IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2021.DOI

Robust Digital Signal Recovery for LEO Satellite Communications Subject to High SNR Variation and Transmitter Memory Effects

QINGYUE CHEN[1,2,3], YUFENG ZHANG[1,2,3], FERIDOON JALILI[3], ZHUGANG WANG[1], YONGHUI HUANG[1], YUBO WANG[1,2,3], YING LIU[4], GERT FRØLUND PEDERSEN[3], (Senior Member, IEEE), and MING SHEN[3], (Member, IEEE)

[1]National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
[3]Department of the Electronic Systems, Aalborg University, Aalborg 9220, Denmark
[4]School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China

Corresponding author: Qingyue Chen (e-mail: qich@es.aau.dk)

ABSTRACT This paper proposes a robust digital signal recovery (DSR) technique to tackle the high signal-to-noise ratio (SNR) variation and transmitter memory effects for broadband power efficient down-link in next-generation low Earth orbit (LEO) satellite constellations. The robustness against low SNR is achieved by concurrently integrating magnitude normalization and noise feature filtering using a filtering block built with one batch normalization (BN) layer and two bidirectional long short-term memory (BiLSTM) layers. Moreover, unlike existing deep neural network-based DSR techniques (DNN-DSR), which failed to effectively take into account the memory effects of radio-frequency power amplifiers (RF-PAs) in the model design, the proposed BiLSTM-DSR technique can extracts the sequential characteristics of the adjacent in-phase (I) and quadrature (Q) samples, and hence can obtain superior memory effects compensation compared with the DNN-DSR technique. Experimental validation results of the proposed BiLSTM-DSR with a 100 MHz bandwidth OFDM signal demonstrate an excellent performance of 11.83 dB and 9.4% improvement for adjacent channel power ratio (ACPR) and error vector magnitude (EVM), respectively. BiLSTM-DSR also outperforms the existing DNN-DSR technique in terms of the ACPR and EVM by 2.4 dB and 0.9%, which provides a promising solution for developing deep learning-assisted receivers for high-throughput LEO satellite networks.

INDEX TERMS
Bidirectional long short-term memory (BiLSTM), robust digital signal recovery (DSR), radio-frequency power amplifiers (RF-PAs), low Earth orbit (LEO), broadband communications.

I. INTRODUCTION

LOW Earth orbit (LEO) satellite constellations have grown drastically in recent years and have attracted great attention from both academia and industry. Unlike geostationary Earth orbit (GEO), medium Earth orbit (MEO), and high Earth orbit (HEO) satellites, LEO satellites are deployed at altitudes between 500 and 2000 km to ensure low latency. Owing to their attractive features of vast global coverage, low propagation latency, and low costs of manufacturing and deployment [1]–[3], LEO satellites are a key enabling technology in expanding the coverage of fifth-generation (5G) networks and are expected to play a significant role in sixth-generation (6G) communication networks [4].

The increasing demands of higher data rates and the resultant use of efficient high-frequency signals pose severe challenges for spaceborne radio-frequency power amplifiers (RF-PAs) while maintaining high power efficiency and linearity. High-frequency efficiency signals with a high peak-to-average power ratio (PAPR), such as OFDM signals, are promising candidates for broadband satellite links [5]. Due to the extremely limited power supplies, RF-PAs prefer to work near the saturation region, where the gain is compressed.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3117517, IEEE Access

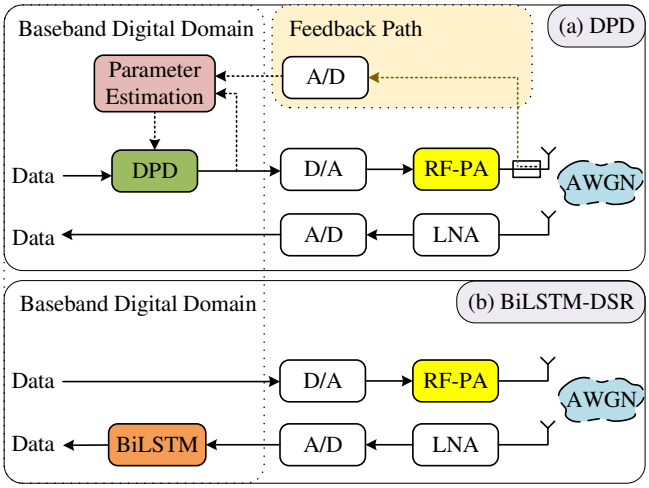


FIGURE 1: System architecture comparison between conventional DPD and the proposed BiLSTM-DSR techniques. (a) DPD technique, (b) BiLSTM-DSR technique.

This causes out-of-band spectral regrowth and in-band nonlinear signal distortion, resulting in an increased adjacent channel power ratio (ACPR) and a deteriorated error vector magnitude (EVM) at ground stations. Hence, mitigating the performance degradation caused by the nonlinear distortion of RF-PAs has become urgent for broadband LEO satellite communication systems. Different from RF-PA linearization techniques for narrow band terrestrial communication systems, the mitigation of nonlinear distortion in broadband LEO satellite systems features a few unique challenges including the remarkably varying SNR of the received signal due to the low obit, the strong memory effects that are uneasy to characterize when signal bandwidth increases, and the stringent requirement of high power efficiency and low system complexity for the satellites.

The power back-off technique developed in previous years for terrestrial systems is not suitable for LEO satellites due to the low efficiency of RF-PAs. Although the digital predistortion (DPD) [6]–[8] technique can achieve excellent results in improving linearity while maintaining the high efficiency of RF-PAs, it requires a feedback path as well as extra power for signal processing, which is a major challenge for LEO communication systems (Fig. 1). According to the Consultative Committee for Space Data Systems (CCSDS) 131.2.B.1 standard, the Application Specific Integrated Circuit (ASIC) design includes a pre-distorter module providing 5th order static pre-distortion [9], which means that there is no feedback path at the transmitter to receive RF-PAs output signals for adjusting the pre-distortion parameters in real-time. The pre-distortion parameter is fixed and the outputs of RF-PAs cannot be tracked. Thus, the conventional DPD techniques with feedback path have not been widely used due to limited spaceborne resources. To overcome this challenge, several techniques have been studied to recover distorted signals at the receiver. In the digital post-distortion technique [10], the RF-PA and the channel are first modeled using a digital detector based on a Volterra model, and the Volterra kernels and the transmitted symbols are estimated using a Kalman filter (KF)-based algorithm. In [11], a partitioned distortion mitigation technique incorporates the phase-DPD technique at the transmitter and the amplitude-post-compensation technique at the receiver. These techniques can improve the overall power efficiency and bit error rate (BER) performance of the transceiver. However, they cannot handle the time-varying signal power caused by the varying transmission distance between LEO satellites and ground stations [12]. Actually, there are several techniques to deal with varying power in the receiver. In [13], a new method based on time varying gain amplifier (TVGA) is proposed for obtaining a wider dynamic range of the received power and an adaptive sweep optimization (ASO) method is applied for increasing the linearity of TVGA. In a novel Automatic gain control (AGC) circuit [14], a flexible gain compensation scheme with hybrid gamma parameters is adapted to accommodate different types of applications. Nevertheless, these techniques cannot cope with high signal-to-noise ratio (SNR) variation well.

Deep neural network-based digital signal recovery (DNN-DSR) is a novel technique allowing spaceborne RF-PAs to operate with high efficiency by compensating for the nonlinear distortion in the signals received at ground stations [12]. The DNN-DSR technique has demonstrated its capability of handling varying power and additive white Gaussian noise (AWGN). This allows LEO spaceborne RF-PAs to operate near their saturation regions for high power efficiency while maintaining satisfactory EVM performance at ground stations. However, the DNN pays little attention to the time series characteristics of the data, which is closely related to the memory effects of RF-PAs and hence leaves much room for improvement.

Long short-term memory (LSTM) networks have a powerful capability of capturing features in time series data, benefiting from their recurrent structure [15]–[18]. Given this advantage, LSTM networks have been applied to the behavioral modeling of RF-PAs and the linearization of RF-PAs in recent years [19]–[22]. However, the correlation between the memory effects of RF-PAs and LSTM networks is not studied deeply and intuitively in these papers, especially for broadband systems. Besides, they are mainly to prevent interference to adjacent channels, and are not EVM quantitative evaluation based RF-PAs nonlinear modeling and linearization. These neural network-based DPD techniques not only require feedback paths, but also need to store more parameters than memory polynomial (MP) methods. Therefore, the system complexity and high power consumption brought by these DPD techniques applied to transmitters are unacceptable for LEO satellites.

In this paper, a novel DSR technique that directly improves the signal quality by reducing the EVM at the receiver based on bidirectional LSTM (BiLSTM) is proposed. As illustrated in Fig. 1, the distorted signal caused by RF-PAs can be

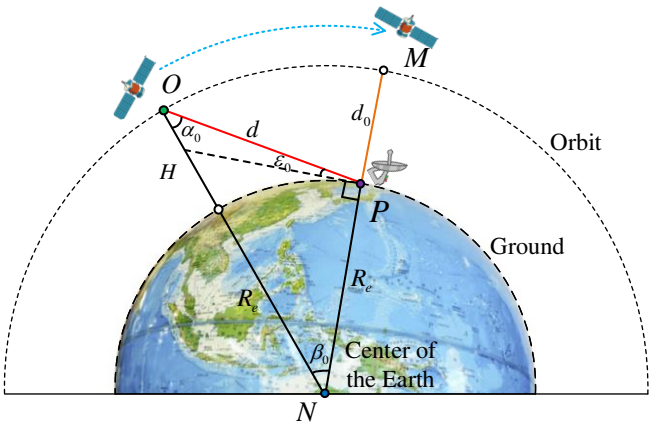


FIGURE 2: Geometry graph of the satellite-to-ground data transmission link.

removed by BiLSTM networks at the receiver side. The intuitive connection between BiLSTM networks and the memory effects of RF-PAs is established by training the BiLSTM model using data measured from the RF-PAs. In addition to retaining the robustness against power variation of the DNN-DSR, the proposed technique achieves better recovery performance in terms of ACPR and EVM. In particular, the better robustness of high SNR variation is obtained. Experiments were conducted to compare the DNN-DSR, LSTM-DSR, and BiLSTM-DSR techniques under both narrowband and broadband conditions for performance validation.

This paper is organized as follows: Section I is the introduction. Section II describes problem formulation. Section III presents the proposed BiLSTM-DSR technique. The experimental results of different techniques are provided in Section IV. Finally, the conclusion of this paper is presented in Section V.

II. PROBLEM FORMULATION

In this section, we pay close attention to the high SNR variation and inherent memory effects of RF-PAs.

A. DYNAMIC SATELLITE-TO-GROUND LINK CONDITIONS AND SNR VARIATION

For the transmission performance of downlink telemetry data, the SNR of the received signal is a very important indicator. This is because the free space loss (FSL) is affected by the distance variation between the satellite and the ground station, thereby affecting the link budget.

As is illustrated in Fig. 2, the satellite-to-ground data transmission link changes dynamically, where ε_0 represents the elevation angle, d is the distance of data transmission, H denotes the orbital altitude of LEO satellite, $R_e = 6371$ km is the earth radius.

When the satellite reaches point O, the elevation angle ε_0 is 5°, and the data transmission link is established. The distance d has the largest value at this point. When the satellite moves to point M, the elevation angle increases to 90° and the distance d_0 is the shortest. In the ΔNOP, there's the geometric relationship expressed as

$$(H + R_e)^2 = {R_e}^2 + d^2 - 2R_e d\cos\left(\frac{\pi}{2} + \varepsilon_0\right), \quad (1)$$

therefore, the distance can be written as a function of ε_0

$$d(\varepsilon_0) = R_e\left[\sqrt{\left(\frac{H+R_e}{R_e}\right)^2 - \cos^2\varepsilon_0} - \sin\varepsilon_0\right]. \quad (2)$$

The FSL is the main loss of the transmission link, and it can be decided by the carrier frequency f and the distance d [23]:

$$L(\varepsilon_0) = \left(\frac{4\pi d(\varepsilon_0)}{\lambda}\right)^2 = \left(\frac{4\pi f d(\varepsilon_0)}{c}\right)^2. \quad (3)$$

In the receiver model, the noise comes from the received background and the receiver itself. These noises are fixed values, regardless of the distance. Therefore, the varying SNR can be represented by

$$\Delta\frac{S}{N}(dB/Hz) = \left(\frac{S}{N}\right)_{5^\circ} - \left(\frac{S}{N}\right)_{90^\circ} = \Delta L(\varepsilon_0). \quad (4)$$

Assuming that the orbital altitude of LEO satellite is 800 km, then the distance variation brings a approximately 12 dB link budget variation in received signal power. Thus, the variation range of SNR would be as big as 12 dB. Furthermore, as the down-link window is typically around 10 to 15 minutes, the drastic change in SNR of the received down-link signal within the short window exhibits great challenge for reliable signal recovery and therefore calls for novel robust DSR techniques.

B. MEMORY EFFECTS IN RF-PAS

With the increasing need for broadband signal formats such as OFDM in LEO satellite communications, the memory effects of RF-PAs tend to be significant and cannot be ignored [24]. To fully understand the nonlinearity of RF-PAs, memory effects must be taken into account. As RF microwave devices, the memory effects of RF-PAs based on the GaN/GaAs process result from dynamic self-heating and dynamic trapping [25]. Generally, memory effects make the output power of RF-PAs depend not only on the current input signal but also on the previous input signal of the RF-PAs. One impact of these "memory" effects is that the gain of the RF-PA is much larger than that of the corresponding RF carrier period on the same time scale. These so-called memory effects can be expressed by several mathematical methods, including the MP [26], Generalized MP (GMP) [27], Volterra [28], and Wiener-Hammerstein [29] methods.

For the narrowband case, the MP model is the most commonly used RF-PA modeling method due to its low complexity. A typical MP of nonlinearity order K and memory depth Q is represented by

$$y_{MP}(n) = \sum_{k=0}^{K-1}\sum_{q=0}^{Q-1} c_{k,q}x(n-q)|x(n-q)|^k, \quad (5)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3117517, IEEE Access

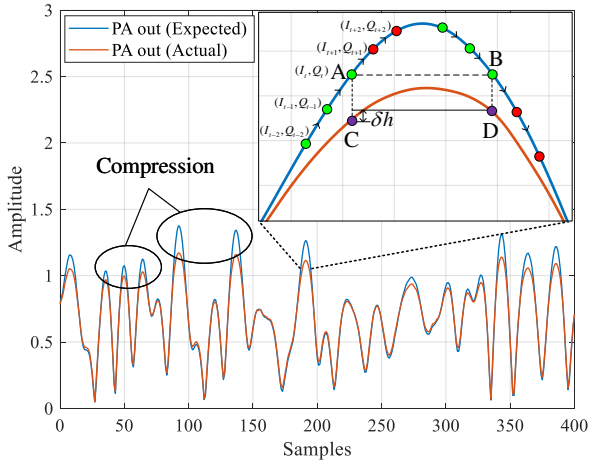
where $x(n)$ and $y_{MP}(n)$ represent the input signal and the output signal of the RF-PAs, respectively. $c_{k,q}$ are the complex polynomial coefficients.

For multi-carrier and wideband signals, the GMP model is required, which can be mathematically expressed as

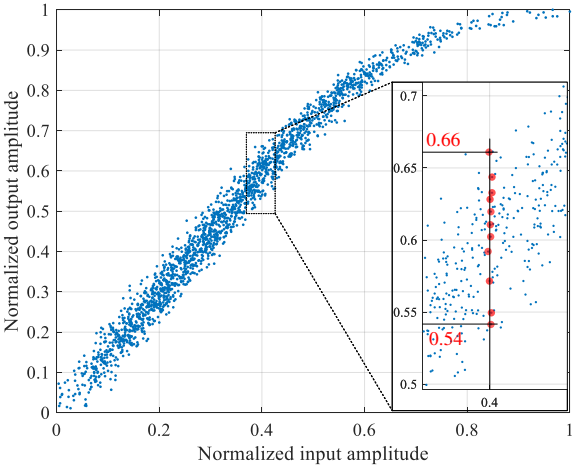
$$\begin{aligned} y_{GMP}(n) = &\sum_{k=0}^{K_a-1}\sum_{l=0}^{L_a-1} a_{kl}x(n-l)|x(n-l)|^k \\ &+\sum_{k=1}^{K_b}\sum_{l=0}^{L_b-1}\sum_{m=1}^{M_b} b_{klm}x(n-l)|x(n-l-m)|^k \\ &+\sum_{k=1}^{K_c}\sum_{l=0}^{L_c-1}\sum_{m=1}^{M_c} c_{klm}x(n-l)|x(n-l+m)|^k. \end{aligned} \tag{6}$$

where $x(n)$ and $y_{GMP}(n)$ are the input signal and the output signal of the RF-PAs, respectively. K_a and L_a are the order and memory depth of the MP sub-model, respectively. (K_b, L_b, M_b) and (K_c, L_c, M_c) are the order, the memory depth, and the maximum orders of the lagging and leading cross-terms, respectively. a_{kl}, b_{klm} and c_{klm} are the model coefficients.

An illustration of the gain compression and baseband electrical memory effects of the large-amplitude signal caused by the nonlinearity of the RF-PAs is shown in Fig. 3(a). Two sets of five adjacent samples are indicated with circles. The actual outputs (i.e., points C and D) are different for the expected outputs of the same value (i.e., points A and B) subject to the memory effects. The difference is indicated as δh in Fig. 3(a). In addition, more different outputs can be obtained throughout the entire transmission waveform. In particular, in a broadband scenario with a high PAPR, such a difference is more prominent [30], resulting in the AM-AM curves of Fig. 3(b). For the same normalized input amplitude of 0.4, the normalized output amplitude ranges from 0.54 to 0.66 due to the different previous states. The transmitter memory effects of RF-PAs have a great impact on the transmission performance.



(a) Gain compression and memory effects manifested in amplitude curves.



(b) Output diffusion due to memory effects.

FIGURE 3: Illustration of (a) the gain compression and memory effects manifested in the amplitude curves, and (b) output diffusion due to memory effects.

III. PROPOSED TECHNIQUE BASED ON BILSTM NETWORK

In this section, we exhibit the complete BiLSTM-based network structure and propose solutions for high SNR variation and transmitter memory effects. Furthermore, we also briefly discuss the training process and complexity analysis.

A. BILSTM-DSR NETWORK STRUCTURE

As illustrated in Fig. 4, the proposed BiLSTM-DSR is a technique involving off-line training phase and on-line implementation. During off-line training phase, the proposed BiLSTM-based network captures the nonlinear features of the RF-PA by extracting the sequential characteristics of adjacent sampled data. With the result of the loss function gradually decreasing to the preset threshold, the training is completed. The structure and parameters of the proposed network are fixed and integrated into the receiver. Subsequently, the LEO satellite with the RF-PA is launched, and the receiver with the proposed network can correct the distortion of the RF-PA in the on-line real-time application scenario. The proposed BiLSTM-DSR architecture consists of five layers, an input layer, two BiLSTM layers, a fully connected layer, and an output layer. Unlike DNN-DSR, which accepts only one-dimensional (1D) inputs, the proposed technique accepts two-dimensional (2D) signals as inputs. Each sample is fed into the network in a 2D data format (i.e., (I_t, Q_t)), preserving the original amplitude and phase information. To capture the memory effects while maintaining a reasonable training time, a memory depth of 3 is chosen for the narrowband case. Thus, the numbers of neurons in the input layer and output layer are set to 5 (i.e., the timestep is set to 5) and 2, respectively. The BiLSTM layer receives a signal sequence consisting of advanced and delayed items.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3117517, IEEE Access

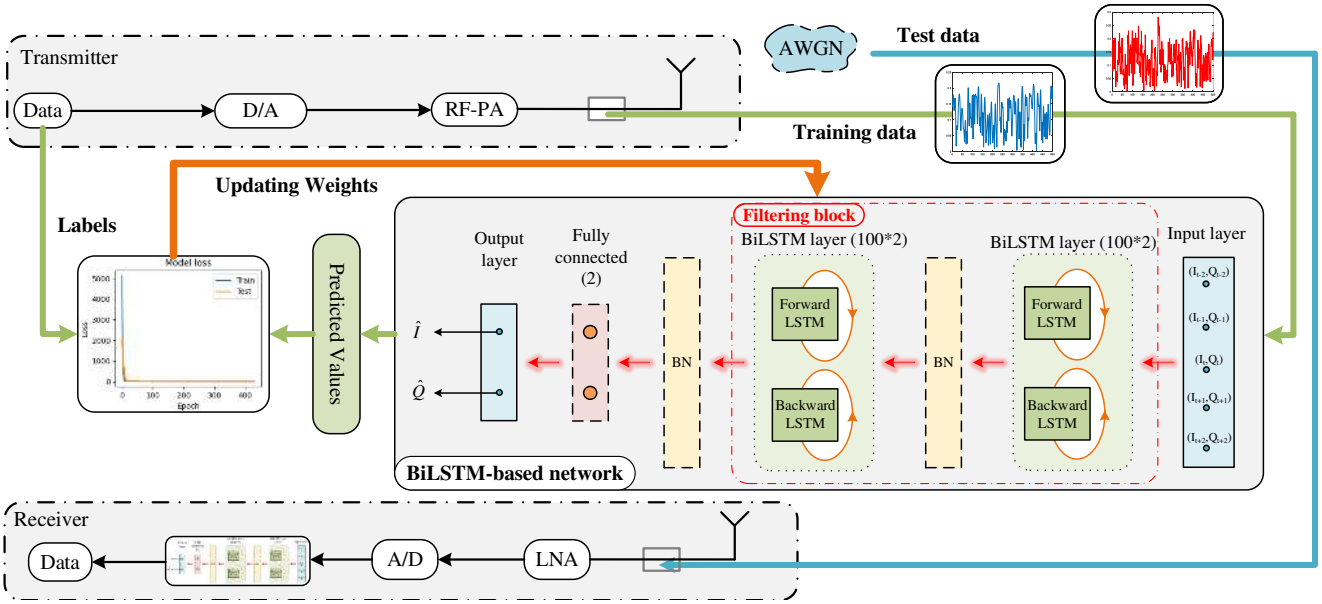


FIGURE 4: The off-line training and on-line implementation process of the proposed BiLSTM-DSR technique.

Hyperbolic tangent and sigmoid functions are used as activation functions in each BiLSTM layer. Then, the outputs from each BiLSTM layer are fed into a batch normalization (BN) layer. The output of the last BN layer is used as the input of the fully connected layer. In the final layer, the output with two neurons provides the desired outputs $(\hat{I}_t, \hat{Q}_t)$.

B. NORMALIZATION AND FILTERING FOR HANDLING HIGH SNR VARIATION

It is the simplest and most effective way to train the network by mapping the highly distorted signal directly to the pure noiseless signal. During the training process, the BiLSTM-based network obtains the anti-noise capability by capturing the long-term and short-term correlation of signals. Therefore, as shown by the filtering block in BiLSTM-based network in Fig. 4, we explicitly design two BiLSTM layers and one BN layer to make the network more robust against AWGN.

The extra BN layer with normalization is adopted to handle the high SNR variation in this work. The formula for the BN layer can be expressed as

$$\hat{\boldsymbol{y}}_i = \gamma \frac{\boldsymbol{y}_i - E[\boldsymbol{y}_i]}{\sqrt{Var[\boldsymbol{y}_i] + \varepsilon}} + \boldsymbol{\beta}, \tag{7}$$

where $\boldsymbol{y}_i$ and $\hat{\boldsymbol{y}}_i$ are the input and output of the BN layer, respectively. γ and $\boldsymbol{\beta}$ represent the scaling and shifting parameters, respectively. Besides, ε is set to 0.001 to prevent the denominator from being zero. Through normalization of the input data by recentering and rescaling [31], BN layer can allows artificial neural networks to quickly and stably respond to the high SNR variation signals.

In the on-line application phase, when the SNR of the received signal is 15 dB, the normalized output amplitude of each layer is plotted in Fig. 5 to show the process of the signal

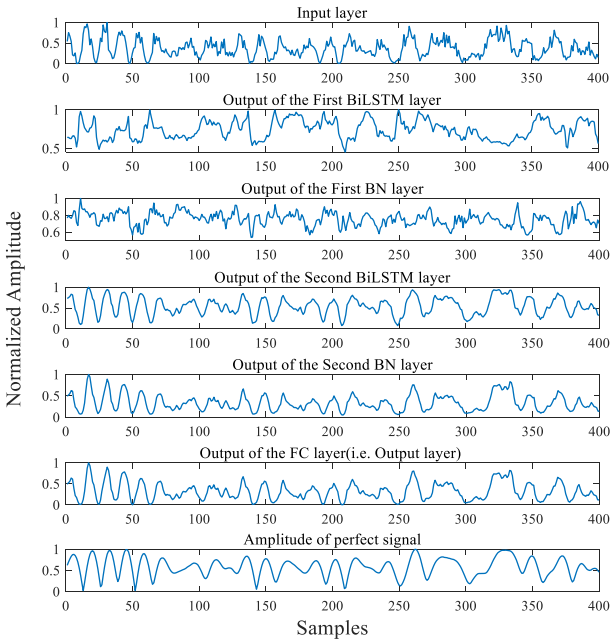


FIGURE 5: Normalized outputs of each layer in the proposed BiLSTM network.

recovery performance. After the filtering block of BiLSTM-based network in Fig. 4, the noise in the received signal is clearly suppressed, and the SNR is increased by 6.63 dB (the four subplots on the top in Fig. 5). At the same time, the useful signal is maintained intact without distortion. With the subsequent BN layer and FC layer, the output of the FC layer matches the perfect signal to the RF-PA well, which means the original signal transmitted by the RF-PA is recovered with high precision (the two subplots at the bottom in Fig. 5).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3117517, IEEE Access

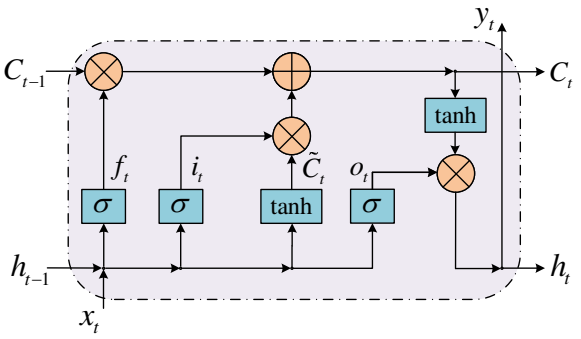


FIGURE 6: Schematic diagram of LSTM cell. The symbols of $\oplus$, $\otimes$, and σ represent the operation of addition, element-wise multiplication, and Sigmoid function, respectively.

C. BILSTM BASED NETWORK FOR TACKLING MEMORY EFFECTS

Unlike traditional methods, an LSTM cell is applied to capture the memory effects of RF-PAs in this work. Additionally, LSTM networks can not only store short-term states but also maintain long-term states. LSTM utilizes a special unit, named a cell, as illustrated in Fig. 6, to overcome the problem of vanishing and exploding gradients in common RNNs. LSTM carries out the preservation or removal of model input information according to its importance by adding a forget gate, input gate, output gate, and nonlinear output. All the gates of the cell calculate the activation (using an activation function) of a weighted sum ($\boldsymbol{wx} + \boldsymbol{b}$), similar to standard neurons, where $\boldsymbol{w}$, $\boldsymbol{x}$, and $\boldsymbol{b}$ stand for weights, inputs, and biases, respectively.

In Fig. 6, the forget gate $\boldsymbol{f}_t$ determines whether to retain the previous data in the cell state $\boldsymbol{C}_{t-1}$. The input gate $\boldsymbol{i}_t$ selectively records new data into the cell state. The output gate $\boldsymbol{o}_t$ decides which data will be outputted. The temporary status $\tilde{\boldsymbol{C}}_t$ denotes a vector of the new candidate values. The specific data flow of the LSTM network is described as follows:

$$\boldsymbol{f}_t = \sigma(\boldsymbol{W}_f[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_f), \tag{8}$$

$$\boldsymbol{i}_t = \sigma(\boldsymbol{W}_i[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_i), \tag{9}$$

$$\tilde{\boldsymbol{C}}_t = \tanh(\boldsymbol{W}_C[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_C), \tag{10}$$

and

$$\boldsymbol{o}_t = \sigma(\boldsymbol{W}_o[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_o), \tag{11}$$

where $\boldsymbol{W}_f$, $\boldsymbol{W}_i$, $\boldsymbol{W}_C$, and $\boldsymbol{W}_o$ represent the weight matrices for the joint vector (i.e., $[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t]$) composed of the current input vector (i.e., $\boldsymbol{x}_t$) and the previous hidden state (i.e., $\boldsymbol{h}_{t-1}$). $\boldsymbol{b}_f$, $\boldsymbol{b}_i$, $\boldsymbol{b}_C$, and $\boldsymbol{b}_o$ are the corresponding bias terms.

The current long-term status $\boldsymbol{C}_t$ and the LSTM cell's output $\boldsymbol{y}_t$ can be calculated by

$$\boldsymbol{C}_t = \boldsymbol{f}_t \otimes \boldsymbol{C}_{t-1} + \boldsymbol{i}_t \otimes \tilde{\boldsymbol{C}}_t \tag{12}$$

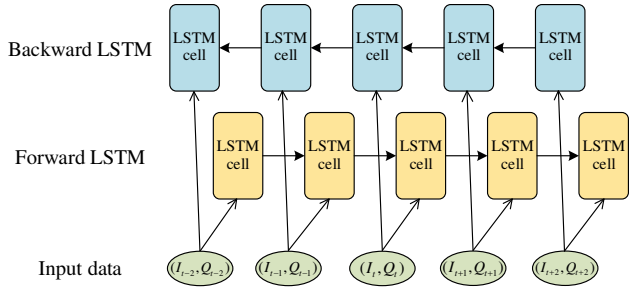


FIGURE 7: The forward and backward LSTM process of the BiLSTM layer.

and

$$\boldsymbol{y}_t = \boldsymbol{h}_t = \boldsymbol{o}_t \otimes \tanh(\boldsymbol{C}_t), \tag{13}$$

where $\boldsymbol{C}_{t-1}$ denotes the previous long-term status. The symbol $\otimes$ represents element-wise multiplication [20]. As is clearly shown, the current output depends not only on the current long-term status and input but also on the previous short-term status.

There is an intuitive timing relation between the marked samples in Fig. 3(a). Due to the memory effects of the RF-PAs, (I_t, Q_t) is affected by the data at previous moments (i.e., (I_{t-2}, Q_{t-2}) and (I_{t-1}, Q_{t-1})) and prominently affects the next data (i.e., (I_{t+1}, Q_{t+1}) and (I_{t+2}, Q_{t+2})). Therefore, to extract the memory features of RF-PAs more completely, we utilize a bidirectional mechanism for the LSTM network. The forward and backward LSTM processes of the BiLSTM layer are shown in Fig. 7. Each training process loops the five input samples forward and backward, and the BiLSTM layer records the timing characteristics of the input data accordingly. In this way, the network can better retrieve memory effects of RF-PAs.

D. NETWORK TRAINING

To conveniently predict the intermediate sequences of both forward and backward LSTM networks, we set the data format of the training data and labels to that shown in Table 1. The characteristic information used for off-line training is extracted from the received signal containing the nonlinearity of the RF-PAs. The labels are allocated according to the I/Q signal of the transmitter constellation trace.

TABLE 1: Training data and Lables

Training data	Lables
$(I_{t-2}, Q_{t-2})(I_{t-1}, Q_{t-1})(I_t, Q_t)(I_{t+1}, Q_{t+1})(I_{t+2}, Q_{t+2})$	$(\hat{I}_t, \hat{Q}_t)$
$(I_{t-1}, Q_{t-1})(I_t, Q_t)(I_{t+1}, Q_{t+1})(I_{t+2}, Q_{t+2})(I_{t+3}, Q_{t+3})$	$(\hat{I}_{t+1}, \hat{Q}_{t+1})$
$(I_t, Q_t)(I_{t+1}, Q_{t+1})(I_{t+2}, Q_{t+2})(I_{t+3}, Q_{t+3})(I_{t+4}, Q_{t+4})$	$(\hat{I}_{t+2}, \hat{Q}_{t+2})$
$(I_{t+1}, Q_{t+1})(I_{t+2}, Q_{t+2})(I_{t+3}, Q_{t+3})(I_{t+4}, Q_{t+4})(I_{t+5}, Q_{t+5})$	$(\hat{I}_{t+3}, \hat{Q}_{t+3})$
.....	.....

According to the proposed network architecture, the back propagation (BP) algorithm is used to train the network. First, we chose Huber loss as the loss function of the BP algorithm, as it can strengthen the robustness of the mean square error (MSE) against outliers, which stand for the abnormal I/Q

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3117517, IEEE Access

TABLE 2: Model training configurations

Maximum Epochs	500
Minimum Batch Size	512
Optimizer	RMSprop
Loss Function	Huber loss
State Activation Function	Tanh
Gate Activation Function	Sigmoid
Initial Learning Rate	0.001
Early Stop Patience	50 epochs
Learn Rate Drop Factor	0.1
Minimum Learning Rate	10^{-5}
Dropout	0.05
Time Step	5
Bias Regularizer	L2(1e-3)
Testing Samples	10^5

samples that are excessively far from the actual samples, as shown in Fig. 10(a). The Huber loss function can be expressed as

$$L_\delta(y, f(x)) = \begin{cases} \frac{1}{2}(y - f(x))^2 & for\ |y - f(x)| \le \delta \\ \delta\,|y - f(x)| - \frac{1}{2}\delta^2 & otherwise \end{cases}, \tag{14}$$

where δ is the parameter of Huber loss, which is set to 1. y and $f(x)$ represent the observed and predicted values, respectively. The root-mean-square prop (RMSprop) is usually an appropriate choice for LSTM network training due to its adaptive learning rate. In this work, we adopted the RMSprop algorithm to calculate the optimal parameters of the proposed network. Additionally, the detailed simulation parameters of the BiLSTM-based network are shown in Table 2.

E. COMPUTATIONAL COMPLEXITY ANALYSIS

In our scenario, the input size is $samples \times timesteps \times M$, where M represents the number of features of one sample, which are 2D data containing I/Q information. For the proposed BiLSTM-based network, the computational complexity is relevant to the number of units in the network. The computational complexity of each BiLSTM layer is $O(2 \times 4 \times (M + n) \times n + n)$, where n denotes the number of neurons of each gate in an LSTM cell and the number 2 represents the forward LSTM and backward LSTM. It should be noted that $timesteps$ affects only the training time, not the computational complexity of the BiLSTM layer.

Then, the computational complexity of the BN layer is $O(4 \times L)$, where L is the number of output units in the previous layer. The number 4 represents the mean, variance, scaling parameter, and shifting parameter. Finally, the computational complexity of the FC layer is $O((s + 1) \times d)$, where s and d stand for the number of output units in the previous layer and the neurons of the current FC layer, respectively. Overall, the computational complexity of the proposed network is $O(a\times(2\times4\times(M+n)\times n+n)+b\times(4\times L)+c\times((s+1)\times d))$, where a, b, and c represent the number of BiLSTM layers, BN layers, and FC layers, respectively.

The only structural difference between the LSTM-based network and the proposed BiLSTM-based network is that the former contains only forward LSTM and no backward LSTM. Based on the detailed analysis above, the computational complexity of the LSTM network is $O(a \times (4 \times (M + n) \times n + n) + b \times (4 \times L) + c \times ((s + 1) \times d))$. The meanings of the symbols are the same as above. Since the activation function layer does not increase the computational complexity, the computational complexity of the DNN is $O(b\times(4\times L)+c\times((s+1)\times d))$. The meanings of the symbols remain the same as above. The computational complexity comparison of the three networks is shown in Table 3.

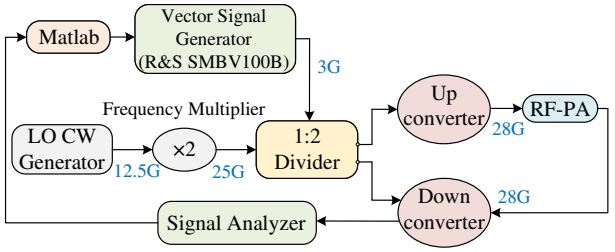


FIGURE 8: Schematic diagram of the measurement setup.

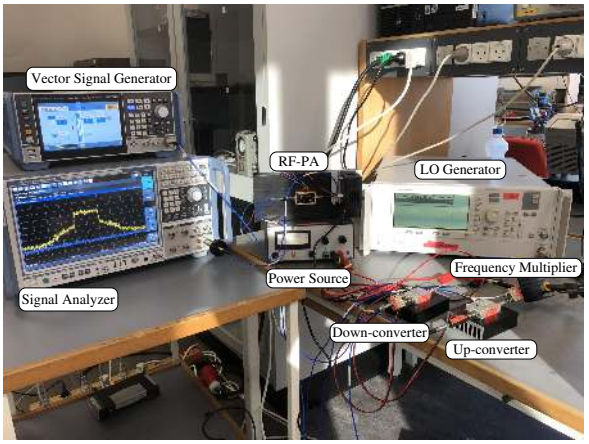


FIGURE 9: Photograph of the measurement setup.

IV. RESULTS BASED ON MEASUREMENTS

In this section, the linearization performance of the proposed BiLSTM-DSR technique is shown and compared with the existing DNN-DSR technique. As BiLSTM-DSR is developed based on LSTM, the DSR results using LSTM (LSTM-DSR) are also included in the comparison.

A. EXPERIMENTAL SETUPS

The experiments include validations with a narrowband, 10 MHz LTE OFDM signal with an 8 W CGH40006P GaN RF-PA operating at 3.5 GHz and a wideband, 100 MHz 5G OFDM signal with a 2 W Ducommun RF-PA operating at 28 GHz. For the narrowband experiment, the data from [12] are used.

The experimental data for broadband validation are collected in this work using the measurement platform illustrated in Fig. 8 and Fig. 9. The input source of the 3 GHz signal for the measurements is generated from a vector signal

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3117517, IEEE Access

TABLE 3: Layers and parameters of different architectures

Architecture	BN-layer	Relu layer	LSTM layer (units of each layer)	FC layer (units of each layer)	Total parameters
DNN-DSR	3	2	0	3([500,350,2])	182,960
LSTM-DSR	2	0	2([200,100])	1(2)	184,202
BiLSTM-DSR	2	0	2([100*2,50*2]))	1(2)	184,202

TABLE 4: Measured EVM and ACPR of different techniques

Bandwidth	10MHz		100MHz	
Scheme	ACPR	EVM	ACPR	EVM
Original signal: PA input	-50.03 dB	-	-46.45 dB	-
PA output without DSR	-32.06 dB	8.03%	-30.08 dB	11.92%
DNN-DSR [12]	-46.43 dB	2.89%	-39.48 dB	3.45%
LSTM-DSR(memoryless)	-47.25 dB	1.89%	-39.99 dB	8.09%
LSTM-DSR	-47.82 dB	1.71%	-40.16 dB	2.95%
BiLSTM-DSR(memoryless)	-47.04 dB	1.88%	-39.79 dB	8.07%
BiLSTM-DSR	**-47.80 dB**	**1.72%**	**-41.91 dB**	**2.52%**

generator (R&S SMBV100B). For both upconversion and downconversion, an unmodulated signal of 12.5 GHz from a local oscillator (LO) continuous wave generator is converted to 25 GHz by a frequency multiplier and fed into a power divider as an LO signal. After upconversion, the 28 GHz signal is fed into the RF-PA. To ensure sufficient nonlinearity for the validation, the RF-PA is driven into a relatively high-compression region at its 1 dB compression point. The EVM is 8.03% and 11.92% for narrowband and wideband signal respectively without linearization, which is a huge challenge for DPD with such strong nonlinearity. After downconversion, the intermediate frequency (IF) signal is captured by signal analyzer and converted to baseband digital domain for Matlab operating.

The PAPRs of the tested narrowband and wideband signals are 10.06 dB and 11.69 dB, respectively. Fig. 10(a) and Fig. 11(a) show the AM-AM and AM-PM curves based on measurements of the 10M bandwidth and 100M bandwidth signals (without DSR), respectively. The nonlinear trajectories of the AM/AM curves reveal the gain compression in different regions of the RF-PA. The obvious diffuse samples in the curves indicate that the RF-PA under test exhibit strong memory effects.

On the other hand, to avoid overfitting and underfitting during the training of the models, an appropriate number of neurons are selected in each layer. In addition to the proposed BiLSTM-DSR, we adopted the DNN-DSR technique and LSTM-DSR technique under the same circumstances. Table 3 provides details about the number of layers and units and the total number of parameters. For fair comparison, the memory depths of DNN-DSR, LSTM-DSR, and BiLSTM-DSR are all set to 3 for the narrowband signal validation case and 4 for the wideband validation case. To demonstrate the obvious advantages of the proposed BiLSTM DSR technique in memory effect mitigation, memoryless BiLSTM-DSR (i.e., with a memory depth of 1) is also included in the comparison. The total number of trainable parameters also remains the same. The training and validation of the proposed technique are implemented using TensorFlow 1.14 through the Keras API in Python 3.7.6. The pre- and post-processing and data plotting are performed in MATLAB 2020b.

B. RECOVERY PERFORMANCE

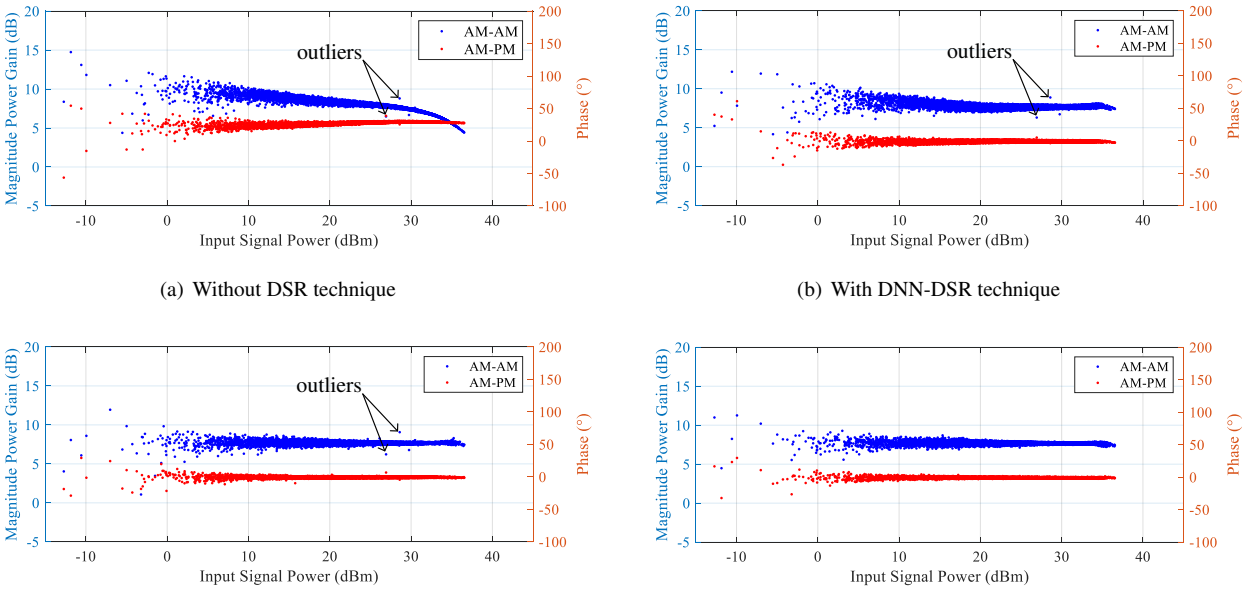
In addition to the 1 dB compression point data, the input and output signals of the RF-PA in wider operating states are collected with a granularity of 1 dBm. Each set of data is trained and tested separately to obtain the relationship between EVM and average output power, as illustrated in Fig. 12. As the average input power increasing, the drain efficiency and the nonlinearity of the RF-PA enhance, while the EVM caused by distortion deteriorates. For the 10 MHz narrowband signals, the EVM performances of LSTM-DSR and BiLSTM-DSR are superior to DNN-DSR by about 1.5% on average. Due to the consistent nonlinearity in the narrowband application, the LSTM-DSR and BiLSTM-DSR techniques with or without memory acquire almost similar results. Whereas, for the 100 MHz wideband signals, the EVM performances of memoryless LSTM-DSR and BiLSTM-DSR both are even 4% to 6% worse than that of DNN-DSR. The RF-PA features non flatten in-band response in the wideband application so that LSTM-DSR and BiLSTM-DSR techniques with memory become more suitable. The EVM performance of BiLSTM-DSR with memory is about 0.5% better than that of LSTM-DSR, and 1% DNN-DSR, respectively.

Since the state of the RF-PA in practical application is set near the 1 dB compression point to maintain a high power efficiency, we mainly focus on the recovery performance at the 1 dB compression point. Under 1 dB compression point, the measured comparative results of the different techniques are summarized in Table 4, the characteristic curve comparisons of AM-AM, AM-PM are shown in Fig. 10 and Fig. 11, and the power spectrum density (PSD) is depicted in Fig. 13. The output of the RF-PA is severely distorted by the nonlinear RF-PA for the narrowband and wideband signals when DSR techniques are not applied.

For the 10 MHz narrowband signals, the ACPR and EVM using BiLSTM-DSR are approximately 1.4 dB and 1.2% better than those of the DNN-DSR technique, respectively. In addition, the recovery effect of BiLSTM-DSR on the outliers, shown in Fig. 10(d), is better than that on the others due to the bidirectional mechanism of BiLSTM, which is of great benefit in boosting the quality of the received signal. Compared with DSR techniques with memory depth, EVM deteriorates by approximately 0.2% for memoryless BiLSTM.

For the 100 MHz wideband signals, the improvements in ACPR and EVM using BiLSTM-DSR are approximately 2.4 dB and 0.9% better than those of the DNN-DSR technique, respectively. In this case, LSTM-DSR is only 0.68 dB and 0.50% better in terms of ACPR and EVM, respectively, in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3117517, IEEE Access



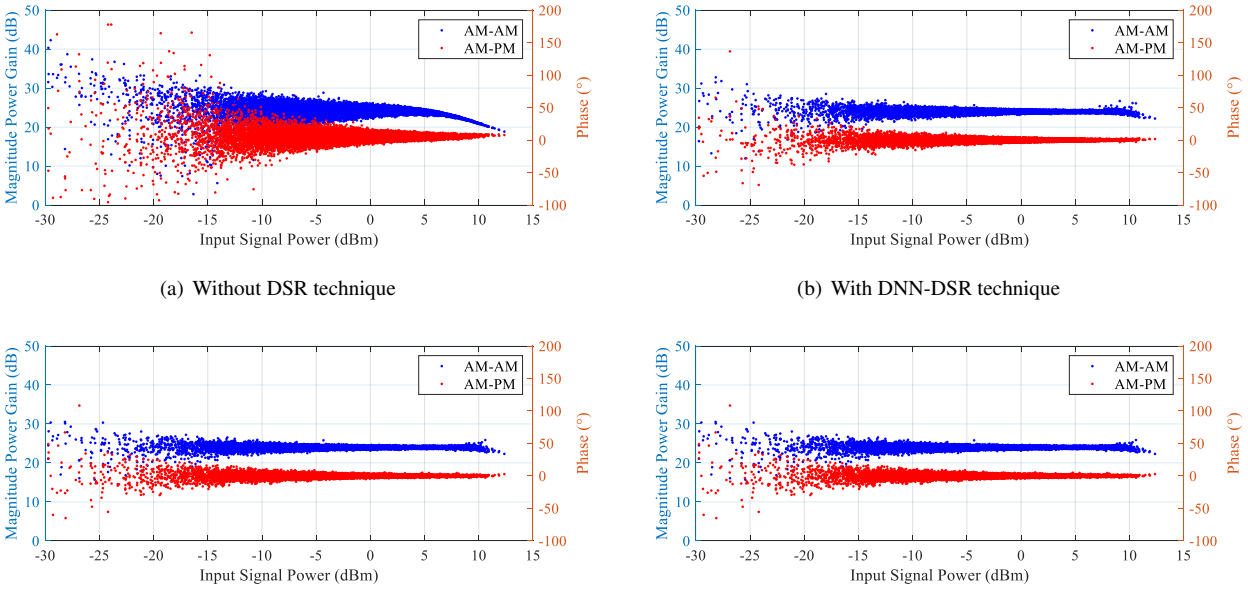
(a) Without DSR technique

(b) With DNN-DSR technique

(c) With LSTM-DSR technique

(d) With BiLSTM-DSR technique

FIGURE 10: AM-AM and AM-PM curves with different DSR techniques for the 10MHz OFDM signal. (a) Without DSR technique, (b) with DNN-DSR technique, (c) with LSTM-DSR technique, (d) with the proposed BiLSTM-DSR technique.



(a) Without DSR technique

(b) With DNN-DSR technique

(c) With LSTM-DSR technique

(d) With BiLSTM-DSR technique

FIGURE 11: AM-AM and AM-PM curves with different DSR techniques for the 100MHz 5G signal. (a) Without DSR technique, (b) with DNN-DSR technique, (c) with LSTM-DSR technique, (d) with the proposed BiLSTM-DSR technique.

comparison with the DNN-DSR technique. The advantages of the BiLSTM-DSR technique in wideband conditions have far-reaching significance for applying 5G signals in satellite communication systems. BiLSTM-DSR with memory depth obtains an excellent EVM reduction of over 5% compared to memoryless BiLSTM-DSR. The capability of the proposed technique to capture features in time series data and effectively compensate for the memory effects in broadband PAs is powerfully demonstrated.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3117517, IEEE Access

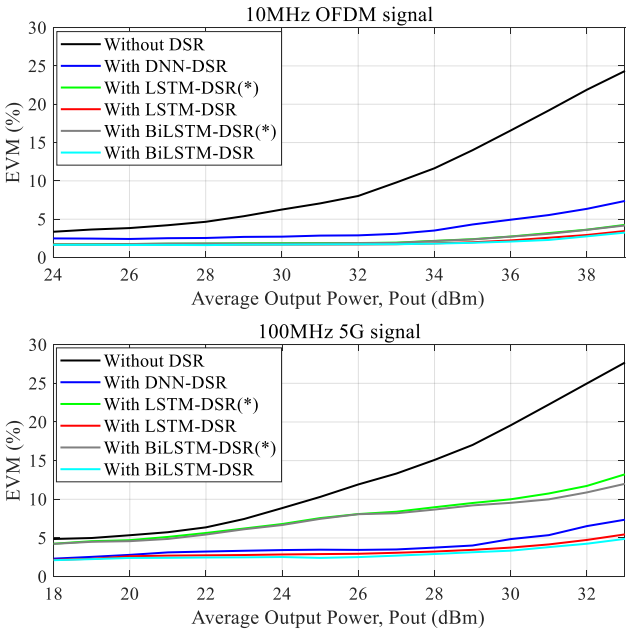


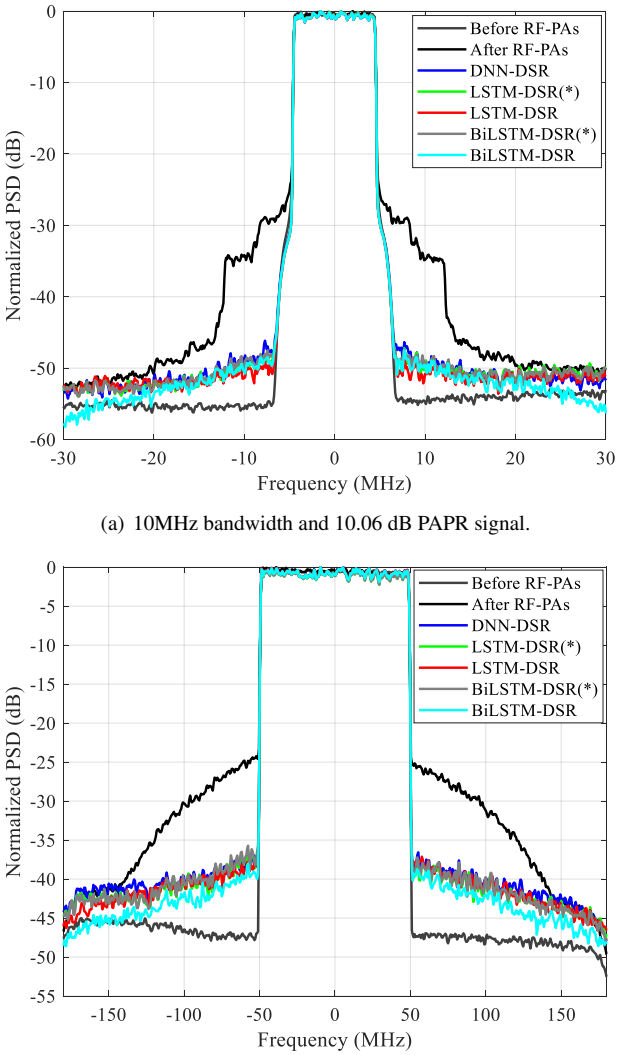
FIGURE 12: Measured EVM vs. Average Output Power curves with different techniques. The symbols of (*) represent memoryless.

C. NOISE ROBUSTNESS ANALYSIS

In online applications, high SNR variation and notable power variation may affect the recovery performance. Therefore, it is essential for the trained models to be robust against these conditions. The impact of power variation was thoroughly discussed in [12], and hence, this work mainly focuses on the impact of high SNR variation.

Fig. 14(a) shows the recovery performance for the narrowband signal with different techniques under the AWGN channel. In the entire test range of the SNR, the ACPR results for the BiLSTM-DSR technique are 1 dB better than those of the other two techniques with the same SNR. If the SNR is lower than 30 dB, the BiLSTM-DSR technique shows excellent performance in terms of EVM. For instance, when the EVM is needed to be 12%, the BiLSTM-DSR technique can improve the SNR performance by approximately 3.5 dB. This means that 3.5 dB of transmission power can thus be saved while maintaining the consistent EVM, with the BiLSTM-DSR technique applied at ground stations. This can greatly increase the satellite power efficiency. Moreover, when the SNR is approximately 20 dB, a more than 3% improvement in the EVM can be acquired by adopting the BiLSTM-DSR technique in comparison with the other two techniques.

The results for the 100 MHz wideband signal are shown in Fig. 14(b). In the entire test range of the SNR, the ACPR results for the BiLSTM-DSR and LSTM-DSR techniques are 2 dB and 1 dB, respectively, better than the DNN-DSR technique with the same SNR. Based on our previous experience [12], we know that the EVM, not the ACPR results, is

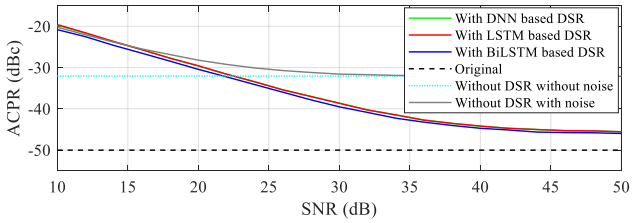


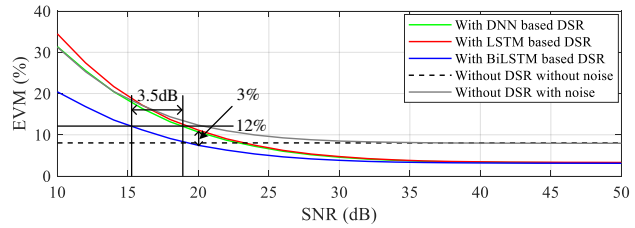
(a) 10MHz bandwidth and 10.06 dB PAPR signal.

(b) 100MHz bandwidth and 11.69 dB PAPR signal.

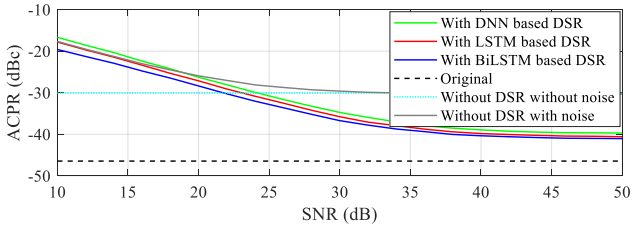
FIGURE 13: PSDs of the output signal with different DSR techniques. The symbols of (*) represent memoryless. (a) 10 MHz narrow band OFDM signal, (b) 100 MHz Wide band 5G signal.

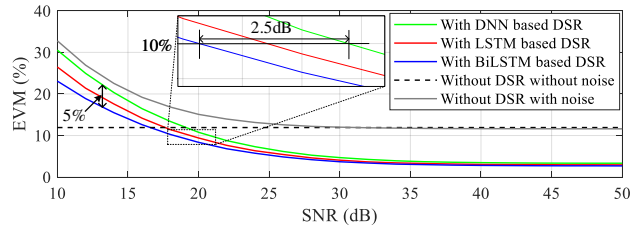
the most important evaluation metric for the recovery performance of DSR techniques, as the EVM is closely correlated to the bit error rate (BER). If the SNR is approximately 13 dB, BiLSTM-DSR and LSTM-DSR show approximately 5% and 3% improvements in EVM in comparison with the DNN-DSR technique, respectively. When the EVM is needed to be 10%, the BiLSTM-DSR technique can equivalently relax the SNR requirement by approximately 2.5 dB (from 20.5 dB to 18 dB). This gain is not as good as that in the test using the narrowband signal. The possible reason for this is that the RF components used in the satellite transmitter do not exhibit sufficiently flat frequency responses within the required bandwidth, which will cause phase and amplitude fluctuations during transmission [32]. Hence, the wideband





(a) 10MHz bandwidth and 10.06 dB PAPR signal.





(b) 100MHz bandwidth and 11.69 dB PAPR signal.

FIGURE 14: ACPR and EVM performance about the noise robustness comparison between different techniques. (a) 10 MHz aarrowband OFDM signal, (b) 100 MHz Wideband 5G signal.

signals show inconsistent nonlinearities in the operation band, which will increase the challenge for signal recovery. However, 2.5 dB is still considered a great improvement, as it is equivalent to an almost halved requirement in the transmitted power for the downlink, meaning a significant improvement in the total power efficiency of the satellites.

While the proposed technique has achieved satisfactory results, it also leaves sufficient room for improvement, such as by taking into account complex channel conditions. In addition, due to the limited capability of the signal generator and signal analyzer, we only validated the proposed technology at a 100 MHz bandwidth in this work. However, we believe the technology has the potential for even broader bandwidths, which could be a topic for further investigation.

V. CONCLUSION

In this paper, we proposed a novel DSR technique for broadband LEO satellite communications. The proposed technique adopts BiLSTM networks to correct the nonlinear distortion caused by spaceborne RF-PAs at ground stations. This allows spaceborne RF-PAs to work in their saturation regions for high power efficiency while maintaining a satisfactory EVM at the ground station. We established a close correlation between our BiLSTM model and the memory effects of RF-PAs and concurrently integrated noise feature filtering and magnitude normalization into a filtering block, which ensures superior signal recovery performance than existing DNN-DSR techniques. The experimental results show that the proposed BiLSTM-DSR technique outperforms the existing DNN-DSR technique in terms of ACPR and EVM by 2.4 dB and 0.9%, respectively, using the same broadband 100 MHz OFDM signal data set. This is equivalent to a 2.5 dB relaxation in the required SNR, from 20.5 dB to 18 dB, to maintain an EVM of the received signal at 10%, which is a big boost of the robustness against low SNR. These results indicate the high potential of the proposed BiLSTM-DSR technique in the development of the emerging LEO satellite communication systems aiming at high throughput and high power efficiency.

In future work, we will extend the BiLSTM-DSR technique to broader bandwidth and higher-order modulation schemes. Moreover, it is challenging to handle the time-varying Doppler effects in real LEO satellite systems, and we will explore more robust networks such as Transformer networks for this issue. The further reduction of SNR requirements is also open for research.

REFERENCES

[1] I. Leyva-Mayorga, B. Soret, M. Röper, D. Wübben, B. Matthiesen, A. Dekorsy, and P. Popovski, "Leo small-satellite constellations for 5g and beyond-5g communications," *IEEE Access*, vol. 8, pp. 184 955–184 964, 2020.

[2] Y. Su, Y. Liu, Y. Zhou, J. Yuan, H. Cao, and J. Shi, "Broadband leo satellite communications: Architectures and key technologies," *IEEE Wireless Communications*, vol. 26, no. 2, pp. 55–61, 2019.

[3] D. He, P. You, and S. Yong, "Comparative handover performance analysis of mipv6 and pmipv6 in leo satellite networks," in *2016 Sixth International Conference on Instrumentation Measurement, Computer, Communication and Control (IMCCC)*, 2016, pp. 93–98.

[4] B. Di, L. Song, Y. Li, and H. V. Poor, "Ultra-dense leo: Integration of satellite access networks into 5g and beyond," *IEEE Wireless Communications*, vol. 26, no. 2, pp. 62–69, 2019.

[5] G. Araniti, I. Bisio, M. De Sanctis, A. Orsino, and J. Cosmas, "Multimedia content delivery for emerging 5g-satellite networks," *IEEE Transactions on Broadcasting*, vol. 62, no. 1, pp. 10–23, 2016.

[6] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Transactions on communications*, vol. 52, no. 1, pp. 159–165, 2004.

[7] C. Tang, B. Lian, and Y. Zhang, "Dual loop feedback pre-distortion in satellite communication," *Journal of Systems Engineering and Electronics*, vol. 24, no. 4, pp. 586–591, 2013.

[8] C. Liu, W. Du, P. Tang, X. Li, and Z. Xiong, "Research and implementation of adaptive digital pre-distortion in high-speed 32-apsk satellite communication system," in *9th International Conference on Communications and Networking in China*, 2014, pp. 321–326.

[9] P. Wertz, B. Hespeler, M. Kiessling, and F.-J. Hagmanns, "Next generation high data rate downlink subsystems based on a flexible apsk modulator

applying sccc encoding," in *2016 International Workshop on Tracking, Telemetry and Command Systems for Space Applications (TTC)*, 2016, pp. 1–7.

[10] M. B. Mabrouk, G. Ferré, E. Grivel, and N. Deltimple, "Interacting multiple model based detector to compensate power amplifier distortions in cognitive radio," *IEEE Transactions on Communications*, vol. 63, no. 5, pp. 1580–1593, 2015.

[11] M. V. Amiri, S. A. Bassam, M. Helaoui, and F. M. Ghannouchi, "Partitioned distortion mitigation in lte radio uplink to enhance transmitter efficiency," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 8, pp. 2661–2671, 2015.

[12] 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 Transactions on Industrial Electronics*, vol. 68, no. 7, pp. 6141–6151, 2021.

[13] A. Almslmany, C. Wang, Q. Cao, and Q. Zeng, "Time varying gain amplifier linearity enhancement for wide dynamic range in radar receiver," in *2015 IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO)*. IEEE, 2015, pp. 1–4.

[14] X. qin Wang, Y. Hei, X. Zhou, and Y. mei Zhou, "Adaptive automatic gain control using hybrid gamma parameters for frame-based ofdm receivers," in *2007 7th International Conference on ASIC*, 2007, pp. 810–813.

[15] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[16] K. Kim, D. Kim, J. Noh, and M. Kim, "Stable forecasting of environmental time series via long short term memory recurrent neural network," *IEEE Access*, vol. 6, pp. 75 216–75 228, 2018.

[17] Y. Chen and K. Wang, "Prediction of satellite time series data based on long short term memory-autoregressive integrated moving average model (lstm-arima)," in *2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP)*, 2019, pp. 308–312.

[18] Y. Tang, F. Yu, W. Pedrycz, X. Yang, J. Wang, and S. Liu, "Building trend fuzzy granulation based lstm recurrent neural network for long-term time series forecasting," *IEEE Transactions on Fuzzy Systems*, pp. 1–1, 2021.

[19] P. Chen, S. Alsahali, A. Alt, J. Lees, and P. J. Tasker, "Behavioral modeling of gan power amplifiers using long short-term memory networks," in *2018 International Workshop on Integrated Nonlinear Microwave and Millimetre-wave Circuits (INMMIC)*, 2018, pp. 1–3.

[20] H. Li, Y. Zhang, G. Li, and F. Liu, "Vector decomposed long short-term memory model for behavioral modeling and digital predistortion for wideband rf power amplifiers," *IEEE Access*, vol. 8, pp. 63 780–63 789, 2020.

[21] J. Sun, W. Shi, Z. Yang, J. Yang, and G. Gui, "Behavioral modeling and linearization of wideband rf power amplifiers using bilstm networks for 5g wireless systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp. 10 348–10 356, 2019.

[22] D. Phartiyal and M. Rawat, "Lstm-deep neural networks based predistortion linearizer for high power amplifiers," in *2019 National Conference on Communications (NCC)*, 2019, pp. 1–5.

[23] S. Cakaj, B. Kamo, A. Lala, and A. Rakipi, "Elevation impact on signal to spectral noise density ratio for low earth orbiting satellite ground station at s-band," in *2014 Science and Information Conference*, 2014, pp. 641–645.

[24] T. Liu, S. Boumaiza, A. B. Sesay, and F. M. Ghannouchi, "Quantitative measurements of memory effects in wideband rf power amplifiers driven by modulated signals," *IEEE Microwave and Wireless Components Letters*, vol. 17, no. 1, pp. 79–81, 2007.

[25] S. S. Islam and A. F. M. Anwar, "Analysis of rf performances of gan mesfets including self-heating and trapping effects," in *IEEE MTT-S International Microwave Symposium Digest, 2003*, vol. 1, 2003, pp. 459–462 vol.1.

[26] L. Ding, G. Zhou, D. Morgan, Z. Ma, J. Kenney, J. Kim, and C. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Transactions on Communications*, vol. 52, no. 1, pp. 159–165, 2004.

[27] D. Morgan, Z. Ma, J. Kim, M. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of rf power amplifiers," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3852–3860, 2006.

[28] M. Schetzen, "The volterra and wiener theories of nonlinear systems," 1980.

[29] Ai Hui Tan, "Wiener-hammerstein modeling of nonlinear effects in bilinear systems," *IEEE Transactions on Automatic Control*, vol. 51, no. 4, pp. 648–652, 2006.

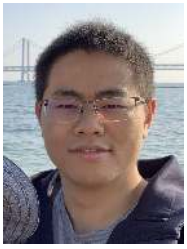
[30] M. A. Chaudhary, Z. A. Memon, J. Lees, J. Benedikt, and P. Tasker, "Investigation of baseband electrical memory effects on the dynamic characteristics of power transistors," in *INMIC*, 2013, pp. 106–109.

[31] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[32] B. Shi, "Digital predistortion linearization of wideband transmitter for high data rate satellite communications," in *2019 IEEE Asia-Pacific Microwave Conference (APMC)*, 2019, pp. 1589–1591.

QINGYUE CHEN was born in Qingdao, China. He received the B.Eng. degree in Shandong University of Science and Technology, Qingdao, China, in 2017. He is currently pursuing the Ph.D. degree in electromagnetic field and microwave technology with the University of Chinese Academy of Sciences (UCAS), Beijing, China. He is also a Guest Ph.D. student with the Department of Electronic Systems, Aalborg University, Aalborg, Denmark.

He worked on the linearization of RF-PAs and channel equalization. His recent research interests include artificial neural networks and its application for communication systems.



YUFENG ZHANG was born in Altay, China. He received the B.Eng. degree in electronic engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2015. He is currently pursuing the Ph.D. degree in electromagnetic field and microwave technology with the University of Chinese Academy of Sciences (UCAS), Beijing, China. He is also a Guest Ph.D. student with the Department of Electronic Systems, Aalborg University, Aalborg, Denmark.

His research interest includes artificial neural networks and its application for communication systems.

FERIDOON JALILI received the M.Sc. degree in electrical engineering from Aalborg University, Aalborg, Denmark, in 1992. Since the graduation, he has been working with RF development in the telecommunication and defense industries in Denmark. Areas of works have been RF front-end design and architecture, level-planning, and performance enhancement techniques and leading as an Expert. He was a Senior Staff Engineer with Intel Corporation, Aalborg, where he contributed to several generations of cellular RF front ends and mobile communications platforms. He has authored or coauthored several patents in this field.

He is currently a Researcher and a Ph.D. Fellow with Aalborg University. His research areas are efficiency improvement of millimeter wave circuits and systems used in the active antenna array for 5G and satellite communications.

ZHUGANG WANG was born in Hangzhou, China. He received the B.Sc. degree in information and electronics from Zhejiang University, Zhejiang, China, in 1995, the M.Sc. degree in engineering from the Department of Computer Science, Tsinghua University, Beijing, China, in 2003, and the Ph.D. degree in satellite communication engineering from the University of Chinese Academy of Sciences (UCAS), Beijing, in 2017.

He is currently a Researcher with the National Space Science Center, Chinese Academy of Sciences, Beijing. Over the past decade, he has hosted or participated in several satellite TT&C communication projects. His current research interests include communications, ranging, and synchronization in satellite constellations.

YONGHUI HUANG was born in Anshan, China. He received the B.Sc. degree in electronics engineering from Tsinghua University, Beijing, China, in 1998, the M.Sc. degree in aero-spacecraft design from the University Chinese Academy of Sciences, Beijing, in 2001, and the Ph.D. degree in wireless communication from Aalborg University, Aalborg, Denmark, in 2008. From 2001 to 2002, he was an Engineer with Datang Mobile, Beijing. From 2002 to 2011, he worked as a Postdoctoral Researcher and a Research Assistant with Aalborg University. He is currently a Professor with the National Space Science Center, Chinese Academy of Science, Beijing.

His current research interests include deep space communication, satellite wireless communication, phased array antenna, and transmitter linearization. He is also a TPC Member of IEEE CCET and IEEE WiSEE.

YUBO WANG was born in Zhengzhou, China. He received the B.Eng. degree in Zhengzhou University (ZZU), Zhengzhou, China, in 2018. He is currently pursuing the master's degree in computer application technology with the University of Chinese Academy of Sciences (UCAS), Beijing, China.

He worked on the satellite communications, heterogeneous parallel computing. His recent research interest includes artificial neural networks and its application for space adaptive scientific exploration and communication.

YING LIU is a professor at the School of Electronic and Information Engineering, Beijing Jiaotong University. She received her M.S. and Ph.D. degrees from Beijing Jiaotong University in 2003 and 2012, respectively.

Her current research interests include network architecture, network security, protocols optimization, wireless communications and cloud computing.

GERT F. PEDERSEN (SM'19) was born in 1965. He received the B.Sc. and E.E. (Hons.) degrees in electrical engineering from the College of Technology in Dublin, Dublin Institute of Technology, Dublin, Ireland, in 1991, and the M.Sc.E.E. and Ph.D. degrees from Aalborg University, Aalborg, Denmark, in 1993 and 2003, respectively. Since 1993, he has been with Aalborg University where he is a Full Professor heading the Antennas, Propagation and Millimeter-wave Systems LAB with 25 researchers. He is also the Head of the Doctoral School on wireless communication with some 40 Ph.D. students enrolled. His research interests include radio communication for mobile terminals especially small antennas, diversity systems, propagation, and biological effects. He has published more than 500 peer reviewed papers, 6 books, 12 book chapters and holds over 50 patents.

He has also worked as a Consultant for developments of more than 100 antennas for mobile terminals including the first internal antenna for mobile phones in 1994 with lowest SAR, first internal triple-band antenna in 1998 with low SAR and high TRP and TIS, and lately various multiantenna systems rated as the most efficient on the market. He has worked most of the time with joint university and industry projects and have received more than 21 M$ in direct research funding. He is currently the Project Leader of the RANGE project with a total budget of over 8 M$ investigating high performance centimetre/millimeter-wave antennas for 5G mobile phones. He has been one of the pioneers in establishing over-the-air measurement systems. The measurement technique is now well established for mobile terminals with single antennas and he was chairing the various COST groups with liaison to 3GPP and CTIA for over-the-air test of MIMO terminals. He is currently involved in MIMO OTA measurements.

MING SHEN (Member, IEEE) was born in Yuxi, China. He received the M.Sc. degree in electrical engineering from the University of Chinese Academy of Sciences (UCAS), Beijing, China, in 2005, and the Ph.D. degree in wireless communications, with the Spar Nord Annual Best Thesis nomination, from Aalborg University, Denmark.

He is currently an Associate Professor in RF and mm-wave circuits and systems with the Department of Electronic Systems, Aalborg University. He has 20 years of experience in RF and millimeter wave circuits and systems, including 12 years of experience in CMOS RF/mixed-signal IC design. He is also the Grant Holder and PI of two Danish national research projects, and the Management Committee Member Substitute from Denmark in the EU COST Action IC1301 with the aim to gather the international efforts and address efficient wireless power transmission technologies. His current research interests include circuits and antennas for 5G and satellite communications, low power CMOS RF and millimeter wave circuits and systems, circuits and systems for biomedical imaging, and artificial intelligence. He is a TPC Member of IEEE NORCAS. He serves as a Reviewer for the IEEE and Kluwer.