Machine Learning Physical-Layer Receivers for DFT-s-OFDM

Jaakko Pihlajasalo, Dani Korpi, *Member, IEEE*, Taneli Riihonen, *Senior Member, IEEE*, Jukka Talvitie, *Member, IEEE*, and Mikko Valkama, *Fellow, IEEE*

***Abstract*—In this work, we propose two alternative machine learning (ML) based physical-layer receivers for discrete Fourier transform (DFT) precoded multicarrier systems. The receivers are designed to provide reliable soft bit estimates under two coexisting transmitter impairments, namely, power amplifier nonlinear distortion and oscillator phase noise. Considering data modulations up to 256-QAM, realistic 5G NR uplink evaluations at 28 GHz demonstrate large performance gains compared to reference receivers, especially when the ML receiver has trainable layers both before and after the inverse DFT. Furthermore, compared to a baseline linear receiver, we show an uplink coverage increase of over 20% with the proposed fully learned receiver.**

***Index Terms*—5G, 6G, coverage, DFT-s-OFDM, deep learning, distortion, physical-layer, power-efficiency, RF impairments.**

I. Introduction

Artificial intelligence (AI) and machine learning (ML) have proven to be efficient methods for learning various physical-layer algorithms directly from data [1], [2], [3], [4]. In fact, it is envisioned that the complete physical-layer of the future 6G radio systems could even be built around ML-based solutions, meaning that 6G could have a completely AI-native air interface [5], [6]. One benefit of ML-based algorithms is their inherent resilience against various hardware impairments, as long as the dominant impairments are properly represented in the training data [7], [8], [9], [10], [11]. This is especially useful for the millimeter-wave (mmWave) bands at frequency range 2 (FR2), where one main challenge is producing sufficient transmit power levels with high enough signal quality [12], [13]. What is more, the orthogonal frequency-division multiple access (OFDM) waveform is known to suffer from high peak-to-average-power ratio (PAPR) [14], and thus from poor power amplifier (PA) efficiency. Hence, the 5G New Radio (NR) uplink supports also a reduced-PAPR waveform variant called discrete Fourier transform (DFT)-spread OFDM, or DFT-s-OFDM, for coverage enhancement purposes [13]. This is also the main technical premise of this article, with specific emphasis on improving the uplink power-efficiency and coverage in future networks through advanced ML-based receivers deployed at base-station.

In general, the use of ML for physical-layer processing in pure OFDM based networks has been rather broadly studied in the literature, see, e.g., [3], [10] and the references therein. However, the corresponding literature on ML-based physical-layer processing in *DFT-s-OFDM* based uplink systems is more scarce. To this end, in [8] and [9], ML-based symbol-to-bit demapping is proposed for DFT-s-OFDM receivers to perform soft bit computations in the presence of phase noise or PA distortion, respectively. These models seek to tolerate the impairment effects present after an ordinary linear equalizer and output the log-likelihood ratios (LLRs) corresponding to different bits. It is shown that the models achieve certain performance gains in terms of block error rate (BLER) under fixed impairment and channel profiles. Furthermore, more complete DFT-s-OFDM physical-layer receiver solutions are proposed in [15] and [16]. The work in [15] reports a dense network approach to recover data symbols prior to the inverse DFT (IDFT). However, the work also neglects the PA nonlinear distortion, and the dense network structure limits the design to one resource grid configuration, unlike the convolutional layers, which can in principle support arbitrary bandwidths. Finally, in [16], a complete neural network receiver solution with embedded IDFT is described with focus on short-range sub-THz experimental system. Consequently, the work does not consider scenarios with higher mobility and neglects PA distortion.

In this work, we propose and describe novel ML-based receiver algorithms for DFT-s-OFDM based mmWave cellular networks under the presence of two dominant and co-existing RF impairments of user equipment (UE) transmitters, namely, the PA distortion and oscillator phase noise. Specific emphasis is on cellular network coverage improvements, hence the receiver systems are trained to detect properly even heavily distorted signals with error vector magnitude (EVM) levels beyond the current 5G specifications [17]. This allows one to push the transmitter power amplifiers much closer to their saturation power level, which will enhance both the power-efficiency of the PAs as well as the resulting effective isotropic radiated power (EIRP) of the transmitter. Compared to state-of-the-art in [8], [9], [15], [16], our receiver design and training procedures allow for co-existing impairments while also generalizing across different PA samples and varying levels of nonlinear distortion, phase noise characteristics and mobile radio channel profiles. We also offer actual uplink coverage analysis. Additionally, and importantly, we harness the power of convolutional neural networks (CNNs) and allow for trainable layers both before and after the IDFT in the receiver. The proposed CNN structure can in principle support arbitrary resource grid configurations, unlike the prior works, e.g., in [15]. Finally, compared to [2] and [3], we focus explicitly on DFT-s-OFDM and craft receiver solutions that utilize efficiently the natural fast Fourier transform (FFT) and inverse FFT (IFFT) of classical receivers, being then further complemented with the learned CNN layers.

It is shown through 5G NR standard-compatible uplink evaluations at 28 GHz that the proposed ML-based receivers will be able to detect the signal accurately, despite the ensuing high level of distortion. This will result in higher coverage and/or throughput of the radio link. The proposed receivers are also explicitly shown to outperform the relevant existing state-of-the-art in [8], [16], particularly, when the full-complexity variant of the proposed receivers is considered, containing neural network layers both before and after the IDFT.

The rest of this paper is organized as follows. The basic system model with the considered RF impairments is stated in Section II. The proposed ML receiver schemes, training methods and data generation are, in turn, described in Section III. Numerical results and their analysis are provided in Section IV, complemented with notes and discussion on processing complexity. The conclusions are drawn in Section V.

Received 20 April 2024; revised 12 November 2024; accepted 19 March 2025. Date of publication 1 April 2025; date of current version 15 August 2025. This work was supported in part by Nokia Corporation, in part by Business Finland under the 6G POWER project, and in part by the Research Council of Finland under Grant 345654 and Grant 357730. The review of this article was coordinated by Dr. Kezhi Wang. *(Corresponding author: Mikko Valkama.)*

Jaakko Pihlajasalo, Taneli Riihonen, Jukka Talvitie, and Mikko Valkama are with the Department of Electrical Engineering, Tampere University, 33100 Tampere, Finland (e-mail: mikko.valkama@tuni.fi).

Dani Korpi is with Nokia Bell Labs, 37007 Espoo, Finland.

This article has supplementary downloadable material available at https://doi.org/10.1109/TVT.2025.3556843, provided by the authors.

Digital Object Identifier 10.1109/TVT.2025.3556843

II. System Model

For generality, we consider an uplink multiple-input multiple-output (MIMO) transmission scenario that contains both digital baseband (BB) precoding as well as RF beamforming, and describe all the signal models using discrete-time baseband equivalents. To this end, assuming S streams and L_{TX} transmit chains, the signal in the ith antenna branch after the RF beamforming stage can be expressed as

$$x_{\text{RF},i}(n) = \sum_{l=1}^{L_{\text{TX}}} w_{\text{TX},i,l}^{\text{RF}} e^{\theta_{\text{TX},l}(n)} \sum_{s=1}^{S} w_{\text{TX},l,s}^{\text{BB}}(n) * x_{\text{DFT},s}(n) \quad (1)$$

where $*$ denotes convolution, $x_{\text{DFT},s}(n)$ is the DFT-s-OFDM waveform corresponding to the sth spatial layer, and $\theta_{\text{TX},l}(n)$ is the phase noise sequence of TX chain l. We denote the DFT-s-OFDM waveform related basic IFFT and FFT sizes by N and N_{sc}, respectively, where N_{sc} also refers to the number of active subcarriers. For notational simplicity, the BB precoding is here expressed in time-domain while it can in practice be implemented in the subcarrier domain. The corresponding antenna signal reads then

$$x_{\text{ANT},i}(n) = \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} a_{p,i}(n) * \left(|x_{\text{RF},i}(n)|^{p-1} x_{\text{RF},i}(n) \right) \quad (2)$$

where $a_{p,i}(n)$ is the pth order response of the assumed memory-polynomial PA model [18].

The antenna signals propagate over the MIMO channel, with the individual antenna-to-antenna channels being both time- and frequency-selective. Assuming $N_{\text{ANT}}^{\text{TX}}$ and $N_{\text{ANT}}^{\text{RX}}$ transmit and receive antennas, respectively, the resulting signal at the kth receiver branch after the receiver RF beamforming reads

$$y_{\text{RX},k}(n) = \sum_{j=1}^{N_{\text{ANT}}^{\text{RX}}} w_{\text{RX},k,j}^{\text{RF}} \left(\sum_{i=1}^{N_{\text{ANT}}^{\text{TX}}} h_{i,j}(n,\tau) * x_{\text{ANT},i}(n) + n_j(n) \right) \quad (3)$$

where $n_j(n)$ denotes additive white Gaussian noise (AWGN) at receiver j while $h_{i,j}(n,\tau)$ is the time-varying radio channel, influenced by the Doppler effect, from TX antenna i to RX antenna j – hence the convolution in (3) is interpreted for the channel delay variable τ. For further details on channel models and Doppler modeling, we refer to [19].

In general, the ultimate uplink coverage of the network is dictated by the single-layer ($S = 1$) transmission performance. Additionally, mmWave UE implementations mostly build on phased-arrays with $L_{\text{TX}} = 1$ or $L_{\text{TX}} = 2$. Hence, in the following, we will focus on the fundamentally important single-layer mmWave phased-array scenario for practical relevance.

III. Proposed ML Receivers and Data Generation

A. Preprocessing and Conventional Receivers

We assume that the fundamental FFT and IDFT sizes in the DFT-s-OFDM receiver are N and N_{sc}, respectively, while the number of multicarrier symbols per transmission time interval (TTI) or slot is N_{symb}. Considering then only the active subcarriers at the output of the FFT, the resulting frequency-domain data within one slot or a TTI can be stacked into a matrix $\mathbf{Y}_{\text{f}} \in \mathbb{C}^{N_{\text{sc}} \times N_{\text{symb}}}$ for further receiver processing.

In classical multicarrier receivers, the frequency-domain data is processed with a linear equalizer – the linear minimum mean-squared error (LMMSE) equalizer being one of the most widely adopted approaches. The corresponding multipath channel estimation builds on the demodulation reference signals (DMRSs) through, e.g., least-squares (LS) model fitting [2]. In DFT-s-OFDM case, the equalizer output is followed by an additional IDFT, after which the LLRs are extracted for the channel decoder. The equalizer output after the IDFT is here denoted by $\mathbf{Y}_{\text{eq}}$. A high-level illustration of the receiver processing is given in Fig. 1, including also the ML receivers proposed in this work.

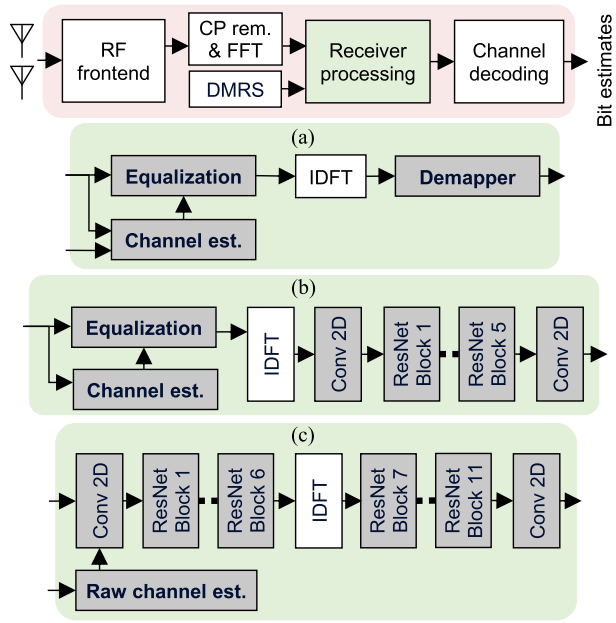

Fig. 1. High-level block diagram of the overall receiver chain, with different alternatives for the receiver processing: (a) conventional receiver, (b) the proposed light DFT-s CNN receiver, and (c) the proposed full DFT-s CNN receiver.

B. Proposed ML Receivers

We propose two ML-based receivers for DFT-s-OFDM signals impaired by phase noise and PA distortion, with the goal of detecting the transmitted bits during a single TTI. The first proposed receiver structure, the so-called *light DFT-s CNN receiver*, is designed to be lighter in depth and complexity, and considers bit detection from conventionally equalized impaired signals. The other proposed receiver system, referred to as the *DFT-s CNN receiver*, has more depth and complexity as it also considers channel estimation and equalization as part of the end-to-end learning task. Specifically, as elaborated more later in the paper, the amount of the trainable parameters in the DFT-s CNN receiver is four times that of the lighter variant.

The architectures of the proposed CNN receivers follow residual network (ResNet) structure [20] with high-level diagrams presented in Fig. 1. They are built with multiple ResNet blocks, which consist of convolutional layers with residual connections, with the number of filters ranging from 64 up to 256. Importantly, convolutional layers are suitable for multiple resource grid configurations, as they are not dependent on the input array size. For a more detailed description of a ResNet block, we refer the reader to [2].

The input to the light DFT-s CNN receiver is the post-IDFT equalized signal $\mathbf{Y}_{\text{eq}}$ transformed to real-valued array $\mathbf{Z}_{\text{eq}} \in \mathbb{R}^{N_{\text{sc}} \times N_{\text{symb}} \times 2}$ by concatenating real and imaginary parts in the third dimension. The main input for the full DFT-s CNN receiver is the received frequency-domain signal $\mathbf{Y}_{\text{f}}$. Since the full DFT-s CNN receiver embeds also implicitly

TABLE I
EVALUATION PARAMETERS FOR TRAINING AND VALIDATION

Parameter	Value	Randomization
Center frequency	28 GHz	None
Channel model	TDL-A to TDL-E	Uniform distribution
Delay spread	Up to 100 ns	Uniform distribution
Maximum Doppler shift	0–1500 Hz	Uniform distribution
PA type and model	Analog Devices HMC943APM5E; 17th order polynomial	Dithered polynomial parameters
PA induced EVM	0.5%–13.5% (64-QAM) 0.3%–4.4% (256-QAM) 0.5%–24.5% (QPSK)	Uniformly randomized PA backoff
Phase noise 3dB bandwidth	1–20 Hz 1–10 Hz (256-QAM)	Uniform distribution
Beamformed SNR	0–30 dB (64-QAM) 17.5–40 dB (256-QAM) 0–15 dB (QPSK)	Uniform distribution
Antenna counts ($N_{\text{ANT}}^{\text{TX}}$, $N_{\text{ANT}}^{\text{RX}}$)	8, 128	None
Beamforming type	DFT codebook	None
Channel bandwidth	50 MHz	None
Number of subcarriers (N_{sc})	792 subcarriers	None
FFT size (N)	1024	None
Subcarrier spacing	60 kHz	None
Maximum CP length (N_{CP})	104	None
TTI length (N_{symb})	14 OFDM symbols	None
DMRS configuration	2 OFDM symb. / TTI	None
Modulation schemes	64-QAM, 256-QAM and QPSK	None

the channel estimation and equalization, raw LS-based DMRS channel estimates $\widehat{\mathbf{H}} \in \mathbb{C}^{N_{\text{sc}} \times N_{\text{symb}}}$, where the non-DMRS resource elements are filled with zeroes, are also given as an additional input. The final input is a concatenation of the real and imaginary parts of these arrays, denoted by $\mathbf{Z}_{\text{f}} \in \mathbb{R}^{N_{\text{sc}} \times N_{\text{symb}} \times 4}$. Both proposed receivers output an array of LLRs, denoted by $\mathbf{L} \in \mathbb{R}^{N_{\text{sc}} \times N_{\text{symb}} \times B}$, where B is the modulation order.

The first CNN receiver has trainable layers only after the IDFT while the latter has trainable layers also before it. Moreover, the full DFT-s CNN receiver performs multiple parallel IDFTs within the model to avoid bottlenecks, similar to [3]. Especially, the output arrays of the previous layer are paired up to form complex arrays to perform separate IDFT for each pair, after which the complex arrays are separated back to real arrays. The total number of IDFTs performed is thus half the number of convolutional filters of the previous layer, our configuration having 64 parallel IDFTs.

Similar to how a conventional receiver performs channel equalization before the IDFT, the neural layers before the IDFTs seek to facilitate bit detection while tolerating the effects of the channel. The two-dimensional filters are also expected to build resilience to the inter-carrier interference (ICI) induced by the Doppler shift and phase noise, as they allow the neural network to simultaneously observe multiple ICI-affected neighboring subcarriers. The layers after the IDFT seek to provide tolerance against the nonlinear effects caused by the impairments while accurately detecting the bits and performing the mapping to LLRs.

C. Training and Data Generation

Training and validation data were generated with Matlab's 5G Toolbox [21] by simulating a standard-compatible DFT-precoded 5G physical uplink shared channel (PUSCH). Simulation and impairment parameters are stated and listed in Table I. In particular, we consider

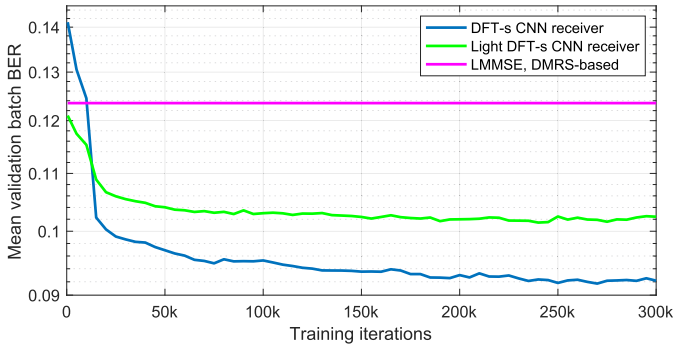

Fig. 2. Examples of mean validation batch BER of the proposed receivers with 64-QAM over the training iterations.

time-varying mobile radio channels defined in [19] and varying levels of impairments for the datasets, while TX and RX beamformers are always assumed to be directed towards the dominant path.

More specifically, to ensure that the neural networks generalize properly, we use separate tapped delay line (TDL) channels [19] and multiple different PA samples for training and validation. To this end, a 17th order polynomial is fit to real RF measurements from a 28 GHz PA module (HMC943APM5E), being then further varied and randomized by dithering its parameters to generate separate PA samples or instances for training and validation. The power backoff values of the PA instances are also deliberately varied, to allow the ML receivers to generalize to different nonlinear distortion conditions. As for the channel models, we use TDL-B, TDL-C, and TDL-D models for training and the remaining two for validation. For further details on utilized PA and TDL models, we refer to [3] and [19], while the true measurement-based polynomial coefficients are also provided as supplementary material along this correspondence. During training, the signal-to-noise ratio (SNR) is drawn from a random uniform distribution for each TTI, while the validation dataset is constructed with predefined SNRs over a uniform grid. For each considered modulation scheme, we generated 630 000 TTIs for training and 170 000 TTIs for validation.

The training loss is defined as the binary cross entropy between the transmitted and estimated bits for both ML receivers, conceptually similar to [2], which we refer to for more details. In Fig. 2, we illustrate the mean validation bit error rate (BER) for the proposed receivers over 300 000 training iterations, while showing also that of the DMRS-based LMMSE receiver for visual reference. This illustrates that the training characteristics of the proposed receivers are well-behaving with decreasing trends.

IV. NUMERICAL RESULTS AND ANALYSIS

The proposed receiver schemes are evaluated for 64-QAM, 256-QAM, and QPSK modulation schemes, with datasets employing varying levels of PA-induced nonlinear distortion and phase noise. For 64-QAM, we perform further analysis to support the architectural design of the DFT-s CNN receiver. For QPSK, we present results with varying levels of nonlinearity from which we also perform actual cellular network coverage analysis. Uncoded BER is considered the main performance criteria and we utilize a conventional LMMSE receiver with LS channel estimation as the baseline. We also note that based on our earlier observations in, for example, [2], analysis can be done equally well using either BER or BLER values – in particular when the main interest is on relative performance differences between the different receiver schemes. For comparison purposes, we include

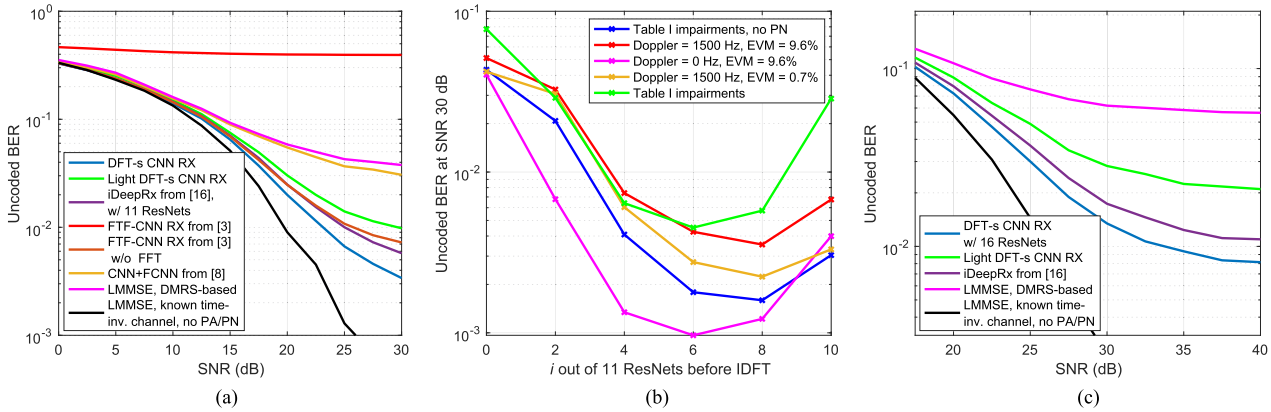

Fig. 3. Uncoded BER performance of (a) all the considered receivers with 64-QAM, (b) DFT-s CNN variants for IDFT location analysis and (c) a subset of the considered receivers with 256-QAM. In (a) and (c), the signals are affected by PA distortion and phase noise according to parameters in Table I. The results in (b) show the impact of the number of ResNets in the proposed DFT-s CNN receiver before and after IDFT in various impairment scenarios noted in the legend. (a) 64-QAM, all considered receivers, (b) 64-QAM, DFT-s CNN receiver, (c) 256-QAM, proposed receivers vs. iDeepRX.

an LMMSE receiver with known time-invariant channel without the nonlinear impairments, as an upper limit of the achievable performance. Moreover, it should be emphasized that the light DFT-s CNN receiver proposed herein can be considered an extended version of the ML-based receiver presented in [9]. The main architectural difference is the use of 2D convolutional filters and residual connections which facilitates the processing of equalized samples with a larger receptive field. This allows the neural network to learn local patterns affecting neighboring resource elements, which is not possible when using the models in [8], [9].

For completeness, however, we include in our validation and assessments the CNN+FCNN receiver scheme from [8], to show its performance in comparison to our proposed receivers. The CNN+FCNN consists of 5 fully-connected layers tasked with symbol-to-bit demapping, with a CNN providing additional inputs related to the impaired constellation. Compared with the FCNN in [9], the CNN+FCNN receiver in [8] has the additional CNN which allows for enhanced processing capabilities – thus the performance of [9] is not explicitly shown. Moreover, we include comparisons with iDeepRx, presented in [16], which also has layers on both sides of the IDFT. Compared to our model, the iDeepRx utilizes up to four IDFT transformations, as opposed to our 64 parallel IDFTs in the proposed DFT-s CNN receiver. The iDeepRx consists of DeepRx and Post-DeepRx nets with a total of 16 ResNet blocks. For fair comparisons, we have modified the iDeepRx to also have 11 ResNet blocks for the results with 64-QAM, to make the two receivers have similar number of trainable parameters. With 256-QAM, in turn, we include the original iDeepRx, while then adding further blocks also to the DFT-s CNN receiver such that both have 16 ResNet blocks.

In addition to the architectural differences, our evaluation parameters have major differences between the works presented in [8], [9], [16], most notably the variation in mobility and PA samples. The considered mobility of the user is a major difference between the works, with the Doppler induced ICI affecting the signals being significant with our assumptions. Likewise, the prior works do not consider any variation on the PA model, which we consider by dithering the PA models.

Finally, we also include comparisons with the FTF-CNN receiver from [3], designed originally for CP-OFDM, as well as its more DFT-s-OFDM compatible variant. Specifically, we modify the FTF-CNN receiver from [3] by removing the additional FFT from within the neural network, such that the output is in the data symbol domain. It is noted that, strictly-speaking, such receiver is not anymore among those proposed in [3], but is anyway included here in the comparisons for the purpose of completeness and fairness.

A. *Results With 64-QAM*

Fig. 3(a) shows the BER performance of all the receivers under PA distortion and coexisting phase noise according to Table I. The lower EVM values represent almost linear PA, while the higher EVM range is already beyond the current 5G specifications [17]. From the results, we see that both proposed receivers outperform the conventional LMMSE receiver. The performance gains can be explained by the post-IDFT layers providing resilience against the PA-induced nonlinear distortion and phase noise. The mutual performance difference between the two proposed receivers is due to the ML-based channel estimation and equalisation within the pre-IDFT layers, which makes the fully learned DFT-s CNN receiver more resilient against ICI and channel aging induced by higher Doppler shifts.

Additionally, these results also indicate that the proposed full DFT-s CNN receiver outperforms those proposed in [8], [16]. The CNN+FCNN from [8] also falls short of the BER achieved by the light DFT-s CNN receiver, which indicates the model's architectural inability to tolerate severe nonlinear distortion and ICI due to the input consisting only of a single resource element. Additionally, the 11 Resnet block variant of iDeepRx from [16] achieves performance in between those of the proposed receivers. This illustrates the benefits of the parallel IDFT transformations utilized by the DFT-s CNN receiver.

Furthermore, the poor performance of FTF-CNN receiver from [3] shows the importance of the IDFT transformation within the neural network. This is because the convolutional layers are not enough to unwrap the DFT. The proposed DFT-s CNN receiver also clearly outperforms the modified FTF-CNN receiver, which is not optimized for DFT-s-OFDM. Specifically, the SNR gain at 1% BER is around 2.5 dB in favor of the DFT-s CNN receiver. Thus, from now on, we focus on further quantification of only the proposed receivers, while still also comparing against the iDeepRx from [16] in the challenging case of 256-QAM in the forthcoming subsection.

To provide further insights on the architecture and number of layers needed for the proposed receivers, we performed additional analysis on the exact position of the IDFT within the ResNets of the DFT-s CNN receiver. For this analysis, we fixed the total number of ResNet blocks to 11 and trained a total of six variants of the DFT-s CNN receiver with different positions of the IDFT block. The ensuing receivers were then

evaluated for various cases as shown in Fig. 3(b), considering scenarios with varying levels of mobility, nonlinear distortion of PA and phase noise. Most notably, the results illustrate that:

- Enhanced resilience against a more challenging time-varying multipath channel and Doppler induced ICI benefits from more layers *before IDFT*;
- Improved tolerance of PA distortion and phase noise benefits with more layers *after IDFT*;
- Having layers *both before and after the IDFT* is largely beneficial to receiver performance.

These results indicate that 6 ResNet blocks before the IDFT and 5 after the IDFT is a favorable architecture when considering 11 ResNets in total. Hence, in the results presented herein (including already Fig. 3(a)), we have adopted this architecture for the DFT-s CNN receiver. As the lighter variant operates only after the IDFT, we adopt the architecture of 5 ResNet blocks as suggested by the analysis, and utilize less convolutional filters to make the network lighter. Finally, we state that yet another variant could be adopted with all the ResNet blocks before the IDFT, after which a conventional demapper could detect the bits. However, since this variant needs a conventional demapper, we have decided not to include it in our analysis.

B. Results With 256-QAM

We next consider the challenging case of 256-QAM, which is currently the highest supported data modulation order in 5G NR uplink [17]. Fig. 3(c) illustrates the BER performance of the 16 ResNet block variant of the DFT-s CNN receiver, light DFT-s CNN receiver and the iDeepRx from [16] according to the impairment parameters in Table I. The results illustrate, that the considered ML-based receivers outperform the conventional LMMSE receiver. Furthermore, we can also clearly observe that the proposed full DFT-s CNN receiver outperforms the light DFT-s CNN receiver by a larger margin, compared to the studies with 64-QAM. This is understandable since the demodulation and LLR generation task with high modulation orders, such as 256-QAM, is more challenging than with lower modulation orders. Importantly, we also observe that the proposed full DFT-s CNN receiver clearly outperforms the iDeepRX reference scheme. At an example target BER point of 1%, the SNR difference between our proposed full DFT-s CNN receiver and the iDeepRX is already more than 5 dB in favor of our proposed receiver.

C. Results With QPSK

Next, we focus on lower modulation order and the actual network uplink coverage with QPSK. Fig. 4 presents the BER performance of the proposed and the LMMSE reference receivers with respect to the EVM induced by the PA – here specifically showing the required receiver SNR to reach the target BER of 1%. Both proposed receivers perform well under the coexisting effects of phase noise and PA distortion. Notably, the proposed DFT-s CNN receiver achieves the target BER of 1% at lower SNR throughout the whole considered EVM range than the conventional receiver operating at the lowest considered EVM. This highlights the proposed receiver's capability to tolerate the transmitter impairment effects.

To further illustrate the performance gains, we next assess the uplink coverage which we define here as the maximum link distance for the different receivers to reach 1% raw BER. We first map the PA induced EVM range of Fig. 4 to the actual physical output powers of a realistic mmWave PA unit through actual RF measurements. Together with UE beamforming and antenna gains, these map to EIRP range shown on the right side of Fig. 4. We further take similar assumptions

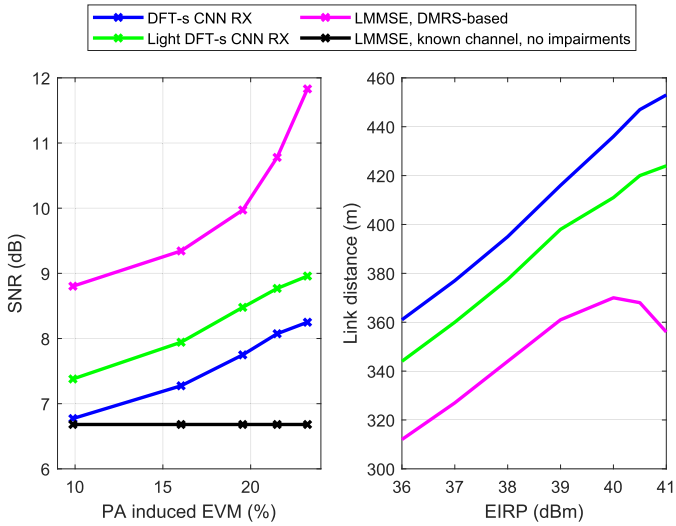

Fig. 4. Left: SNR performance of the considered receivers with impaired QPSK signals with varying levels of PA nonlinearity at 1% uncoded BER target. Right: Achievable maximum link distances of impaired QPSK signals with respect to transmitter EIRP in UMi NLoS deployment scenario.

to [3] about base-station receiver antenna gain and noise figure, and assume Urban micro (UMi) scenario and non-line-of-sight (NLoS) conditions for link budget calculations while following the pathloss models in [19]. As shown in Fig. 4, both proposed ML-based receivers provide extended coverage, as they can reliably detect bits when the PA is pushed deeper into saturation, thereby facilitating higher transmit powers. In this deployment example, the proposed receivers improve the uplink coverage by 14.6% and 22.4%, respectively, compared to the conventional LMMSE receiver.

D. Notes on Complexity

Some insight into the relative complexities of the different ML-based receivers can be obtained by considering their respective numbers of trainable parameters. The proposed fully learned DFT-s CNN receiver has about 703 000 trainable parameters, when considering 11 ResNet blocks. Meanwhile, the light DFT-s CNN receiver has only a quarter of that, with about 170 000 trainable parameters and 5 ResNet blocks. The iDeepRx with 11 ResNet blocks has 654 000 trainable parameters, while additionally performing significantly less IDFT transformations within the network. Moreover, for our implementation of the receiver proposed in [8], which has similar inputs and outputs as the light DFT-s CNN receiver, we used 5 fully-connected layers with 64 neurons. The model utilizes a FCNN for each resource element, which can be interpreted as a 1x1 CNN, as opposed to the 3x3 convolutional filters used in this work.

Although the relative complexities of the different ML receivers depend also on other factors beyond the number of parameters, in this case the architectures are sufficiently similar to allow for a rough comparison, especially considering the large differences in the parameter counts. Therefore, it can be concluded that while the ML-based receiver models considered in this work are likely to achieve higher accuracy than the solutions in [8], they achieve this with a larger number of parameters, meaning that they also incur a higher processing complexity. Likewise, the parallel IDFTs utilized in the DFT-s CNN receiver will somewhat increase the overall computational cost, in comparison to other receiver schemes utilizing only one IDFT.

Focusing then on the run-time inference complexity of the DFT-s CNN receiver, its convolutional layers have an asymptotic complexity of $\mathcal{O}(N_{\text{sc}}N_{\text{symb}})$. Additionally, the IFFT is known to have asymptotic complexity of $\mathcal{O}(N_{\text{sc}}\log(N_{\text{sc}})N_{\text{symb}})$. Thus the overall asymptotic complexity consists of these two terms. However, it should be noted that this only reflects how the complexity of the DFT-s CNN receivers scales with respect to the slot size. The true computational complexity of the ML algorithms can only be evaluated by considering the inference hardware, whose efficiency determines the overall energy consumption. It is possible that, for example, the proposed ML-based receivers consume more energy per time unit than a conventional non-ML receiver, while providing greatly improved detection accuracy in return. It is then up to the exact implementation requirements whether the added complexity is a reasonable investment or not. However, since such hardware aspects are not in the scope of our paper, we leave any further analysis as future work.

V. Conclusion

In this correspondence, we proposed two ML-based physical-layer receivers for detecting distorted DFT-s-OFDM waveforms under mmWave scenarios. In particular, both power amplifier induced non-linear distortion as well as oscillator phase noise were considered in the system model and the corresponding receiver designs. The performance of the receivers was evaluated in the context of 5G NR uplink at 28 GHz. The obtained results show that the ML-based receivers can effectively detect the signals even when the transmit signal error vector magnitude is high, and that the receivers can well generalize across different power amplifier realizations, phase noise bandwidths and mobile radio channel profiles when properly trained. When considering the proposed full DFT-s CNN receiver with trainable layers both before and after the receiver IDFT, this was shown to translate to an uplink coverage increase of over 20% compared to a conventional linear receiver, when utilizing QPSK modulation. Our future work will focus on developing neural processing receiver systems for modulation orders beyond 256-QAM while also considering non-QAM type of modulation schemes.

References

[1] M. K. Shehzad, L. Rose, M. M. Butt, I. Z. Kovács, M. Assaad, and M. Guizani, "Artificial intelligence for 6G networks: Technology advancement and standardization," *IEEE Veh. Technol. Mag.*, vol. 17, no. 3, pp. 16–25, Sep. 2022.

[2] M. Honkala, D. Korpi, and J. M. J. Huttunen, "DeepRx: Fully convolutional deep learning receiver," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3925–3940, Jun. 2021.

[3] J. Pihlajasalo et al., "Deep learning OFDM receivers for improved power efficiency and coverage," *IEEE Trans. Wireless Commun.*, vol. 22, no. 8, pp. 5518–5535, Aug. 2023.

[4] S. Yun, J.-M. Kang, I.-M. Kim, and J. Ha, "Deep artificial noise: Deep learning-based precoding optimization for artificial noise scheme," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3465–3469, Mar. 2020.

[5] International Telecommunication Union (ITU), "Future technology trends of terrestrial international mobile telecommunications systems towards 2030 and beyond," ITU, Geneva, Switzerland, Tech. Rep. M.2516-0, Nov. 2022.

[6] J. Hoydis, F. A. Aoudia, A. Valcarce, and H. Viswanathan, "Toward a 6G AI-native air interface," *IEEE Commun. Mag.*, vol. 59, no. 5, pp. 76–81, May 2021.

[7] J. Pihlajasalo, D. Korpi, T. Riihonen, J. Talvitie, M. A. Uusitalo, and M. Valkama, "Detection of impaired OFDM waveforms using deep learning receiver," in *Proc. IEEE 23rd Int. Workshop Signal Process. Adv. Wireless Commun.*, 2022, pp. 1–5.

[8] H. Farhadi and M. Sundberg, "Machine learning empowered context-aware receiver for high-band transmission," in *Proc. IEEE Globecom Workshops*, 2020, pp. 1–6.

[9] H. Farhadi, J. Haraldson, and M. Sundberg, "A deep learning receiver for non-linear transmitter," *IEEE Access*, vol. 11, pp. 2796–2803, 2023.

[10] A. Mohammadian, C. Tellambura, and G. Y. Li, "Deep learning LMMSE joint channel, PN, and IQ imbalance estimator for multicarrier MIMO full-duplex systems," *IEEE Wireless Commun. Lett.*, vol. 11, no. 1, pp. 111–115, Jan. 2022.

[11] Z. Zhao, M. C. Vuran, F. Guo, and S. D. Scott, "Deep-waveform: A learned OFDM receiver based on deep complex-valued convolutional networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2407–2420, Aug. 2021.

[12] C. Fager, T. Eriksson, F. Barradas, K. Hausmair, T. Cunha, and J. C. Pedro, "Linearity and efficiency in 5G transmitters: New techniques for analyzing efficiency, linearity, and linearization in a 5G active antenna transmitter context," *IEEE Microw. Mag.*, vol. 20, no. 5, pp. 35–49, May 2019.

[13] *5G Technology: 3GPP New Radio*, H. Holma, A. Toskala, and T. Nakamura, Eds., 1st ed. Hoboken, NJ, USA: Wiley, 2019.

[14] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1567–1592, Fourth Quarter 2013.

[15] Y. Wu, F. Lemic, C. Han, and Z. Chen, "Sensing integrated DFT-spread OFDM waveform and deep learning-powered receiver design for terahertz integrated sensing and communication systems," *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 595–610, Jan. 2023.

[16] W. Qi, C. Ye, D. J. Korpi, C. Gong, Y. Jin, and T. Yang, "iDeepRx enabled 100 Gb/s DFT-s-OFDM data transmission over 220 GHz testbed," in *Proc. IEEE Int. Conf. Commun.*, 2023, pp. 3684–3689.

[17] 3rd Generation Partnership Project (3GPP), "NR; User Equipment (UE) Radio Transmission and Reception; Part 2: Range 2 Standalone (Release 18)," 3GPP, Sophia Antipolis, France, Tech. Specification 38.101-2, Mar. 2024.

[18] A. Brihuega et al., "Piecewise digital predistortion for mmWave active antenna arrays: Algorithms and measurements," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 9, pp. 4000–4017, Sep. 2020.

[19] 3rd Generation Partnership Project (3GPP), "Study on Channel Model for Frequencies From 0.5 to 100GHz," (*Release 18*) 3GPP, Sophia Antipolis, France, Tech. Rep. 38.901 Mar. 2024.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.

[21] Mathworks, Inc., "Matlab 5G toolbox," 2020. [Online]. Available: https://www.mathworks.com/products/5G.html