# ON AUTOREGRESSIVE AND NEURAL METHODS FOR MASSIVE-MIMO CHANNEL DE-NOISING

Dmitry Artemasov[1], Alexander Blagodarnyi[2], Alexander Sherstobitov[3], Vladimir Lyashev[2]

[1]Center for Next Generation Wireless and IoT (NGW), Skolkovo Institute of Science and Technology, Moscow, Russia,
[2]The Department of Multimedia Technology and Telecommunication, Moscow Institute of Physics and Technology, Moscow, Russia,
[3]Moscow Radio Transmission Technology Laboratory, Moscow, Russia

NOTE: Corresponding author: Dmitry Artemasov, d.artemasov@skoltech.ru

***Abstract*** – *In modern wireless communication systems, the Multiple-Input Multiple-Output (MIMO) technology allows to greatly increase power efficiency, the serving area, and the overall cell throughput through the use of the antenna array beamforming. Nevertheless, the MIMO systems require accurate channel state knowledge to apply correct precoding. In 5G Time Division Duplex (TDD) systems, the Channel State Information (CSI) is obtained via Sounding Reference Signals (SRS) transmitted by the User Equipment (UE). UEs have limited power capabilities and thus cannot achieve high Uplink (UL) Signal-to-Noise Ratio (SNR) on gNodeB (gNB) in large bandwidth. There are multiple techniques that can be applied to improve the accuracy of Channel Estimation (CE) in noisy conditions. In this paper, we describe a classical method, namely the Vector Autoregression (VAR) with adaptive model order estimation, as well as a modern Deep Neural Network (DNN) approach for the massive-MIMO channel estimation de-noising problem. The developed methods and signal pre and post-processing steps are described, followed by their performance evaluation in a set of realistic simulations. The designed algorithms provide results outperforming the baseline spatio-temporal windowing approaches by $\approx 2dB$ effective Downlink (DL) Signal-to-Interference-plus-Noise Ratio (SINR) metric in single and multi-user MIMO scenarios. Extensive simulation results demonstrate the robustness of the developed methods to the dynamic channel conditions.*

**Keywords** – 5G, channel estimation, CSI, deep learning, de-noising, DFT, massive MIMO, neural networks, sparse channel representation, SRS, vector autoregression

## 1. INTRODUCTION

The quality of radio Channel Estimation (CE) plays a crucial role in the performance of Multiple-Input Multiple-Output (MIMO) communication systems. The gain of MIMO systems lies in the spatial domain multiplexing technique. Using the radio channel state knowledge, gNodeB (gNB) can construct the proper precoding matrix to steer the radiation pattern of the antenna array in the desired direction.

The degradation of the communication system's performance can be caused by the imprecise estimation of the channel state. This issue occurs when the pilot signals are received at a low Signal-to-Noise Ratio (SNR). The User Equipment (UE) radio transmission power is limited by battery, power amplifier efficiency, and maximum power radiation constraints. Thus, the UE cannot transmit sounding signals on large bandwidths while retaining high SNR levels on the gNB side. In order to overcome such constraints advanced processing techniques of channel estimation and noise suppression are applied at the gNB side.

Nowadays, a wide variety of approaches are used for channel estimation. Among them, Machine Learning (ML) approaches are becoming popular. In this paper, we investigate the performance of vector autoregression and the convolutional Deep Neural Network (DNN) model for a massive MIMO channel estimation de-noising task.

Autoregressive (AR) models are an important class of statistical methods for describing time-varying processes. In its generalized version, Vector Autoregression (VAR) is widely used to capture multivariate relationships rather than univariate ones. This class of models is often used for time-series forecasting. There are a number of works describing the AR models application for the radio channel state prediction [1, 2, 3, 4]. These models can be applied for multivariate time-series de-noising as well. Recently, VAR models have been used in application to seismic data [5], audio recordings [6, 7], and medical data de-noising [8, 9].

Recently, the deep learning approach has become popular for channel de-noising in MIMO systems [10, 11], due to the possibility of achieving higher performance than classical algorithms. However, deep learning models tend to have high computational complexity. Those facts make learning models valuable for further research. In this paper, we consider one of the classical deep-learning models used for channel de-noising called DnCNN [12] modified to exploit signal processing in its sparse representation.

In this paper, we exploit one of the classical approaches for MIMO systems to perform signal processing in the spatio-temporal domains, which helps to utilize the sparse nature of the received signal and improve system performance [13, 14, 15, 16]. These methods often use a Fourier basis for obtaining the channel representation in

terms of beams and delay taps, followed by thresholding or other adaptive filtering methods.

The goal of this paper is to investigate the applicability of the classical and modern machine learning methods for low-SNR UL MIMO channel estimation (de-noising). The paper is organized as follows. In Section 2, the system model and the signal preprocessing pipeline are described. Section 3 introduces the VAR model and describes its application for beam domain Channel State Information (CSI) de-noising. Section 4 is devoted to the DNN model architecture description. The complexity of the developed algorithms is evaluated in Section 5. Numerical results and performance comparisons are provided in Section 6. Achieved results and possible improvements are discussed in Section 7.

## 2. SYSTEM MODEL

### 2.1 MIMO channel

We consider the Uplink (UL) Sounding Reference Signals (SRS) transmission for channel state information estimation in 5G New Radio (NR) Time Division Duplex (TDD). Let $N_{UE}$ be the total number of UEs in the system. Each UE has $N_u$ antenna elements and gNB has $N_{BS}$ antenna elements.

We assume that the UEs are transmitting sounding reference signals in a comb-$n$ manner, which implies that each $n$ subcarrier is used for pilot transmission[1]. Let $N_f$ be the total number of active subcarriers. Then, the channel snapshot of user $k$ can be represented as a tensor

$$\mathcal{H}^{(k)} \in \mathbb{C}^{N_{BS} \times N_u \times N_f}, \qquad (1)$$

which is sliced by the frequency dimension

$$\mathcal{H}^{(k)} = [\mathbf{H}_f^{(k)}], \forall f \in [1, N_f], \qquad (2)$$

or by the UE antenna dimension

$$\mathcal{H}^{(k)} = [\mathbf{H}_u^{(k)}], \forall u \in [1, N_u]. \qquad (3)$$

TDD systems share the same frequency band for the uplink and downlink transmissions. Thus, the radio channel is assumed to be reciprocal, implying that $\mathbf{H}_{DL} = \mathbf{H}_{UL}^H$, where $(\cdot)^H$ denotes the Hermitian transpose.

We assume that the received by gNB signal is corrupted by noise and, thus, the channel estimation is defined as follows

$$\widehat{\mathbf{H}}_f^{(k)} = \mathbf{H}_f^{(k)} + \mathbf{n}_f^{(k)}, \forall f, k. \qquad (4)$$

Here $\widehat{\mathbf{H}}_f$ is the CSI estimated by SRS pilots, $\mathbf{H}_f$ is the actual CSI, and $\mathbf{n}_f$ is the Additive White Gaussian Noise (AWGN) with complex normal distribution $\mathcal{CN}(0, \sigma^2)$.

---

[1] 5G UL pilots structure is described in section 6 of [17].

In what follows, we assume that the de-noising operator $\mathcal{D}(\cdot)$ is applied to the channel estimation

$$\bar{\mathbf{H}}_f^{(k)} = \mathcal{D}(\widehat{\mathbf{H}}_f^{(k)}). \qquad (5)$$

The objective function of the MIMO system can be expressed in terms of the system capacity [18]

$$\begin{aligned} \underset{\mathbf{W}_f}{\text{maximize}} \quad & \sum_{k \in N_{UEs}} \mathbb{E}_f \Big[ \log \big( 1 + \text{SINR}_f^{(k)}(\mathbf{W}_f) \big) \Big], \\ \text{subject to} \quad & \mathbf{W}_f = \mathcal{S}([\bar{\mathbf{H}}_f^{(1)}, ..., \bar{\mathbf{H}}_f^{(k)}]), \forall f, \end{aligned} \qquad (6)$$

where $\mathbf{W}_f$ is a Multiuser (MU) Downlink (DL) precoding matrix, obtained via the function $\mathcal{S}(\cdot)$ from the de-noised CSIs $\bar{\mathbf{H}}_f$. Function $\mathcal{S}(\cdot)$ implies zero forcing-based MU precoder design. Signal-to-Interference-plus-Noise ratio (SINR) of $k$th UE is defined as

$$\text{SINR}_f^{(k)} = \frac{|\mathbf{H}_f^{(k)H} \mathbf{W}_f^{(k)}|^2}{\sum\limits_{\substack{l=1 \\ l \neq k}}^{N_{UE}} |\mathbf{H}_f^{(k)H} \mathbf{W}_f^{(k)}|^2 + \sigma^2}. \qquad (7)$$

In order to keep the following description compact, we omit the superscript $(k)$.

### 2.2 Sparse channel representation

It is possible to apply several transformations to represent the CSI slices $\mathbf{H}_u \in \mathbb{C}^{N_{BS} \times N_f}, \forall u$ *(3)* in a sparse form.

The antenna domain of the channel matrix describes the superposition of the planar waves received by each antenna element. The straightforward de-noising in the antenna domain is not optimal. The noise and the target signal are spread across all antenna elements (see Fig.1a).

The antenna domain can be transformed to the beam (angular) domain by projecting CSI slices $\widehat{\mathbf{H}}_u$ on the 2D-DFT beamspace $\mathbf{F}_b$. The projection matrix $\mathbf{F}_b$ depends on the structure of the antenna array and can be constructed as follows:

$$\mathbf{F}_b = \mathbf{F}_{N_R} \otimes \mathbf{F}_{N_C}, \qquad (8)$$

where $\mathbf{F}_{N_R}$ and $\mathbf{F}_{N_C}$ are the DFT matrices of size equal to the number of rows $N_R$ and columns $N_C$ of the antenna array, respectively. The symbol $\otimes$ denotes the Kronecker product.

The channel angular representation is sparse since it characterizes signal Angles of Arrival (AoA). While the noise component is spread across all angular directions, the target signal AoA is defined by the location of the UEs and by the scatters in the environment of signal propagation.
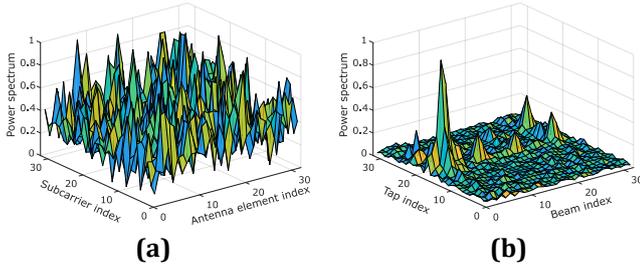
**Fig. 1** – Channel estimation matrix representation: (a) antenna-frequency domain, (b) sparse beam-delay domain.

The frequency domain of the channel matrix can be transformed into the delay domain by applying the inverse FFT. Let us denote by $\mathbf{F}_d \in \mathbb{C}^{N_f \times N_f}$ the DFT matrix. Then the CSI slices of *(3)* in the beam-delay domains are defined as

$$\widetilde{\mathbf{H}}_u = \mathcal{F}(\widehat{\mathbf{H}}_u) = \mathbf{F}_b \widehat{\mathbf{H}}_u \mathbf{F}_d^H, \forall u. \qquad (9)$$

For the following description, we denote the taps set by $\tau \in [1, N_\tau]$ and the beams set by $b \in [1, N_b]$. Applying *(9)* to CSI we expect the CE matrix $\widetilde{\mathbf{H}}_u$ to be sparse in the beam and delay domains (see Fig. 1b).

The beam-delay domains CSI representation describes the MIMO channel in terms of the amplitude pulses, which are received from different angles and defined by the locations of the scatters in the environment of signal propagation. The target signal is localized by the AoA and by the propagation delays, while the noise remains distributed uniformly.

The wide frequency band of the SRS signal transmission allows a high resolution to be achieved in the delay domain. The power distribution of channel taps in the delay domain is a Rayleigh-Rice random process and may be defined by the Power-Delay Profile (PDP). Channel PDP taps distribution has good localization properties inside the symbol (usually signal taps occupy a small part at the start of the symbol). Once the PDP is estimated, the rectangular window can be applied for the PDP de-noising. The angular spectrum of the CSI is a random process as well, and it describes the channel properties in the space domain (beams). The resolution of the CSI in the angular domain is defined by the aperture of the antenna array and the pattern design of each antenna port. The angular spectrum has an irregular distribution and, thus, it cannot be localized as simply as the PDP (see Fig. 1b) [19].

## 3. VECTOR AUTOREGRESSION FOR CSI DE-NOISING

### 3.1 VAR model

Autoregressive models are associated with time series processing. The channel tensor *(1)* is defined for some time snapshot $t$. In what follows by $t$ we denote the current time sample and by $p$ the model order. Let us consider the VAR time series applied on the beam dimension of the channel vector $\tilde{\mathbf{h}}_{\tau,u}(t)$

$$\bar{\mathbf{h}}'_{\tau,u}(t) = \mathbf{a}_0 + \sum_{i=1}^{p} \mathbf{A}_i \tilde{\mathbf{h}}_{\tau,u}(t-i) + \epsilon_t, \forall \tau, u, \qquad (10)$$

where $\tilde{\mathbf{h}}_{\tau,u}(t) \in \mathbb{C}^{N_b}$, $\tilde{\mathbf{h}}_{\tau,u}(t) \subset \widetilde{\mathbf{H}}_u(t)$ is a vector of CSI, which corresponds to a tap $\tau$ at time sample $t$ received from UE antenna $u$ *(2)*, $\mathbf{a}_0 \in \mathbb{C}^{N_b}$ is the fixed bias vector, $\mathbf{A}_i \in \mathbb{C}^{N_b \times N_b}, \forall i \in [1, p]$ are the AR coefficients matrices and $\epsilon_t \in \mathbb{C}^{N_b}$ - error term [20].

The vector autoregressive process can be estimated by the Least Squares (LS) method. For the following description by matrix $\mathbf{B} = [\tilde{\mathbf{h}}_{\tau,u}(t-1),..,\tilde{\mathbf{h}}_{\tau,u}(t-p)] \in \mathbb{C}^{N_b \times p} \forall \tau, u$ we denote the stacked beam-domain CSI vectors corresponding to the time snapshots $T \in [t-p, t-1]$. In that notation, the VAR-based filter is constructed as

$$\bar{\mathbf{h}}'_{\tau,u}(t) = \underbrace{\mathbf{B}(\mathbf{B}^H\mathbf{B} + \gamma\mathbf{I})^{-1}\mathbf{B}^H}_{\text{VAR filter}} \tilde{\mathbf{h}}_{\tau,u}(t), \forall \tau, u, \qquad (11)$$

where $\bar{\mathbf{h}}'_{\tau}(t)$ is the de-noised CSI vector in beam domain, $\gamma$ is a regularization coefficient and $\mathbf{I}$ is an identity matrix.

VAR de-noising is applied on the beam dimension of CSI, i.e., CSI is processed in a tap-by-tap manner. After the VAR de-noising, the CSI matrix is transformed back to the antenna-frequency domains.

$$\bar{\mathbf{H}}_u = \mathcal{F}^{-1}(\bar{\mathbf{H}}'_u) \quad \forall u. \qquad (12)$$

The proposed de-noising approach is summarized in Algorithm 1 and schematically depicted in Fig. 2.

---

**Algorithm 1** Spatio-temporal VAR de-noising

---

**Input:** $\widehat{\mathbf{H}}_u(t) \in \mathbb{C}^{N_{BS} \times N_f}$ - CSI matrices, $u \in [1, N_u]$
**Output:** $\bar{\mathbf{H}}_u(t) \in \mathbb{C}^{N_{BS} \times N_f}$ - de-noised CSI matrices
    *Transform to beam-delay domains (9)*:
1:  $\widetilde{\mathbf{H}}_u(t) = \mathcal{F}(\widehat{\mathbf{H}}_u(t))$, $\forall u \in [1, N_u]$
    *Tap-by-tap VAR de-noising*:
2:  **for** $u = 1$ to $N_u$ **do**
3:    **for** $\tau = 1$ to $N_\tau$ **do**
4:      $\mathbf{B} = [\tilde{\mathbf{h}}_{\tau,u}(t-p),...,\tilde{\mathbf{h}}_{\tau,u}(t-1)]$
5:      $\bar{\mathbf{h}}'_{\tau,u}(t) = \mathbf{B}(\mathbf{B}^H\mathbf{B} + \gamma\mathbf{I})^{-1}\mathbf{B}^H\tilde{\mathbf{h}}_{\tau,u}(t)$
6:    **end for**
7:  **end for**
8:  $\bar{\mathbf{H}}'_u(t) = [\bar{\mathbf{h}}'_1(t),..,\bar{\mathbf{h}}'_\tau(t)]$, $\forall u$
    *Transform to antenna-frequency domain*:
9:  $\bar{\mathbf{H}}_u(t) = \mathcal{F}^{-1}(\bar{\mathbf{H}}'_u(t))$
10: **return** $\bar{\mathbf{H}}_u(t)$

---
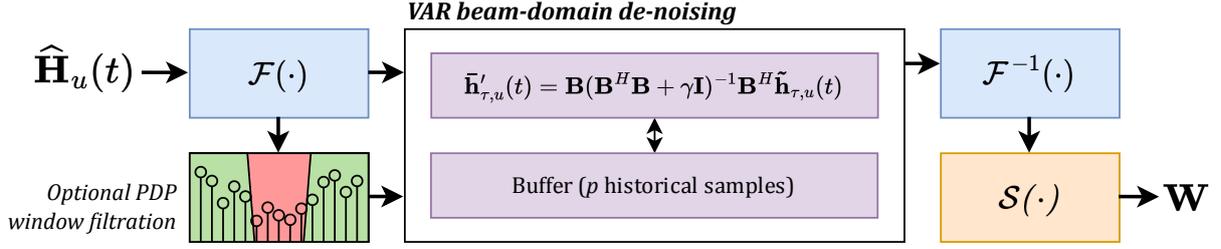
**Fig. 2** – Proposed VAR-based de-noising system architecture.

## 3.2 Adaptive VAR model order selection

The optimal VAR model order selection depends on the specific radio channel conditions. If the channel is static (speed of UEs close to zero), the optimal noise suppression approach would imply averaging over the infinite number of CSI samples. In the high-dynamic channel conditions (mobile UEs), the CSI history used for the model construction shall not exceed the channel coherence time.

The coherence time $T_c$ of a wireless channel is defined (in an order of magnitude sense) as the interval over which the channel state information changes significantly. For the Clarke's model, it is defined as

$$T_c = \frac{1}{4D_s},\qquad (13)$$

where $D_s$ is the Doppler spread [18].

This is a somewhat imprecise relation, since in the real case scenario the Doppler spread is not guaranteed to match the U-shape form and the Doppler shift harmonics with the largest frequency may belong to paths that are too weak to make a difference. Depending on the channel multipath properties, the factor in the denominator of *(13)* can differ from 4. In this paper, we consider the adaptive scheme, which implies the fine-tuning of the coherence time $T_c$ factor, for the following denoted by $\alpha$.

$$T_c = \frac{1}{\alpha f_{D_m}},\qquad (14)$$

where the $\alpha$ coefficient is supposed to be approximated during the gNB configuration stage and $f_{D_m}$ Doppler spectrum harmonic with maximum power.

Nevertheless, the important thing is to recognize that the major effect in determining the time coherence $T_c$ is the Doppler spread and that the relationship is reciprocal: the larger the Doppler spread is, the smaller the time coherence and vice versa.

For the adaptive VAR-based MIMO channel de-noising, we consider the estimation of the model order, which would utilize the historical samples

$$p = \left\lfloor \frac{T_c}{T_p} \right\rfloor,\qquad (15)$$

where $T_p$ denotes the pilot transmission period.

For a known UE velocity and angular spectrum, the Doppler frequency can be calculated as

$$f_D = \frac{\mathbf{r}^T \mathbf{v}}{\lambda_0},\qquad (16)$$

where $\mathbf{r} = \begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix}$ is the spherical unit vector with azimuth arrival angle $\phi$ and elevation arrival angle $\theta$, $\mathbf{v} = v\begin{bmatrix} \sin\theta_v\cos\phi_v & \sin\theta_v\sin\phi_v & \cos\theta_v \end{bmatrix}^T$ is the user velocity vector with speed $v$, travel azimuth angle $\phi_v$ and elevation angle $\theta_v$, $\lambda_0$ is the carrier wavelength.

Since the true Doppler spread is not available to the gNB, we consider Doppler spectrum estimation from the historical measurements. Let us define the user $k$ de-noised channel estimation tensor time series as follows

$$\bar{\mathcal{H}}(T_D) = [\bar{\mathcal{H}}(t)] \in \mathbb{C}^{N_{BS}\times N_u \times N_f \times N_D},\qquad (17)$$

where $\bar{\mathcal{H}}(T_D)$ is the set of channel tensors estimation $\bar{\mathcal{H}}$ at the time moments $T_D \in [t-1, t-N_D]$. In the alternative form, $\bar{\mathcal{H}}(T_D)$ consist of the channel vectors $\bar{\mathbf{h}}_D \in \mathbb{C}^{N_D}$

$$\bar{\mathcal{H}}(T_D) = [\bar{\mathbf{h}}_D]_{b,u,f}, \quad \forall b, u, f$$
$$\forall b \in [1, N_{BS}], \quad \forall u \in [1, N_u], \quad \forall f \in [1, N_f],\qquad (18)$$

where $T_D$ denotes the historical channel estimation time indices.

To calculate the Doppler spectrum components of the channel tensor *(17)*, the discrete Fourier transform $\Phi(\cdot)$ is applied over the time domain samples $T_D$

$$[\check{\mathbf{h}}_D]_{b,u,f} = \Phi([\bar{\mathbf{h}}_D]_{b,u,f}), \forall b, u, f,\qquad (19)$$

and then its power spectrum is averaged over the frequency and antenna domains

$$\mathbf{p}_{\check{\mathbf{h}}_D} = \mathbb{E}_{b,u,f}[\check{\mathbf{h}}_D \check{\mathbf{h}}_D^H]_{jj}, \forall j \in [1, N_D],\qquad (20)$$

where $\mathbb{E}(\cdot)$ is an expectation over indices $b, u, f$.
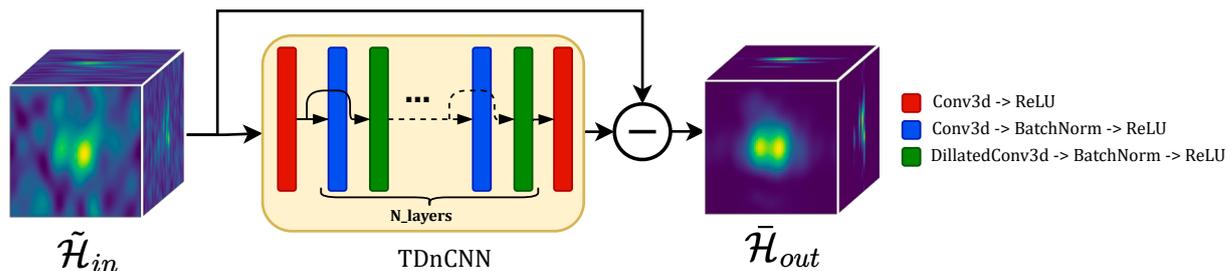
**Fig. 3** – Tensor DnCNN architecture scheme in beam-delay domain.

In this paper, we follow the simple idea of the VAR model order estimation based on the frequency of the Doppler harmonic with maximum power, which is obtained as

$$f_{D_m} = \arg\max_{\tau_D \in [1, N_D]} \mathbf{p}_{\breve{\mathbf{h}}_D},$$ *(21)*

where $f_{D_m}$ is the estimated Doppler frequency harmonic with maximum power.

Recalling equation *(15)*, the optimal VAR de-noising model order can be approximated as

$$p = \left\lfloor \frac{T_c}{T_p} \right\rfloor = \left\lfloor \frac{1}{\alpha f_{D_m} T_p} \right\rfloor$$ *(22)*

## 4. DEEP LEARNING CSI DE-NOISING

### 4.1 Model description

A De-noising Convolutional Neural Network (DnCNN) is a popular deep learning model used for image de-noising, which doesn't use prior information about image content. This is achieved via a learning model to predict Gaussian noise rather than image content. The classical DnCNN is a residual network, where each block contains a convolution layer padded to have the same output size as input, batch normalization, and activation function. For blind de-noising, it contains 20 such residual blocks [12]. Recently, DnCNN and its variations showed promising performance for MIMO channel de-noising. Such variations include complex-valued DnCNN adapted to perform on complex numbers [10] [21], combining DnCNN with compressed sensing algorithms, which utilizes the channel sparsity property [22] and ensembles with other de-noising architectures or feature engineering techniques [23].

In our research, as a baseline a complex-valued DnCNN [10] was used, which was additionally adapted to

process input data as a channel tensor $\tilde{\mathcal{H}}$ in the spatio-temporal domain. Such a tensor model is called TDnCNN [24] and it utilizes 3D convolutions instead of 2D ones. The scheme of the model is presented in Fig. 3. The motivation for using this model includes the fact that a fully convolution neural network can be considered as a set of Finite Impulse Response (FIR) filters connected

via a non-linear function. To utilize the channel sparsity in the angular and delay domains we apply the reshaping operator $\mathcal{R}(\cdot)$:

$$\mathcal{R} : \tilde{\mathcal{H}} \to \tilde{\mathcal{H}}_{in},$$ *(23)*

where $\tilde{\mathcal{H}} \in \mathbb{C}^{N_b \times N_u \times N_\tau}$ and $\tilde{\mathcal{H}}_{in} \in \mathbb{C}^{N_u \times N_R \times N_C \times N_\tau}$. Here $N_R$ and $N_C$ denote the beam dimension shape ($N_R \cdot N_C = N_b$), which describe the azimuth and elevation directions respectively. The role of the TDnCNN model is to predict the tensor of noise, which is subtracted from the noisy channel estimation to obtain the de-noised model output $\bar{\mathcal{H}}_{out}$. The resulting estimation $\bar{\mathcal{H}}$ with an initial shape can be obtained via the inverse reshaping operator $\mathcal{R}^{-1}(\cdot)$ applied to $\bar{\mathcal{H}}_{out}$:

$$\mathcal{R}^{-1} : \bar{\mathcal{H}}_{out} \to \bar{\mathcal{H}}$$ *(24)*

where $\bar{\mathcal{H}}_{out} \in \mathbb{C}^{N_u \times N_R \times N_C \times N_\tau}$ and $\bar{\mathcal{H}} \in \mathbb{C}^{N_b \times N_u \times N_\tau}$.

To increase the receptive field of the model and estimate the beams more accurately, the trainable convolutions with dilation equal to three were applied after each block of classical convolution. The kernel size of each convolution was set to $3 \times 3 \times 3$, and the total number of de-noising layers to 15.

### 4.2 Training setup

For the training procedure, the Mean Squared Error (MSE) loss function was utilized.

$$\mathcal{L}_{MSE} = \frac{1}{N_b N_u N_f} \sqrt{\sum_{b,u,f} ([\mathbf{n}]_{b,u,f} - [\bar{\mathbf{n}}]_{b,u,f})^2},$$ *(25)*

where $\mathbf{n}$ and $\bar{\mathbf{n}}$ are actual and estimated noise samples respectively.

Although MSE is good for training models via backpropagation, it was observed that such metrics as MSE and its variations cannot be physically interpreted to define and compare the quality of the model implemented in the wireless system. Therefore MSE was used for model training and the SINR metric *(7)* for its evaluation. The effective SINR has a strong physical meaning and it can be well interpreted.

**Table 1** – Proposed algorithms complexity evaluation

| Step \ Algorithm | VAR | DnCNN |
|---|---|---|
| Antenna-to-beam transform | $\mathcal{O}\Big(N_u N_f \cdot 2^{\lceil \log_2 N_{BS} \rceil} \cdot (\lceil \log_2 N_{BS} \rceil - 1)\Big)$ | |
| Frequency-to-PDP transform | $\mathcal{O}\Big(N_u N_b \cdot 2^{\lceil \log_2 N_f \rceil} \lceil \log_2 N_f \rceil\Big)$ | |
| CE de-noising | $\mathcal{O}\Big(N_u N_\tau [N_b p + 2 N_b p^2 + N_b^2 p + p^3]\Big)$ | $\mathcal{O}\Big(N_u N_b N_\tau k^3 N_l\Big)$ |
| PDP-to-frequency transform | $\mathcal{O}\Big(N_u N_b \cdot 2^{\lceil \log_2 N_f \rceil} \lceil \log_2 N_f \rceil\Big)$ | |
| Beam-to-antenna transform | $\mathcal{O}\Big(N_u N_f \cdot 2^{\lceil \log_2 N_{BS} \rceil} \cdot (\lceil \log_2 N_{BS} \rceil - 1)\Big)$ | |

As for the dataset, it was fully generated with a QuaDRiGa channel model [25]. The configuration of the channel scenario is discussed in Section 6. The generated dataset contains a set of UE channels with different random seeds, which determine the position of reflectors in the environment. This dataset was split into three subsets: training (100 channels) used only for model weight adjustment, validation (20 channels) utilized for model selection, and test (20 channels) used for final model evaluation. The model was trained offline for 300 epochs with $10^{-3}$ learning rate of Adam optimizer and a batch size of three (channel realizations). The choice of offline training mode is motivated by the fact that the noise tensor was used as a target value in the loss function, which is practically impossible to measure in online mode. Moreover, the model architecture is assumed to be independent of the signal scenario, since it predicts the noise, which is subtracted from initial noisy measurements. For better convergence, the learning rate was decreased with a decay factor of 10 after every 100 iterations, when no target metric improvement could be observed.

## 5. COMPLEXITY EVALUATION

In this section, we evaluate the complexity of the proposed algorithms. For both VAR and DNN-based approach processing starts with the channel transformation into the sparse form *(9)*. This procedure implies: (i) transforming the antenna domain of the channel matrix to the beam domain, and (ii) transforming the frequency domain of the channel matrix to PDP.

We consider an antenna array with two polarizations, which are processed independently. Thus, the complexity of antenna-to-beam 2D-DFFT is

$$\mathcal{O}\Big(N_u N_f \cdot 2 \cdot 2^{\lceil \log_2 \frac{N_{BS}}{2} \rceil} \cdot \log_2 2^{\lceil \log_2 \frac{N_{BS}}{2} \rceil}\Big) = \\ \mathcal{O}\Big(N_u N_f \cdot 2^{\lceil \log_2 N_{BS} \rceil} \cdot (\lceil \log_2 N_{BS} \rceil - 1)\Big). \quad (26)$$

To represent the channel in the PDP domain IDFFT is applied

$$\mathcal{O}\Big(N_u N_b \cdot 2^{\lceil \log_2 N_f \rceil} \lceil \log_2 N_f \rceil\Big). \quad (27)$$

The complexity of the VAR de-noising algorithm applied for a single UE is

$$\mathcal{O}\Big(N_u N_\tau [N_b p + 2 N_b p^2 + N_b^2 p + p^3]\Big). \quad (28)$$

The complexity of the TDnCNN forward pass for single UE is

$$\mathcal{O}\Big(N_u N_b N_\tau k^3 N_l\Big), \quad (29)$$

where $k$ is a convolution kernel size, and $N_l$ is the number of layers. An additional complexity reduction can be done via beam selection before passing the neural network de-noiser.

After the de-noising channel is transformed back to the antenna-frequency domains. The complexity of these steps is equivalent to the forward transforms *(26)* – *(27)*. The complexity of the proposed MIMO de-noising pipelines is summarized in Table 1.

## 6. NUMERICAL RESULTS

We compare the proposed VAR and DNN de-noising algorithms with two baseline approaches. The combinations

**Table 2** – Summary of algorithms combination

| | PDP de-noising | Beam de-noising |
|---|---|---|
| Baseline | Windowing | ✗ |
| Baseline[Ext] | Windowing | Windowing |
| VAR 1 | ✗ | VAR |
| VAR 2 | Windowing | VAR |



- selected taps
- suppressed taps
- selected beams
- suppressed beams
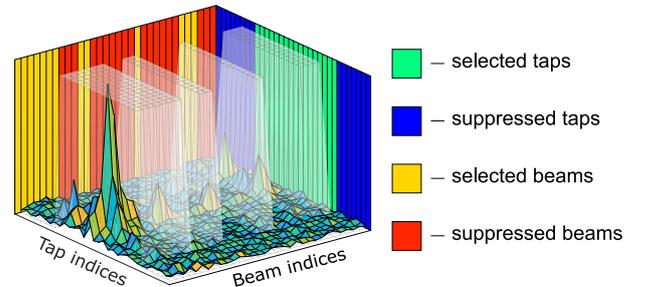
**Fig. 4** – Delay and beam domains window filtration. The colored box faces illustrate rectangular filters for delay and beam domains. The white boxes define the positions of the selected taps/beams in the sparse channel matrix.

SU-MIMO, $\nu = 5$ km/h
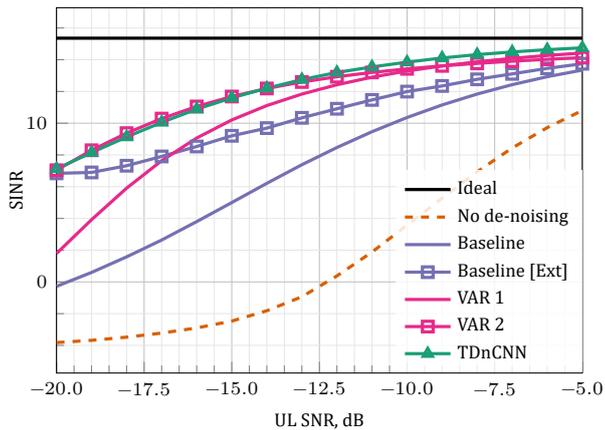


**Fig. 5** – CE de-noising performance: SU-MIMO 5 km/h.

SU-MIMO, $\nu = 15$ km/h
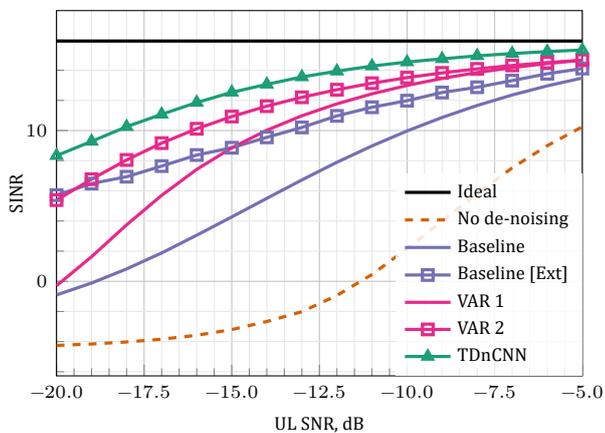


**Fig. 6** – CE de-noising performance: SU-MIMO 15 km/h.

MU-MIMO, $\nu = 5$ km/h
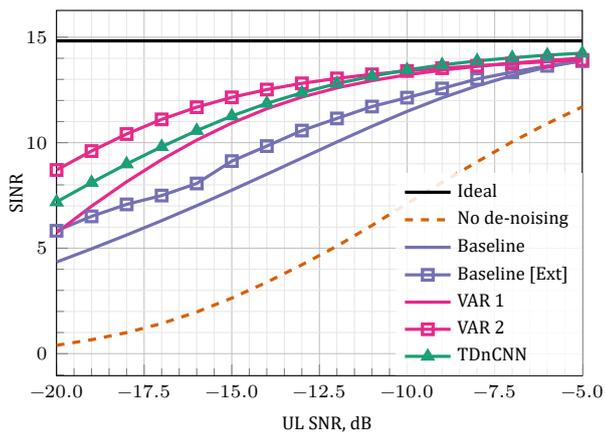


**Fig. 7** – CE de-noising performance: MU-MIMO 5 km/h.

**Table 3** – Main simulation parameters

| Parameter type | Parameter value | |
|---|---|---|
| MIMO mode | SU-MIMO 1 UE, 4 layers | MU-MIMO 4 UE, 1 layer |
| Channel model type | QuaDRiGa Berlin NLoS UMa [25] | |
| Antenna configuration [BS/UE] | 64/4 | |
| Bandwidth, MHz | 20 | |
| SRS comb type | 2 | |
| Subcarrier spacing. kHz | 30 | |
| SRS period, ms | 5 | |
| Central frequency, GHz | 3.5 | |
| UL SNR range, dB | -20 ÷ -5 | |
| VAR/MA order | 4 | |
| gNB height, m | 25 | |
| Distance to UEs, m | 50 | |
| UEs speed, km/h | 5 ÷ 15 | |
| Coherence time, ms | 15.4 ÷ 5.1 | |
| Number of averaging snapshots | 100 | |

MU-MIMO, $\nu = 15$ km/h
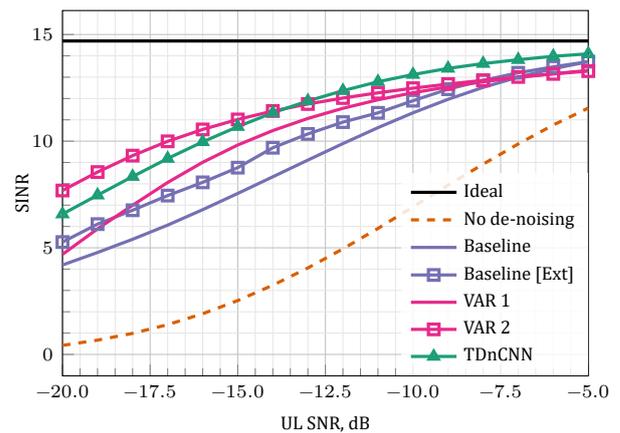


**Fig. 8** – CE de-noising performance: MU-MIMO 15 km/h.

tration (selecting taps/beams from the 'green'/'yellow' ranges Fig. 4).

For the PDP taps filtration the predefined rectangular window, optimized to cover $\approx 95\%$ of the delay spread energy in the noiseless channel estimation, is utilized. The beam domain windowing procedure implies two-step processing: (i) pre-SINR estimation from signal/noisy taps power ratio separately on each of the beams; (ii) selection of beams with pre-SINR metric exceeding the fine-tuned threshold.

Two MIMO scenarios were considered: SU-MIMO (1 UE, 4 layers transmission) and MU-MIMO (4 UE, 1 layer transmission). For the algorithms' performance evaluation, the MIMO channel was generated in a 3GPP TR 38.901 [26] compliant QuaDRiGa generator [25]. The "Berlin UMa NLoS" channel scenario was selected. Its parame-

of the considered de-noising algorithms are summarized in Table 2. By *Baseline* we denote the CSI de-noising by the PDP tap window filtration applied for all the beams (selecting taps from the 'green' range Fig.4), by *Baseline[Ext]* we denote the joint PDP and beam domain window fil-

ters were estimated from the field test measurements described in [27]. The "Berlin UMa NLoS" has a rich multipath and a wider angle spread in comparison to other channel models, so we consider it as the realistic and complicated scenario for the algorithm's performance evaluation. Simulations were performed for two UE velocities: 5 and 15 km/h. Recalling eq. *(14)* and considering $\alpha = 4$, 5 km/h velocity corresponds to $T_c = 15.4$ ms, which is far beyond the pilot transmission period $T_p = 5$ ms. 15 km/h UE velocity corresponds to $T_c = 5.1$ ms, which approaches the coherence interval edge. So, such a scenario allows us to analyze the performance of the considered algorithms in the critical regime. The average SINR *(7)* was used as the performance metric. A detailed description of the scenario parameters is summarized in Table 3.

Several outcomes can be highlighted from the simulation results:
The preliminary PDP rectangular window filtration gives a performance boost for the de-noising algorithms (*Baseline, Baseline[Ext], VAR 2*).

The *VAR 2* algorithm provides the best de-noising performance among the discussed baselines. In the SU-MIMO scenario, the *VAR 1* algorithm outperforms baselines in the SNR region higher than $-15$ dB and provides a performance lower than *Baseline[Ext]* in the UL SNR region below $-15$ dB. The *VAR 1* worse de-noising performance can be explained by the insufficiency of beam filtration in the low SNR region. Apart from beam selection, the *Baseline[Ext]* algorithm performs the PDP de-noising, while *VAR 1* implies the VAR beam domain filtration only.

The DNN-based de-noising provides a performance comparable to the *VAR 2* algorithm in the SU-MIMO 5km/h scenario and best results in SU-MIMO 15 km/h, which is achieved by better channel adaptability properties of the neural network. In the 15 km/h scenario the VAR model struggles to fit the outdated CSI samples to the new fast-varying channel measurements, while the non-linear nature of TDnCNN allows it to inherit the channel properties in a better way.

In the MU-MIMO regime, the VAR de-noising approach outperforms the DNN-based method. Due to the non-linearity of TDnCNN, the residual noise in the DNN output tensor $\bar{\mathcal{H}}'$ is colored, which leads to the increase of the interlayer interference, induced by the zero-forcing procedure.

It is worth noticing that the performance of the *Baseline[Ext]* algorithm degrades fast with decreasing UL SNR in comparison to the VAR models. This is caused by the non-robust threshold-based beam filtration approach. The optimal beam selection threshold may vary with changing wireless channel conditions. The VAR de-noising can be interpreted as an adaptive filtering with memory since it does not require beam suppression threshold tuning.

The *Baseline[Ext]* algorithm with hard beams suppression performs better in the SU-MIMO scenarios (figures 5–6) rather than in MU-MIMO (figures 7–8). This is caused by the diversity of AoA of multiple users. When hard beam selection is performed, the low-power beams are completely suppressed, which results in interference leakage during the zero-forcing procedure in the construction of the MU-MIMO precoder *(6)*.

# 7. CONCLUSION

The de-noising approaches described in the paper operate in the sparse channel representation and allow us to raise the signal component above the noise level by utilizing the signal resolution in the spatio-temporal domain. The paper provided comparison and performance evaluation results of VAR and DNN-based massive MIMO de-noising methods.

VAR de-noising was applied on the beam domain for the following reasons:

- The channel power delay profile has an almost static shape. Thus, as it is estimated once, the low-complexity PDP filtration can be performed by the Hadamard product with the rectangular window.

- The angle representation is sparse, but unlike the PDP, it heavily depends on the UE location. It is not possible to estimate and reuse the single window filter for the beam domain.

- The threshold-based beam de-noising approach is not robust, since the optimal threshold value depends on many parameters: number of propagation paths, channel angular spread, SNR level, number of UEs, etc.

When the VAR-based de-noising filter is able to detect the beam dynamic correlation in the subsequent time samples, it fits the time series coefficients, hence keeping the strong components. In the opposite case, when the beam dynamic in $p$ subsequent samples is uncorrelated, the VAR filter cannot fit the time series. The model weights are distributed uniformly among the observation samples, which results in the elimination of noisy components.

The TDnCNN model provides results outperforming the set of baseline algorithms and achieves the best results among all considered in this paper on the edge of the time coherence interval.

On one side, TDnCNN has a relatively large complexity to be implemented in the hardware of communication devices; on the other side it outperforms VAR in some cases, which makes it a prospective from a research point of view and motivates us to continue researching deep learning models, analyze the source of provided benefits and find ways to reduce the complexity.

Also, we note, that in recent years a direction of time series processing neural network models has been developing rapidly [28]. We highlight the application of time-series state-of-the-art NN architectures for MIMO CE denoising as the potential research area.

## ACKNOWLEDGEMENT

## REFERENCES

[1] Darina Jarinová. "On autoregressive model order for long-range prediction of fast fading wireless channel". In: *Telecommunication Systems* 52.3 (Mar. 2013), pp. 1533–1539. ISSN: 1572-9451. DOI: 10.1007/s11235-011-9520-6. URL: https://doi.org/10.1007/s11235-011-9520-6.

[2] Chi Wu, Xinping Yi, Yiming Zhu, Wenjin Wang, Li You, and Xiqi Gao. "Channel Prediction in High-Mobility Massive MIMO: From Spatio-Temporal Autoregression to Deep Learning". In: *IEEE Journal on Selected Areas in Communications* 39.7 (2021), pp. 1915–1930. DOI: 10.1109/JSAC.2021.3078503.

[3] Weihang Liu, Ziyu Chen, and Xiang Gao. "Massive MIMO Channel Prediction in Real Propagation Environments Using Tensor Decomposition and Autoregressive Models". In: *2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. 2022, pp. 849–855. DOI: 10.1109/PIMRC54779.2022.9977664.

[4] Yadan Zheng, Mingke Dong, Ye Jin, and Jianjun Wu. "CQI Prediction for Shadow Fading in LTE-compatible GEO Mobile Satellite Communications System". In: *Proceedings of the 2012 2nd International Conference on Computer and Information Application (ICCIA 2012)*. Atlantis Press, 2014/05, pp. 623–626. ISBN: 978-94-91216-41-1. DOI: https://doi.org/10.2991/iccia.2012.150. URL: https://doi.org/10.2991/iccia.2012.150.

[5] Mostafa Naghizadeh and Mauricio Sacchi. "Vector AR filters: Extending f-x random noise attenuation to the multicomponent case". In: *SEG Technical Program Expanded Abstracts 2011*. SEG, 2012, pp. 3617–3621. DOI: 10.1190/1.3627952. eprint: https://library.seg.org/doi/pdf/10.1190/1.3627952.

[6] Florian Heese, Richard Steinbiss, Peter Jax, and Peter Vary. "Noise Reduction in the Time Domain Using ARMA Filtering". In: *Proceedings of the 12th ITG Symposium on Speech Communication, Paderborn, Germany, October 5-7, 2016*. IEEE, 2016, pp. 1–5. URL: https://ieeexplore.ieee.org/document/7776204/.

[7] Niedźwiecki Maciej and Ciołek Marcin. "Sparse vector autoregressive modeling of audio signals and its application to the elimination of impulsive disturbances". In: *IFAC-PapersOnLine* 48.28 (2015). 17th IFAC Symposium on System Identification SYSID 2015, pp. 1208–1213. ISSN: 2405-8963. DOI: https://doi.org/10.1016/j.ifacol.2015.12.296. URL: https://doi.org/10.1016/j.ifacol.2015.12.296.

[8] Thomas Schanze. "Denoising of signals and noise extraction by sparse autoregression". In: *AUTOMED* 1.1 (2020). Proceedings on Automation in Medical Engineering. DOI: https://doi.org/10.18416/AUTOMED.2020. URL: https://journals.infinite-science.de/index.php/automed/article/view/346.

[9] Sijia Chen, Zhizeng Luo, and Tong Hua. "Research on AR-AKF Model Denoising of the EMG Signal". In: *Computational and Mathematical Methods in Medicine* 2021 (Nov. 2021), p. 9409560. ISSN: 1748-670X. DOI: 10.1155/2021/9409560. URL: https://doi.org/10.1155/2021/9409560.

[10] Shicong Liu, Zhen Gao, Jun Zhang, Marco Di Renzo, and Mohamed-Slim Alouini. "Deep Denoising Neural Network Assisted Compressive Channel Estimation for mmWave Intelligent Reflecting Surfaces". In: *IEEE Transactions on Vehicular Technology* (Jan. 2020). DOI: 10.1109/TVT.2020.3005402. URL: https://hal.science/hal-03020379.

[11] Hongyuan Ye, Feifei Gao, Jing Qian, Hao Wang, and Geoffrey Ye Li. *Deep Learning based Denoise Network for CSI Feedback in FDD Massive MIMO Systems*. 2020. arXiv: 2004.07576 [eess.SP].

[12] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising". In: *IEEE Transactions on Image Processing* 26.7 (July 2017), pp. 3142–3155. DOI: 10.1109/tip.2017.2662206. URL: https://doi.org/10.1109%2Ftip.2017.2662206.

[13] Ali Osman Kalayci and Gokhan M. Guvensen. "An Efficient Beam and Channel Acquisition via Sparsity Map and Joint Angle-Delay Power Profile Estimation for Wideband Massive MIMO Systems". In: *CoRR* abs/1910.05815 (2019). arXiv: 1910.05815. URL: http://arxiv.org/abs/1910.05815.

[14] An-An Lu, Yan Chen, and Xiqi Gao. *2D Beam Domain Statistical CSI Estimation for Massive MIMO Uplink*. 2022. DOI: 10.48550/ARXIV.2207.04695. URL: https://arxiv.org/abs/2207.04695.

[15] Xin Xiong, Xiaodong Wang, Xiqi Gao, and Xiaohu You. "Beam-Domain Channel Estimation for FDD Massive MIMO Systems With Optimal Thresholds". In: *IEEE Transactions on Wireless Communications* 16.7 (2017), pp. 4669–4682. DOI: 10.1109/TWC.2017.2701371.

[16] Alexander Osinsky, Andrey Ivanov, and Dmitry Yarotsky. "Spatial Denoising for Sparse Channel Estimation in Coherent Massive MIMO". In: *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*. 2021, pp. 1–5. DOI: 10.1109/VTC2021-Fall52928.2021.9625153.

[17] 3GPP. *5G NR Physical channels and modulation*. Technical Specification (TS) 38.211. Version 16.10.0. 3rd Generation Partnership Project (3GPP), 2022. URL: https://www.3gpp.org/ftp/Specs/archive/38%5C_series/38.211/.

[18] David Tse and Pramod Viswanath. *Fundamentals of Wireless Communication*. Wiley series in telecommunications. Cambridge University Press, 2005. ISBN: 9780521845274. URL: https://books.google.com/books?id=66XBb5tZX6EC.

[19] Kan Zheng, Long Zhao, Jie Mei, Bin Shao, Wei Xiang, and Lajos Hanzo. "Survey of Large-Scale MIMO Systems". In: *IEEE Communications Surveys & Tutorials* 17.3 (2015), pp. 1738–1760. DOI: 10.1109/COMST.2015.2425294.

[20] Helmut Lütkepohl. *New Introduction to Multiple Time Series Analysis*. Springer Books 978-3-540-27752-1. Springer, Feb. 2005. DOI: 10.1007/978-3-540-27752-1.

[21] Tianyi Zeng, Junyao Li, Mengshi Hu, Shuai Hou, and Qi Zhang. "Toward Higher Performance for Channel Estimation With Complex DnCNN". In: *IEEE Communications Letters* 24.1 (2020), pp. 198–201. DOI: 10.1109/LCOMM.2019.2953704.

[22] Hengtao He, Rui Wang, Weijie Jin, Shi Jin, Chao-Kai Wen, and Geoffrey Ye Li. *Beamspace Channel Estimation for Wideband Millimeter-Wave MIMO: A Model-Driven Unsupervised Learning Approach*. 2021. arXiv: 2006.16628 [eess.SP].

[23] Wenhan Shen, Zhijin Qin, and Arumugam Nallanathan. "Deep Learning for Super-Resolution Channel Estimation in Reconfigurable Intelligent Surface Aided Systems". In: *IEEE Transactions on Communications* 71.3 (2023), pp. 1491–1503. DOI: 10.1109/TCOMM.2023.3239621.

[24] Alexander Blagodarnyi, Roman Bychkov, Stanislav Krikunov, and Andrey Ivanov. "Tensor-Assisted CNN to Estimate Channel in Massive MIMO". In: *2022 IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*. 2022, pp. 20–24. DOI: 10.1109/SIBIRCON56155.2022.10016907.

[25] Stephan Jaeckel, Leszek Raschkowski, Kai Börner, and Lars Thiele. "QuaDRiGa: A 3-D Multi-Cell Channel Model With Time Evolution for Enabling Virtual Field Trials". In: *IEEE Transactions on Antennas and Propagation* 62.6 (2014), pp. 3242–3256. DOI: 10.1109/TAP.2014.2310220.

[26] 3GPP. *5G Study on channel model for frequencies from 0.5 to 100 GHz*. Technical Report (TR) 38.901. Version 17.0.0. 3rd Generation Partnership Project (3GPP), 2022. URL: https://www.3gpp.org/ftp/Specs/archive/38_series/38.901/.

[27] Stephan Jaeckel. "Quasi-deterministic channel modeling and experimental validation in cooperative and massive MIMO deployment topologies". en. PhD thesis. Ilmenau, Aug. 2017. URL: http://uri.gbv.de/document/gvk:ppn:896596648.

[28] Bryan Lim and Stefan Zohren. "Time-series forecasting with deep learning: a survey". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* (2021).

## AUTHORS

**Dmitry Artemasov** received his B.Sc. degree in telecommunications from Ural Federal University and his M.Sc. degree in computer science from Skolkovo Institute of Science and Technology (Skoltech) in 2021 and 2023 respectively. Currently, he is obtaining a PhD degree in Skoltech. Dmitry's research interests include MIMO systems, coding theory, information theory, deep learning, IoT, and digital signal processing.

**Alexander Blagodarnyi** received a B.S. degree in radio-physics from Southern Federal University, Rostov-on-Don, Russia, in 2020 and an M.S. degree in information science and technology from Skolkovo Institute of Science and Technology, Moscow, Russia, in 2022. He is currently pursuing a Ph.D. degree in information science and technology at Moscow Institute of Physics and Technology, Moscow, Russia.

**Alexander Sherstobitov** received an M.Sc. degree in radio engineering from the South-Russian State University of Economics and Service, Shakhty, in 2004, and a Ph.D. degree in radio engineering from Southern Federal University, Taganrog, Russia, in 2008. He is currently a principal engineer with the RTT Laboratory, Huawei Technologies, Moscow, Russia. His current research interests include wireless algorithms design for MIMO systems, wireless channel precoding and equalization, signal processing, and computational efficiency.

**Vladimir Lyashev** received bachelor's and M.E. degrees in electrical engineering from the Taganrog State University of Radio Engineering, Taganrog, Russia, in 2002 and 2004, respectively, and a Ph.D. degree in numerical mathematics and mathematical modeling from Southern Federal University, Taganrog, Russia, in 2007. He took part in the Virtual Test Bed (VTB) Project as a model and system developer. He has also participated in developing ICS and DCS systems. He was a visiting researcher with the Fraunhofer Institute for Algorithms and Scientific Computing, Sankt Augustin, Germany, from 2003 to 2004. In 2011, he had joined the Huawei Research Center, Moscow, Russia, and involved in algorithm design for LTE advanced receiver. Later, in 2014, he lead the Laboratory of Radio Transmission Technologies, Moscow Research Center, Moscow. His main research interests include computing limited wireless communication systems in intelligent electromagnetic environment.