A Two-stage Deep Learning Receiver for High Throughput Power Efficient LEO Satellite System with Varied Operation Status

Nielsen, Martin Hedegaard; De Carvalho, Elisabeth; Shen, Ming

Published in:
IEEE Access

DOI (link to publication from Publisher):
10.1109/ACCESS.2022.3180055

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 12, 2022, accepted May 31, 2022, date of publication June 3, 2022, date of current version June 13, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3180055

A Two-Stage Deep Learning Receiver for High Throughput Power Efficient LEO Satellite System With Varied Operation Status

MARTIN H. NIELSEN, ELISABETH DE CARVALHO, (Senior Member, IEEE), AND MING SHEN, (Senior Member, IEEE)
Department of Electronic Systems, Aalborg University, 9220 Aalborg, Denmark

Corresponding author: Martin H. Nielsen (mhni@es.aau.dk)

This work was supported by the Innovation Fund Denmark as part of the Research Project under Grant MARS2.

ABSTRACT Low Earth orbit satellites are expected to be one of the biggest suppliers of wireless communication within the coming years. For this to happen 5G and 6G networks, are crucial to be implemented in satellite communication. This comes with the problem of power-efficient transmissions. This paper exploits recent advances in complex-valued deep learning to cope with this challenge. The proposed approach is based on the autoencoder structure, where a legacy orthogonal frequency division modulation (OFDM) transmitter is used as an encoder and a deep complex convoluted network (DCCN) is used as decoder/receiver. Different from other state-of-the-art receiver architectures based on one-stage trained neural networks, our proposed DCCN adopts a two-stage training scheme, where the first stage is trained using AWGN channel and a fixed non-linear front end. The second stage uses a transfer learning to adapt to the flat fading channels and the front end model can be changed to compensate for different front ends, significantly reducing training time.This allows for power-efficient transmission at different operation statuses (e.g. radiated power levels and steering angles) without compromising the bit error rate in both average white Gaussian noise (AWGN) and flat fading channels. A K-band (28 GHz) active phased array in package (AiP) transmitting a 5G NR OFDM signal with a bandwidth of 100 MHz was used as the main front end test vehicle for validating the proposed DCCN. Satisfactory bit error rates were achieved while the AiP was driven into saturation with high power efficiency at different power levels and steering angles. This work demonstrates, for the first time, the promising capability of deep neural networks in processing varied operation staged non-linear OFDM waveforms in the form of an auto-decoder receiver.

INDEX TERMS Deep learning, OFDM, receiver, non-linear, power efficiency, LEO satellite.

I. INTRODUCTION

Satellite communication has been a topic of interest for many years, but recently it has been developing rapidly with the advancements in low Earth orbit (LEO) satellites that are usually deployed at an altitude of 500–1500 km to ensure low latency in established communication links [1]. Dependent on the altitude deployed the satellite will over the course of a transmission window different distances to the ground making the communication channel dynamic and varying free space path loss. This is conventionally overcome at the transmitter side [2].

Currently, the satellite sector is changing its service focus from TV and maritime applications towards broadband Internet services using the LEO satellites and 5G signals [3]–[5]. With higher data rates needed for Internet access, highly bandwidth-efficient modulation schemes, such as multi-sub-carrier orthogonal frequency division modulation (OFDM), are wanted. Multi-subcarrier OFDM signals can provide the necessary bandwidths for future space communication and are already being used in some satellite systems [6]–[8]. However, this comes with the cost of a high peak-to-average power ratio (PAPR) and low efficiency in the transmitter front end, since to ensure linearity back off is typically used.

For many years the industry and research practice for increasing power efficiency is use pre-distortion to deal with non-linear amplifier responses. Pre-distortion can be implemented in a plethora of ways all from look up tables using memory polynomial models to different deep learning

The associate editor coordinating the review of this manuscript and approving it for publication was Nan Wu.

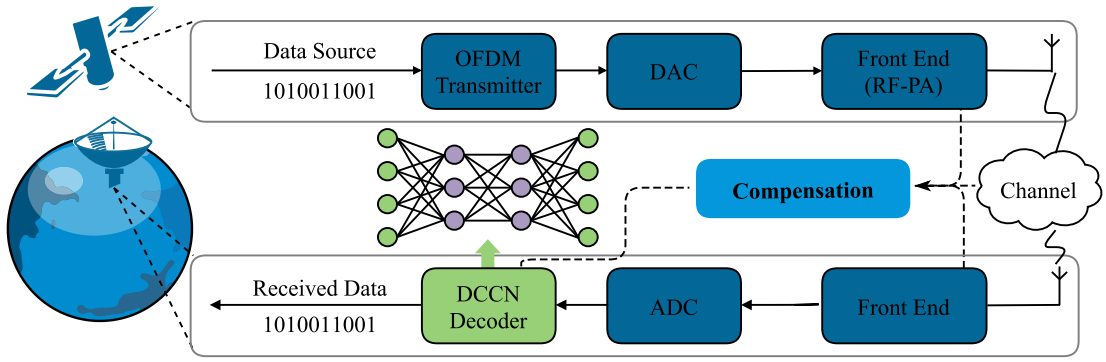

FIGURE 1. Concept figure of the proposed system architecture where a DCCN receiver is used to compensate for the front end.

techniques [6], [8]–[23]. Memory polynomial unlike deep learning uses an equation to describe non-linear behavior. It is based on a feedback loop between the input and output of the power amplifier (PA) to achieve pre-distortion. It uses the inverse polynomial to change the input signal such that the wanted efficiency/linearity of the system is achieved [9]–[11].

Implementing pre-distortion in satellites is not as easy, since the optimization can be costly for the satellites [12]. When a highly integrated active phased array is used, it is very difficult to establish the feedback loop. Further, the PAPR change in the input signal caused by pre-distortion can in some cases not be realized [12]. Commonly the deep learning approach is based on autoencoders, where both the transmitter and receiver are changed to neural networks. Some research reduces the PAPR [14], [20], while some gaining better bit error rate (BER) performance [13], [15], [16]. However to our knowledge no autoencoders have gained both high BER performance and are able to combat high non-linearity without changing the transmitter completely [13]–[20]. However, using deep learning comes with challenges [4], [5], [24]. With the wanted integration of satellites into the 5G network, multiple satellites can and will have to connect to the same ground station in these new mega-constellations [4], [24]. This can cause issues having to deal with multiple slightly different front ends. Another inconvenience is that the satellite has limited power on board and using pre-distortion or an autoencoder can be a become resource problem.

Instead of relying on changing the on-board communication system, this paper proposes to train a receiver that can compensate for the different non-linear behavior present in the received signal. Then only the power output from the transmitter can be increased and the satellite can achieve the performance increases that conventional pre-distortion gives a front-end transmitter by only changing the receiver side. This allows for multiple satellites to be serviced by the same ground station without having to change the satellite architecture. The idea was initially proposed in [21]. It was shown possible to recover the received signal and compensate for the non-linear distortion before the actual receiver and BER performance was later shown in [22]. However, it lacked more realistic channels and results for how it compensates for different amplifiers, steering angles, and power levels without retraining the model. A fading channel, unlike AWGN, distort and corrupts the signal. Hence, using pre-processing for OFDM signals is both time-consuming and inconvenient. Instead using complex-valued neural networks as receivers has the potential to solve these issues faster and as reliable as legacy receivers [23]. Complex valued networks are similar to deep neural networks but they can process and understand complex numbers without pre-processing. This paper proposes to combine the use of a complex-valued neural network [23] with the idea of recovering signals [21], [22] to provide a new receiver structure that can both decode and compensate for the non-linearity in a single complex-valued neural network for an OFDM system. This paper proposes a two-stage Deep Complex Convolution Neural Network (DCCN) which is trained in two steps to achieve the wanted performance. The special two-stage DCCN comes with the benefit that it is very robust to variations in the non-linear signal. Further, it can quickly adapt to different front ends with only one retraining of step two. This makes quick adaptation to different satellites very fast at the ground station. Different operation modes for the same front end are possibly compensated for. This is possible without retraining for operation modes such as different power saturation, different bandwidths, and different steering angles, these are all compensated for due to the two-stage DCCN.

To train the neural network this paper uses an online end-to-end numeric simulator that uses a legacy OFDM transmitter, together with real experimental data gathered from the front end to train and test the DCCN.

The paper is structured as follows. In Sec. II the problem and communication system are described, Sec. II-A the legacy OFDM transmitter is described, Sec. II-B the end-to-end simulator is detailed and Sec. II-C describes the channel models used. Sec. III describes the acquisition and results of the measured front-end models together with the results of the end-to-end simulator. In Sec. IV the deep learning receiver architecture is elaborated and the training process is

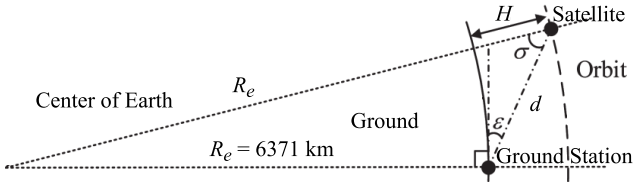

FIGURE 2. Geometrical illustrating of the satellite to ground data transmission link.

described. The results of the deep learning receiver are the presentation in Sec. VI and discussed in Sec. VII. The paper is concluded in Sec. VIII.

II. COMMUNICATION SYSTEM & PROBLEM FORMULATION

The satellite communication system starts when the satellite establishes a communication link with the ground station [2], [21], [22], [25]–[28]. The window for an LEO satellite starts when the elevation angle ϵ is greater than 5°. The satellite will then travel with a constant velocity towards the ground station where the elevation angle is 90°. The satellite will then travel from the ground station until the elevation angle is 5°. The distance to the ground station will not be constant and can be geometrically described as shown in Fig. 2.

The differentiation in the distance from the ground station causes the transmission distance to vary. The link will therefore change over time. This can be described mathematically.

$$(H + R_e)^2 = R_e^2 + d^2 - 2R_e d\cos(\frac{\pi}{2} + \epsilon_0), \tag{1}$$

where ϵ represents the elevation angle, d is the distance of data transmission, H denotes the orbital altitude of LEO satellite, R_e is the earths radius. This makes it possible to denote the distance as a function of ϵ

$$d(\epsilon) = R_e[\sqrt{(\frac{H + R_e}{R_e})^2 - \cos^2(\epsilon)} - \sin(\epsilon)]. \tag{2}$$

Then the free space path loss can be expressed as

$$L(\epsilon_0) = (\frac{4\pi d(\epsilon)}{\lambda})^2 = (\frac{4\pi f d(\epsilon)}{c})^2, \tag{3}$$

since the noise can is fixed the variation in SNR regardless of distance is

$$\Delta\frac{S}{N}(dB/Hz) = (\frac{S}{N}_{5^\circ}) - (\frac{S}{N}_{90^\circ}) = \Delta L(\epsilon). \tag{4}$$

Based on the equation the variation for a satellite in LEO at an altitude of 800 km is approximately 12 dB in the link budget between 5° and 90° [22]. To have the possibility of an efficient as possible satellite downlink, the transmitter should be able to vary transmission power such that the power received is constant. However to keep the transmitter efficient the transmitter should be driven non-linearly.

A. OFDM TRANSMITTER

5G NR is an evolution of LTE signal that is proposed to become the future communication signal for satellites [6], [7]. OFDM is the basis for 5G and LTE signals therefore, this paper only considers the OFDM signal since any channel coding/precoding associated with LTE or 5G NR happens before OFDM transformation, see Fig. 3 a. OFDM is based on a frame structure where a communication slot is composed of multiple OFDM symbols. The notations related to the OFDM slot are as follows: an OFDM symbol contains N subcarriers, where N is the size of DFT/IDFT. Among the N subcarriers, a total of G nullified guard subcarriers are placed at the center (DC guard band) and the edge guard band. A subcarrier in an OFDM symbol is referred to as a resource element (RE). A collection of REs is referred to as a Resource block as shown in Fig. 3. A communication slot contains F consecutive OFDM symbols, in which P and D REs are allocated to pilot and data, respectively. The length of a time-domain full OFDM symbol is $S = N + N_{cp}$ where N_{cp} is the length of CP. Under m-ary modulation, an IQ sample carries m bits, and the size of the constellation is 2^m.

B. END-TO-END SIMULATOR

The end-to-end simulator is built based on Fig. 3. It uses a random bit generator to produce the bit sequence $\mathbf{b}$ that is encoded such that $\mathbf{b} \in \pm 1$. The encoded bits are then converted to complex-valued in-phase and quadrature (IQ) data by mapping to a constellation on the IQ plane. From the IQ data, an OFDM frequency-domain symbol $\mathbf{X}$ is is created by inserting pilot signals and guard bands into the IQ data, and then $\mathbf{X}$ is transformed to a time-domain OFDM symbol, $\mathbf{x}$, via an N-point IDFT and a subsequent parallel to serial (P/S) conversion.

Next, the Cyclic Prefix (CP) is prepended to $\mathbf{x}$ to create a time-domain full OFDM symbol, $\mathbf{x}_{cp}$. The baseband signal $\mathbf{x}_{cp}$ is then upconverted to RF and sent to the RF front end model. The RF front end model is created based on the popular Memory Polynomial model that can describe the non-linear distortion given the input and output of a non-linear system. It is given as

$$y(t) = \sum_{n=1}^{N}\sum_{m=0}^{M} a_{nm}x(t - \tau_m)|x(t - \tau_m)|^{n-1}, \tag{5}$$

where N is the polynomial order while M is the memory effect, i.e, the number of previous samples that have effect on the current output. To determine the amplifier coefficients a_{nm} the input and output of the device need to be captured, and reverse modeling is used to find them. The coefficients are dependent on the signal power levels and how high the PAPR of the signal was at the input and output baseband signals.

The now distorted signal is propagated over a channel model and the received IQ samples are then represented as OFDM-time domain samples $\mathbf{y}_{cp}$. Then, the CP is removed from $\mathbf{y}_{cp}$ and the rest of the IQ samples, $\mathbf{y}$, are transformed to the frequency-domain OFDM symbol, $\mathbf{Y}$, via DFT. Based

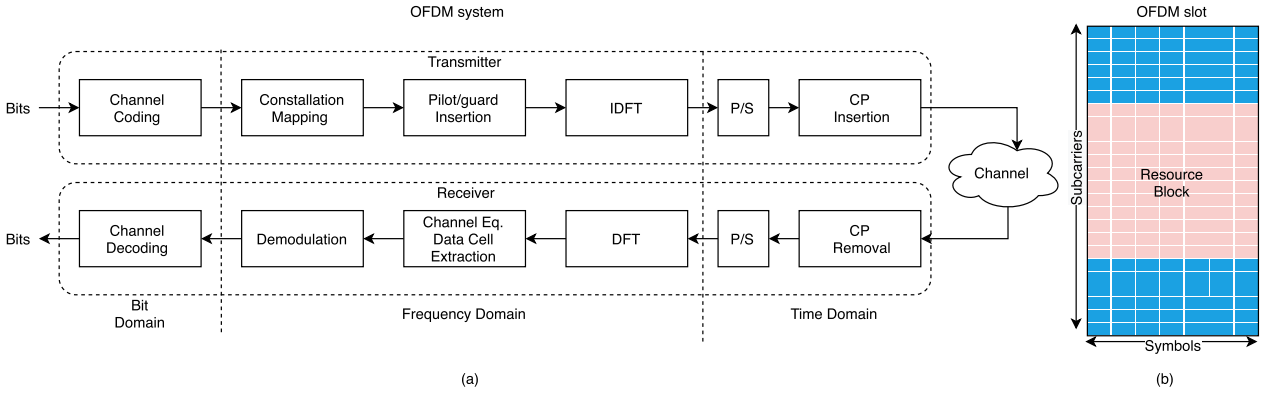

FIGURE 3. a) Block diagram of a PHY layer OFDM communication system. b) OFDM slot illustration, subcarriers on y axis and symbols on x axis. A resource block is shown to be 12 subcarriers and 7 symbols.

on $\mathbf{Y}$, a channel equalizer outputs the estimated transmit frequency-domain IQ data $\hat{\mathbf{X}}$ which is then demodulated to soft bits (log-likelihood) $\tilde{\mathbf{b}}$ which are converted to binary outputs bits via hard decision to become $\hat{\mathbf{b}}$. The end-to-end simulator can then calculate the Bit Error Rate depending on how much difference there is between $\hat{\mathbf{b}}$ and the ground truth $\mathbf{b}$.

C. LEO CHANNEL

According to [29] and [30], the LEO communication channel cannot be treated as an AWGN channel. It is stated that it can be seen as a fading environment. We therefore model our channel after the well described flat fading model [31] since it is a good baseline model:

$$\mathbf{y} = \mathbf{x} * \mathbf{h} + \mathbf{n}, \mathbf{Y} = \mathbf{X} \odot \mathbf{H} + \mathbf{N}, \tag{6}$$

where $\mathbf{x}, \mathbf{y}, \mathbf{n}$ are the time-domain transmitted and received signals and white noise, respectively and $\mathbf{n}$ is the channel impulse response. $\mathbf{Y}, \mathbf{X}, \mathbf{H}, \mathbf{N}$ are the frequency domain transforms. $*$ is the convolution and $\odot$ is the elementwise product.

III. MEASUREMENT OF NON-LINEAR HARDWARE

To have a practical RF front end model in the end-to-end simulator, an active phased array in package (AiP) is used for measuring the non-linear behavior of the front end. The AMOTECH A0404 AiP uses 4 Anokiwave AWMF-0158 beam forming devices and a 4 by 4 patch antenna array for transmitting [32]. In this paper, the changes happening in the transmission window i.e. different input power is given and a variation in the steering angle are measured and investigated.

To capture the data for use in the hardware models the AiP is measured over the air using a 28 GHz 5G NR signal with a bandwidth of 100 MHz, compliant with the 3GPP downlink specification for 5G NR OFDM modulated with a peak to average power ratio of 10.6 dB generated by the R&S SMBV100B signal generator. The 5G NR signal from the generator is up-converted from 3 GHz. A continuous-wave (CW) signal has been multiplied by two into 25 GHz for up-conversion and down-conversion as shown in Fig. 4. A pre-amplifier is used to push the AiP into compression, the pre-amplifier is in backoff to ensure linearity. This setup is shown in Fig. 4.

The data is captured by the observation horn antenna placed 42 wavelengths away (44 cm) and aligned with the main beam which is connected to and analyzed in the spectrum analyzer (R&S FSW 67 GHz). The procedure for taking measurements is as follows:

1) I and Q waveform for the 5G NR signal is uploaded using the R&S VISA tool from the PC to the vector signal generator.
2) The APA is driven into the non-linear region.
3) The I and Q at the receiver are then captured from the signal analyzer using the R&S VISA tool.
4) The signal is then time aligned with the input signals samples post-process to ensure each sample corresponds to the correct time sample of the previously recorded input signal.

A. MEASURED RESULTS

The measurement results are shown in Fig. 5. Fig. 5a the power spectrum density is shown. The 38 dB output power signal shows the AiPs non-linear behavior and the high adjacent channel power present. Fig. 5b shows the amplitude-to-amplitude (AM/AM) and amplitude-to-phase (AM/PM) behavior of the AiP. The non-linear saturation and memory effects can be seen both in the spread of points and the steep drop after 30 dB for AM/AM and the rotation in phase for symbols in the AM/PM. Looking at the power added efficiently (PAE) curve, Fig. 5c it is clear that the AiP is not very efficient. To have the AiP be as efficient as possible it is better to have an output power level of around 38 dBm at all times.

B. SIMULATION RESULTS ON MEASURED AiP MODEL

To determine the non-linearity effect on the communication system the simulator is tested in three different scenarios, AiP distortion only, with AWGN and AiP distortion, and

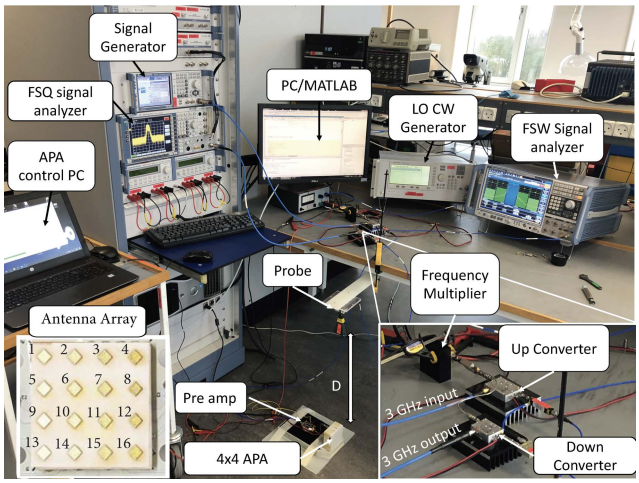

FIGURE 4. Measurement setup of the AiP with all instruments used labeled.

with fading and AiP distortion. The noisy baseband signal results from the three tests are shown in Fig. 6. The resulting constellation patterns show that without proper modifications to understand the non-linearity and fading it is not possible to decode the symbols given.

IV. DCCN ARCHITECTURE

The neural network-based OFDM receiver uses similar signal processing modules as legacy OFDM receivers with layers functioning differently. The proposed architecture is shown in Fig. 7 which includes recommendations founds in [23].

The proposed architecture uses the input signal, $\mathbf{y}_{cp}$, and uses the first complex dense layer to remove the CP. The DFT like C-Conv layer converts the signal $\mathbf{y}$ to the frequency domain $\mathbf{Y}$. The channel estimator uses four dense layers to both compensate for the AiP and estimate the channel response, $\hat{\mathbf{H}}$. After the channel estimator, it then does element-wise complex division to the equalization so that it can estimate $\hat{\mathbf{X}}$. To do the estimation the first dense layer of $\mathbb{C}_{FNP}$, where F is OFDM symbols per coherence slot, N is number subcarriers which are proportional to the DFT/ODFT size, and P number of pilots. The dense layer is designed to locate pilots and estimate channel coefficients on pilots $\hat{H}^P_{LS}$, where LS is least-square similar to LMSSE. The operation can be shown as equal to that of LMSSE which matematically can be described as,

$$\hat{\mathbf{X}} = \frac{\mathbf{Y}}{\hat{H}_{LS}}, \quad \hat{H}_{LS} = \mathcal{L}(\frac{\mathbf{Y}_P}{\mathbf{X}_P}), \tag{7}$$

where $\hat{\mathbf{X}}$ is the estimated signal, $\hat{H}_{LS}$ contains the LS channel estimates, $\mathbf{X}_P$ and $\mathbf{Y}_P$ are transmitted and received pilots, and $\mathcal{L}()$ is the interpolation operation. To obtain $\hat{H}$ an interpolation of $\hat{H}^P_{LS}$ to the entire slot and channel estimation is done in the next three dense layers and a 2D filter of size (F, N). The 2D filter can be mathematically described as follows

$$\hat{\mathbf{H}}_{LRA} = \mathbf{U}\mathbf{D}_p U^H \hat{H}_{LS}, \tag{8}$$

where D_p is a diagonal matrix with trainable parameters $\delta_k = \frac{\lambda_k}{\lambda_k + \frac{\beta}{\alpha}}$, U is a unitary matrix containing the singular vectors of the frequency domain covariance matrix of channel realizations denoted $R_{HH} = E\{H, H^H\}$. Instead of setting the parameters explicitly they are learned directly from data. As the non-linearity is treated at the receiver side to be part of the channel a separate layer is not needed. However, through testing, it has been shown that having three dense instead of only two has better generalization capabilities than using two [23].

To gain the time domain signal the network uses a complex dense to take out the Resource Elements (RE) so we get $\hat{\mathbf{X}}_D$ and then using an IDFT like C-Conv layer it is possible to transform $\hat{\mathbf{X}}$ to time-domain $\hat{\mathbf{x}}$. The rest of the forward network converts IQ samples to soft bits, where an input IQ sample $\mathbb{C}$ is treated as a vector of 2 real numbers, $\mathbb{R}^2$. The extracted IQ vector and its non-linear (Leaky ReLU) activation, A0, are concatenated to a tensor of shape [B, D, 4], where B is the number of slots in a batch of input signals. A0 is fed to a small dense layer of $\mathbb{R}_{42^m}$ followed by another Leaky ReLU activation, A1, of which the output tensor is reshaped to [B, D,m, 2] and then activated by a softmax function along its last dimension to produce a soft bit–a vector of likelihoods of ± 1.

A. TRAINING

The training of the DNN receiver as illustrated in Fig. 8. The training is done online using the end-to-end simulator which generates a random binary stream $\mathbf{b}$. The stream is turned into time-domain OFDM symbols, $\mathbf{x}_{cp}$ by the OFDM transmitter. The transmitted signal $\mathbf{y}_{cp}$ is then distorted by a non-linear AiP model together with a channel model that adds noise and fading to $\mathbf{x}_{cp}$. Thus the received signal, $\mathbf{y}_{cp}$ is the training data, and $\mathbf{b}$ is the labels. The output of the receiver is given as soft bits, $\tilde{\mathbf{b}}$ and after the hard decision the $\hat{\mathbf{b}}$. The loss function is a weighted sum of the cross entropy loss and the regularization loss

$$\ell(\mathbf{b}, \tilde{\mathbf{b}}, \Phi) = -\ln\left(\frac{e^{\mathbf{b}_k}}{\sum_i e^{\tilde{\mathbf{b}}_i}}\right) + \epsilon \ell_{reg}(\Phi), \tag{9}$$

where k is the index of the target bit, and $\epsilon < 1$ is a small constant. The Adam optimizer that is doing the back-propagation is randomly initializing Φ of DCCN receivers and updates the loss function accordingly.

It is difficult to train the DCCN for fading channels due to the severe distortion that will be implemented on top of the already non-linear distortion introduced by the front-end model. Therefore a two-stage training method is implemented. In stage 1 the neural network is trained with the APA model and AWGN channel. In stage 2, the DCCN is retrained with flat fading and the AiP model for better BER in flat fading channels. This is done by transfer learning. The flow graph of the TensorFlow session freezes the CP remover, data extraction, demodulator, and channel decoder and only keeps training the 1D complex convolution and

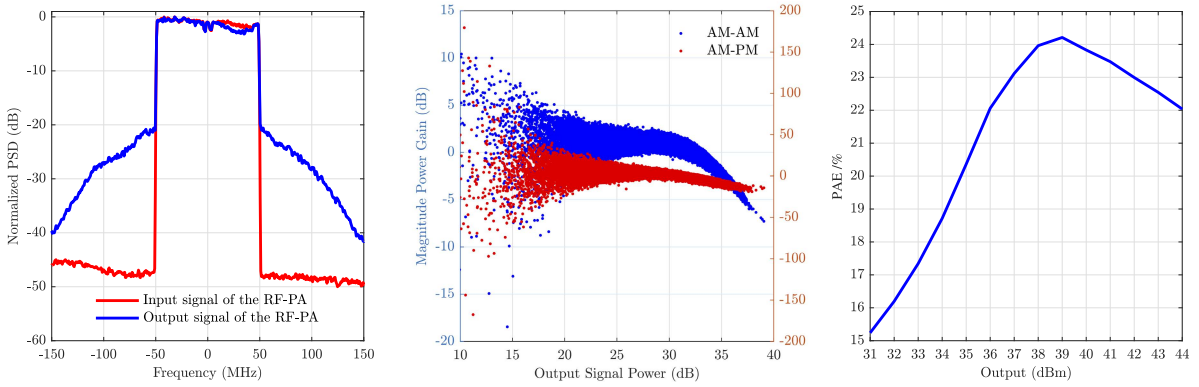

FIGURE 5. Measurement of the AiP. a) Measured power density response of the 100MHz 5G NR response. b) Measured AM/AM AM/PM responds. c) AiP power added efficiently vs input power.

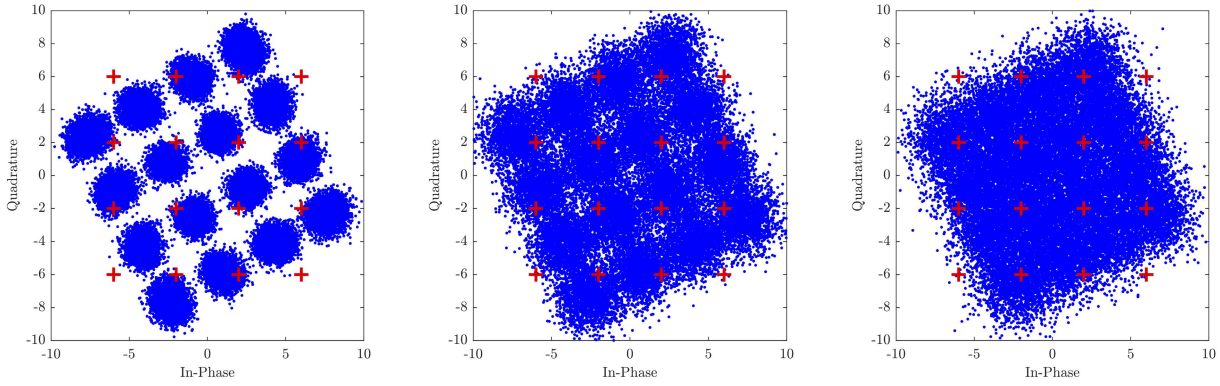

FIGURE 6. Effects of the non-linear distortion from the AiP and AWGN and fading on the constellation with SNR at 30 dB. a) Constellation of the non-linear distorted signal without the channel. b) Constellation of the non-linear distorted signal with the AWGN. c) Constellation of the non-linear distorted signal with flat fading.

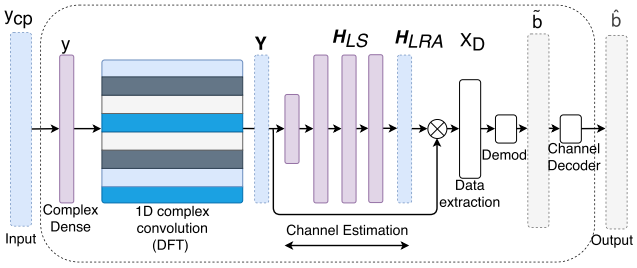

FIGURE 7. Neural network receiver architecture to deal with non-linear amplifier and different fading channels.

the channel estimator. Thus the already pre-existing weights in a new session can be used to further enhance the neural network for flat fading and non-linearity. The loss function is the same throughout stages 1 and 2. The graph editing technique enables back-propagation when the second half of the forward network is frozen. This two-stage training approach can increase the data efficiency by reusing the same pre-trained receiver in stage 2 for different fading settings and different active phased array (APA) models.

This increased training efficiency since multiple different channels can be trained with the same initial model. This also allows for more fine-tuning training to compensate for the APA and given channel. To optimize training all operations are vectorized and all data are fed as tensors in the code. Loops are avoided and the learning rate is decayed exponentially for fine-tuning as the training proceeds. For training data, all data is generated online. Hence iterations are used instead of epochs. Early stopping is implemented to make sure that when the BER has not improved for 20 iterations the DCNN model is saved for testing.

Determining SNR during training is not straightforward, however, [23] provides a recommendation of using fixed SNR when only considering AWGN channels and a mix of high and low SNR during flat fading. We choose to use an SNR of 5 dB during AWGN in stage 1 and a variational SNR during flat fading training, stage 2. In stage two different SNRs are used with a mix of both high and low SNR with about a 30/70 split, this is because the channel estimator relies more heavily on the channel statistics and a variation of SNR is therefore needed. This variation is not needed for AWGN since it is a

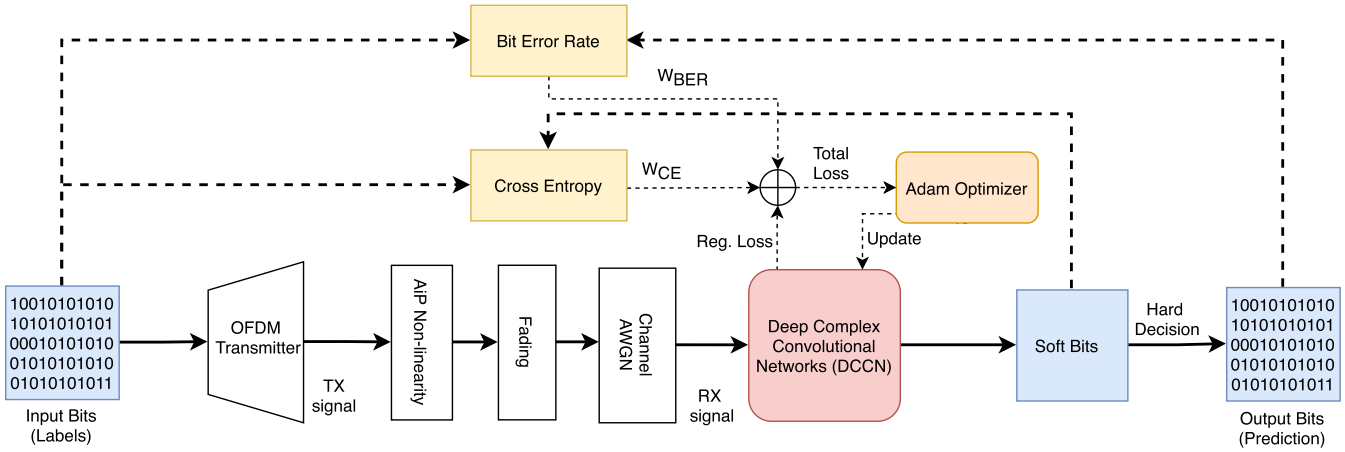

FIGURE 8. Training diagram of the proposed DCCN receiver.

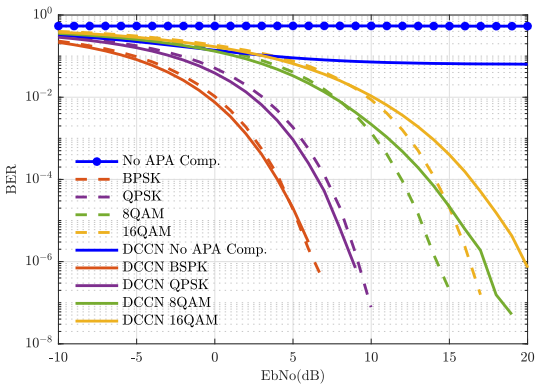

FIGURE 9. Bit error rate plot with the trained decoder over AWGN channel(solid) and LMMSE BER results for comparison(dotted).

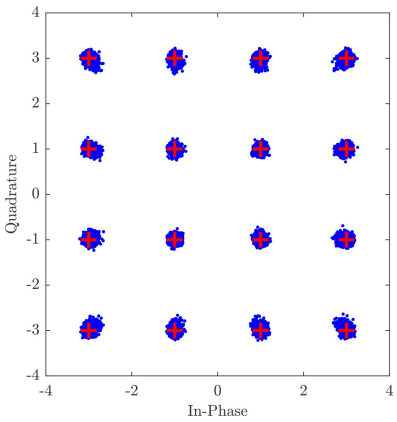

FIGURE 10. Constallation of what the soft bit output of the DCCN receiver. Red crosses is a normal 16QAM decoder scheme. SNR = 20 dB smaller SNR will cause more spread of blue dots.

simpler channel model and the DCCN can better generalize for the AWGN using a low SNR of around 3-5 dB.

V. RESULTS

The proposed DCCN receiver is evaluated using numeric results. For comparison, legacy OFDM receivers like LMMSE, ALMMSE, etc. are used. Both Average White Gaussian Noise channel (AWGN) and a Flat fading channel are used with and without training for the non-linear APA model. The selected APA model used for training and evaluation is the AiP at 38 dBm output power 100MHz with steering angle 0. It is highly non-linear as shown in Fig. 6. APA models for different output powers (31-39 dBm) are used. To further evaluate the DCCN an APA model based on [21] is also used to determine how well the DCCN handles never before seen APA models. Fig. 9 shows that the DCCN is capable of handling AWGN and non-linear distortion at the same time for different modulation forms. The blue solid line is for the DCCN with no APA model during training. It shows that if the APA model is not known a-priory the DCCN cannot handle the non-linear distortion very well. The blue dotted line is legacy receivers with non-linear distortion. It shows it is not possible to handle non-linear distortion on its own. The striped lines are all LMMSE cases for AWGN with no implementation of an APA model. Hence, if the non-linear distortion is handled at the transmitter side the baseline for legacy receivers is slightly better in higher modulation orders as shown. It is possible to drive the AiP at the highest PAE of 24% while maintaining the BER performance of legacy receivers in different modulations using the DCCN receiver. For other AiPs, APAs, and amplifiers this potential benefit can be higher.

A. SOFT BIT RESULTS

Fig. 10 shows some spread is still present in the DCCN predicted soft bits, but compared to Fig. 6c there is a big difference. The previously shown phase shift is now removed. Further small variations are shown but have no impact on the hard decision used on the soft bit prediction. The small variations are due to the non-linearity and noise not being fully eliminated.

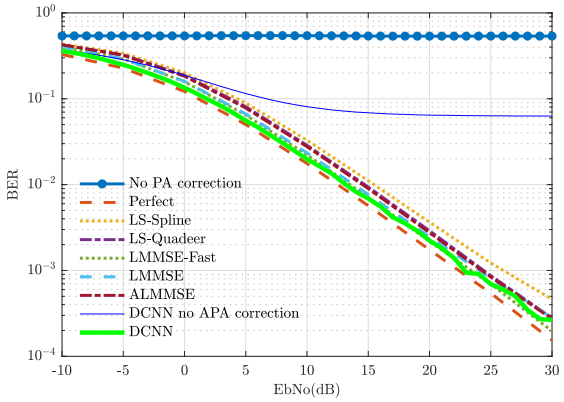

FIGURE 11. Bit error rate plot with the trained decoder over flat fading channel vs conventional OFDM receiver codes.

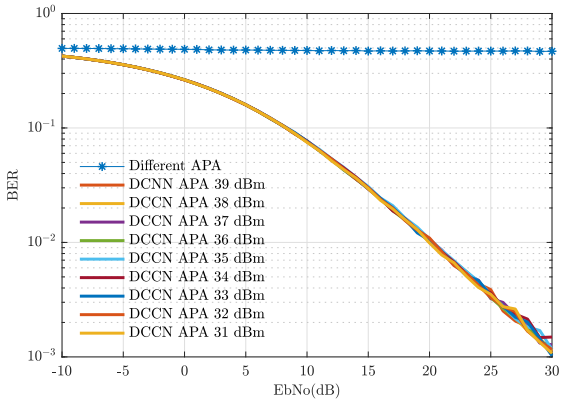

FIGURE 12. Bit error rate plot with the trained decoder over flat fading channel with different power outputs from the APA.

B. FLAT FADING

To test the DCCN in flat fading with the AiP it is compared to different legacy receivers, like LMSSE, ALMSSE, Spline, etc. [33]. The transmitter is once again without an APA model. Fig. 11 shows that the DCCN can compensate for the non-linearity while simultaneous deals just as well as legacy receivers with the flat fading channel. Both LMMSE and DCCN can not compensate for the non-linearity caused by the AiP if they are not aware of it, similar to the AWGN performance. The DCCN does slightly better only in low SNR. This is due to high noise blocking the non-linearity from the DCCN. When trained with the AiP model the DCCN works very well. It is only a few BER worse than conventional approaches like ALMMSE and the DCCN is equal to LMMSE in performance.

C. OUTPUT POWER, AND STEERING ANGLE DIFFERENCE

It is evaluated how well the DCCN compensates for different non-linear models from both the same AiP and different front ends.

The DCCN shows that it is capable of compensating for different power levels without changing the model. If the non-linear frontend model is changed to a different AiP or

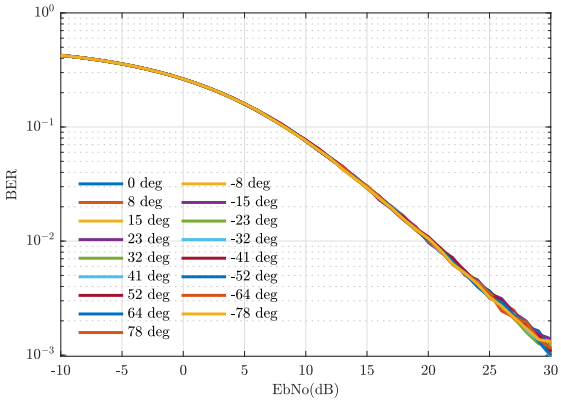

FIGURE 13. Bit error rate plot with the trained decoder over flat fading channel with steering angle changing.

APA the DCCN cannot any longer function as shown in Fig. 12. The dotted line represents the changed front-end model and the DCCN does not function properly. However, since a two-stage training method is used the front end model can be changed in step 2 which significantly eliminates this problem. This makes training time significantly less costly for reapplying this model to different APA models and channel environments since training time is reduced by one step.

VI. DISCUSSION

The method given in this paper shows that the receiver can handle different power levels without having to change the front end model. Further, it shows that the given DCCN receiver can be retrained for any given front end model while only changing the model used in step two of the training process. This is possible due to the equalizer used for channel selective fading and is the most important aspect to achieve high BER performance with non-linear amplifiers. Further, this method makes it possible to not change our transmitter architecture. Thus this can be readily implemented in satellite to ground transmissions. And since only the DCCN needs to be changed at ground station it can service a lot of satellites. For the physical implementation of the system, the model can be pre-trained before deployment and then uploaded to the ground station. If the BER drifts significantly the model can be retrained for the new channel and front end situation. Since the model only needs a synthetic model of the transmitter front end any changes to the whole system can be directly updated and retrained off the actual deployed system.

The impact of steering angle for the AiP can change the non-linearity behavior slightly, but as shown in 13 this variation is negligible and the DCCN performs as expected. The complex convolution layer that is doing the DFT-like operation together with the equalization layers is why it is possible to adapt to the change in the non-linearity. Since the AiP at different power outputs does not dramatically change the non-linear behavior the DCCN is capable of equalizing for it without re-training. This is due to the strong generalization capabilities of first the DFT-like convoluted layer and second

the equalizer. The convoluted layer makes it possible for the DCCN to understand IQ symbols. Any change and variation due to the front end can be estimated by the channel equalizer since the front end disturbances are seen to be part of the channel matrix H. This comes with the cost of generalization for different front ends. However for the same AiP the generalization is very good. Both output power variation and steering angle variaiton is possible to be compensated for at the bit prediction side meaning a good BER is maintained even with high distortion in the signal.

With these findings it is capable that a single ground station can have multiple DCCN models trained for different satellites and compensate for them all. If a multiple satellites are to be serviced by the same ground station and they use similar front ends no retraining is needed for the DCCN model. Only if the front end changes completely a retrain is needed.

Since a legacy transmitter is used, the DCCN is a very robust receiver that can use legacy OFDM transmitters with non-linearity. To use this receiver a single measurement of the given front end with a known signal is needed. Hence for mobile communication systems, the proposed system could be difficult to use due to multiple users all with different APAs. For future work, it is wanted to investigate possible solutions for adapting to different APAs. Hence, the limitation of the proposed DCCN is how many APA models it can compensate for without retraining. It should be noted that the proposed DCCN does not suppress the adjacent channel power at the transmitter side and it could be a problem for terrestrial radio systems. But this is not an issue as band-pass filtering is usually already included before feeding the signal to antenna in most satellite transmitters to fulfill the interference regulations of the International Telecommunication Union (ITU).

To help fast-track evaluation of different APAs, using the over-the-air measurement setup shown in Fig.4, it is wanted to implement the trained receiver into a software-defined radio setup. This is left for future work.

VII. CONCLUSION

In this paper, we presented a fully trainable two-stage deep learning receiver that addressed particular non-linearity issues in the satellite transmitter. The deep learning receiver can achieve equal BER to that of conventional approaches and compensate for non-linearity in a dynamically changing transmission link without the need for pre-distortion techniques in the transmitter. This is due to the unique two-stage training scheme proposed that integrates the estimation for the channel and non-linear front end into one. Using an online end-to-end simulator both for training and testing it is possible to numerically validate the proposed method.

The DCCN can deal with different steering angles and output power levels from the same AiP without retraining. It is only needed to include a new AiP in step two, significantly reducing training time for different transmitters to the same ground station. Numerical validation results also show that it is possible to almost double the power efficiency of the chosen AiP without sacrificing BER. This is different from conventional OFDM receivers like ALMMSE and LMMSE since no architecture change at the transmitter side is needed.

REFERENCES

[1] J. Wang, L. Li, and M. Zhou, "Topological dynamics characterization for LEO satellite networks," *Comput. Netw.*, vol. 51, no. 1, pp. 43–53, Jan. 2007.
[2] S. Egami, "Closed-loop transmitting power control system for K-band satellite communications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-19, no. 4, pp. 577–584, Jul. 1983.
[3] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
[4] O. Kodheli, E. Lagunas, N. Maturo, S. K. Sharma, B. Shankar, J. F. M. Montoya, J. C. M. Duncan, D. Spano, S. Chatzinotas, S. Kisseleff, J. Querol, L. Lei, T. X. Vu, and G. Goussetis, "Satellite communications in the new space era: A survey and future challenges," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 70–109, 4th Quart., 2021.
[5] F. Fourati and M.-S. Alouini, "Artificial intelligence for satellite communication: A review," *Intell. Converged Netw.*, vol. 2, no. 3, pp. 213–243, Sep. 2021.
[6] L. Siyang, Q. Fei, G. Zhen, Z. Yuan, and H. Yizhou, "LTE-satellite: Chinese proposal for satellite component of IMT-advanced system," *Commun., China*, vol. 10, no. 10, pp. 47–64, Oct. 2013.
[7] K. Rahil, "System and method for satellite-long term evolution (S-LTE) air interface," U.S Patent 2010 0 068 993, Mar. 18, 2010.
[8] M. N. Sweeting, "Modern small satellites-changing the economics of space," *Proc. IEEE*, vol. 106, no. 3, pp. 343–361, Mar. 2018.
[9] Y. Guo, C. Yu, and A. Zhu, "Power adaptive digital predistortion for wideband RF power amplifiers with dynamic power transmission," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 11, pp. 3595–3607, Nov. 2015.
[10] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.
[11] A. Molina, K. Rajamani, and K. Azadet, "Digital predistortion using lookup tables with linear interpolation and extrapolation: Direct least squares coefficient adaptation," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 3, pp. 980–987, Mar. 2017.
[12] V. Karunanithi, C. Verhoeven, and W. Lubbers, "High efficiency transmitter architecture compatible with CCSDS and ECSS standards for nano-satellite missions," in *Proc. 32nd AIAA Int. Commun. Satell. Syst. Conf.*, 2014, p. 4252.
[13] A. Felix, S. Cammerer, S. Dörner, J. Hoydis, and S. ten Brink, "OFDM-autoencoder for End-to-End learning of communications systems," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2018, pp. 1–5.
[14] M. Kim, W. Lee, and D.-H. Cho, "A novel PAPR reduction scheme for OFDM system based on deep learning," *IEEE Commun. Lett.*, vol. 22, no. 3, pp. 510–513, Mar. 2018.
[15] J. Kim, B. Lee, H. Lee, Y. Kim, and J. Lee, "Deep learning-assisted multi-dimensional modulation and resource mapping for advanced OFDM systems," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2018, pp. 1–6.
[16] T. Wada, T. Toma, M. Dawodi, and J. Baktash, "A denoising autoencoder based wireless channel transfer function estimator for OFDM communication system," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, 2019, pp. 530–533.
[17] T. Van Luong, Y. Ko, N. A. Vien, M. Matthaiou, and H. Q. Ngo, "Deep energy autoencoder for noncoherent multicarrier MU-SIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3952–3962, Jun. 2020.
[18] T. J. Oshea, T. Roy, N. West, and B. C. Hilburn, "Physical layer communications system design over-the-air using adversarial networks," in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 529–532.
[19] V. Raj and S. Kalyani, "Backpropagating through the air: Deep learning at physical layer without channel models," *IEEE Commun. Lett.*, vol. 22, no. 11, pp. 2278–2281, Nov. 2018.
[20] B. Zhu, J. Wang, L. He, and J. Song, "Joint transceiver optimization for wireless communication PHY using neural network," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1364–1373, Jun. 2019.

[21] Y. Zhang, Z. Wang, Y. Huang, W. Wei, G. F. Pedersen, and M. Shen, "A digital signal recovery technique using DNNs for LEO satellite communication systems," *IEEE Trans. Ind. Electron.*, vol. 68, no. 7, pp. 6141–6151, Jul. 2021.
[22] Q. Chen, Y. Zhang, F. Jalili, Z. Wang, Y. Huang, Y. Wang, Y. Liu, G. F. Pedersen, and M. Shen, "Robust digital signal recovery for LEO satellite communications subject to high SNR variation and transmitter memory effects," *IEEE Access*, vol. 9, pp. 135803–135815, 2021.
[23] 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.
[24] P. A. Oche, G. A. Ewa, and N. Ibekwe, "Applications and challenges of artificial intelligence in space missions," *IEEE Access*, early access, Dec. 3, 2021, doi: 10.1109/ACCESS.2021.3132500.
[25] F. Alagoz and G. Gur, "Energy efficiency and satellite networking: A holistic overview," *Proc. IEEE*, vol. 99, no. 11, pp. 1954–1979, Nov. 2011.
[26] M. Ibnkahla, Q. M. Rahman, A. I. Sulyman, H. A. Al-Asady, J. Yuan, and A. Safwat, "High-speed satellite mobile communications: Technologies and challenges," *Proc. IEEE*, vol. 92, no. 2, pp. 312–339, Feb. 2004.
[27] A. Sridhar and A. Ephremides, "Comparison of satellite and cellular architectures for downlink broadcast data transmission," *Int. J. Satell. Commun. Netw.*, vol. 25, no. 1, pp. 27–43, Jan. 2007, doi: 10.1002/sat.844.
[28] A. Sridhar and A. Ephremides, "SAT04–1: Satellite vs. cellular broadcasting : A multi-criteria comparison," in *Proc. IEEE Globecom*, Nov. 2006, pp. 1–5.
[29] M. Rahman, T. Walingo, and F. Takawira, "Impact of varying wireless channel on the performance of LEO satellite communication system," in *Proc. AFRICON*, Sep. 2015, pp. 1–6.
[30] M. Yang, F. Meng, S. Shi, and Q. Guo, "Markov chain based two-state satellite mobile channel model," in *Proc. IEEE 73rd Veh. Technol. Conf. (VTC Spring)*, May 2011, pp. 1–5.
[31] P. Pradhan, O. Faust, S. Patra, and B. Chua, "Channel estimation algorithms for OFDM systems," in *Proc. Int. J. Signal Imag. Syst. Eng.*, vol. 5, Jan. 2011, pp. 15–33.
[32] *Datasheet AMO AiP AAiPK428GC-A0404-XXX*, document eVB-R05 190212 b, AMOTECH, Feb. 2019.
[33] O. Edfors, M. Sandell, J. J. van de Beek, S. K. Wilson, and P. O. Börjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, Jul. 1998.

MARTIN H. NIELSEN was born in Aalborg, Denmark. He received the B.Eng. degree in electronic systems and the M.Eng. degree in wireless communications from Aalborg University, Denmark, in 2017 and 2019, respectively, where he is currently pursuing the Ph.D. degree, researching AI, RF, and mm-wave systems. His research interests include artificial intelligence, circuits and systems for 5G and 6G, and satellite communication.

ELISABETH DE CARVALHO (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Telecom Paris Tech, France. She was a Postdoctoral Fellow at Stanford University, Standford, CA, USA, and then with industry in the field of DSL and wireless LAN. Since 2005, she has been an Associate Professor with Aalborg University, where she has led several research projects in wireless communications. Her main research interests include signal processing for MIMO communications, with recent focus on massive MIMO, including channel measurements, channel modeling, beamforming, and protocol aspects.

MING SHEN (Senior Member, IEEE) was born in Yuxi, China. He received the Ph.D. degree in wireless communications from Aalborg University, Denmark. His Ph.D. thesis was the Spar Nord Annual Best Thesis nomination. He is currently an Associate Professor of RF and mm-wave systems with Aalborg University. His research interests include circuits and systems for 5G and Satcom, biomedical sensing, and artificial intelligence. He received Grants from several Danish national research projects. He is a TPC member of IEEE Nordic Circuits and Systems Conference (NORCAS).