IEEE Wireless Comm. Lett.*, vol. 11, no. 11, pp. 2295-2299, Nov. 2022.
- [26] A. Abdallah *et al.*, "Deep Learning-Based Frequency-Selective Channel Estimation for Hybrid mmWave MIMO Systems," *IEEE Trans. Wireless Comm.*, vol. 21, no. 6, pp. 3804-3821, Jun. 2022.
- [27] A. Abdallah *et al.*, "RIS-Aided mmWave MIMO Channel Estimation Using Deep Learning and Compressive Sensing," *IEEE Trans. Wireless Comm.*, vol. 22, no. 5, pp. 3503-3521, May 2023.
- [28] J. Kang, C. Chun and I. Kim, "Deep Learning Based Channel Estimation for MIMO Systems with Received SNR Feedback," *IEEE Access*, vol. 8, pp. 121162-121181, 2020.
- [29] Z. Liu, M. del Rosario and Z. Ding, "A Markovian Model-Driven Deep Learning Framework for Massive MIMO CSI Feedback," *IEEE Trans. Wireless Comm.*, vol. 21, no. 2, pp. 1214-1228, Feb. 2022.
- [30] W. Ma, C. Qi, Z. Zhang and J. Cheng, "Sparse Channel Estimation and Hybrid Precoding Using Deep Learning for Millimeter Wave Massive MIMO," *IEEE Trans. Comm.*, vol. 68, no. 5, pp. 2838-2849, May 2020.
- [31] S. Srivastava *et al.*, "OTFS Transceiver Design and Sparse Doubly-Selective CSI Estimation in Analog and Hybrid Beamforming Aided mmWave MIMO Systems," *IEEE Trans. Wireless Comm.*, vol. 21, no. 12, pp. 10902-10917, Dec. 2022.
- [32] M. Payami and S. D. Blostein, "Deep Learning-based Channel Estimation for Massive MIMO-OTFS Communication Systems," 23rd Wireless Telecomm. Symp. (WTS), Oakland, CA, USA, Apr. 2024.
- [33] N. Y. Yu, "Multiuser Activity and Data Detection via Sparsity-Blind Greedy Recovery for Uplink Grant-Free NOMA," *IEEE Comm. Lett.*, vol. 23, no. 11, pp. 2082-2085, Nov. 2019.
- [34] J. Salo *et al.*, "MATLAB implementation of the 3GPP Spatial Channel Model (3GPP TR 25.996)," Jan. 2005. [Online]. Available: <http://www.tkk.fi/Units/Radio/scm/>
- [35] P. Raviteja, K. T. Phan, Y. Hong and E. Viterbo, "Interference Cancellation and Iterative Detection for Orthogonal Time Frequency Space Modulation," *IEEE Trans. Wireless Comm.*, vol. 17, no. 10, pp. 6501-6515, Oct. 2018.

- [36] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "QuaDRiGa: A 3-D Multi-cell Channel Model with Time Evolution for Enabling Virtual Field Trials," *IEEE Trans. Antennas Propag.*, vol. 62, no. 6, pp. 3242–3256, Jun. 2014.
- [37] F. Hlawatsch and G. Matz, *Wireless Communications Over Rapidly Time-varying Channels*. New York, NY, USA: Academic, 2011.
- [38] A. Vaswani *et al.*, "Attention is All You Need," *31st Conf. Advances Neural Info. Process. Sys. (NIPS) 2017*, USA, 2017.
- [39] W. Dai and O. Milenkovic, "Subspace Pursuit for Compressive Sensing Signal Reconstruction," *IEEE Trans. Info. Theory*, vol. 55, no. 5, pp. 2230-2249, May 2009.
- [40] M. Sadeghi and M. Babaie-Zadeh, "Iterative Sparsification-Projection: Fast and Robust Sparse Signal Approximation," *IEEE Trans. Signal Process.*, vol. 64, no. 21, pp. 5536–5548, Nov. 2016.
- [41] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, "A Fast Approach for Overcomplete Sparse Decomposition based on Smoothed ℓ^0 Norm," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 289–301, Jan. 2009.
- [42] F. Ghayem *et al.*, "Sparse Signal Recovery Using Iterative Proximal Projection," *IEEE Trans. Signal Process.*, vol. 66, no. 4, pp. 879-894, Feb. 2018.
- [43] E. C. Marques *et al.*, "Nonlinear Functions in Learned Iterative Shrinkage-Thresholding Algorithm for Sparse Signal Recovery," *2019 IEEE Int. Workshop Signal Process. Sys. (SiPS)*, China, 2019, pp. 324-329.
- [44] A. Monk, R. Hadani, M. Tsatsanis, and S. Rakib, "OTFS - Orthogonal Time Frequency Space: A Novel Modulation Technique meeting 5G High Mobility and Massive MIMO Challenges," 2016, arXiv:1608.02993.
- [45] M. Payami, "Sparse Signal Recovery Neural Network," <https://github.com/MostafaPayami/SSRnet>, 2024.



Mostafa Payami received the M.Sc. degree in electrical engineering from Iran University of Science and Technology, Tehran, Iran in 2012. He is currently working toward a Ph.D. degree at Queen's University, Kingston, ON, Canada, where he conducts research at the Advanced Wireless Communications Laboratory. His research focuses on deep learning-based approaches for channel estimation and hybrid beamforming for high-mobility massive MIMO systems. He has over seven years of professional experience in wireless communications engineering. His research interests include mmWave massive MIMO systems, 6G mobile networks, and AI-driven wireless communications.



Steven D. Blostein (S'82-M'88-SM'96) received the B.S. degree in electrical engineering from Cornell University, Ithaca, NY, USA, in 1983, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois, Urbana-Champaign, IL, USA, in 1985 and 1988, respectively. Since 1988, he has been on the Faculty in the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON, Canada, and currently holds the position of Professor. From 2004 to 2009, he was Department Head. He has also been a consultant to industry and government in the areas of image compression, target tracking, radar imaging, and wireless communications.

His current interests include wireless communications systems, including detection and estimation, signal processing, energy efficiency, MIMO, dynamic access, and dense deployments. He was Chair of IEEE Kingston Section, Chair of the Biennial Symposium on Communications, Publications Chair IEEE ICASSP, Associate Editor of IEEE Transactions on Image Processing and IEEE Transactions on Wireless Communications and served on numerous Technical Program Committees for the IEEE Communications Society conferences. He is a registered Professional Engineer in Ontario.