

Received 13 June 2022, accepted 28 June 2022, date of publication 4 July 2022, date of current version 11 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3188111

The logo consists of a series of vertical bars of varying heights on the left, followed by the word "SURVEY" in a bold, blue, sans-serif font inside a rounded rectangular border.

A Survey on Deep Learning Based Channel Estimation in Doubly Dispersive Environments

ABDUL KARIM GIZZINI¹, (Member, IEEE), AND MARWA CHAFII^{2,3}, (Member, IEEE)

¹ETIS, UMR8051, ENSEA, CNRS, CY Cergy Paris Université, 95000 Cergy, France

²Engineering Division, New York University (NYU) Abu Dhabi, Abu Dhabi, United Arab Emirates

³NYU WIRELESS, NYU Tandon School of Engineering, Brooklyn, NY 11201, USA

Corresponding author: Abdul Karim Gizzini (abdulkarim.gizzini@ensea.fr)

This work was supported by the CY INEX for Project through the ASIA Chair of Excellence under Grant PIA/ANR-16-IDEX-0008.

ABSTRACT Wireless communications systems are impacted by multi-path fading and Doppler shift in dynamic environments, where the channel becomes doubly-dispersive and its estimation becomes an arduous task. Only a few pilots are used for channel estimation in conventional approaches to preserve high data rate transmission. Consequently, such estimators experience a significant performance degradation in high mobility scenarios. Recently, deep learning has been employed for doubly-dispersive channel estimation due to its low-complexity, robustness, and good generalization ability. Against this backdrop, the current paper presents a comprehensive survey on channel estimation techniques based on deep learning by deeply investigating different methods. The study also provides extensive experimental simulations followed by a computational complexity analysis. After considering different parameters such as modulation order, mobility, frame length, and deep learning architecture, the performance of the studied estimators is evaluated in several mobility scenarios. In addition, the source codes are made available online in order to make the results reproducible.

INDEX TERMS Artificial neural networks, channel estimation, convolutional neural networks, deep learning, Doppler effect, feedforward neural networks, frequency-selective channels, long short term memory, time-varying channels.

I. INTRODUCTION

With the commercialization of fifth generation networks globally, research into sixth generation (6G) networks has been initiated to address the demands for high data rates and low latency mobile applications, including unmanned aerial vehicles [1], high-speed railway [2], and vehicular communications [3]. Mobile wireless communications systems offer the freedom to move around without being disconnected from the network. However, the mobility feature is ridden with several challenges that have a severely adverse impact on the communication reliability, such as fast and frequent handovers [4], carrier frequency offset [5], inter-carrier interference [6], high penetration loss [7], and fast time-varying wireless channel [8].

The associate editor coordinating the review of this manuscript and approving it for publication was Gerard-Andre Capolino.

In wireless environment, transmitted signals are known to propagate via a multitude of paths, each entailing a different attenuation and delay in addition to the Doppler shift effect stemming from the motion of network nodes along with the surrounding environment. As a result, the wireless channel becomes frequency-selective and time-varying. Given that a precisely estimated channel response influences the follow-up equalization, demodulation, and decoding operations at the receiver, the accuracy of the channel estimation influences the system performance. Therefore, ensuring communication reliability via accurate channel estimation in such environments is highly important.

In the extant literature, a vast body of work has been carried out to address the problem of doubly-dispersive channels. While some works have focused on investigating the waveform design [8]–[11], we are interested in this paper in the channel estimation task. In general, channel estimators

can be classified into two main categories: (i) symbol-by-symbol (SBS) channel estimators: the channel is estimated for each received symbol separately using only the previous and current received pilots [12]–[14] (ii) frame-by-frame (FBF) channel estimators: the previous, existing, as well as future pilots are employed in the channel estimation for each received symbol [15]. It is possible to achieve a higher channel estimation accuracy by utilizing FBF estimators, since the channel estimation of each symbol benefits from the combined knowledge of all allocated pilots within the frame. However, the conventional estimators' performance mainly relies on the allocated reference training pilots within the transmitted frames. The majority of standards allocate a few pilots to maintain a good transmission data rate. Therefore, these pilots are insufficient for accurately tracking the doubly-dispersive channel, because they are not spaced closely enough to capture the variation of the channel in the frequency domain. Consequently, conventional estimators are primarily based on the demapped data subcarriers, besides pilot subcarriers to update the channel estimate for each received symbol. This procedure called data-pilot aided (DPA) channel estimation is regarded as unreliable because the demapping error gets enlarged from one symbol to another, which leads to another additional error in the estimation process, especially in highly dynamic time-varying channels. Moreover, other conventional estimators like the linear minimum mean square error (LMMSE) [16] estimator rely on many assumptions that limit their performance in highly dynamic time-varying channels. Moreover, linear conventional estimators are impractical solutions in real case scenarios as they rely on statistical models and require high implementation complexity, in addition, they lack robustness in highly dynamic environments. Therefore, investigating estimators with a good trade-off complexity vs. performance is a crucial need for improving the channel estimation accuracy while preserving good data rate as well as maintaining affordable computational complexity.

As a prevailing approach to AI, deep learning (DL) is an efficient method to analyze data by identifying patterns and learning underlying structures, denoting an effective approach to problems faced in various scientific fields. DL algorithms have been integrated into the physical layer of wireless communications systems [17]–[19], including channel estimation [20]–[25]. In turn, this is attributable to the great success in enhancing the overall system performance, particularly when used in addition to conventional estimators, where coarse channel estimation is derived from conventional estimators, following which DL is employed to achieve a fine estimation. Therefore, DL-based channel estimators are capable of significantly enhancing the performance while preserving low computational complexity. In addition, the GPU-based distributed processing allows the DL employment in real-time applications, as a result of which DL can overcome the limitations of traditional channel estimation through robust, low-complexity, and generalized solutions that improve the performance of wireless systems.

Motivated by these advantages, DL algorithms have been integrated in frequency-selective [23]–[25] and doubly-dispersive channel estimation. In this survey, we examine the recently proposed DL-based channel estimation schemes in doubly-dispersive environments, where DL algorithms are utilized in two different manners: (i) feed-forward neural networks (FNNs) with different architectures and configurations are employed on top of the conventional SBS channel estimators [26]–[28]. (ii) convolutional neural networks (CNNs) processing is employed where the estimated channel for the entire frame is modeled as a 2D low-resolution noisy image, whereas CNN-based processing is implemented as super resolution and denoising techniques [29]–[31].

The majority of surveys conducted in the literature [32], [33] lack intensive simulations in the performance evaluation and complexity analysis of the studied channel estimators. Moreover, they do not cover both SBS and FBF based estimators. In addition, [32] compares the performance of different DL architectures used after the least squares (LS) and the LMMSE estimators without considering several conventional channel estimation schemes, whereas [33] provides a general overview of several channel estimators without any performance evaluation. Given this context, to the best of our knowledge, this is the first survey that presents a comprehensive study on the recently proposed DL-based SBS and FBF estimators in doubly-dispersive environments, while presenting intensive simulations evaluating the system performance in different scenarios, providing a detailed complexity analysis, as well as the source codes to reproduce all the presented results. We believe that this survey is a very relevant reference to initiate researches pertaining to the domain of deep learning based channel estimation in doubly dispersive channels. The contributions of this paper can be summarized in the following manner

- Comprehensive study on the recently proposed DL-based channel estimation techniques for doubly-dispersive channels.
- Overview of the DL networks, especially those used in the studied channel estimators, such as FNN, long short-term memory (LSTM), super resolution CNN (SR-CNN), and denoising CNN (DN-CNN).
- Performance analysis of different channel estimation schemes and a fair comparison between them in terms of normalized mean-squared error (NMSE) and bit error rate (BER) for different mobility scenarios and frame length, and modulation order.
- Detailed computational complexity analysis for the studied channel estimators concerning the overall required real-valued operations.
- Simulation source code for various channel estimation schemes to reproduce all the comparison results presented in this paper [34].

The remainder of this paper is organized as follows: Section II elucidates the system model, illustrating signal transmission over a doubly-dispersive channel. Section III provides a brief overview of the main DL networks employed

in this survey. The recently proposed DL-based SBS and FBF channel estimation schemes are thoroughly investigated and discussed in Sections IV and V, respectively. In Section VI, different modulation orders are used to present simulation results, wherein the performance of the studied estimators is examined in terms of BER, and NMSE. Detailed computational complexity analysis is provided in Section VII. Finally, Section VIII concludes this study.

Notations: Throughout the paper, vectors are defined with lowercase bold symbols \mathbf{x} whose k -th element is $\mathbf{x}[k]$. Time and frequency domain vectors are represented by \mathbf{x} and $\tilde{\mathbf{x}}$ respectively. Matrices are written as uppercase bold symbols \mathbf{X} . $\mathbb{E}[\cdot]$ denotes the expectation operator. The trace of a square matrix \mathbf{X} is $\text{trace}[\mathbf{X}]$. The notation \odot and \oslash refer to the element-wise multiplication and division operations, respectively. Finally, the pseudo inverse and conjugate matrices of \mathbf{X} are signified by \mathbf{X}^\dagger and \mathbf{X}^H , respectively.

II. SYSTEM MODEL

Consider a frame comprising I orthogonal frequency division multiplexing (OFDM) symbols. The i -th transmitted frequency-domain OFDM symbol $\tilde{\mathbf{x}}_i[k]$, is denoted by

$$\tilde{\mathbf{x}}_i[k] = \begin{cases} \tilde{\mathbf{x}}_{d_i}[k], & k \in \mathcal{K}_d. \\ \tilde{\mathbf{x}}_{p_i}[k], & k \in \mathcal{K}_p. \end{cases} \quad (1)$$

where $0 \leq k \leq K - 1$. $\tilde{\mathbf{x}}_{d_i}[k]$ and $\tilde{\mathbf{x}}_{p_i}[k]$ represent the modulated data symbols and the predefined pilot symbols allocated at a set of subcarriers denoted \mathcal{K}_d and \mathcal{K}_p , respectively. $\mathbf{x}_i[k]$ is converted to the time domain by applying the inverse discrete Fourier transform (IDFT), such that

$$\mathbf{x}_i[n] = \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} \tilde{\mathbf{x}}_i[k] e^{j2\pi \frac{nk}{K}}. \quad (2)$$

A cyclic prefix (CP) of length larger than the delay spread is added. Therefore, after passing via the doubly-dispersive channel and removing the CP, the received OFDM symbol $\mathbf{y}_i[n]$ can be expressed as follows

$$\begin{aligned} \mathbf{y}_i[n] &= \sum_{l=0}^{L-1} \mathbf{h}_i[l, n] \mathbf{x}_i[n-l] + \mathbf{v}_i[n] \\ &= \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} \tilde{\mathbf{h}}_i[k, n] \tilde{\mathbf{x}}_i[k] e^{j2\pi \frac{nk}{K}} + \mathbf{v}_i[n]. \end{aligned} \quad (3)$$

$\mathbf{h}_i[l, n]$ denotes the delay-time response of the discrete linear time-variant (LTV) channel of L taps at the i -th OFDM symbol, whereas $\tilde{\mathbf{h}}_i[k, n] = \sum_{l=0}^{L-1} \mathbf{h}_i[l, n] e^{-j2\pi \frac{lk}{K}}$ refers to the frequency-time response. Moreover, \mathbf{v}_i signifies the additive white Gaussian noise (AWGN) of variance σ^2 . The i -th received frequency-domain OFDM symbol is derived from (3) via discrete Fourier transform (DFT), and thus

$$\tilde{\mathbf{y}}_i[k] = \frac{1}{K} \sum_{q=0}^{K-1} \tilde{\mathbf{x}}_i[q] \sum_{n=0}^{K-1} \tilde{\mathbf{h}}_i[q, n] e^{-j2\pi \frac{n(k-q)}{K}} + \tilde{\mathbf{v}}_i[k]. \quad (4)$$

It is noteworthy that index k is used in (3) to express the channel delay-time response in terms of the channel frequency-time response. While the change of index into q in (4) is used to express the i -th received symbol in frequency domain. This, in turn, better illustrates the DFT transform. Moreover, $\tilde{\mathbf{h}}_i[q, n]$ refers to time-variant at the scale of the OFDM symbol duration (the index i) and within the symbol itself (the index n). Accordingly,

$$\tilde{\mathbf{h}}_i[q, n] = \sum_{l=0}^{L-1} e^{-j2\pi \frac{lq}{K}} \int_{v=-v_d}^{v=v_d} \bar{h}(l, v) e^{j2\pi v n_i} e^{j2\pi v n} dv, \quad (5)$$

where $\bar{h}(l, v) = \sum_n h[l, n] e^{-j2\pi n v}$ signifies the channel delay-Doppler response, v refers to the normalized Doppler frequency, $n_i = i(K + K_{cp}) + K_{cp}$. And $v_d = \frac{f_d}{F_s}$ represents the maximum Doppler frequency. Let

$$\begin{aligned} \bar{h}_i[l, v] &= \frac{1}{K} \sum_{q=0}^{K-1} \sum_{n=0}^{K-1} \tilde{\mathbf{h}}_i[q, n] e^{-j2\pi \frac{nv}{K}} e^{j2\pi \frac{ql}{K}} \\ &= \int_{v=-v_d}^{v=v_d} \bar{h}(l, v) e^{j2\pi v n_i} \sum_{n=0}^{K-1} e^{-j2\pi (v - \frac{v}{K}) n} dv, \end{aligned} \quad (6)$$

be the discrete delay-Doppler response at the i -th OFDM symbol. For the sake of simplicity, $\bar{h}(l, v)$ is assumed to be uncorrelated in both domains [35], such that $\mathbb{E}[\bar{h}(l, v) \bar{h}^*(l', v')] = S_h(l, v) \delta(l-l') \delta(v-v')$, where $S_h(l, v)$ is the delay-Doppler spectrum [35], and $\delta(x)$ denotes the Dirac delta function. Using (6), we have

$$\begin{aligned} \mathbb{E}[\bar{h}_i[l, v] \bar{h}_i^*[l, v']] &= \int_{v=-v_d}^{v=v_d} S_h(l, v) \sum_{n=0}^{K-1} \sum_{n'=0}^{K-1} e^{-j2\pi v(n-n')} e^{-j2\pi \frac{v'-v}{K} n} dv. \end{aligned} \quad (7)$$

This correlation that is independent of the index i can be approximated as follows

$$\begin{aligned} \mathbb{E}[\bar{h}_i[l, v] \bar{h}_i^*[l, v']] &\approx K^2 \rho[l, v] \delta[v - v'], \\ \text{where } \rho[l, v] &= S_h(l, \frac{v}{N}). \end{aligned} \quad (8)$$

The time selectivity of the channel depends on the mobility. In very low mobility, where $f_d \approx 0$, $\tilde{\mathbf{h}}_i[q, n] = \tilde{\mathbf{h}}[q]$ is constant during the whole frame. For moderate to high mobility, the channel variation within the duration of one OFDM symbol is negligible, and therefore, $\tilde{\mathbf{h}}_i[q, n] = \tilde{\mathbf{h}}_i[q]$. At very high mobility, the channel becomes variant within a single OFDM symbol. In this instance, $\tilde{\mathbf{h}}_i[q, n] = \tilde{\mathbf{h}}_i[q] + \tilde{\boldsymbol{\epsilon}}_i[q, n]$, where

$$\tilde{\mathbf{h}}_i[q] = \frac{1}{K} \sum_{n=0}^{K-1} \tilde{\mathbf{h}}_i[q, n], \quad \text{and } \tilde{\boldsymbol{\epsilon}}_i[q, n] = \tilde{\mathbf{h}}_i[q, n] - \tilde{\mathbf{h}}_i[q]. \quad (9)$$

Replacing this in (4), we get

$$\tilde{\mathbf{y}}_i[k] = \tilde{\mathbf{h}}_i[k] \tilde{\mathbf{x}}_i[k] + \tilde{\boldsymbol{\epsilon}}_{i,d}[k] + \tilde{\mathbf{v}}_i[k], \quad k \in \mathcal{K}_{on}. \quad (10)$$

The term $\tilde{e}_{i,d}[k]$ denotes the Doppler interference given by

$$\begin{aligned} \tilde{e}_{i,d}[k] &= \frac{1}{K} \sum_{q=0}^{K-1} \sum_{\substack{n=0 \\ q \neq k}}^{K-1} \tilde{h}_i[q, n] e^{-j2\pi \frac{n(k-q)}{K}} \tilde{x}_i[q] \\ &= \frac{1}{K} \sum_{\substack{q \in \mathcal{K}_{\text{on}} \\ q \neq k}}^{L-1} \sum_{l=0}^{L-1} \tilde{h}_i[l, k-q] e^{-j2\pi \frac{lq}{K}} \tilde{x}_i[q]. \end{aligned} \quad (11)$$

The Doppler interference destroys the orthogonality of the subcarriers within the received OFDM symbol, leading to a significant degradation in the overall system performance [36]. Assuming that the subcarriers are uncorrelated with power E_q , i.e. $E[\tilde{x}_i[q]\tilde{x}_i^*[q']] = E_q\delta[q-q']$ and using (8) then

$$\begin{aligned} E[\tilde{e}_{i,d}[k]\tilde{e}_{i,d}^*[k']] &= \sum_{l=0}^{L-1} \sum_{\substack{q \in \mathcal{K}_{\text{on}} \\ q \neq k}} E_q \rho[l, k-q] \delta[k-k'] \\ &= \sigma_d^2[k] \delta[k-k']. \end{aligned} \quad (12)$$

Thus, it is assumed that the Doppler interference is uncorrelated. However, the variance $\sigma_d^2[k] = E[|\tilde{e}_{i,d}[k]|^2]$ depends on the subcarrier index. Noting that

$$\tilde{h}_i[k] = \frac{1}{K} \sum_{l=0}^{L-1} \tilde{h}_i[l, 0] e^{-j2\pi \frac{kl}{K}}, \quad (13)$$

the channel gain and Doppler interference are uncorrelated, i.e. $E[\tilde{h}_i[k]\tilde{e}_{i,d}^*[k]] = 0$. Moreover, it is possible to estimate the $\tilde{h}_i[k]$ from L uncorrelated taps defined by $\tilde{h}_i[l, 0]$.

III. DL TECHNIQUES OVERVIEW

This section discusses the DL networks employed in the studied DL-based channel estimation schemes, providing the mathematical representation of each network.

A. FNN

Neural networks are one of the most popular machine learning algorithms [37]. Initially, neural networks are inspired by the neural architecture of a human brain. For this reason, the basic building block is called a neuron as is the case with a human brain. Its functionality is similar to that of a human neuron, i.e. it takes in some inputs and then fires an output. In purely mathematical terms, a neuron denotes a placeholder for a mathematical function whose job is to yield an output by applying the function on the given inputs. Neurons are stacked together to form a layer. The neural network comprises at least one layer; in case multiple layers are employed, the neural network is referred to as deep FNN.

Consider a FNN architecture shown in Figure 1. Here \mathcal{L} represents the number of layers, including one input layer, $\mathcal{L} - 2$ hidden layers, as well as one output layer. The l -th hidden layer of the network consists of J_l neurons where $2 \leq l \leq \mathcal{L} - 1$. Moreover, each neuron in the l -th hidden layer is denoted by j where $j_1 \leq j \leq J_l$. The FNN inputs \mathbf{i}

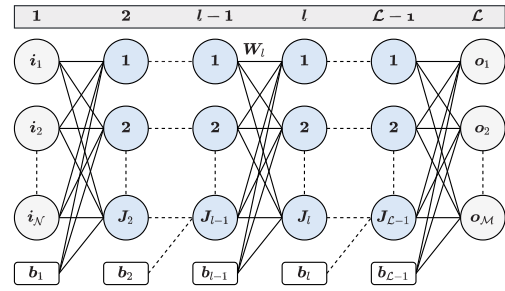


FIGURE 1. FNN architecture showing the input, output, and hidden layers.

and outputs \mathbf{o} are expressed as $\mathbf{i} = [i_1, i_2, \dots, i_N]^T \in \mathbb{R}^{N \times 1}$ and $\mathbf{o} = [o_1, o_2, \dots, o_M]^T \in \mathbb{R}^{M \times 1}$, where N and M refer to the number of FNN inputs and outputs, respectively. $\mathbf{W}_l \in \mathbb{R}^{J_l \times J_{l-1}}$, and $\mathbf{b}_l \in \mathbb{R}^{J_l \times 1}$ are used to express the weight matrix and the bias vector of the l -th hidden layer, respectively.

Each neuron $n_{(l,j)}$ performs a nonlinear transform of a weighted summation of the preceding layer's output values. This nonlinear transformation is represented by the activation function $f_{(l,j)}$ on the neuron input vector $\mathbf{i}_{(l)} \in \mathbb{R}^{J_{l-1} \times 1}$ using its weight vector $\boldsymbol{\omega}_{(l,j)} \in \mathbb{R}^{J_{l-1} \times 1}$, and bias $b_{(l,j)}$, respectively. The neuron's output $o_{(l,j)}$ is

$$o_{(l,j)} = f_{(l,j)}(b_{(l,j)} + \boldsymbol{\omega}_{(l,j)}^T \mathbf{i}_{(l)}). \quad (14)$$

The deep neural network (DNN) overall output of the l -th hidden layer is signified by the vector form

$$\mathbf{o}_{(l)} = \mathbf{f}_{(l)}(\mathbf{b}_{(l)} + \mathbf{W}_{(l)} \mathbf{i}_{(l)}), \quad \mathbf{i}_{(l+1)} = \mathbf{o}_{(l)}, \quad (15)$$

where $\mathbf{f}_{(l)}$ is a vector resulting from the stacking of the n_l activation functions.

After the selection of the FNN architecture, the parameter $\theta = (\mathbf{W}, \mathbf{B})$ representing the total FNN weights and biases must be estimated via the learning procedure applied during the FNN training phase. As well known, θ estimation is obtained by minimizing a loss function $\text{Loss}(\theta)$. The loss function measures how far apart the predicted FNN outputs $(\mathbf{o}_{(\mathcal{L})}^{(P)})$ are from the true outputs $(\mathbf{o}_{(\mathcal{L})}^{(T)})$. Therefore, the FNN training phase carried over N_{train} training samples can be explained in two steps: (i) calculate the loss, and (ii) update θ . This process is repeated until convergence, so that the loss becomes very small. Accordingly, various optimization algorithms can be used for minimizing $\text{Loss}(\theta)$ by iteratively updating the parameter θ , i.e., stochastic gradient descent [37], root mean square prop [38], and adaptive moment estimation (ADAM) [39].

The final step after FNN training is to test the trained FNN on new data in order to evaluate its performance. An elaborate comprehensive analysis of FNN different principles is presented in [40].

B. LSTM

Another well-known DL tool is available in the form of LSTM networks that essentially deal with sequential data

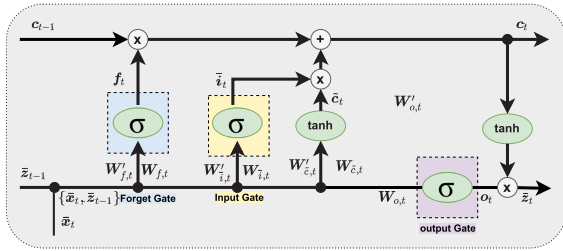


FIGURE 2. LSTM unit architecture [41] showing the forget, input, and output gates.

where the order of the data matters and a correlation exists between the previous and the future data. In this context, LSTM networks are defined with a special architecture capable of learning the data correlation over time, which enables the LSTM network to predict the future data based on prior observations.

The LSTM unit, as shown in Figure 2, contains computational blocks referred to as gates, which are responsible for controlling and tracking the information flow over time. The LSTM network mechanism can be explicated in four major steps:

1) FORGET THE IRRELEVANT INFORMATION

In general, the LSTM unit classifies the input data into relevant and irrelevant information. The first processing step entails eliminating the irrelevant information that is not important for predicting the future. This can be undertaken through the forget gate that decides which information the LSTM unit should retain, and which information can be deleted. The forget gate processing is defined as follows

$$f_t = \sigma(W_{f,t}\bar{x}_t + W'_{f,t}\bar{z}_{t-1} + \bar{b}_{f,t}), \quad (16)$$

where σ denotes the sigmoid function, $W_{f,t} \in \mathbb{R}^{P \times K_{in}}$, $W'_{f,t} \in \mathbb{R}^{P \times P}$ and $\bar{b}_{f,t} \in \mathbb{R}^{P \times 1}$ are the forget gate weights and biases at time t , $\bar{x}_t \in \mathbb{R}^{K_{in} \times 1}$ and \bar{z}_{t-1} represents the LSTM unit input vector of size K_{in} , and the previous hidden state of size P , respectively.

2) STORE THE RELEVANT NEW INFORMATION

After classifying the relevant information, the LSTM unit applies some computations on the selected information via the input gate

$$\bar{i}_t = \sigma(W_{i,t}\bar{x}_t + W'_{i,t}\bar{z}_{t-1} + \bar{b}_{i,t}), \quad (17)$$

$$\bar{c}_t = \tanh(W_{c,t}\bar{x}_t + W'_{c,t}\bar{z}_{t-1} + \bar{b}_{c,t}). \quad (18)$$

3) UPDATE THE NEW CELL STATE

Next the LSTM unit is supposed to update the current cell state c_t based on the two previously-mentioned steps such that

$$c_t = f_t \odot c_{t-1} + \bar{i}_t \odot \bar{c}_t. \quad (19)$$

where \odot denotes the Hadamard product.

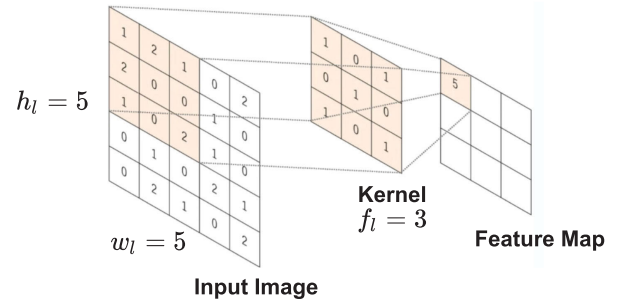


FIGURE 3. Numerical example showing the convolution operation performed by the CNN convolutional layer [44].

4) GENERATE THE LSTM UNIT OUTPUT

Updating the hidden state and generating the output by the output gate is the final processing step. The output is considered to be a cell state filtered version and can be computed such that

$$o_t = \sigma(W_{o,t}\bar{x}_t + W'_{o,t}\bar{z}_{t-1} + \bar{b}_{o,t}), \quad (20)$$

$$\bar{z}_t = o_t \odot \tanh(c_t). \quad (21)$$

In literature, there exists several LSTM architecture variants, where the interactions between the LSTM unit gates are modified. Greff *et al.* [42] provide a detailed comparison of popular LSTM architecture variants.

C. CNN

Another type of deep learning is CNN model. This is commonly used for processing data with grid patterns, such as images [43]. Thus, CNN has generally become the state of the art for several visual applications such as image classification, due to its demonstrated ability to extract patterns from the input image. CNN can be seen as a set of several layers stacked together to accomplish the requisite task. These layers include

- **Input layer:** It represents the 2D or 3D input image. For the sake of simplicity, let us consider a 2D image input to the l -th CNN layer denoted by $X_l \in \mathbb{R}^{h_l \times w_l}$, where h_l and w_l denote the height and the width of the X_l input image.
- **Convolutional layer:** refers to a specialized type of linear operation used for feature extraction, where predefined filters referred to as kernels scan the input matrix to fill the output matrix denoted as feature map, which is shown in Figure 3. We note that different kernels can be considered as different feature extractors.

Two key hyper parameters define the CNN convolutional layer, namely, the size and number of kernels denoted by f_l and n_l , respectively. The typical kernel size is 3×3 , but sometimes 5×5 or 7×7 . The number of kernels is arbitrary and determines the depth of output feature maps. It is possible to tune these parameters according to the application type. Furthermore, the process of training a CNN model regarding the convolution layer involves identifying the kernels values that

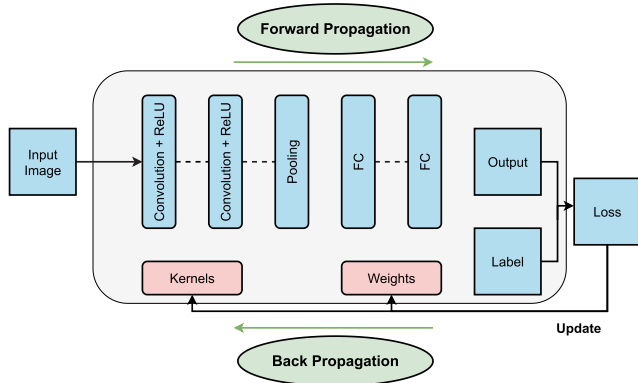


FIGURE 4. CNN classical architecture [44] showing the concatenation of convolutional, pooling, and FC layers.

work optimally for a particular task based on a given training dataset. In the convolution layer, the kernels are the only automatically learned parameters during the training process. Mathematically speaking, for a given input image X_I and kernel $K_I \in \mathbb{R}^{h_i \times w_i \times 1}$, we consider one kernel for simplicity, the generated feature map $Y_I \in \mathbb{R}^{(h_i-f+1) \times (w_i-f+1)}$ can be expressed as follows

$$Y_I[x, y] = \sum_{i=1}^{h_i} \sum_{j=1}^{w_i} K_I[i, j] X_I[x + i - 1, y + j - 1]. \tag{22}$$

- **Activation layer:** The outputs of a linear operation such as convolution pass through a nonlinear activation function. This activation function introduces non-linear processing to the CNN architecture given that the input-output CNN pairs relation could be non-linear. While several non-linear activation functions exist such as sigmoid or hyperbolic tangent (tanh) function, the most common presently used function is the rectified linear unit (ReLU).
- **Pooling layer:** This layer is employed to decrease the number of parameters when the images are too large. Pooling operation is also referred to as sub-sampling or down-sampling. This reduces the dimensionality of all feature maps but does manage to retain significant information. Notably, none of the pooling layers contains any learnable parameter. The most popular form of pooling operation is max pooling, which extracts patches from the input feature maps, outputs the maximum value in each patch, and then discards all the other values. However, there are other pooling operations such as global average pooling [45].
- **Fully connected layer (FC):** This layer forms the last block of the CNN architecture and is mainly employed in the classification problems. It is a simple feed-forward neural network layer that comprises at least one hidden layer; its role is to transform the 2D CNN layer output into a 1D vector. In classification problems, the final outputs of the CNN network represent the probabilities

for each class, where the final fully-connected layer typically has the same number of output nodes as the number of classes.

- **Batch normalization:** It is used to increase the CNN stability of the output by normalizing each layer’s output. Moreover, batch normalization layer reduces overfitting and accelerates the CNN training.
- **Output layer:** This layer is configured in accordance with the studied problem. For instance, in classification problems the CNN output layer is a fully connected layer with softmax activation function. On the other hand, in regression problems, the CNN output does not use any activation function.

Figure 4 illustrates the classical CNN architecture. As seen in this figure, the only trainable parameters within the CNN network are the kernels and the fully connected layer weights. Similar to all other DL techniques, CNN network updates its trainable parameters by minimizing the CNN loss function that measures how far the inputs are from the outputs. Thereafter, the CNN kernels and weights are updated in the back propagation operation [46]. Finally, the performance of the trained CNN model is examined in the testing phase where new unobserved images are fed to the trained CNN model.

It is noteworthy that there are special CNN architectures such as SR-CNN [47], DN-CNN [48], and super resolution convolutional long short-term memory (SR-ConvLSTM) [49] that are mainly used for regression problems. SR-CNN is used for enhancing the quality of the input image, where it takes the low-resolution image as the input and outputs the high-resolution one. DN-CNN uses another methodology to improve the image quality by separating the noise from the input noisy image employing residual learning [50]. The input noisy image is then subtracted from the extracted noise, resulting in the denoised image. Furthermore, SR-ConvLSTM combines both LSTM and CNN networks together where time correlation across the whole input image is learned, thus leading to a better estimation accuracy.

IV. DL-BASED SBS CHANNEL ESTIMATION

In DL-based SBS channel estimation, FNN and LSTM networks are primarily integrated with conventional estimation schemes in the following two manners: (i) FNN is implemented as a post-processing module after conventional DPA, spectral temporal averaging (STA), and time domain reliable test frequency domain interpolation (TRFI) estimators. (ii) LSTM network gets implemented as a pre-processing unit before conventional DPA estimation to minimize the DPA demapping error iteratively. Both implementations are helpful in improving the channel estimation’s accuracy, particularly in high mobility scenarios. However, the LSTM-based estimation illustrates a considerable superiority over the FNN-based estimation as demonstrated in Section VI. Hereafter, the steps applied in each DL-based SBS estimator are presented.

A. DPA-FNN

The DPA estimation [12] utilizes the demapped data subcarriers of the previously received OFDM symbol for estimating the channel for the existing OFDM symbol such that

$$\tilde{\mathbf{d}}_i[k] = \mathfrak{D}\left(\frac{\tilde{\mathbf{y}}_i[k]}{\hat{\mathbf{h}}_{\text{DPA}_{i-1}}[k]}\right), \quad \hat{\mathbf{h}}_{\text{DPA}_0}[k] = \hat{\mathbf{h}}_{\text{LS}}[k], \quad (23)$$

where $\mathfrak{D}(\cdot)$ refers to the demapping operation to the nearest constellation point in accordance with the employed modulation order. $\hat{\mathbf{h}}_{\text{LS}}$ signifies the LS estimated channel at the received preambles, such that

$$\hat{\mathbf{h}}_{\text{LS}}[k] = \frac{\sum_{u=1}^P \tilde{\mathbf{y}}_u^{(p)}[k]}{P\tilde{\mathbf{A}}[k]}, \quad k \in \mathcal{K}_{\text{on}}, \quad (24)$$

where $\tilde{\mathbf{A}}$ denotes the frequency domain predefined preamble sequence. Thereafter, the final DPA channel estimates are updated in the following manner

$$\hat{\mathbf{h}}_{\text{DPA}_i}[k] = \frac{\tilde{\mathbf{y}}_i[k]}{\tilde{\mathbf{d}}_i[k]}. \quad (25)$$

DPA estimation suffers from two main limitations. First, it is based on the basic $\hat{\mathbf{h}}_{\text{LS}}$ estimation suffering from noise enhancement. Second, the demapping step in DPA leads to a significant demapping error primarily in low signal-to-noise ratio (SNR) region stemming from the noise imperfections and doubly-dispersive channel variations. This demapping error is enlarged in high mobility scenarios employing high modulation orders. In addition, since the DPA estimated channels are updated iteratively over the received frame, the demapping error propagates via the frame that results in a significant degradation in performance. In order to address these limitations, the DPA-FNN scheme [26] has been proposed to compensate the DPA estimation error, where $\hat{\mathbf{h}}_{\text{DPA}_i}[k]$ is fed to a three-hidden-layers FNN with 40 – 20 – 40 neurons, as shown in Figure 5. Using the FNN in addition to the DPA scheme yields good performance but it is not sufficient, because it ignores the time and frequency correlation between successive received OFDM symbols. Also, the employed FNN architecture can be optimized to reduce the computational complexity of channel estimation.

B. STA-FNN

To improve the conventional DPA estimation, the STA estimator [12] has been proposed where frequency and time-domain averaging are applied on top of the DPA estimated channel as follows

$$\hat{\mathbf{h}}_{\text{FD}_i}[k] = \sum_{\lambda=-\beta}^{\lambda=\beta} \omega_\lambda \hat{\mathbf{h}}_{\text{DPA}_i}[k + \lambda], \quad \omega_\lambda = \frac{1}{2\beta + 1}. \quad (26)$$

$$\hat{\mathbf{h}}_{\text{STA}_i}[k] = \left(1 - \frac{1}{\alpha}\right) \hat{\mathbf{h}}_{\text{STA}_{i-1}}[k] + \frac{1}{\alpha} \hat{\mathbf{h}}_{\text{FD}_i}[k]. \quad (27)$$

STA estimator performs well in the low SNR region. However, it suffers from a considerable error floor in high SNR

regions due to the large DPA demapping error. Importantly, in [12], the values of the frequency and time averaging coefficients are fixed to $\alpha = \beta = 2$. Thus, the final STA estimated channel is a linear combination between the previously estimated channel (27) and the frequency averaged channel estimates (26). However, this linear combination leads to a significant performance degradation in real case scenarios due to the doubly-dispersive channel non-linear imperfections. Here, FNN is utilized as a post non-linear processing unit after the conventional STA scheme [27]. STA-FNN captures more the time-frequency correlations of the channel samples, apart from correcting the conventional STA estimation error. Furthermore, the optimized STA-FNN architecture performs better than the DPA-FNN with a significant computational complexity decrease, as elucidated in Section VII.

C. TRFI-FNN

TRFI estimation scheme [14] is another methodology used for improving the DPA estimation in (25). Assuming that the time correlation of the channel response between two adjacent OFDM symbols is high, TRFI define two sets of subcarriers such that: (i) \mathcal{RS}_i set: that includes the reliable subcarriers indices, and (ii) \mathcal{URS}_i set: which contains the unreliable subcarriers indices. The estimated channels for the \mathcal{URS}_i are then interpolated using the \mathcal{RS}_i channel estimates by means of the frequency-domain cubic interpolation. This procedure can be expressed in the following manner

- Equalize the previously received OFDM symbol by $\hat{\mathbf{h}}_{\text{TRFI}_{i-1}}[k]$ and $\hat{\mathbf{h}}_{\text{DPA}_i}[k]$, such that
$$\tilde{\mathbf{d}}'_{i-1}[k] = \mathfrak{D}\left(\frac{\tilde{\mathbf{y}}_{i-1}[k]}{\hat{\mathbf{h}}_{\text{DPA}_i}[k]}\right), \quad \tilde{\mathbf{d}}''_{i-1}[k] = \mathfrak{D}\left(\frac{\tilde{\mathbf{y}}_{i-1}[k]}{\hat{\mathbf{h}}_{\text{TRFI}_{i-1}}[k]}\right). \quad (28)$$

- According to the demapping results, the subcarriers are grouped as follows

$$\begin{cases} \mathcal{RS}_i \leftarrow \mathcal{RS}_i + k, & \tilde{\mathbf{d}}'_{i-1}[k] = \tilde{\mathbf{d}}''_{i-1}[k] \\ \mathcal{URS}_i \leftarrow \mathcal{URS}_i + k, & \tilde{\mathbf{d}}'_{i-1}[k] \neq \tilde{\mathbf{d}}''_{i-1}[k]. \end{cases} \quad (29)$$

- Finally, frequency-domain cubic interpolation is employed to estimate the channels at the \mathcal{URS}_i as follows

$$\hat{\mathbf{h}}_{\text{TRFI}_i}[k] = \begin{cases} \hat{\mathbf{h}}_{\text{DPA}_i}[k], & k \in \mathcal{RS}_i \\ \text{Cubic Interpolation}, & k \in \mathcal{URS}_i. \end{cases} \quad (30)$$

Performing frequency-domain interpolation in addition to the DPA estimation enhances the performance. However, TRFI still suffers from the demapping and interpolation errors as the number of reliable subcarriers (RS) subcarriers is inversely proportional to the channel variations. Additionally, the condition where $\tilde{\mathbf{d}}'_{i-1}[k] \neq \tilde{\mathbf{d}}''_{i-1}[k]$ is more dominant in high mobility scenarios. It is for this reason that only

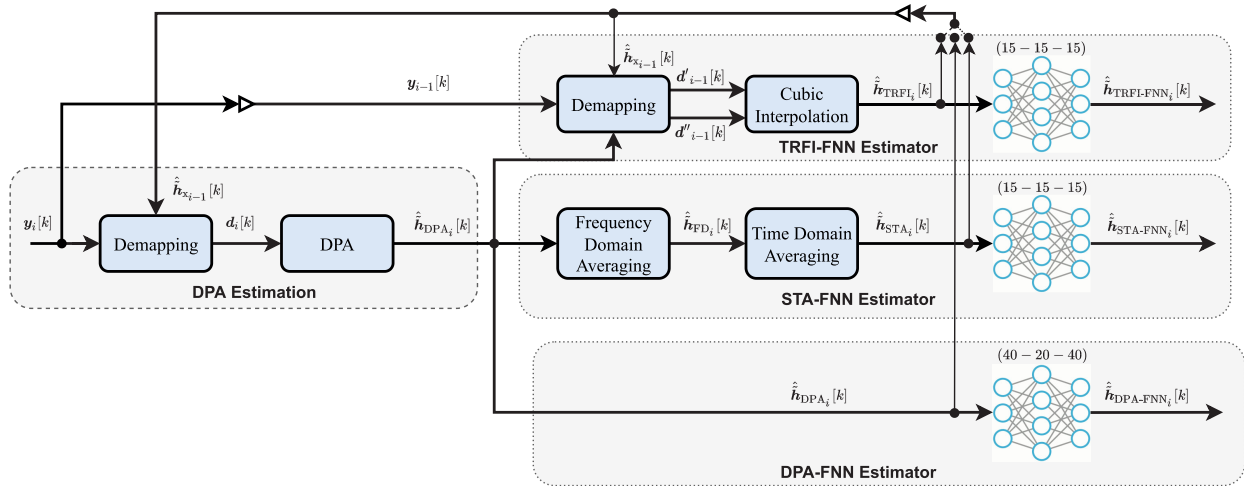


FIGURE 5. Block diagram of the studied DNN-based SBS estimators.

a few RS subcarriers will be selected and the employed cubic interpolation performance will be degraded.

Inspired by the work undertaken in STA-FNN, the authors in [28] used the same optimized FNN architecture as in [27], albeit with $\hat{h}_{\text{TRFI}}[k]$ as an input instead of $\hat{h}_{\text{STA}}[k]$. TRFI-FNN corrects the cubic interpolation error and also learns the channel frequency domain correlation, thus leading to an improved performance in high SNR regions.

D. LSTM-FNN-DPA

Unlike the FNN-based estimators, where the DL processing is employed following the conventional estimators, the work carried out in [51] shows that employing the DL processing prior to the conventional estimator, specifically the DPA estimation, could lead to a significant improvement in the overall performance. In this context, the authors have proposed to use two cascaded LSTM and FNN networks for both channel estimation as well as noise compensation, as shown in Figure 6.

The LSTM-FNN-DPA estimator employs the previous and current pilot subcarriers besides the LSTM-FNN estimated channel employed in the DPA estimation, such that

$$\tilde{d}_{\text{LSTM-FNN},i,d}[k] = \mathcal{D}\left(\frac{\tilde{y}_{i,d}[k]}{\hat{h}_{\text{LSTM-FNN},i-1,d}[k]}\right),$$

$$\hat{h}_{\text{LSTM}_0}[k] = \hat{h}_{\text{LS}}[k], \quad (31)$$

$$\hat{h}_{\text{DL},i,d}[k] = \frac{\tilde{y}_{i,d}[k]}{d_{\text{LSTM},i,d}[k]}. \quad (32)$$

While this estimator can outperform the FNN-based estimators, it experiences a considerable computational complexity arising from the employment of two DL networks.

E. LSTM-DPA-TA

Gizzini *et al.* [41] propose to use only LSTM network instead of two as implemented in the LSTM-FNN-DPA estimator. In addition, noise compensation is made possible by applying time averaging (TA) processing as shown in Figure 6.

This methodology only requires the previous pilots besides the LSTM estimated channel as an input. Then, the LSTM estimated channel is employed in the DPA estimation as follows

$$\tilde{d}_{\text{LSTM}_i}[k] = \mathcal{D}\left(\frac{\tilde{y}_i[k]}{\hat{h}_{\text{LSTM}_{i-1}}[k]}\right),$$

$$\hat{h}_{\text{LSTM}_0}[k] = \hat{h}_{\text{LS}}[k], \quad (33)$$

$$\hat{h}_{\text{LSTM-DPA}_i}[k] = \frac{\tilde{y}_i[k]}{\tilde{d}_{\text{LSTM}_i}[k]}. \quad (34)$$

Finally, to alleviate the impact of the AWGN noise, TA processing is applied to the $\hat{h}_{\text{LSTM-DPA}_i}[k]$ estimated channel, such that

$$\hat{h}_{\text{DL-TA},i,d} = \left(1 - \frac{1}{\alpha}\right)\hat{h}_{\text{DL-TA},i-1,d} + \frac{1}{\alpha}\hat{h}_{\text{LSTM-DPA}_i,d}. \quad (35)$$

Here, α denotes the utilized weighting coefficient. In [41], the authors use a fixed $\alpha = 2$ for simplicity. Therefore, the TA applied in (35) reduces the AWGN noise power σ^2 iteratively within the received OFDM frame according to the ratio

$$R_{\text{DL-TA}_q} = \left(\frac{1}{4}\right)^{(q-1)} + \sum_{j=2}^q \left(\frac{1}{4}\right)^{(q-j+1)} = \frac{4q-1+2}{3 \times 4^{q-1}}. \quad (36)$$

This corresponds to the AWGN noise power ratio of the estimated channel at the q -th estimated channel, where $1 < q < I + 1$ and $R_{\text{DL-TA}_1} = 1$ denotes the AWGN noise power ratio at $\hat{h}_{\text{LS}}[k]$. From the derivation of $R_{\text{DL-TA}_q}$, it can be seen that the noise power decreases over the received OFDM frame, i.e., the SNR increases, resulting in an overall improved performance. Moreover, the input dimension reduction, coupled with the simple TA processing, significantly lowers the overall computational complexity. Intensive experiments reveal that the performance of DL networks is strongly related to the SNR considered in the training [52].

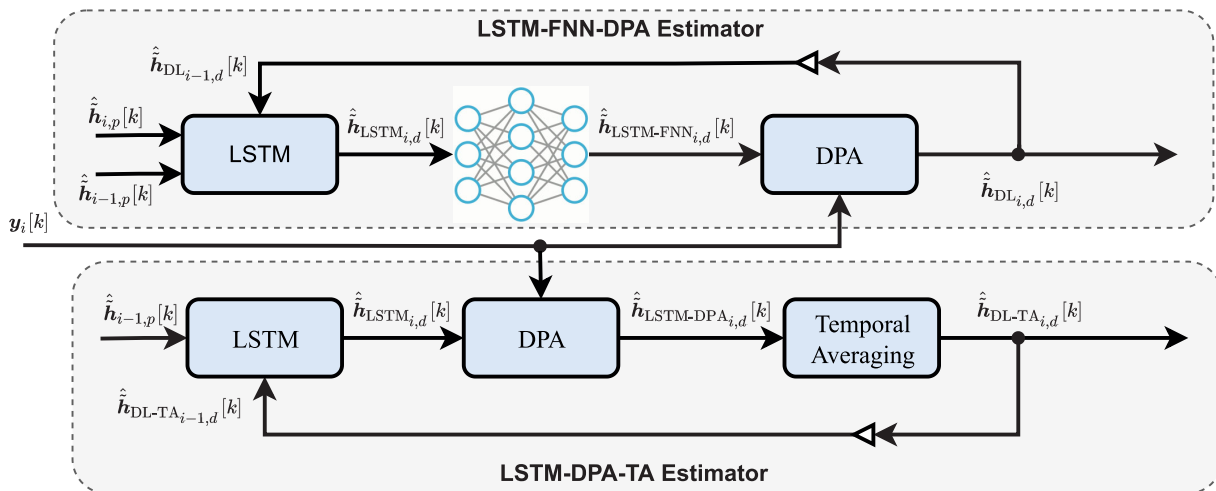


FIGURE 6. Block diagram of the studied LSTM-based SBS estimators.

TABLE 1. Parameters of the studied DL-based SBS channel estimators.

| Parameter | Values |
|---|--------------|
| DPA-FNN (Hidden layers; Neurons per layer) | (3;40-20-40) |
| STA-FNN (Hidden layers; Neurons per layer) | (3;15-15-15) |
| TRFI-FNN (Hidden layers; Neurons per layer) | (3;15-15-15) |
| LSTM (Hidden layers; Neurons per layer) | (1;128) |
| Activation function | ReLU |
| Number of epochs | 500 |
| Training samples | 800000 |
| Testing samples | 200000 |
| Batch size | 128 |
| Optimizer | ADAM |
| Loss function | MSE |
| Learning rate | 0.001 |
| Training SNR | 40 dB |

The training undertaken at the highest SNR value provides the best performance. In fact, the DL network is able to learn better the channel when the training is performed at a high SNR value because the impact of the channel is higher than the impact of the noise in this SNR range. Owing to the robust generalization properties of DL, trained networks can still estimate the channel even if the noise increases, i.e., at low SNR values. Therefore, FNN and LSTM based estimators training is performed using SNR = 40 dB to attain the best performance. Moreover, intensive experiments are performed using the grid search algorithm [53] to select the most suitable FNN and LSTM hyper parameters in terms of performance as well as complexity. Figures 5 and 6 illustrate the block diagram of the FNN and LSTM based estimators. Furthermore, Table 1 presents their parameters.

V. DL-BASED FBF CHANNEL ESTIMATION SCHEMES

This section presents the DL-based FBF estimators introduced to improve the channel estimation accuracy,

particularly in very high mobility scenarios, where the channel variation is found to be severe. Similar to the DL-based SBS estimators, the DL-based FBF estimators apply first conventional estimation followed by means of CNN processing.

A. CHANNELNET

In [29], the authors use forward a CNN-based channel estimator denoted as channel network (ChannelNet) scheme, where 2D radial basis function (RBF) interpolation is implemented as an initial channel estimation. The underlying motivation of the 2D RBF interpolation is to approximate multidimensional scattered unknown data from their surrounding neighbors known data by employing the radial basis function. In order to achieve the purpose, the distance function is calculated between every data point to be interpolated and its neighbours, where closer neighbors are assigned higher weights. Thereby, the RBF interpolated frame is considered a low resolution image, where SR-CNN is utilized to obtain an improved estimation. Finally, to ameliorate the effect of noise within the high resolution estimated frame, DN-CNN is implemented leading to a high resolution and noise alleviated estimated channels. The ChannelNet estimator considers sparsely allocated pilots within the IEEE 802.11p frame and initially applies the LS estimation to the pilot subcarriers within the received OFDM frame. Subsequently, the 2D RBF interpolation is derived by the weighted summation of the distance between each data subcarrier to be interpolated as well as all the pilot subcarriers in the received OFDM frame, such that

$$\hat{\mathbf{H}}_{\text{RBF}}[k, i] = \sum_{j=1}^{K_p I} \omega_j \Phi(|k - \mathcal{K}_f[j]|, |i - \mathcal{K}_t[j]|). \quad (37)$$

$\mathcal{K}_f = [\mathcal{K}_{p_1}, \dots, \mathcal{K}_{p_I}] \in \mathbb{R}^{1 \times K_p I}$ and $\mathcal{K}_t = [(1)_{\times K_p}, \dots, (I)_{\times K_p}] \in \mathbb{R}^{1 \times K_p I}$ represent the frequency and time indices vectors of the allocated pilot subcarriers within the received OFDM frame, respectively. ω_j is the RBF weight multiplied

AQ:4 **TABLE 2. Main characteristics and features of the studied DL-based channel estimators.**

| Estimator Type | Estimator Reference | Conventional Estimation | DL-based Method | Complexity | BER Performance | Robustness | Pros and Cons |
|----------------|---------------------|-------------------------|-------------------|------------|-----------------|------------|--|
| SBS | [26] | DPA | FNN | ++ | ++ | ++ | + Significant performance superiority over conventional estimators. - Ignore the time and frequency correlation between successive received OFDM symbols. - Complex FNN architecture to compensate the conventional DPA demapping error. |
| | [27] | STA | | + | +++ | ++ | + STA averaging ameliorate the impact of the AWGN noise in low SNR regions. + Optimized FNN architecture. - Fixed averaging coefficients. - Performance degradation in high mobility scenarios. |
| | [28] | TRFI | | + | ++++ | +++ | + Cubic Interpolation enhances the performance in the entire SNR region. + Optimized FNN architecture. - Assume high correlation between successive OFDM symbols. - Lack of robustness in very high mobility scenarios. |
| | [51] | DPA | LSTM and FNN | +++ | ++++ | ++++ | + Outperform FNN-based estimators. + Improved estimation since LSTM is implemented before DPA estimation - Employ LSTM and FNN in the same architecture. |
| | [41] | DPA and TA | LSTM | +++ | ++++ | ++++ | + TA processing results in a considerable decline in the AWGN noise. + Employ only one optimized LSTM unit. + Reduced input dimension. |
| FBF | [29] | 2D RBF | SR-CNN and DN-CNN | +++++ | ++ | ++ | - 2D RBF interpolation high computational complexity. - The 2D RBF function and scale factor should be optimized in accordance with the channel variations. - Employ two high-complexity CNN architectures. |
| | [30] | ADD-TT | SR-ConvLSTM | +++++ | +++ | +++ | + Outperform ChannelNet estimator [29]. - Fixed ADD-TT Averaging coefficients. - High computational complexity owing to the integration of both LSTM and CNN architectures. |
| | [31] | WI | SR-CNN or DN-CNN | +++ | ++++ | ++++ | + Adaptive frame structure according to the mobility condition. + Reduced buffering time at the receiver. + Transmission data rate gain. + Optimized CNN architectures. |

by the RBF interpolation function $\Phi(\cdot)$ between the (k, i) data subcarrier and the $(\mathcal{K}_f[j], \mathcal{K}_t[j])$ pilot subcarrier. In [29], the RBF gaussian function is applied, such that

$$\Phi(x, y) = e^{-\frac{(x+y)^2}{r_0}}. \quad (38)$$

r_0 refers to the 2D RBF scale factor that varies based on the used RBF function. Notably, altering the value of r_0 alters the shape of the interpolation function. Moreover, the RBF weights $\mathbf{w}_{\text{RBF}} = [\omega_1, \dots, \omega_{K_p I}] \in \mathbb{R}^{K_p I \times 1}$ are calculated using the following relation:

$$\mathbf{A}_{\text{RBF}} \mathbf{w}_{\text{RBF}} = \tilde{\mathbf{h}}_{\text{LS}}. \quad (39)$$

Here, $\mathbf{A}_{\text{RBF}} \in \mathbb{R}^{K_p I \times K_p I}$ is the RBF interpolation matrix of the pilots subcarriers, with entries $a_{i,j} = \Phi(\mathcal{K}_f[i], \mathcal{K}_t[j])$ where $i, j = 1, \dots, K_p I$. It is observed that, $\tilde{\mathbf{h}}_{\text{LS}} = \text{vec} \left\{ \hat{\mathbf{H}}_{\text{LS}} \right\} \in \mathbb{C}^{K_p I \times 1}$ is a vector that contains the LS estimated channels at all the pilot subcarriers within the received OFDM frame. This is expressed as

$$\hat{\mathbf{H}}_{\text{LS}}[k, i] = \frac{\tilde{\mathbf{Y}}[k, i]}{\tilde{\mathbf{P}}[k, i]}, \quad k \in \mathcal{K}_p, \quad 1 \leq i \leq I, \quad (40)$$

with $\tilde{\mathbf{P}}[k, i]$ is the frequency-domain pre-defined pilot subcarriers, and \mathcal{K}_p refers to the allocated sparse pilots indices within the received OFDM symbol. After computing \mathbf{W}_{RBF} , it is possible to calculate the RBF estimated channel for every data subcarriers within the received OFDM frame, as shown in (37). Finally, the RBF interpolation estimated frame $\hat{\mathbf{H}}_{\text{RBF}}$ is fed as an input to SR-CNN and DN-CNN to improve the channel estimation accuracy and reduce the noise impact.

The ChannelNet estimator limitations lie in: (i) 2D RBF interpolation high computational complexity arising from the computation of (39) for the channel estimation of all data subcarriers. (ii) The 2D RBF function and scale factor needs to be optimized in accordance with the channel variations. (iii) The integrated SR-CNN and DN-CNN architectures have significant computational complexity. Notably, the ChannelNet estimator uses a fixed RBF function and scale factor, thus experiencing a considerable degradation in performance, particularly in low SNR regions, where the noise impact remains dominant, as well as high mobility vehicular scenarios, where the channel varies swiftly within the OFDM frame.

B. TS-ChannelNet

Temporal spectral ChannelNet (TS-ChannelNet) [30] is based on applying average decision-directed with time truncation (ADD-TT) interpolation to the received OFDM frame. Thereafter, accurate estimation is achieved by implementing SR-ConvLSTM network to track doubly-dispersive channel variations by learning the vehicular channel's time and frequency correlations. It is observed that the ADD-TT interpolation is an SBS estimator, where DPA estimation is initially applied as explained in (23) and (25). Thereafter, the enlarged DPA demapping error is reduced by applying time domain

TABLE 3. Parameters of the studied DL-based FBF channel estimators.

| Parameter | Values |
|--|------------------------------------|
| Input/Output dimensions | $2K_{\text{on}} \times I \times 1$ |
| SR-CNN (Hidden layers - n_l, f_l) | (3 - 9,64; 1,32; 5,1) |
| DN-CNN (Hidden layers - n_l, f_l) | (18 - 64, 3) |
| Optimized SR-CNN (Hidden layers - n_l, f_l) | (3 - 9,32; 1,16; 5,1) |
| Optimized DN-CNN (Hidden layers - n_l, f_l) | (7 - 16, 3) |
| SR-ConvLSTM (Hidden layers - n_l, f_l) | (3 - 9,64; 1,32; 5,1) |
| Activation function | ReLU |
| Number of epochs | 250 |
| Training samples | 8000 |
| Testing samples | 2000 |
| Batch size | 128 |
| Optimizer | ADAM |
| Loss function | MSE |
| Learning rate | 0.001 |
| Training SNR | 40 dB |

truncation in the following manner

$$\hat{\mathbf{h}}_{\text{DPA}_i} = \mathbf{F}_K^H \hat{\mathbf{h}}_{\text{DPA}_i}, \quad (41)$$

where $\mathbf{F}_K \in \mathbb{C}^{K \times K}$ denotes the K -DFT matrix, and $\hat{\mathbf{h}}_{\text{DPA}_i}$ represents the time-domain DPA estimated channel. Thereafter, $\hat{\mathbf{h}}_{\text{DPA}_i}$ truncation is applied to the significant L channel taps, such that

$$\hat{\mathbf{h}}_{\text{DPA}_{i,L}} = \hat{\mathbf{h}}_{\text{DPA}_i}(1 : L). \quad (42)$$

Next, $\hat{\mathbf{h}}_{\text{DPA}_{i,L}}$ is converted back to the frequency domain such that

$$\hat{\mathbf{h}}_{\text{TT}_i} = \mathbf{F}_K \hat{\mathbf{h}}_{\text{DPA}_{i,L}}, \quad (43)$$

Implementing the average time truncation operation to $\hat{\mathbf{h}}_{\text{DPA}_i}[k]$ lowers the effect of noise and enlarged demapping error. Moreover, $\hat{\mathbf{h}}_{\text{TT}_i}[k]$ estimated channel is further enhanced by applying frequency and time-domain averaging consecutively as follows

$$\hat{\mathbf{h}}_{\text{FTT}_i}[k] = \sum_{\lambda=-\beta}^{\lambda=\beta} \omega_\lambda \hat{\mathbf{h}}_{\text{TT}_i}[k + \lambda], \quad \omega_\lambda = \frac{1}{2\beta + 1}. \quad (44)$$

The final ADD-TT channel estimates are updated using time averaging between the previously ADD-TT estimated channel and the frequency averaged channel in (44), such that

$$\hat{\mathbf{h}}_{\text{ADD-TT}_i}[k] = (1 - \alpha) \hat{\mathbf{h}}_{\text{ADD-TT}_{i-1}}[k] + \alpha \hat{\mathbf{h}}_{\text{FTT}_i}[k]. \quad (45)$$

The doubly-dispersive channel can be modeled as a time-series forecasting problem. Here, historical data can be utilized to forecast future observations [54]. Motivated by this possibility, Zhu *et al.* [30] apply SR-ConvLSTM network in addition to the ADD-TT interpolation, where convolutional layers get added to the LSTM network to capture more doubly-dispersive channel features. Consequently, this

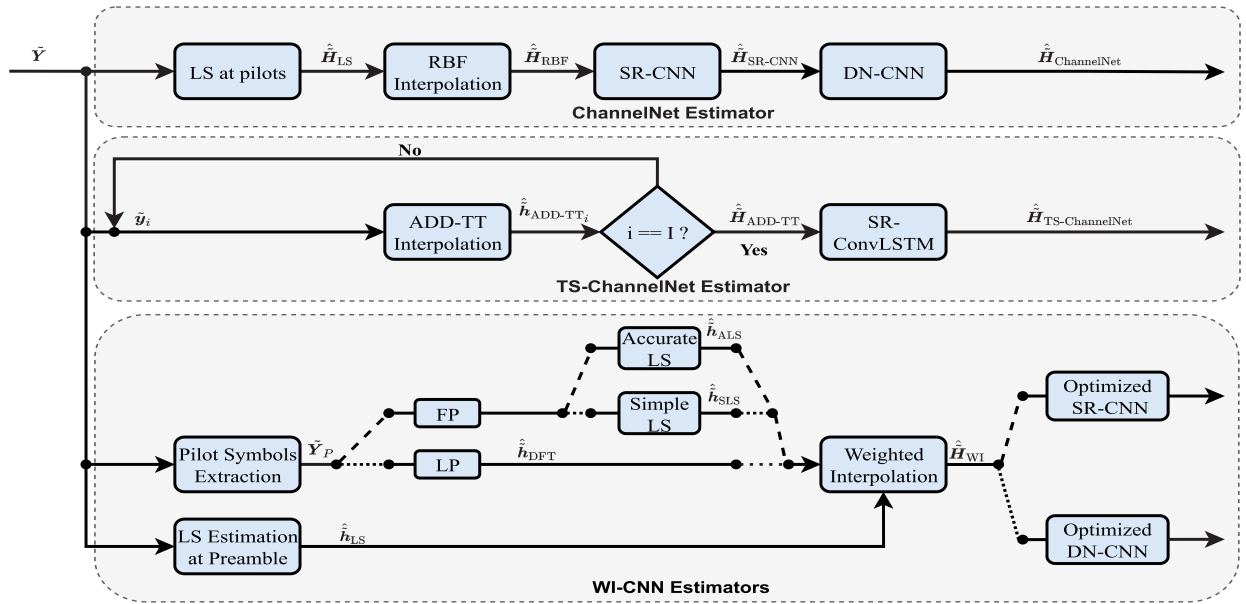


FIGURE 7. Block diagram of the studied CNN-based FBF channel estimators.

improves the estimation performance. Accordingly, the ADD-TT estimated channel for the entire received frame is modeled as a low resolution image. Next, the SR-ConvLSTM network is used after the ADD-TT interpolation. Unlike ChannelNet estimator where two CNNs are employed, TS-ChannelNet estimator uses only one SR-ConvLSTM network, which relatively reduces the overall computational complexity. However, TS-ChannelNet continues to be ridden with high computational complexity due to the integration of LSTM and CNN in a single network.

C. WI-CNN

To overcome the limitations of the ChannelNet and TS-ChannelNet estimators, weighted interpolation (WI)-CNN estimator has been proposed in [31]. In this method, the frame structure is adapted in accordance with the mobility condition employing varied pilot allocation schemes. Particularly, only P pilot OFDM symbols are required in the transmitted frame, such that $\tilde{Y}_P = [\tilde{y}_1^{(p)}, \dots, \tilde{y}_q^{(p)}, \dots, \tilde{y}_P^{(p)}] \in \mathbb{C}^{K_{on} \times P}$. The index $1 \leq q \leq P$ refers to the location of the OFDM pilot symbol in the frame. The other $I_d = I - P$ OFDM data symbols are employed for data transmission purposes. As per the employed pilots allocation scheme, the channel is estimated at the inserted pilot symbols, after which WI is applied to estimate the channel at the OFDM data symbols. The estimated frame is then modeled as a 2D noisy image where optimized SR-CNN and DN-CNN are utilized for noise elimination. Against this backdrop, the WI-CNN proceeds as follows

- Channel estimation at the pilot symbols: Two pilot allocation schemes are defined. The full pilot allocation (FP) where K pilots are inserted within all pilot symbols and LS estimation is applied to estimate the channel for each

inserted pilot symbol, such that

$$\hat{h}_{SLS_q}[k] = \frac{\tilde{y}_q^{(p)}[k]}{\tilde{p}[k]} \tag{46}$$

$\hat{h}_{SLS_q}[k]$ represents the simple LS (SLS) estimation at the q -th inserted pilot symbol. In addition, the accurate LS (ALS) that can be obtained by implementing the DFT interpolation of estimated channel impulse response at the q -th received pilot symbol $\hat{h}_{q,L}$, such that

$$\hat{h}_{ALS_q} = F_K \hat{h}_{q,L}, \quad \hat{h}_{q,L} = F_K^\dagger \hat{h}_{LS_q} \tag{47}$$

ALS relies on the fact that $\tilde{h}_q = F_K h_{q,L}$, where $h_{q,L} \in \mathbb{C}^{L \times 1}$ signifies the channel impulse response at the q -th received pilot symbol that can be estimated by employing the pseudo inverse matrix of F_K , namely, $F_K^\dagger = [(F_K^H F_K)^{-1} F_K^H]$. However, if the number of doubly dispersive-channel taps L remains known, only $K_p = L$ pilot subcarriers are sufficient in each inserted pilot symbol. Accordingly, (47) can be rewritten as

$$\hat{h}_{DFT_q} = F_K \hat{h}_{q,L}, \quad \hat{h}_{q,L} = F_p^\dagger \hat{h}_{LS_q} \tag{48}$$

$F_p^\dagger = [(F_p^H F_p)^{-1} F_p^H]$ denotes the pseudo inverse matrix of $F_p \in \mathbb{C}^{K_p \times L}$ referring to the truncated DFT matrix obtained by selecting K_p rows, and L columns from the K -DFT matrix.

- Channel estimation at data symbols: The estimated channels of the P pilot symbols are first grouped into P matrices to estimate the channel for each received OFDM data symbol, such that

$$\hat{H}_q = [\hat{h}_{q-1}, \hat{h}_q], \quad q = 1, \dots, P. \tag{49}$$

$\hat{\mathbf{h}}_0 = \hat{\mathbf{h}}_{\text{LS}}$ refers to the LS estimated channel at the beginning of the received frame (24). Thus, the received frame can be divided into P sub-frames, where f refers to the sub-frame index, such that $1 \leq f \leq P$. Therefore, the estimated channel for the i -th received OFDM symbol within each f -th sub-frame can be expressed as follows

$$\hat{\mathbf{H}}_{\text{WI}_f} = \hat{\mathbf{H}}_f \mathbf{C}_f. \quad (50)$$

$\hat{\mathbf{H}}_f \in \mathbb{C}^{K \times 2}$ denotes the LS estimated channels at the pilot symbols within the f -th sub-frame, and $\mathbf{C}_f \in \mathbb{R}^{2 \times I_f}$ the interpolation weights of the I_f OFDM data symbols within the f -th sub-frame. The interpolation weights \mathbf{C}_f are calculated by minimizing the mean squared error (MSE) between the ideal channel $\tilde{\mathbf{H}}_f$, and the LS estimated channel at the OFDM pilot symbols $\hat{\mathbf{H}}_f$ as obtained in [55] and expressed in (51), as shown at the bottom of the page. There, $J_0(\cdot)$ is the zeroth order Bessel function of the first kind, T_s signifies the received OFDM data symbol duration, whereas E_q denotes the overall noise of the estimated channel at the q -th pilot symbol.

- CNN-based Processing: The final step in the WI-CNN estimators is to apply CNN processing to further improve the WI estimated channels. Optimized SR-CNN and DN-CNN are employed in this context. The investigations conducted in [31] reveal that both SR-CNN and DN-CNN networks have similar performance in low mobility scenarios, whereas DN-CNN outperforms SR-CNN in high mobility scenarios. Figure 7 and Table 3 illustrate the block diagram as well as configured parameters of the studied CNN-based channel estimators, respectively. Furthermore, the salient features of the studied DL-based channel estimators are summarized in Table 2. Notably, robustness feature alludes to the ability of the studied estimation to maintain good performance as the variation of the doubly-dispersive channel increases.

VI. SIMULATION RESULTS

This section illustrates the performance evaluation of the studied DL-based SBS and FBF estimators in relation to BER, NMSE employing varied metrics and mobility scenarios.

A. CONFIGURATION SETUP

To simulate doubly-dispersive channels, vehicular communications is considered a simulation case study, where three tapped delay line (TDL) channel models [56] are defined as follows

- Low mobility: where VTV Urban Canyon (VTV-UC) vehicular channel model is considered. This channel model is measured between two vehicles moving in a dense urban traffic environment at $V = 45$ Kmph equivalent to $f_d = 250$ Hz.
- High and very high mobility: These scenarios measure the communication channel between two vehicles moving on a highway having center wall between its lanes at $V = 100$ Kmph and 200 Kmph equivalent to $f_d = 500$ Hz and $f_d = 1000$ Hz, respectively. This vehicular channel model is referred to as VTV Expressway Same Direction with Wall (VTV-SDWW).

The employed channel models are generated after the wide-sense stationary uncorrelated scattering (WSSUS) model [57]. Thus, we have

- Each path $h_l(t)$ is a zero mean Gaussian complex process, $E\{h_l(t)\} = 0, \forall t$, and the mean of each path is independent of the time variations. Moreover, the time correlation function $r_{h_l}(t_1, t_2) = E\{h_l(t_1)h_l^*(t_2)\}$ can only be written with the difference $\Delta(t) = (t_1 - t_2)$, such that

$$r_{h_l}(t_1, t_2) = r_{h_l}(\Delta_t). \quad (52)$$

Then, each path $h_l(t)$ is the wide sense stationary (WSS).

- Uncorrelated scattering (US) implies that the paths are uncorrelated, so for $l_1 \neq l_2$ we have

$$E[h_{l_1}(t)h_{l_2}^*(t)] = 0. \quad (53)$$

Table 4 illustrates the main characteristics of the defined TDL channel models.

The OFDM simulation parameters are based on the IEEE 802.11p standard as illustrated in Table 5. These simulations are implemented using QPSK and 16QAM modulation orders, the SNR range is $[0, 5, \dots, 40]$ dB. In addition, the performance evaluation is made according to: (i) modulation order, (ii) mobility, (iii) frame length, and (iv) DL architecture.

Finally, it is observed that the conventional 2D LMMSE estimator [16] is included in the performance evaluation of the DL-based FBF estimators as a lower bound

$$\begin{aligned} \mathbf{C}_f &= E \left[\tilde{\mathbf{H}}_f \hat{\mathbf{H}}_f^H \right] \left[E \left[\hat{\mathbf{H}}_f \hat{\mathbf{H}}_f^H \right] \right]^{-1} = \left[E \left[\tilde{\mathbf{H}}_f \hat{\mathbf{h}}_q^H \right] E \left[\tilde{\mathbf{H}}_f \hat{\mathbf{h}}_{q+1}^H \right] \right] \begin{bmatrix} E \left[\left\| \tilde{\mathbf{h}}_q \right\|^2 \right] + E_q & E \left[\tilde{\mathbf{h}}_q \tilde{\mathbf{h}}_{q+1}^H \right] \\ E \left[\tilde{\mathbf{h}}_{q+1} \tilde{\mathbf{h}}_q^H \right] & E \left[\left\| \tilde{\mathbf{h}}_{q+1} \right\|^2 \right] + E_{q+1} \end{bmatrix}^{-1} \\ &= \left[J_0(2\pi f_d(f-1)T_s) \ J_0(2\pi f_d(I_f+1-f)T_s) \right] \begin{bmatrix} 1 + E_{\Phi_q} & J_0(2\pi f_d I_f T_s) \\ J_0(2\pi f_d I_f T_s) & 1 + E_{q+1} \end{bmatrix}^{-1}. \end{aligned} \quad (51)$$

TABLE 4. Main characteristics of the employed vehicular channel models following Jake’s Doppler spectrum.

| Channel model | Channel taps | Vehicle velocity [kmph] | Doppler shift [Hz] | Average path gains [dB] | Path delays [ns] |
|---------------|--------------|-------------------------|--------------------|--|--|
| VTV-UC | 12 | 45 | 250 | [0, 0, -10, -10, -10, -17.8, -17.8, -17.8, -21.1, -21.1, -26.3, -26.3] | [0, 1, 100, 101, 102, 200, 201, 202, 300, 301, 400, 401] |
| VTV-SDWW | 12 | 100-200 | 500-1000 | [0, 0, -11.2, -11.2, -19, -21.9, -25.3, -25.3, -24.4, -28, -26.1, -26.1] | [0, 1, 100, 101, 200, 300, 400, 401, 500, 600, 700, 701] |

performance limit. The 2D LMMSE estimator almost achieves a similar performance as the ideal channel, but is ridden with high computational complexity. This renders it impractical in terms of real-time applications.

B. DL-BASED SBS ESTIMATION SCHEMES

1) MODULATION ORDER

For QPSK modulation order, we can notice from Figure 8, and Figure 9 that conventional SBS estimators witness a considerable performance degradation in different mobility scenarios primarily due to the enlarged DPA demapping error, particularly under very high mobility. Nevertheless, employing DL techniques in the channel estimation process results in a significant improvement in overall performance. To begin with, the FNN-based estimators, where FNN is employed as a post-processing unit after conventional estimators, are discussed. As observed, FNN can implicitly learn the channel correlations apart from preventing a high demapping error arising from conventional DPA-based estimation, while STA-FNN and TRFI-FNN outperform conventional STA and TRFI estimators by at least 15 dB gain in terms of SNR for BER = 10⁻³. Meanwhile, STA-FNN estimator outperforms DPA-FNN estimator by around 5 dB gain in terms of SNR for BER = 10⁻³. However, STA-FNN suffers from error floor beginning from SNR = 20 dB, particularly in very high mobility scenarios. This is attributed to the fact that conventional STA estimation outperforms DPA in low SNR region due to the frequency and time averaging operations that can alleviate the impact of noise and demapping error in low SNR regions. On the other hand, the averaging operations are not useful in high SNR regions since the impact of noise is low, and the STA averaging coefficients are fixed. Therefore, TRFI-FNN is used to improve the performance at high SNRs to compensate for the STA-FNN performance degradation in high SNR region. Importantly, STA-FNN and TRFI-FNN can be employed in an adaptive manner where STA-FNN and TRFI-FNN are used in low and high SNR regions, respectively.

For the LSTM-based estimators, employing LSTM as a preprocessing unit rather than a simple FNN in the channel estimation has shown to bring about a significant improvement in the overall performance. This is because LSTM is capable of efficiently learning the time correlations of the channel by taking the advantage of the previous output apart from the current input in order to estimate the current output.

TABLE 5. Simulation parameters of the IEEE 802.11p physical layer.

| Parameter | IEEE 802.11p |
|--------------------------------|--------------|
| Bandwidth | 10 MHz |
| Guard interval duration | 1.6 μs |
| Symbol duration | 8 μs |
| Short training symbol duration | 1.6 μs |
| Long training symbol duration | 6.4 μs |
| Total subcarriers | 64 |
| Pilot subcarriers | 4 |
| Data subcarriers | 48 |
| Subcarrier spacing | 156.25 KHz |

LSTM-FNN-DPA estimator [51] outperforms STA-FNN and TRFI-FNN estimators by approximately 4 dB gain in terms of SNR for BER = 10⁻³. However, this estimator is not impervious to high computational complexity, as discussed in the next section, due to the utilization of two DL networks, i.e, LSTM followed by FNN. On the other hand, the LSTM-DPA-TA estimators performance gain in various scenarios can be explained by employing the TA processing, which significantly alleviates the noise impact aside from the strong ability of the LSTM in learning the channel time correlations compared with a simple FNN architecture. The LSTM-DPA-TA estimator outperforms the LSTM-FNN-DPA estimator by around 4 dB gain in terms of SNR for BER = 10⁻⁴. When adopting high modulation order (16QAM), the LSTM-DPA-TA estimator outperforms the other estimators by at least 7 dB and 3 dB gains in terms of SNR for BER = 10⁻³ in high as well as very high mobility scenarios, respectively, as illustrated in Figure 8b.

2) MOBILITY

The degraded performance with the increased mobility of all the studied schemes can be observed from Figure 8. However, the time diversity gain increases when there is an increase in the Doppler spread, as evidenced by comparing the case of the DL-based estimators in high mobility ($f_d = 500$) and very high mobility ($f_d = 1000$). This behavior can be explained by the ability of DL networks to reduce the estimation error stemming from the AWGN noise and the DPA demapping error. By contrast, the net gain from the time diversity is influenced by the AWGN noise

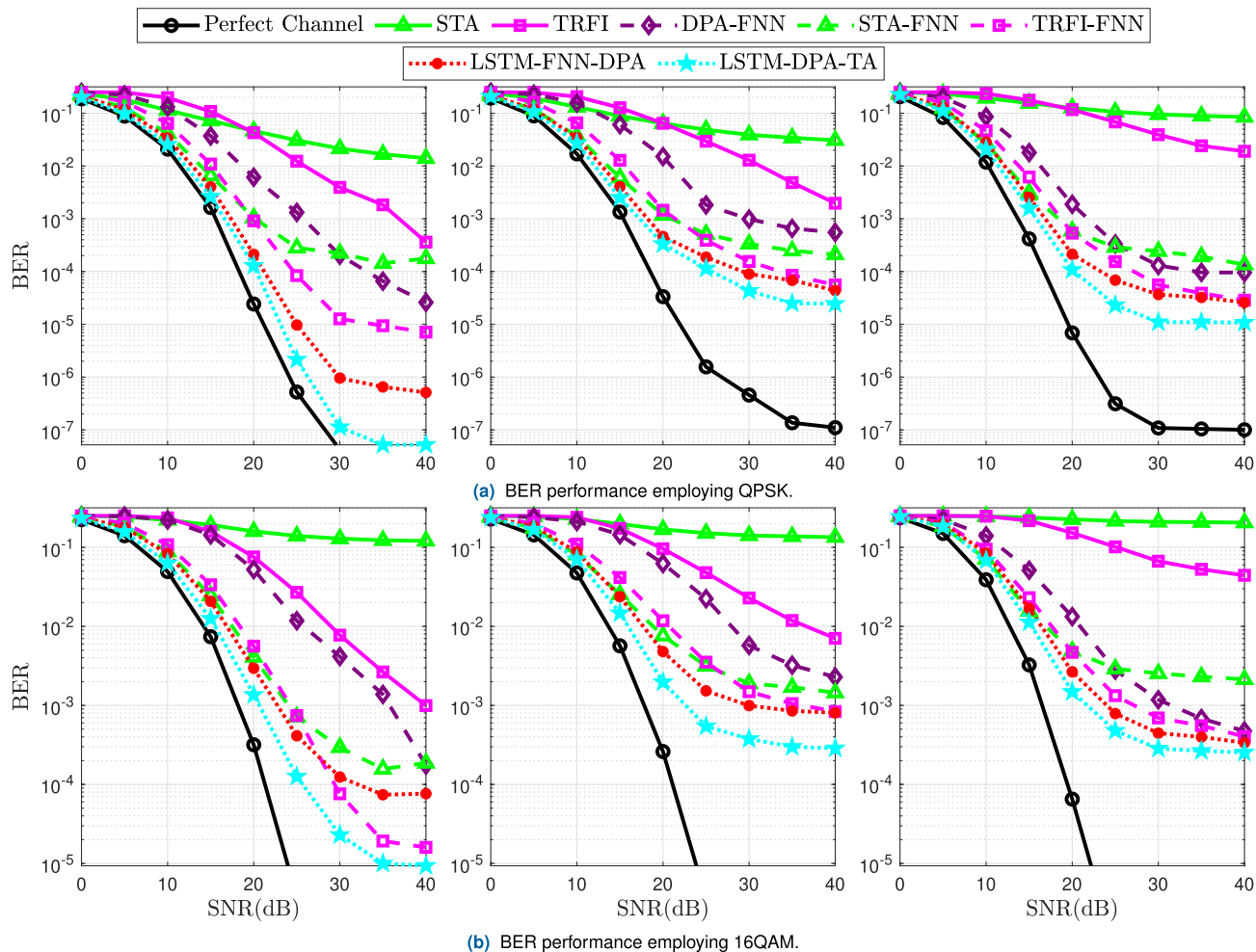


FIGURE 8. BER for $I = 100$, mobility from left to right: low ($v = 45$ Kmph, $f_d = 250$ Hz), high ($v = 100$ Kmph, $f_d = 500$ Hz), very high ($v = 200$ Kmph, $f_d = 1000$ Hz).

and DPA demapping error, as is the case in conventional SBS estimators. The performance degradation is attributed as the mobility increases since the impact of the AWGN noise and DPA demapping error is much more dominant than the time diversity gain. This observation is also valid for high modulation orders such as 16QAM, as evidenced in Figure 8b.

3) FRAME Length

The impact of frame length is illustrated in Figure 10. As can be seen, the performance of the conventional estimators strongly depends on the frame length, given that employing short frame $I = 10$ results in a negligible accumulated DPA demapping error. By contrast, the DL-based estimators are found to be more robust against the changes in the employed frame length. However, in the case of a long frame ($I = 100$), the performance gain of the DL-based estimators is significantly remarkable. This behavior is mainly attributed to the time diversity negligible gain when short frame is employed and vice versa.

To conclude, it can be surmised that increasing the frame length increases the time diversity gain. Additionally, the codeword becomes longer with a longer frame. Therefore, the time diversity is capable of compensating for Doppler error, particularly in very high mobility scenario as illustrated in Figure 10

4) DL ARCHITECTURE

The DPA-FNN estimator integrates three hidden layer FNN in addition to the conventional DPA estimation with $40 - 20 - 40$ neurons. However, as can be observed in Figure 8, correcting the estimation error of the DPA estimation is insufficient even after the inclusion of more neurons in the FNN hidden layers, because it merely corrects the demapping error, neglecting the received symbols' requery and time correlation. Meanwhile, the STA-FNN and TRFI-FNN estimators have better optimized three hidden layers FNN architecture where $15 - 15 - 15$ neurons are used. Consequently, the overall computational complexity is considerably lowered when compared to the DPA-FNN, while attaining performance superiority. This is due to the fact that

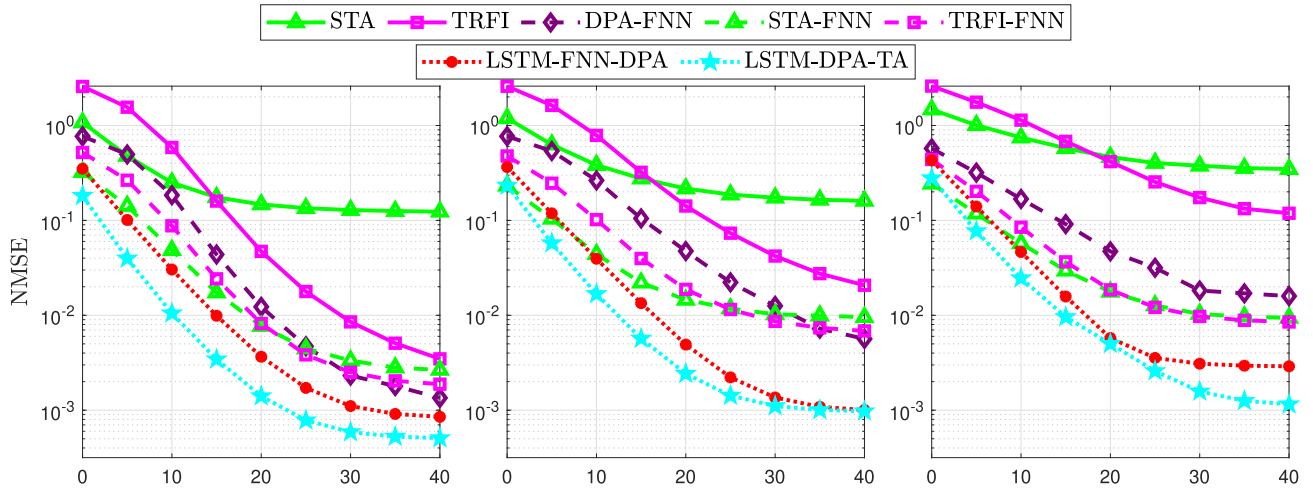


FIGURE 9. NMSE for $I = 100$, mobility from left to right: low ($v = 45$ Km/h, $f_d = 250$ Hz), high ($v = 100$ Km/h, $f_d = 500$ Hz), very high ($v = 200$ Km/h, $f_d = 1000$ Hz).

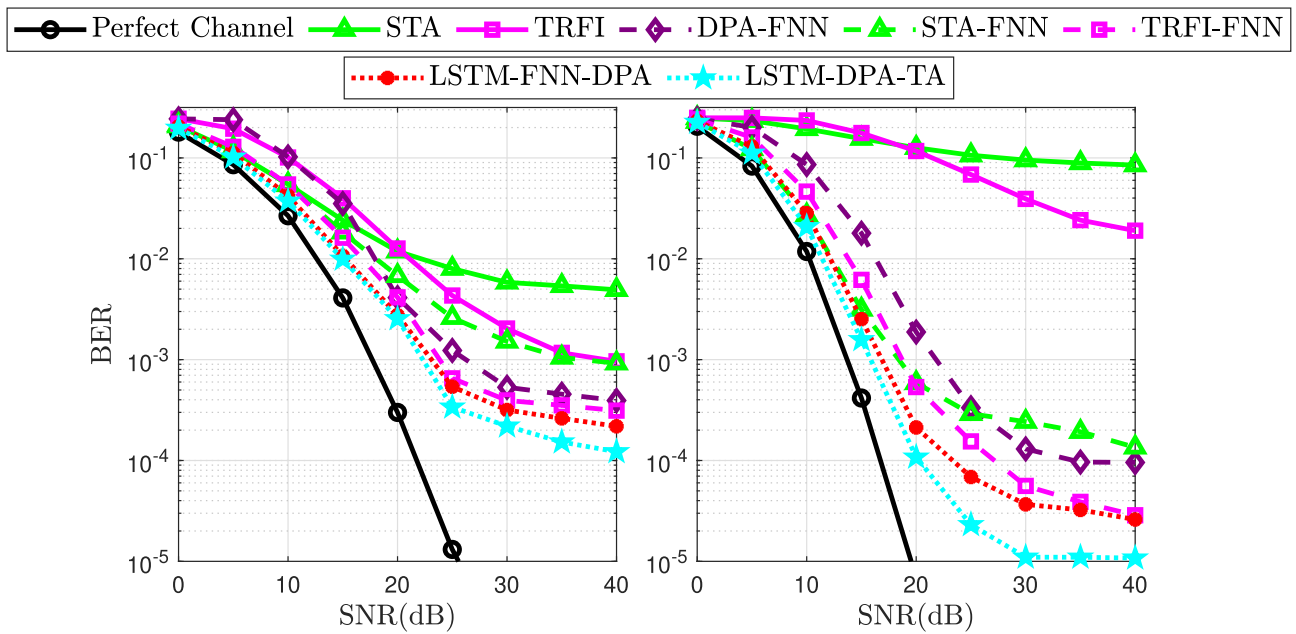


FIGURE 10. BER for QPSK, very high mobility ($v = 200$ Km/h, $f_d = 1000$ Hz) from left to right: $I = 10$, $I = 100$.

STA considers frequency as well a time correlation between the received OFDM symbols, while the conventional TRFI estimator employs frequency-domain cubic interpolation to make further improvements in the DPA estimation.

Therefore, it can be concluded that the pre-estimation should be good enough in order for the FNN processing to be more useful. Put differently, with an increased accuracy of the pre-estimation, low-complexity FNN architecture can be taken advantage of while recording a significant performance gain. On the contrary, if the pre-estimation is poor, employing FNN processing with high-complexity architecture results in a limited performance gain while increasing the overall computational complexity. As is the case with LSTM-based estimators, employing the TA processing in the

LSTM-DPA-TA estimator to ameliorate the AWGN noise impact results in a less complex architecture in comparison to the LSTM-FNN-DPA estimator, where two DL networks are employed.

C. DL-BASED FBF ESTIMATION SCHEME

1) MODULATION ORDER

Figure 11 illustrates the BER performance of the studied DL-Based FBF estimators employing QPSK and 16QAM modulation orders. The 2D LMMSE uses the channel and noise statistics in the estimation, thus leading to comparable performance in terms of the ideal case. However, the 2D-LMMSE is ridden with high computational complexity.

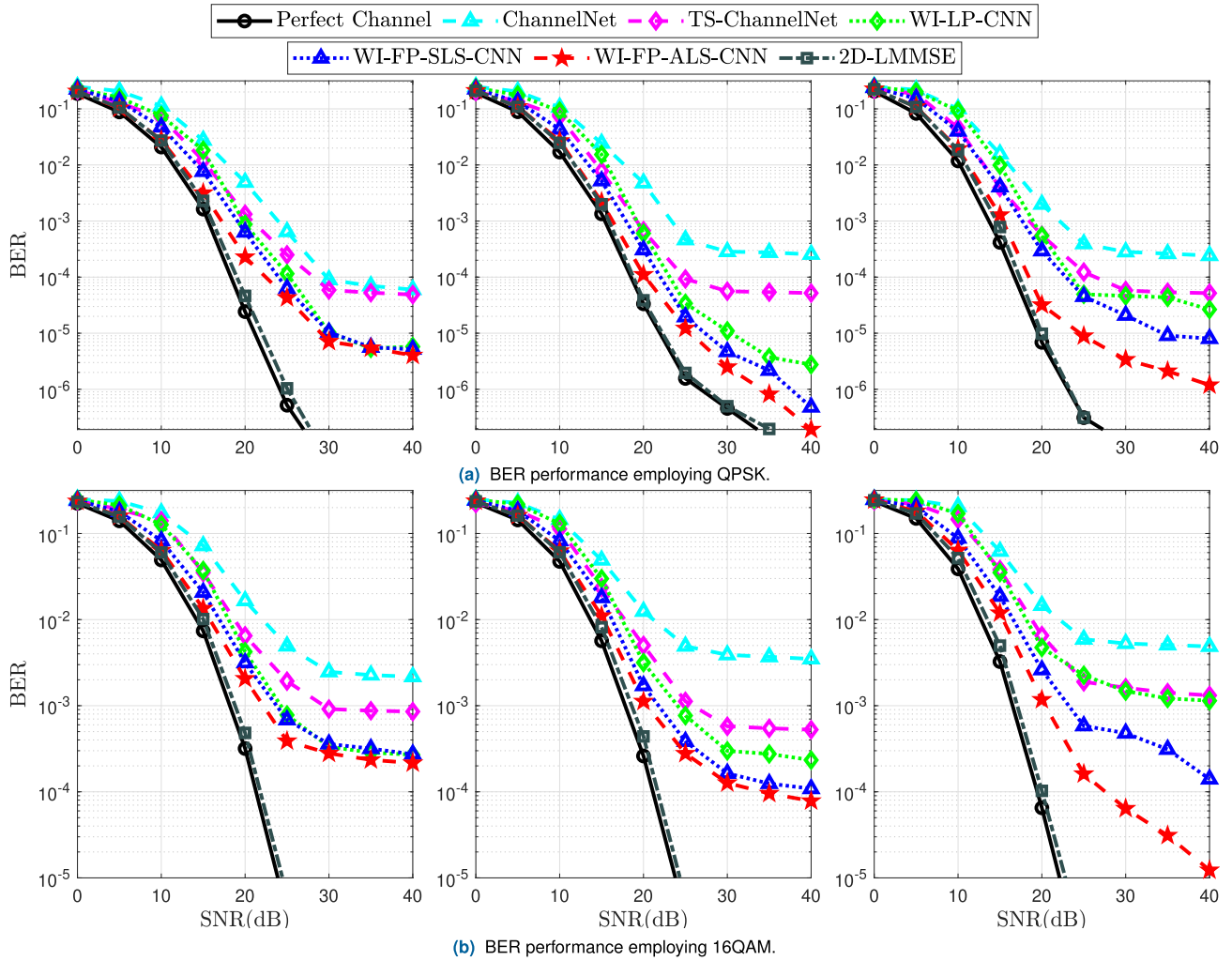


FIGURE 11. BER for $l = 100$, mobility from left to right: low ($v = 45$ Kmph, $f_d = 250$ Hz), high ($v = 100$ Kmph, $f_d = 500$ Hz), very high ($v = 200$ Kmph, $f_d = 1000$ Hz). The CNN refers to SR-CNN and DN-CNN in low and high/very high) mobility scenarios, respectively.

Moreover, the significant BER performance superiority of the WI-CNN estimators can be observed where FP-ALS-CNN outperforms the ChannelNet as well as TS-ChannelNet estimators by at least 6 dB and 3 dB gain in terms of SNR for a $BER = 10^{-3}$. Importantly, ChannelNet and TS-ChannelNet estimators suffer from a considerable performance degradation that is dominant in very high mobility scenarios. This is because their performance accounts for the predefined fixed parameters in the applied interpolation scheme, where it is important to update the RBF interpolation function and the ADD-TT frequency and time averaging parameters in real-time. Furthermore, the ADD-TT interpolation employs only the previous and the current pilot subcarriers for the channel estimation at each received OFDM symbol. By contrast, there are no fixed parameters in the WI-CNN estimators. The time correlation between the previous and the future pilot symbols is considered in the WI interpolation matrix (51), whereas the estimated channel is considered in the overall estimation at all channel taps. These aspects lead to the superior performance of WI-CNN estimators performance, where

a significant robustness is shown against high mobility with varied performance gain according to the employed pilot allocation scheme, i.e FP or LP. In addition, WI-CNN estimators employ optimized SR-CNN and DN-CNN in accordance with the mobility condition, wherein SR-CNN is utilized in low mobility scenarios, whereas, DN-CNN is employed in high and very high mobility scenarios.

2) MOBILITY

A degradation is observed in the overall performance of ChannelNet and TS-ChannelNet estimators as the mobility increases, while the WI-CNN estimators reveal a robustness against high mobility, as illustrated in Figure 11. This is primarily attributed to the accuracy of the WI interpolation, coupled with optimized SR-CNN and DN-CNN. Although CNN processing is implemented in the ChannelNet and TS-ChannelNet, this post CNN processing is unable to perform well due to the high estimation error of the 2D RBF and ADD-TT interpolation techniques in the initial estimation.

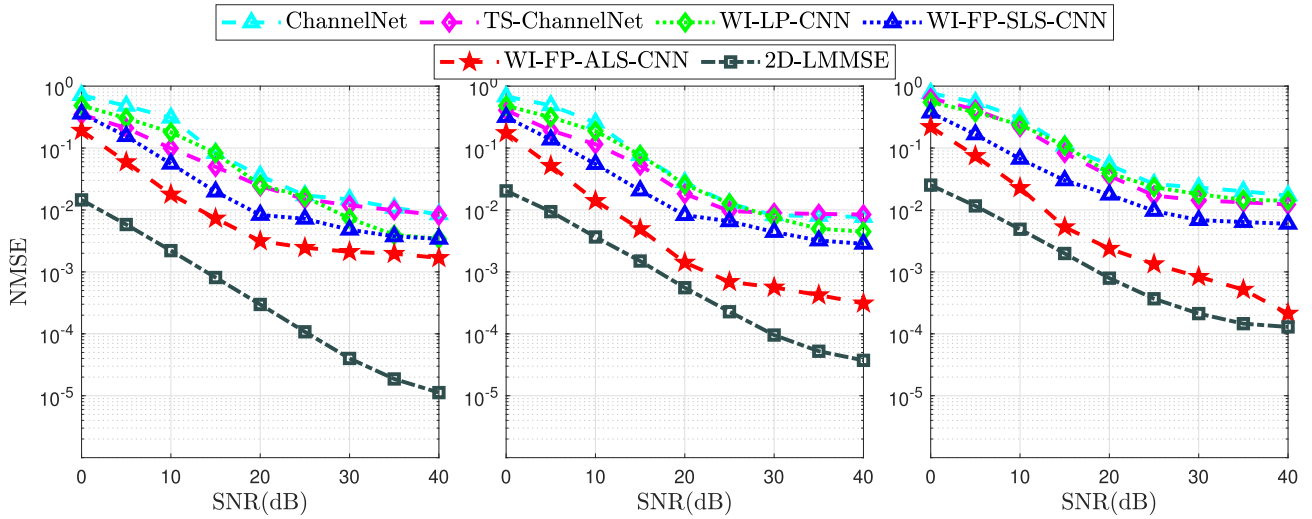


FIGURE 12. NMSE for $l = 100$, mobility from left to right: low ($v = 45$ Kmph, $f_d = 250$ Hz), high ($v = 100$ Kmph, $f_d = 500$ Hz), very high ($v = 200$ Kmph, $f_d = 1000$ Hz).The CNN refers to SR-CNN and DN-CNN in low and high/very high) mobility scenarios, respectively.

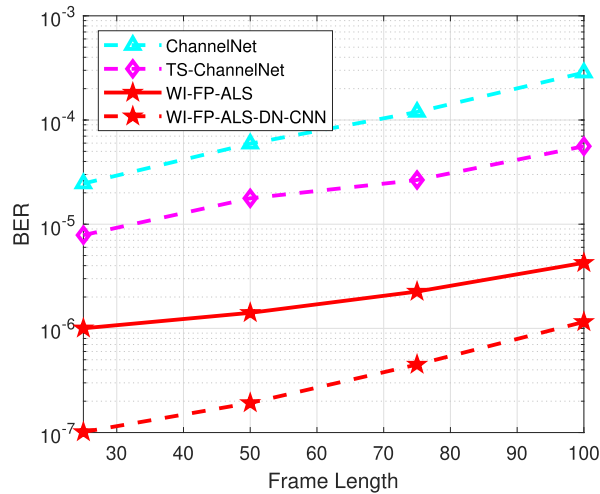


FIGURE 13. BER performance of VTV-SDWW high mobility vehicular channel model employing QPSK modulation and different frame lengths.

Therefore, it can be concluded that employing robust initial estimation as the WI interpolation schemes allows the CNN to better learn the channel correlation with lower complexity, thereby enhancing the channel estimation.

3) FRAME LENGTH

Figure 13 illustrates the BER performance of high mobility vehicular scenario employing QPSK modulation and different frame lengths. As can be clearly observed, the WI-FP-ALS estimator outperforms ChannelNet and TS-ChannelNet for different frame lengths without any post CNN processing. This is because of the long codeword that shows the robustness of the WI-FP-ALS estimator, unlike the 2D RBF and ADD-TT interpolation techniques that suffer from a significant estimation error even when considering a short frame. Moreover, employing the optimized DN-CNN after

the WI-FP-ALS estimator significantly enhances the BER performance.

However, although CNN processing is applied in the ChannelNet and TS-ChannelNet, this post CNN processing is unable to perform well. This is attributed to the high estimation error of the 2D RBF and ADD-TT interpolation techniques in the initial estimation. Thus, we can conclude that employing robust initial estimation as the WI interpolation schemes enable the CNN to better learn the channel correlation with lower complexity, thereby enhancing the channel estimation, as shown in Figure 12.

D. CNN ARCHITECTURE

The ChannelNet estimator employs SR-CNN and DN-CNN following the 2D RBF interpolation. The employed SR-CNN comprises three convolutional layers with ($v_1 = 9; f_1 = 64$), ($v_2 = 1, f_2 = 32$) and ($v_3 = 5, f_3 = 1$), respectively. Moreover, the DN-CNN depth is $D = 18$ with $3 \times 3 \times 32$ kernels in each layer. Meanwhile, SR-ConvLSTM network comprises three ConvLSTM layers of ($v_1 = 9; f_1 = 64$), ($v_2 = 1, f_2 = 32$) and ($v_3 = 5, f_3 = 1$), respectively, and is integrated after the ADD-TT interpolation in the TS-ChannelNet estimator. The SR-ConvLSTM network combines both the CNN and the LSTM networks [49], thus increasing the overall computational complexity, as shall be discussed later. By contrast, the employed optimized SR-CNN and DN-CNN significantly reduces the complexity due to the WI estimators' accuracy. Put succinctly, the complexity of the employed CNN decreases as the accuracy of the pre-estimation increases, because low-complexity architectures can be utilized and vice versa.

E. DL-BASED SBS VS. DL-BASED FBFB ESTIMATION SCHEME

This section further examines the performance assessment of the studied estimators, where only the best DL-based SBS and

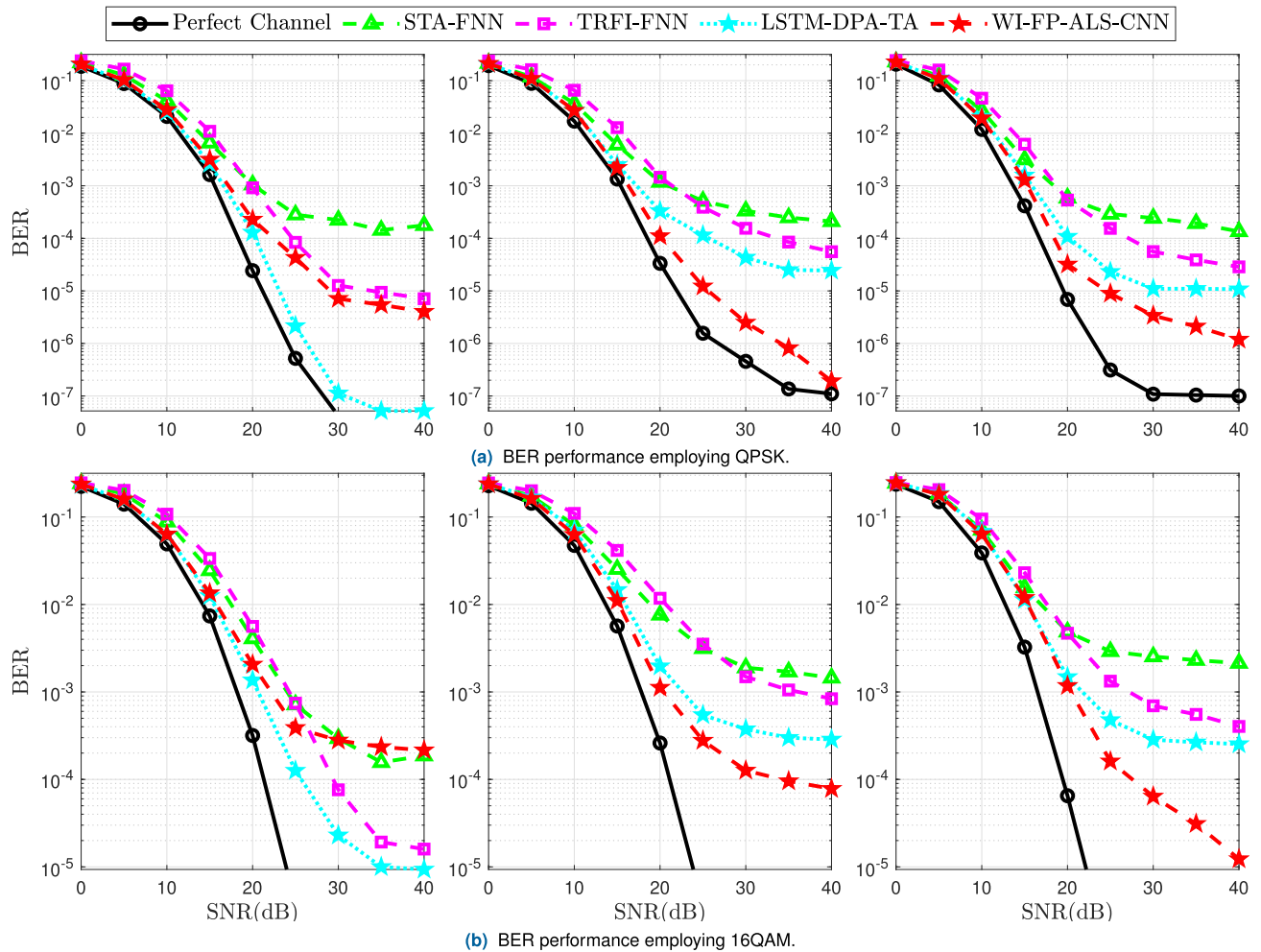


FIGURE 14. BER performance employing three scenarios: (i) first column - low mobility ($v = 45$ Km/h, $f_d = 250$ Hz) (ii) second column - high mobility ($v = 100$ Km/h, $f_d = 500$ Hz) (iii) third column - very high mobility ($v = 200$ Km/h, $f_d = 1000$ Hz). The CNN refers to SR-CNN and DN-CNN in low and high/very high mobility scenarios, respectively.

FBF estimators are compared. Figures 14 and 15 illustrate the BER and NMSE performance of the investigated DL-based estimators in low, high, and very high mobility scenarios, employing QPSK and 16QAM modulation orders.

In low-mobility scenario, the LSTM-DPA-TA SBS estimator outperforms the WI-FP-ALS-SR-CNN FBF estimator. This can be explained by the ability of LSTM to better learn the channel time correlation than the SR-CNN, since Doppler error is somehow negligible in low mobility scenario. However, in high and very high mobility scenarios, WI-FP-ALS-DN-CNN shows a significantly improved performance, outperforming the LSTM-DPA-TA SBS estimator by 3 dB gain in terms of SNR for a BER = 10^{-4} . In high mobility scenarios, where the Doppler error impact is high, LSTM suffers from some performance degradation as learning the time correlation between successive samples is not achievable in the low mobility scenario case. Meanwhile, DN-CNN network can significantly alleviate the impact of noise and Doppler error, where it records at least 5 dB gain in terms of SNR for a BER = 10^{-4} . To conclude, it can

be inferred that employing LSTM network rather than FNN and DN-CNN networks leads to improved performance in low-mobility scenarios. In contrast, DN-CNN is more useful in high as well as very high mobility scenarios because DN-CNN uses the entire pilot subcarriers within the received frame. To summarize, the time correlation between successive received OFDM symbols decreases as the mobility increases. Therefore, the performance of LSTM suffers from performance degradation when compared with CNN. On the other hand, the CNN-based estimators become more useful than the LSTM-based estimators in high mobility scenarios.

Finally, it is observed that DL-based FBF estimators suffer from high buffering time at the receiver, because it is necessary to receive the full frame before the channel estimation begins leading to high latency. However, this buffering time is lowered in the WI-CNN estimators after dividing the received frame into sub frames so that the channel estimation process commences prior to the full frame reception. Moreover, the WI-CNN estimators also help increase the transmission data rate as fewer pilots are inserted into the transmitted frame.

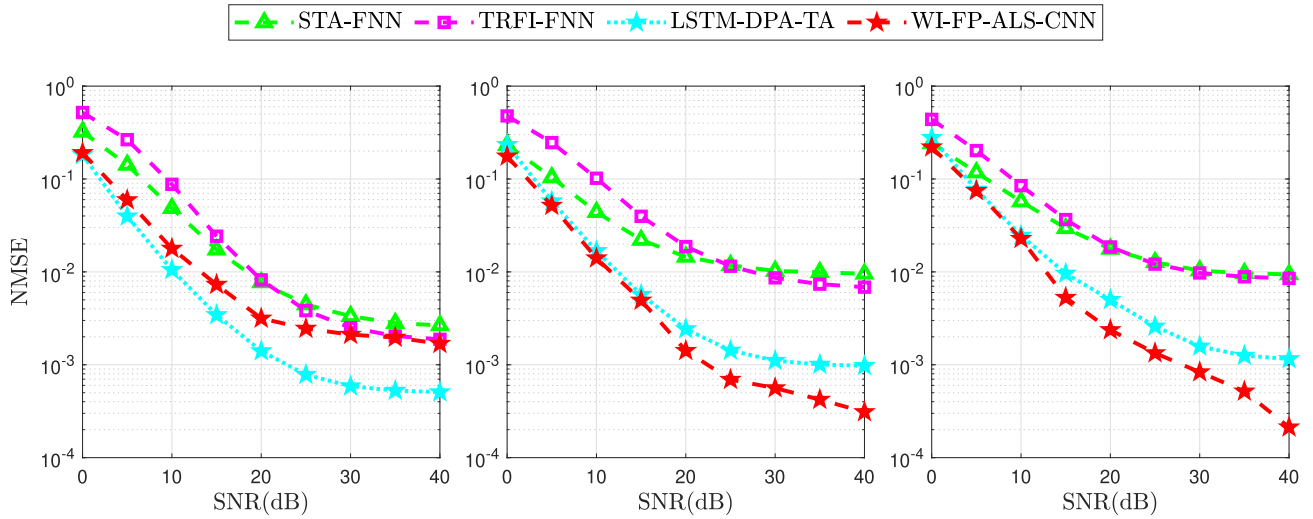


FIGURE 15. NMSE performance employing three scenarios: (i) first column - low mobility ($v = 45$ Kmph, $f_d = 250$ Hz) (ii) second column - high mobility ($v = 100$ Kmph, $f_d = 500$ Hz) (iii) third column - very high mobility ($v = 200$ Kmph, $f_d = 1000$ Hz). The CNN refers to SR-CNN and DN-CNN in low and high/very high mobility scenarios, respectively.

VII. COMPLEXITY ANALYSIS

This section provides a detailed computational complexity analysis of the studied DL-based SBS and FBF estimators. The computational complexity analysis is performed in accordance with the number of real-valued arithmetic operations, multiplication/division and summation/subtraction necessary to estimate the channel for one received OFDM frame. Each complex-valued division requires 6 real-valued multiplications, 2 divisions, 2 summations, and 1 subtraction. In addition, each complex-valued multiplication is performed by 4 real-valued multiplications and 3 summations.

A. DL-BASED SBS ESTIMATORS

The DPA estimation implemented in the DL-based SBS estimators as an initial step needs two equalization steps (23), and (25). Each equalization step comprises K_{on} complex-valued divisions. Moreover, it needs the LS estimated channel at the preamble computed by $2K_{on}$ summation and $2K_{on}$ divisions. Hence, the overall computational complexity of the DPA estimation is $16K_{on}$ multiplications/divisions and $6K_{on}$ summations/subtractions.

The STA estimator applies frequency as well as time-domain averaging in addition to DPA. The frequency-domain averaging (26) coefficient is fixed ($\beta = 2$). Thus, each sub-carrier requires 5 complex-valued summations multiplied by a real-valued weight, which, in turn, are equivalent to 10 real-valued summations, and 2 real-valued multiplications. Consequently, the STA frequency-domain averaging step requires $10K_d$ real-valued summations, and $2K_d$ real-valued multiplications. The STA time-domain averaging step (27) requires $4K_{on}$ real-valued divisions, and $2K_{on}$ real-valued summations. For this reason, the accumulated overall computational complexity of STA estimator is $22K_{on} + 2K_d$ multiplications/divisions and $10K_{on} + 10K_d$ summations/subtractions.

The TRFI estimator implements another two equalization steps after the DPA estimation (28). Thereafter, it applies

cubic interpolation as the last step. Based on the analysis performed in [27], the computational complexity of TRFI is $34K_{on} + 26K_{int}$ multiplications/divisions and $14K_{on} + 30K_{int}$ summations/subtractions, where K_{int} represents the number of unreliable subcarriers in each received OFDM symbol.

1) FNN-BASED ESTIMATORS

For the FNN-based estimators, the DPA-FNN architecture [26] consists of three hidden layers with $J_1 = J_5 = 2K_{on}$, $J_2 = J_4 = 40$, and $J_3 = 20$ neurons, respectively. Therefore, the DPA-FNN requires $4K_{on}J_2 + 2J_2J_3$ multiplications, and $2K_{on} + 2J_2 + J_3$ summations. The computational complexity of LS and the DPA estimation are accumulated for DPA-FNN computational complexity resulting in total of $178K_{on} + 1600$ multiplications and $168K_{on} + 1600$ summations/subtractions.

$$CC_{FNN} = 2 \sum_{l=0}^{L+1} N_{l-1}N_l, \text{ where } N_0 = N_{L+1} = 2K_{on}. \quad (54)$$

The STA-FNN and TRFI-FNN estimators employ a three-hidden layer FNN architecture consisting of 15 neurons each. This FNN architecture requires $4K_{on}J_2 + 2J_2^2$, and $2K_{on} + 3J_2$ summations. This architecture is less complex when compared with the DPA-FNN one. Thus, the STA-FNN overall computational complexity is $82K_{on} + 2K_d + 450$ multiplications, and $70K_{on} + 10K_d + 450$ summations/subtractions. Furthermore, the TRFI-FNN needs $94K_{on} + 26K_{int} + 450$ multiplications, and $74K_{on} + 30K_{int} + 450$ summations/subtractions. The TRFI-FNN estimator reduces the number of multiplications as well as summations by 48% and 56%, respectively, when compared with DPA-FNN, while its computational complexity is similar to that of STA-FNN.

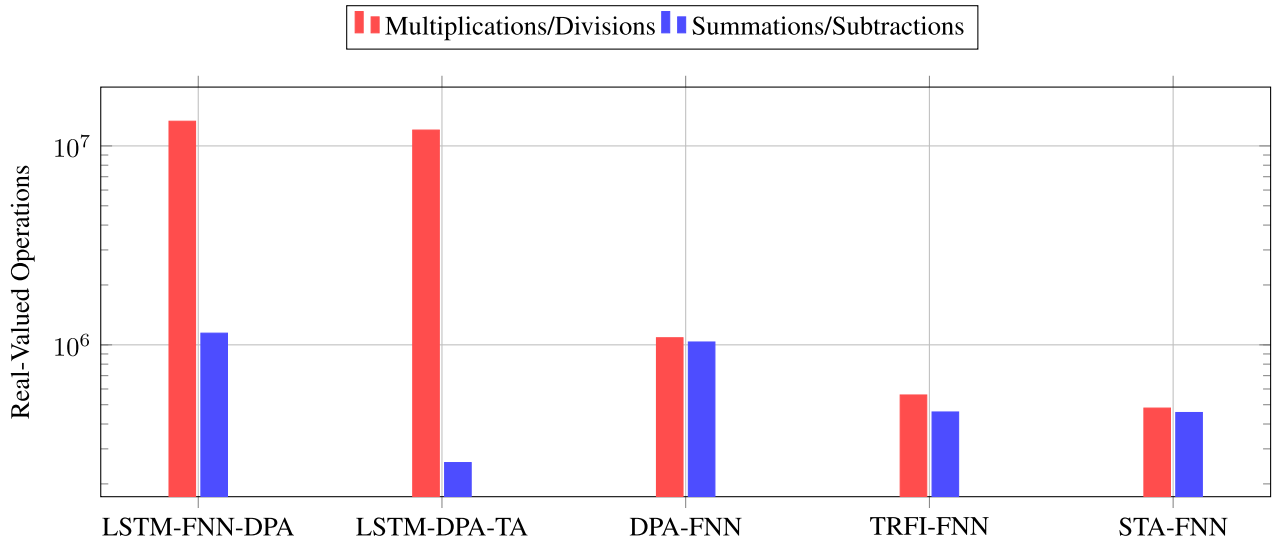


FIGURE 16. Computational complexity comparison of the studied DL-based SBS estimators.

TABLE 6. Detailed computation complexity of the studied DL-based SBS estimators.

| Estimator | Mul./Div. | Sum./Sub. |
|------------------------------|---|---|
| FNN(J_2 - J_3 - J_4) | $2K_{on}J_2 + J_2J_3 + J_3J_4 + 2K_{on}J_4$ | $2K_{on}J_2 + J_2J_3 + J_3J_4 + 2K_{on}J_4$ |
| LSTM (P) | $P^2 + 3P + PK_{in}$ | $4P + K_{in} - 2$ |
| Overall channel estimation | | |
| STA-FNN | $82K_{on} + 2K_d + 450$ | $70K_{on} + 10K_d + 450$ |
| TRFI-FNN | $94K_{on} + 26K_{int} + 450$ | $74K_{on} + 30K_{int} + 450$ |
| DPA-FNN | $178K_{on} + 1600$ | $168K_{on} + 1600$ |
| LSTM-FNN-DPA | $512K_{in} + 98K_d + 71040$ | $4K_{in} + 88K_d + 6776$ |
| LSTM-DPA-TA(64) | $514K_{on} + 18K_d + 16576$ | $10K_{on} + 8K_d + 824$ |
| LSTM-DPA-TA(128) | $1026K_{on} + 18K_d + 65920$ | $10K_{on} + 8K_d + 1656$ |

2) LSTM-BASED ESTIMATORS

The computational complexity of the LSTM unit can be calculated with respect to the number of operations performed by its four gates. Each gate applies $P^2 + PK_{in}$ real-valued multiplications and $3P + K_{in} - 2$ summations apart from $3P$ multiplications, and P summations required by (19), and (21). As a result, the overall computational complexity for the LSTM becomes

$$CC_{LSTM} = 4(P^2 + PK_{in} + 3P + K_{in} - 2) + 4P. \quad (55)$$

Notably, FNN-based estimators need less computation than LSTM, thus achieving lower complexity. The LSTM-FNN-DPA estimator employs one LSTM unit with $P = 128$ and $K_{in} = 112$, followed by one hidden layer FNN network with $N_1 = 40$ neurons. In addition, the LSTM-FNN-DPA estimator implements the DPA estimation that requires $18K_d$ real-valued multiplication/division and

$8K_d$ summation/subtraction. Thus, the overall computational complexity of the LSTM-FNN-DPA estimator is $512K_{in} + 98K_d + 71040$ multiplication/division and $4K_{in} + 88K_d + 6776$ summation/subtraction.

The LSTM-DPA-TA utilizes one LSTM unit with $P = 128$ as LSTM-FNN-DPA estimator. It also uses $K_{in} = 2K_{on}$, and applies TA as a noise alleviation technique to the $\hat{h}_{LSTM-DPA_{i,d}}$ estimated channel, that requires only $2K_{on}$ real-valued multiplication/division and $2K_{on}$ summation/subtraction. Hence, the LSTM-DPA-TA estimator requires $4P^2 + P(8K_{on} + 3) + 18K_d + 2K_{on}$ real-valued multiplication/division and $13P + 10K_{on} + 8K_d - 8$ summation/subtraction. As per this analysis, the LSTM-DPA-TA estimator achieves less computational complexity in comparison to the LSTM-FNN-DPA estimator. It records 9.73% and 77.63% computational complexity decline in the required real-valued multiplication/division and summation/subtraction, respectively.

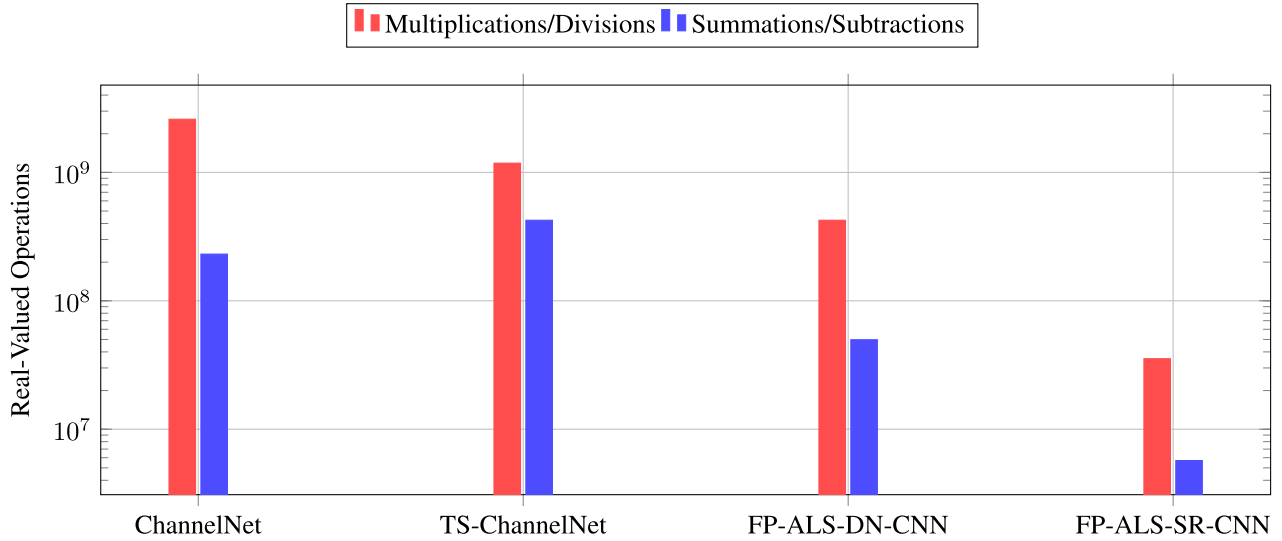


FIGURE 17. Computational complexity comparison of the studied DL-based FBF estimators.

Importantly, replacing the FNN network by the TA to achieve noise alleviation is the primary factor in reducing the overall computational complexity. Moreover, the LSTM-DPA-TA estimator outperforms the LSTM-FNN-DPA estimator while recording lower computational complexity. As a matter of fact, employing the LSTM-DPA-TA LSTM-based estimators as opposed to the FNN-based estimators results in 89.10% and 62.18% increase in the necessary multiplication/division and summation/subtraction, respectively. Nevertheless, it is possible to achieve a significant performance gain. Table 6 and Figure 16 reveal a detailed summary of the computational complexities for the various examined DL-based SBS estimators.

B. DL-BASED FBF ESTIMATORS

1) ChannelNet ESTIMATOR

The ChannelNet estimator utilizes the RBF interpolation followed by SR-CNN and DN-CNN networks. Therefore, the overall computational complexity of the ChannelNet estimator can be expressed as follows

$$CC_{\text{ChannelNet}} = CC_{\text{RBF}} + CC_{\text{SR-CNN}} + CC_{\text{DN-CNN}}. \quad (56)$$

The calculation of $\hat{\mathbf{H}}_{\text{LS}}$ requires $2K_p I$ divisions. The computation of \mathbf{w}_{RBF} requires $4K_p^2 I^2$ multiplications/divisions and $5K_p^2 I^2 - 2K_p I$ summations/subtractions. Meanwhile, $\hat{\mathbf{H}}_{\text{RBF}}$ requires $K_d I(K_p^2 I^2 + 3K_p I)$ multiplications/divisions and $5K_d K_p I^2$ subtractions/summations. Thus, the total computational complexity of the RBF interpolation can be expressed by $K_p^2 I^2(4 + K_d I) + K_p I(2 + 3K_d I)$ multiplications/divisions and $K_p I(5K_p I + 5K_d I - 2)$ summations/subtractions. Subsequently, the ChannelNet estimator applies SR-CNN followed by DN-CNN in addition to the RBF interpolation.

$CC_{\text{SR-CNN}}$ and $CC_{\text{DN-CNN}}$ can be computed as follows

$$\begin{aligned} CC_{\text{SR-CNN}} &= \sum_{l=1}^{\mathcal{L}} h_l w_l d_l v_l^2 f_l + h_l w_l d_l f_l \\ &= \sum_{l=1}^{\mathcal{L}} h_l w_l d_l f_l (v_l^2 + 1). \end{aligned} \quad (57)$$

$$CC_{\text{DN-CNN}} = \sum_{l=1}^{\mathcal{L}} h_l w_l d_l f_l (v_l^2 + 1) + \sum_{j=1}^D 4h_j w_j d_j. \quad (58)$$

\mathcal{L} signifies the number of employed CNN layers. It can be noted that the second term in $CC_{\text{DN-CNN}}$ signifies the number of operations required by the batch normalization employed in the DN-CNN network. Thus, the SR-CNN employed in the ChannelNet estimator needs $16064K_{\text{on}} I$ multiplications/divisions as well as $4288K_{\text{on}} I$ summations/subtractions, while the ChannelNet DN-CNN computations require $334080K_{\text{on}} I$ multiplications/divisions and $38144K_{\text{on}} I$ summations/subtractions.

2) TS-ChannelNet ESTIMATOR

The TS-ChannelNet estimator applies the ADD-TT interpolation followed by the SR-ConvLSTM network. Hence, the overall computational complexity of the TS-ChannelNet estimator can be expressed in the following manner:

$$CC_{\text{TS-ChannelNet}} = CC_{\text{ADD-TT}} + CC_{\text{SR-ConvLSTM}}. \quad (59)$$

The ADD-TT interpolation first applies the DPA estimation requiring $18K_{\text{on}}$ multiplications/divisions and $8K_{\text{on}}$ summations/subtractions. The time-domain truncation operation applied in (43) requires $4LK_{\text{on}}$ multiplications as well as $5K_{\text{on}} L - 2K_{\text{on}}$ summations. In the ADD-TT interpolation, the frequency-domain averaging (44) requires $10K_{\text{on}}$ summations and $2K_{\text{on}}$ multiplications. Furthermore, the time-domain averaging step (45) requires $4K_{\text{on}}$ real valued divisions,

TABLE 7. Detailed computation complexity of the studied CNN-based FBF estimators.

| Scheme | Interpolation | | CNN | |
|---------------|--|------------------------------------|-------------------|------------------|
| | Mul./Div. | Sum./Sub. | Mul./Div. | Sum./Sub. |
| ChannelNet | $K_p^2 I^2 (4 + K_d I)$ $+ K_p I (2 + 3K_d I)$ | $K_p I (5K_p I$ $+ 5K_d I - 2)$ | $350144K_{on} I$ | $42432K_{on} I$ |
| TS-ChannelNet | $24K_{on} I + 4LK_{on} I$ | $18K_{on} I$ $+ 5K_{on} IL$ | $226880K_{on} I$ | $81472K_{on} I$ |
| FP-SLS-SR-CNN | $2K_{on} P + 2K_{on}$ $+ 4K_{on} I_d$ | $2K_{on}$ $+ 2K_{on} I_d$ | $7008K_{on} I_d$ | $1120K_{on} I_d$ |
| FP-ALS-SR-CNN | $4K_{on}^2 P + 2K_{on} P$ $+ 2K_{on} + 4K_{on} I_d$ | $5K_{on}^2 P$ $+ 2K_{on} I_d$ | | |
| LP-SR-CNN | $2LP + 4K_{on} LP$ $+ 2K_{on} + 4K_{on} I_d$ | $5K_{on} LP$ $+ 2K_{on} I_d$ | | |
| FP-SLS-DN-CNN | $2K_{on} P + 2K_{on}$ $+ 4K_{on} I_d$ | $2K_{on}$ $+ 2K_{on} I_d$ | $84096K_{on} I_d$ | $9856K_{on} I_d$ |
| FP-ALS-DN-CNN | $4K_{on}^2 P + 2K_{on} P$ $+ 2K_{on} + 4K_{on} I_d$ | $5K_{on}^2 P$ $+ 2K_{on} I_d$ | | |
| LP-DN-CNN | $2LP + 4K_{on} LP + 2K_{on}$ $+ 4K_{on} I_d$ | $5K_{on} LP$ $+ 2K_{on} I_d$ | | |

and $2K_{on}$ real valued summations. Thus, the overall computational complexity of the ADD-TT interpolation for the whole received OFDM frame requires $24K_{on} I + 4LK_{on} I$ real-valued multiplications/divisions, and $18K_{on} I + 5K_{on} IL$ summations/subtractions. The total computational complexity is expressed with respect to the overall operations implemented in the input, forget, and output gates of the SR-ConvLSTM network, such that

$$CC_{ConvLSTM} = \sum_{l=1}^{\mathcal{L}} h_l w_l d_l f_l (8v_l^2 + 30). \quad (60)$$

Based on (60), the SR-ConvLSTM network employed in the TS-ChannelNet estimator requires $226880K_{on} I$ multiplications/divisions as well as $81472K_{on} I$ summations/subtractions. TS-ChannelNet estimator is less complicated than the ChannelNet estimator, because it employs only one CNN in addition to the ADD-TT interpolation, unlike the ChannelNet estimator where both SR-CNN and DN-CNN are employed.

3) WI-CNN ESTIMATORS

The WI-CNN estimators computational complexity primarily depends on the selected frame structure, the pilot allocation scheme, as well as the selected optimized CNN. Thus, the overall computational complexity of the WI-CNN estimators

can be expressed as follows

$$CC_{WI} = CC_{\hat{H}_{WI}} + CC_{O-CNN}. \quad (61)$$

When full pilot symbols are inserted, two options are taken into consideration. The first option is the SLS estimator, which is performed using $2K_{on} P + 2K_{on}$ divisions, and $2K_{on}$ summations. The second option entails employing the ALS estimator with $2K_{on} P + 2K_{on}$ divisions. This is followed by $4K_{on}^2 P$ multiplications, and $5K_{on}^2 P$ summations. In the instance where $K_p = L$ pilots are inserted with each pilot symbol, the LS estimation requires $2LP + 2K_{on}$ divisions, $4K_{on} LP$ multiplications, and $5K_{on} LP$ summations. In a similar manner, for employing only $K_p = 4$ pilot subcarriers, the WI-CP estimator needs $8P + 2K_{on}$ divisions, $16K_{on} P$ multiplications, as well as $20K_{on} P$ summations. Following the selection of the required frame structure and pilot allocation scheme, the WI-CNN estimators apply the weighted interpolation as demonstrated in (50). The channel estimation for each received OFDM frame needs $4K_{on} I_d$ divisions and $2K_{on} I_d$ summations. Finally, the optimized SR-CNN is utilized in low-mobility scenario and needs $7008K_{on} I_d$ multiplications/divisions and $1120K_{on} I_d$ summations/subtractions. For high-mobility scenarios, the optimized DN-CNN is employed, requiring $84096K_{on} I_d$ multiplications/divisions and $9856K_{on} I_d$ summations/subtractions.

The WI-FP-ALS records the higher computational complexity among the other WI estimators in all mobility scenarios, due to W_{ALS} calculation in (47), whereas, the WI-FP-SLS estimator refers to the simplest one.

Table 7 shows the studied estimators' overall computational complexity with respect to real valued operations. It is noteworthy that the WI estimators achieve significant computational complexity decrease in comparison to ChannelNet and TS-ChannelNet estimators. Figure 17 depicts the computational complexity of the studied DL-based FBF estimators. The ChannelNet and TS-ChannelNet estimators are 70 and 39 times more complex than the FP-ALS-SR-CNN, respectively. In addition, the WI-CNN estimators achieve a minimum of 7027.35 times less complexity than the 2D LMMSE estimator, with an acceptable BER performance, which makes them a feasible alternative to the 2D LMMSE. It is also observed that FP-ALS-DN-CNN is 12 times more complex than FP-ALS-SR-CNN since the optimized DN-CNN architecture complexity employed in high and very high scenarios is higher than the optimized SR-CNN architecture, which, in turn, is employed in low mobility scenarios.

VIII. CONCLUSION

This survey sheds light on the recently proposed DL-based SBS and FBF channel estimators in doubly-dispersive environments. First, we have defined the problem of signal propagation in doubly-dispersive channels. Subsequently, a review of different DL architectures employed in the doubly-dispersive channel estimation has been undertaken, followed by a detailed presentation of the studied DL-based estimators. Finally, the studied estimators have been evaluated with respect to NMSE, BER, and computational complexity, clearly demonstrating a significant improvement of employing DL in the channel estimation across different mobility conditions. We have shown that, while the LSTM and CNN based estimators do outperform the FNN based estimator, more computational complexity is necessary where the LSTM-based SBS estimator is 23.6 times more complex than the FNN-based SBS estimators. Nevertheless, the complexity of the CNN-based FBF estimator exceeds the complexity of LSTM-based SBS estimator by approximately 3450 times because of the significant difference in terms of required operations between the CNN and LSTM networks. Finally, we have observed that the choice of the channel estimator is primarily related to the applications requirements as well as affordable computational complexity. SBS estimators are more useful when the application is sensitive to latency, whereas FBF estimators can be employed if some latency can be accepted. To summarize, a trade-off between the required performance, computational complexity, and the accepted latency must first be defined to select what is the most suitable channel estimator to be employed.

REFERENCES

- [1] H. Chang, C.-X. Wang, Y. Liu, J. Huang, J. Sun, W. Zhang, and X. Gao, "A novel nonstationary 6G UAV-to-ground wireless channel model with 3-D arbitrary trajectory changes," *IEEE Internet Things J.*, vol. 8, no. 12, pp. 9865–9877, Jun. 2021.
- [2] L. Wang, B. Ai, K. Guan, D. He, Z. Zhong, L. Tian, and J. Dou, "Stochastic channel modeling for high-speed railway viaduct scenario at 93.2 GHz," in *Proc. 12th Eur. Conf. Antennas Propag. (EuCAP)*, 2018, pp. 1–4.
- [3] B. Turan and S. Coleri, "Machine learning based channel modeling for vehicular visible light communication," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 9659–9672, Oct. 2021.
- [4] R. Chen, W. Yang, F. Wu, and M. Sun, "Fast handover for high-speed railway via NDN," in *Proc. 1st IEEE Int. Conf. Hot Inf.-Centric Netw. (HotICN)*, Aug. 2018, pp. 167–172.
- [5] S. Wang and Q. Zhang, "A joint time-frequency domain frequency offset estimation algorithm for burst communication," in *Proc. IEEE 3rd Int. Conf. Electron. Technol. (ICET)*, May 2020, pp. 1–5.
- [6] T. Ma, X. Jiang, Y. Wang, and F. Li, "A novel inter-carrier interference cancellation scheme in highly mobile environments," *China Commun.*, vol. 17, no. 12, pp. 194–205, Dec. 2020.
- [7] K. Saito, Q. Fan, N. Keerativoranan, and J.-I. Takada, "4.9 GHz band outdoor to indoor propagation loss analysis in high building environment using unmanned aerial vehicle," in *Proc. 13th Eur. Conf. Antennas Propag. (EuCAP)*, Mar. 2019, pp. 1–4.
- [8] R. Bomfin, M. Chafii, A. Nimr, and G. Fettweis, "A robust baseband transceiver design for doubly-dispersive channels," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 4781–4796, Aug. 2021.
- [9] R. Bomfin, A. Nimr, M. Chafii, and G. Fettweis, "A robust and low-complexity Walsh–Hadamard modulation for doubly-dispersive channels," *IEEE Commun. Lett.*, vol. 25, no. 3, pp. 897–901, Mar. 2021.
- [10] A. Nimr, M. Chafii, M. Matthe, and G. Fettweis, "Extended GFDM framework: OTFS and GFDM comparison," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.
- [11] M. Chafii, J. Palicot, R. Gribonval, and F. Bader, "Adaptive wavelet packet modulation," *IEEE Trans. Commun.*, vol. 66, no. 7, pp. 2947–2957, Jul. 2018.
- [12] J. A. Fernandez, K. Borries, L. Cheng, B. V. Kumar, D. D. Stancil, and F. Bai, "Performance of the 802.11p physical layer in vehicle-to-vehicle environments," *IEEE Trans. Veh. Technol.*, vol. 61, no. 1, pp. 3–14, Jan. 2012.
- [13] Z. Zhao, M. Wen, C.-X. Wang, X. Cheng, and B. Jiao, "Channel estimation schemes for IEEE 802.11p standard," *IEEE Intel. Transport. Syst. Mag.*, vol. 5, no. 4, pp. 38–49, Oct. 2013.
- [14] Y.-K. Kim, J.-M. Oh, Y.-H. Shin, and C. Mun, "Time and frequency domain channel estimation scheme for IEEE 802.11p," in *Proc. 17th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2014, pp. 1085–1090.
- [15] S. Ehsanfar, M. Chafii, and G. P. Fettweis, "On UW-based transmission for MIMO multi-carriers with spatial multiplexing," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 5875–5890, Sep. 2020.
- [16] Y. Choi, J. H. Bae, and J. Lee, "Low-complexity 2D LMMSE channel estimation for OFDM systems," in *Proc. IEEE 82nd Veh. Technol. Conf. (VTC)*, Sep. 2015, pp. 1–5.
- [17] T. Wang, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, "Deep learning for wireless physical layer: Opportunities and challenges," *China Commun.*, vol. 14, no. 11, pp. 92–111, 2017.
- [18] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cognit. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec. 2017.
- [19] M. Chafii, F. Bader, and J. Palicot, "Enhancing coverage in narrow band-LoT using machine learning," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2018, pp. 1–6.
- [20] Y. Yang, F. Gao, X. Ma, and S. Zhang, "Deep learning-based channel estimation for doubly selective fading channels," *IEEE Access*, vol. 7, pp. 36579–36589, 2019.
- [21] X. Ma, H. Ye, and Y. Li, "Learning assisted estimation for time-varying channels," in *Proc. 15th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2018, pp. 1–5.
- [22] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.
- [23] J. Yuan, H. Q. Ngo, and M. Matthaiou, "Machine learning-based channel prediction in massive MIMO with channel aging," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 2960–2973, May 2020.
- [24] H. Kim, S. Kim, H. Lee, C. Jang, Y. Choi, and J. Choi, "Massive MIMO channel prediction: Kalman filtering vs. Machine learning," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 518–528, Jan. 2021.
- [25] C. Wu, X. Yi, Y. Zhu, W. Wang, L. You, and X. Gao, "Channel prediction in high-mobility massive MIMO: From spatio-temporal autoregression to deep learning," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 1915–1930, Jul. 2021.

- [26] S. Han, Y. Oh, and C. Song, "A deep learning based channel estimation scheme for IEEE 802.11p systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.
- [27] A. K. Gizzini, M. Chafii, A. Nimr, and G. Fettweis, "Deep learning based channel estimation schemes for IEEE 802.11p standard," *IEEE Access*, vol. 8, pp. 113751–113765, 2020.
- [28] A. K. Gizzini, M. Chafii, A. Nimr, and G. Fettweis, "Joint TRFI and deep learning for vehicular channel estimation," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2020, pp. 1–6.
- [29] M. Soltani, V. Pourahmadi, A. Mirzaei, and H. Sheikhzadeh, "Deep learning-based channel estimation," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 652–655, Apr. 2019.
- [30] X. Zhu, Z. Sheng, Y. Fang, and D. Guo, "A deep learning-aided temporal spectral channelNet for IEEE 802.11p-based channel estimation in vehicular communications," *EURASIP J. Wireless Commun. Netw.*, vol. 2020, no. 1, pp. 1–15, Dec. 2020.
- [31] A. K. Gizzini, M. Chafii, A. Nimr, R. M. Shubair, and G. Fettweis, "CNN aided weighted interpolation for channel estimation in vehicular communications," *IEEE Trans. Veh. Technol.*, vol. 70, no. 12, pp. 12796–12811, Dec. 2021.
- [32] H. A. Le, T. Van Chien, T. H. Nguyen, H. Choo, and V. D. Nguyen, "Machine learning-based 5G-and-beyond channel estimation for MIMO-OFDM communication systems," *Sensors*, vol. 21, no. 14, p. 4861, Jul. 2021. [Online]. Available: <https://www.mdpi.com/1424-8220/21/14/4861>
- [33] F. Tang, B. Mao, N. Kato, and G. Gui, "Comprehensive survey on machine learning in vehicular network: Technology, applications and challenges," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 2027–2057, 3rd Quart., 2021.
- [34] A. K. Gizzini. (2022). *DL-Based-Channel-Estimation-in-Doubly-Dispersive-Environments*. [Online]. Available: <https://github.com/abdulkarimgizzini/DL-based-Channel-Estimation-in-Doubly-Dispersive-Environments/>
- [35] G. Matz and F. Hlawatsch, "Chapter 1—Fundamentals of time-varying communication channels," in *Wireless Communications Over Rapidly Time-Varying Channels*, F. Hlawatsch and G. Matz, Eds. Cambridge, MA, USA: Academic Press, 2011, pp. 1–63. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780123744838000017>
- [36] N. D. Ricklin, "Time varying channels: Characterization, estimation, and detection," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. California, San Diego, CA, USA, 2010.
- [37] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Oct. 2015, doi: 10.1016/j.neunet.2014.09.003.
- [38] S.-I. Amari, "Backpropagation and stochastic gradient descent method," *Neurocomputing*, vol. 5, nos. 4–5, pp. 185–196, 1993.
- [39] S. De, A. Mukherjee, and E. Ullah, "Convergence guarantees for RMSProp and ADAM in non-convex optimization and an empirical comparison to Nesterov acceleration," 2018, *arXiv:1807.06766*.
- [40] S. Ruder, "An overview of multi-task learning in deep neural networks," 2017, *arXiv:1706.05098*.
- [41] A. K. Gizzini, M. Chafii, S. Ehsanfar, and R. M. Shubair, "Temporal averaging LSTM-based channel estimation scheme for IEEE 802.11p standard," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–7. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-03365697>
- [42] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space Odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017.
- [43] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–6.
- [44] A. K. Gizzini, "Advanced linear and deep learning based channel estimation techniques in doubly dispersive environments," M.S. thesis, Dept. Sci., Technol., Inf., Commun., Cergy Paris CY Univ., Cergy, France, Dec. 2021. [Online]. Available: <https://hal.archives-ouvertes.fr/tel-03482053>
- [45] M. Sun, Z. Song, X. Jiang, J. Pan, and Y. Pang, "Learning pooling for convolutional neural network," *Neurocomputing*, vol. 224, pp. 96–104, Feb. 2017.
- [46] Z. Zhang, "Derivation of backpropagation in convolutional neural network (CNN)," Dept. Comput. Eng., Univ. Tennessee, Knoxville, TN, USA, Tech. Rep., 2016, pp. 23–30, vol. 22.
- [47] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2015.
- [48] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
- [49] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W. K. Wong, and W. C. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 1–9.
- [50] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [51] J. Pan, H. Shan, R. Li, Y. Wu, W. Wu, and T. Q. Quek, "Channel estimation based on deep learning in vehicle-to-everything environments," *IEEE Commun. Lett.*, vol. 25, no. 6, pp. 1891–1895, Jun. 2021.
- [52] A. K. Gizzini, M. Chafii, A. Nimr, and G. Fettweis, "Enhancing least square channel estimation using deep learning," in *Proc. IEEE 91st Veh. Technol. Conf. (VTC-Spring)*, May 2020, pp. 1–5.
- [53] F. J. Pontes, G. F. Amorim, P. P. Balestrassi, A. P. Paiva, and J. R. Ferreira, "Design of experiments and focused grid search for neural network parameter optimization," *Neurocomputing*, vol. 186, pp. 22–34, Apr. 2016.
- [54] C. Chatfield, *Time-Series Forecasting*. Boca Raton, FL, USA: CRC Press, 2000.
- [55] Y. R. Zheng and C. Xiao, "Channel estimation for frequency-domain equalization of single-carrier broadband wireless communications," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 815–823, Feb. 2009.
- [56] G. Acosta-Marum and M. A. Ingram, "Six time- and frequency-selective empirical channel models for vehicular wireless LANs," *IEEE Veh. Technol. Mag.*, vol. 2, no. 4, pp. 4–11, Dec. 2007.
- [57] G. Acosta-Marum, "Measurement, modeling, and OFDM synchronization for the wideband mobile-to-mobile channel," Ph.D. dissertation, Dept. Elect. Comput. Eng., Georgia Inst. Technol., Atlanta, GA, USA, 2007.



ABDUL KARIM GIZZINI (Member, IEEE) received the B.E. and M.E. degrees in computer and communication engineering from the IUL University of Lebanon, in 2015 and 2017, respectively, and the Ph.D. degree in telecommunications engineering from CY Cergy Paris Université, France, in 2021. His master thesis was hosted by the Lebanese National Council for Scientific Research (CNRS). He is currently a Postdoctoral Researcher with ETIS (UMR8051), which is a joint research unit between ENSEA, CY University, and CNRS, France. His research interests include artificial intelligence for wireless communications and image processing.



MARWA CHAFII (Member, IEEE) received the master's degree in advanced wireless communication systems (SAR) and the Ph.D. degree in electrical engineering from CentraleSupélec, France, in 2013 and 2016, respectively. From 2014 and 2016, she was a Visiting Researcher at the Poznań University of Technology, Poland, the University of York, U.K., Yokohama National University, Japan, and the University of Oxford, U.K. She joined as a Research Group Leader with the Technical University of Dresden, Germany, in 2018, and as an Associate Professor with ENSEA, France, in 2019, where she held the Chair of Excellence in artificial intelligence of the CY Initiative. Since September 2021, she has been an Associate Professor with New York University (NYU) Abu Dhabi, and NYU WIRELESS, NYU Tandon School of Engineering. Her research interests include advanced waveform design, machine learning for wireless communications, and indoor localization. She received the Prize of the Best Ph.D. in signal, image, and vision in France, and she has been nominated in the top ten Rising Stars in Computer Networking and Communications by N2Women, in 2020. She was an Associate Editor at IEEE COMMUNICATIONS LETTERS, from 2019 to 2021, where she received the Best Editor Award, in 2020. From 2018 to 2021, she was a Research Lead at Women in AI Organization. She is currently an Associate Editor with the IEEE TRANSACTIONS ON COMMUNICATIONS and serving as the Vice-Chair for the IEEE ComSoc ETI on Machine Learning for Communications and for leading the Education Working Group of the IEEE ComSoc ETI on Integrated Sensing and Communications.

...