IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received February 15, 2022, accepted February 28, 2022, date of publication March 3, 2022, date of current version March 22, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3156605

Deep Learning-Based Joint Optimization of Modulation and Power for Nonlinearity-Constrained System

ZHIYUAN LIU AND MENG MA

School of Electronics, Peking University, Beijing 100871, China

Corresponding author: Meng Ma (mam@pku.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2021YFA1600302, and in part by the Program for Core Technology Tackling Key Problems of Dongguan City under Grant 2019622109010.

ABSTRACT For wireless communication systems with a long distance or severe interference, the insufficient transmit power limits the system performance. In this case, the maximum transmit power depends on the nonlinearity and the saturation region of the power amplifier (PA), which is referred to as a nonlinearity-constrained problem in this paper. To increase the transmit power as high as possible in a nonlinearity-constrained system, this paper proposes an autoencoder-based system to jointly optimize the modulation scheme and transmit power. The optimal solution can achieve a tradeoff between increasing the transmit power and reducing the nonlinear distortion. Meanwhile, the optimized signal constellation and the neural network-based receiver can effectively improve the capacity against nonlinear distortion. The simulation results indicate that the proposed method outperforms conventional methods in terms of symbol error rate (SER) and transmit power, and the SER of the proposed method is close to the SER lower bound of the nonlinear PA.

INDEX TERMS High power, single carrier frequency division multiple access, autoencoder, symbol error rate.

I. INTRODUCTION

In wireless communication systems, the transmit power is usually adjusted to ensure that the received power of each user is no more than the minimum level needed for demodulation to reduce the energy consumption and interference to adjacent users[1], [2]. Such systems can be modeled as power-constrained systems. However, in certain communication systems with long-distance or severe interference, the insufficient receive power is a crucial factor limiting the system performance, and the transmit power should be as high as possible. In this case, the constraint condition can be regarded as the nonlinear and saturation property of the power amplifier (PA), and the system is referred to as a nonlinearity-constrained system in this paper. For example, in satellite communication, due to the long distance to be covered from the on-ground station to the satellite, a strong path-loss of hundreds of dB would be introduced to the satellite communication link. To overcome this problem, satellites are equipped with high power amplifiers (HPA) that may operate close to saturation [3]. Though the use of the maximum power of PA can improve the received signal power, it leads to severe signal distortion and a significant performance loss. Therefore, the optimal strategy for the nonlinearity-constrained system should tradeoff between increasing the signal power and reducing the signal distortion. To improve the performance, two kinds of techniques should be jointly considered: (1) modulation and demodulation techniques to improve the capability against nonlinearity; (2) power control techniques to achieve an optimal transmit power.

So far, the above-mentioned two kinds of techniques have only been studied separately. Over the past decades, a great number of studies investigated the methods to improve the ability to overcome nonlinearity, such as feedforward linearization techniques [4]–[6], feedback linearization/analog predistortion techniques [4], [7]–[9], digital predistortion (DPD) techniques [7], [10]–[14], post-distortion techniques [14], [16], and constellation shaping [17]. In the feedforward linearization technique, the PA input signal is first subtracted from the PA output signal to obtain the distortion introduced by the PA. Then, the distortion is amplified by an auxiliary PA and subtracted from the output signal. To compensate for the distortion, feedback linearization/analog predistortion is implemented by a closed-loop configuration.

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

Cartesian feedback is often used for feedback linearization. In the DPD technique, a digital-domain predistorter is added before the PA to compensate for the nonlinearity of the PA. In [14], a post-distortion algorithm is proposed for orthogonal frequency division multiplexing (OFDM) systems by employing the nonlinear components of the received signal to improve detection performance. Another symbol detection improving method is proposed in [15], in this paper, the author proposes a deep neural network-based (DNN-based) symbol detector for highly dynamic channels (HDCs) and a basis expansion model to reduce the network size dramatically while achieving similar performance. Both of the proposed model have better performance for HDCs compared with the traditional minimum mean square error (MMSE) method. Some other post-distortion algorithms are realized by using a learning method. For example, in [16], the effect of nonlinear distortion on Quadrature Amplitude Modulation (QAM) signals is investigated, and a bit-level demodulator network (BLDnet) using DNN is proposed for nonlinear compensation. BLDnet consists of a fully connected neural network, and it is performed at the receiver to reduce the binary cross-entropy function, thus improving the bit error rate (BER) or symbol error rate (SER) performance. Furthermore, nonlinear interference models such as fiber channel models can be embedded with an autoencoder-based DNN. Through optimizing the mean square error (MSE) between the input and output of the autoencoder, the geometric constellation shapes can be learned under an average power constraint [17]. In [18], the method of overcoming the nonlinearity of PA by optimizing the amplitude and phase of desired beams is proposed, in which the PA nonlinearity is first considered in beamformers. Therefore, the method is useful in redesigning the beamformers for nonlinearity-constrained systems.

Some other methods against nonlinearity focus on reducing the peak-to-average power ratio (PAPR) through coding and modulation schemes. In [19], an autoencoder-based block error rate (BLER) reduction network is proposed for OFDM systems. Through gradient-based training, the network can achieve lower BLER with a coding method. In [20], an autoencoder-based DNN encoder is proposed to reduce PAPR in OFDM systems. However, since the loss function consists of the MSE between the input and output of the autoencoder, and the PAPR of the transmitted signals, this method can achieve lower PAPR and BER.

On the other hand, the transmit power control methods, such as power back off (PBO) enable the PA to work in the linear region. These methods are widely used in narrowband communication systems, such as the Global System for Mobile (GSM) communication system. However, with the increase of PAPR, the power efficiency of the PBO method dramatically decreases, thus leading to a lower power efficiency of the PA.

It should be noted that the above-mentioned methods have their own limitations in the optimization for nonlinearity-constrained system. The PBO method only optimizes the power control without optimizing signal design, while the other methods optimize the signal design without optimizing the power control.

In this paper, the method of optimizing the modulation and power control for the nonlinearity-constrained system simultaneously is proposed. This paper addresses the nonlinearity-constraint system design problem by jointly optimizing the modulator, the transmit power, and the demodulator. Theoretically, signal waveform design can be considered as a vector set design problem, and its complexity increases exponentially as the vector length increases. To facilitate the modulator design, this paper focuses on the constellation design problem in a single carrier frequency division multiple access (SC-FDMA) system with a low PAPR. As shown in Fig. 1, the signal waveform of SC-FDMA can be regarded as an orthogonal time-division multiplexing signal. Therefore, its robustness against nonlinearity is mainly affected by the constellation shaping of the symbols, so the waveform optimization problem can be simplified to a constellation shape optimization problem. To realize joint optimization of constellation shaping, transmit power, and demodulator, an Autoencoder is proposed, which consists of a constellation mapping module, a power control unit, and a neural network-based receiver.

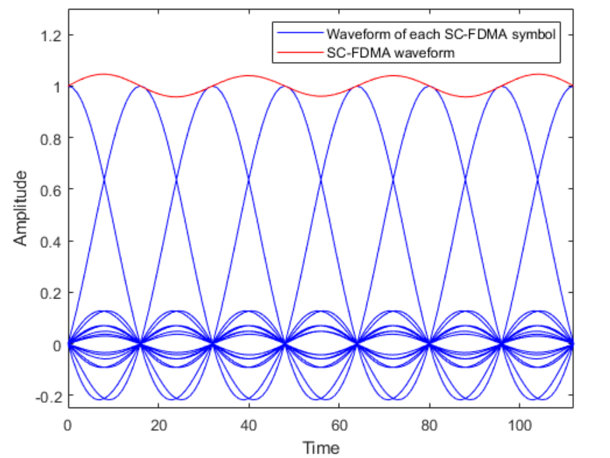


FIGURE 1. The time-domain waveform of SC-FDMA.

The rest of this paper is organized as follows. Section II describes the system model. Section III describes the algorithm of the SC-FDMA Autoencoder. Its performance is evaluated in Section IV, and Section Vconcludes this paper.

II. SYSTEM MODEL

The structure of the proposed nonlinear-constrained communication system based on the SC-FDMA Autoencoder is shown in Fig. 2.

First, the information source is transformed into complex modulated symbols through the designed signal constellation. Then, using a serial to parallel converter, the modulated symbols are grouped into blocks with a length of symbols, i.e.,

$$\mathbf{s}_1 = [s_1(0), s_1(1), \ldots, s_1(N_0-1)]^T . \tag{1}$$

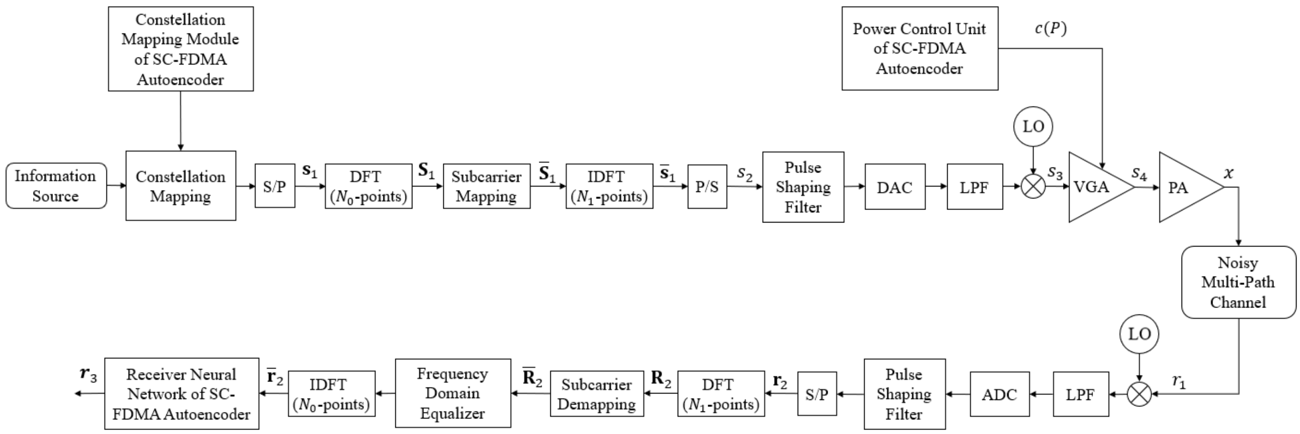


FIGURE 2. The structure of the nonlinear communication system based on the SC-FDMA autoencoder.

Afterward, each block is passed through an N_0-point discrete Fourier transform (DFT) to produce the frequency-domain signal vector

$$\mathbf{S}_1 = \left[S_1(0), S_1(1), \ldots, S_1\left(N_0 - 1\right)\right]^T, \tag{2}$$

and its elements can be expressed as

$$S_1(k) = \frac{1}{\sqrt{N_0}} \sum_{n'=0}^{N_0-1} s_1\left(n'\right) e^{-j\frac{2\pi k n'}{N_0}}, \quad 0 \le k \le N_0 - 1. \tag{3}$$

In localized frequency domain multiple access (LFDMA), $\mathbf{S}_1$ is fed into a set of consecutive subcarriers from N_0 subcarriers ($N_1 > N_0$) to obtain

$$\bar{\mathbf{S}}_1 = \left[\bar{S}_1(0), \bar{S}_1(1), \ldots, \bar{S}_1\left(N_1 - 1\right)\right]^T, \tag{4}$$

and its elements can be expressed as

$$\bar{\mathrm{S}}_1(n) = \begin{cases} \mathrm{S}_1(n), & f_0 \le n \le f_0 + N_0 - 1, \\ 0, & \text{others}, \end{cases} \tag{5}$$

where f_0 is the starting index.

The frequency-domain signal $\bar{\mathbf{S}}_\mathbf{1}$ are then passed through an N_1-point inverse discrete Fourier transform (IDFT) to produce $\bar{\mathbf{s}}_\mathbf{1}$ before parallel to serial conversion, which can be expressed as

$$\bar{\mathbf{s}}_\mathbf{1} = \left[\bar{s}_1(0), \bar{s}_1(1), \ldots, \bar{s}_1\left(N_1 - 1\right)\right]^T. \tag{6}$$

The elements of $\bar{\mathbf{s}}_\mathbf{1}$ can be expressed as

$$\bar{s}_1(n) = \frac{1}{\sqrt{N_1}} \sum_{k=0}^{N_0-1} \bar{\mathrm{S}}_1(k) e^{j\frac{2\pi f_k n}{N_1}}, \quad 0 \le n \le N_1 - 1. \tag{7}$$

The discrete symbols are then processed by a shaping filter and converted into the analog domain for radio frequency (RF) up-conversion. The up-converted signal $s_3(t)$ is then amplified by variable gain amplifier (VGA) to obtain $s_4(t) = \sqrt{P} \cdot s_3(t)$, where P is the power of the digital-domain signal; $c(P)$ is the analog or digital control signal, which is adjusted by the output of the power control unit of SC-FDMA Autoencoder. s_4 is then fed into the PA to obtain the transmit signal $x(t)$.

The channel is modeled as a multi-path fading channel. The received signal $r_1(t)$ is filtered and converted into a digital domain after down-conversion. After the DFT demodulation and subcarrier demapping, the signal $\overline{\mathbf{R}}_2$ is equalized by a zero-forcing (ZF) or MMSE frequency-domain equalizer (FDE). After passing through the IDFT demodulator, $\mathbf{r}_2$ is fed into a neural network-based demapping module, and $\mathbf{r}_3$ is finally obtained in the form of a M-dimensional vector, where M is the modulation order.

III. THE ALGORITHM OF SC-FDMA AUTOENCODER

The proposed SC-FDMA Autoencoder consists of three modules: a constellation mapping module, a power control unit, and a neural network-based demapping module. In the training process of the SC-FDMA Autoencoder, the training data is encoded as a one-hot vector with M elements. Then, the one-hot vector is fed into the power control unit and the constellation mapping module of the SC-FDMA Autoencoder.

A. THE OPTIMIZATION OF POWER CONTROL

The power control unit of SC-FDMA autoencoder is shown as Fig.3. The goal of the power control unit is to find the optimal transmit power for each symbol. After an M-dimensional one-hot vector $\mathbf{s}_{\text{in}}$ is fed into the power control unit, the transformation $P = \mathrm{p}\left(\mathbf{s}_{\text{in}}\right)$ is applied to generate the output P of the power control unit. The function of the power control unit can be expressed as

$$\mathrm{p}\left(\mathbf{s}_{\text{in}}\right) = \sigma\left[W_2\left[\sigma\left(W_1 \mathbf{s}_{\text{in}} + b_1\right)\right] + b_2\right], \tag{8}$$

where W_q and b_q are the parameters of the q-th linear fully connected layer, and $\sigma(u) = \max(0, u)$ is the function of the ReLU activation layer [22].

In the backpropagation process for optimizing the trainable power control unit, assuming the loss of the SC-FDMA Autoencoder is l, then the gradient of the loss to the power P can be expressed as

$$\frac{\partial l}{\partial P} = \sum_{i=1}^{N_{\text{batch}}} \frac{\partial l}{\partial s_4(i)} \left[\frac{\partial s_4(i)}{\partial P}\right]$$

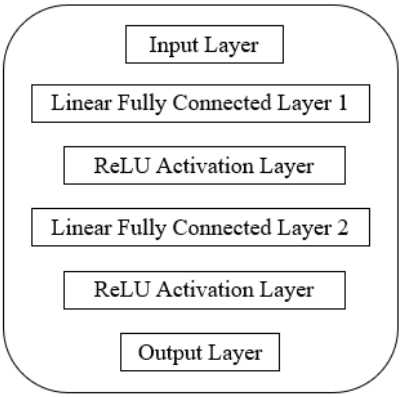


FIGURE 3. The structure of the power control unit of SC-FDMA autoencoder.

$$= \sum_{i=1}^{N_{\text{batch}}} \frac{\partial l}{\partial s_4(i)} \cdot s_3(i) \cdot \frac{1}{2 \cdot \sqrt{P}}, \tag{9}$$

where N_{batch} is the size of each training batch; $s_3(i)$ and $s_4(i)$ respectively denote the value of $s_3(t)$ and $s_4(t)$ of the i-th sample in the training batch; $\frac{\partial l}{\partial s_4(i)}$ is the gradient of the loss to $s_4(i)$, and it is determined by the operation function of the following modules and the loss function of the SC-FDMA Autoencoder. After the gradient of loss to P is calculated, the transmit power P can be optimized by using the adaptive moment estimation (Adam) optimization method [21]. The optimization steps of the transmit power P can be expressed as (10)–(14), shown at the bottom of the page, where χ_τ and v_τ are the first-order momentum term and the second-order momentum term at iteration τ, respectively; the hyperparameters β_1 and β_2 are the dynamic values of the first-order momentum term and the second-order momentum term, respectively; $\hat{\chi}_\tau$ and $\hat{v}_\tau$ are the estimated value of first-order momentum term and the second-order momentum term at iteration τ, respectively; ϵ_0 is a very small number added to prevent the denominator from being 0; lr is the learning rate of the optimizer, and P_τ is the trainable transmit power at iteration τ.

B. THE OPTIMIZATION OF MODULATION

The constellation mapping module and the receiver neural network of the SC-FDMA Autoencoder is shown in Fig. 4.

$$\begin{aligned} \chi_\tau &= \beta_1 \chi_{\tau-1} + (1-\beta_1) \frac{\partial l}{\partial P} \\ &= \beta_1 \chi_{\tau-1} + (1-\beta_1) \sum_{i=1}^{N_{\text{batch}}} \frac{\partial l}{\partial s_4(i)} \cdot s_3(i) \cdot \frac{1}{2 \cdot \sqrt{P}}, \end{aligned} \tag{10}$$

$$\begin{aligned} v_\tau &= \beta_2 v_{\tau-1} + (1-\beta_2) \left(\frac{\partial l}{\partial P}\right)^2 \\ &= \beta_2 v_{\tau-1} + (1-\beta_2) \left[\sum_{i=1}^{N_{\text{batch}}} \frac{\partial l}{\partial s_4(i)} \cdot s_3(i) \cdot \frac{1}{2 \cdot \sqrt{P}}\right]^2, \end{aligned} \tag{11}$$

$$\begin{aligned} \hat{\chi}_\tau &= \frac{\chi_\tau}{1-{\beta_1}^\tau} \\ &= \frac{\beta_1 \chi_{\tau-1} + (1-\beta_1) \sum_{i=1}^{N_{\text{batch}}} \frac{\partial l}{\partial s_4(i)} \cdot s_3(i) \cdot \frac{1}{2 \cdot \sqrt{P}}}{1-{\beta_1}^\tau}, \end{aligned} \tag{12}$$

$$\begin{aligned} \hat{v}_\tau &= \frac{v_\tau}{1-{\beta_2}^\tau} \\ &= \frac{\beta_2 v_{\tau-1} + (1-\beta_2) \left[\sum_{i=1}^{N_{\text{batch}}} \frac{\partial l}{\partial s_4(i)} \cdot s_3(i) \cdot \frac{1}{2 \cdot \sqrt{P}}\right]^2}{1-{\beta_2}^\tau}, \end{aligned} \tag{13}$$

$$\begin{aligned} P_\tau &= P_{\tau-1} - \frac{\widehat{\chi_\tau}}{\sqrt{\hat{v}_\tau} + \epsilon_0} \cdot lr \\ &= P_{\tau-1} \\ &\quad - \frac{\dfrac{\beta_1 \chi_{\tau-1} + (1-\beta_1) \sum_{i=1}^{N_{\text{batch}}} \frac{\partial l}{\partial s_4(i)} \cdot s_3(i) \cdot \frac{1}{2 \cdot \sqrt{P_{\tau-1}}}}{1-\beta_1^\tau}}{\sqrt{\dfrac{\beta_2 v_{\tau-1} + (1-\beta_2) \left[\sum_{i=1}^{N_{\text{batch}}} \frac{\partial l}{\partial s_4(i)} \cdot s_3(i) \cdot \frac{1}{2 \cdot \sqrt{P_{\tau-1}}}\right]^2}{1-{\beta_2}^\tau}} + \epsilon_0} \\ &\quad \cdot lr, \end{aligned} \tag{14}$$

The goal of the constellation mapping module is to find a two-dimensional representation of the input M-dimensional one-hot vector. The value of each dimension of the two-dimensional output refers to the in-phase and quadrature component respectively. After an M-dimensional one-hot vector $\mathbf{s}_{\text{in}}$ is fed into the constellation mapping module, the transformation $\mathbf{s}_{\text{out}} = \mathrm{f}(\mathbf{s}_{\text{in}})$ is applied to generate the output signal $\mathbf{s}_{\text{out}}$. It should be noted that the power of $\mathbf{s}_{\text{out}}$ is normalized to 1. The function of the constellation mapping module can be expressed as

$$\mathrm{f}(\mathbf{s}_{\text{in}}) = \mathrm{Norm}\left[\sigma\left(W_4\left[\mathrm{Norm}\left[\sigma\left(W_3\mathbf{s}_{\text{in}} + b_3\right)\right]\right] + b_4\right)\right], \tag{15}$$

where $\mathrm{Norm}(u_i)$ is the function of the batch normalization layer [23], and it can be expressed as

$$\mathrm{Norm}(u_i) = \gamma \frac{u_i - \frac{1}{N_{\text{batch}}}\sum_{i=1}^{N_{\text{batch}}} u_i}{\sqrt{\frac{1}{N_{\text{batch}}}\sum_{i=1}^{N_{\text{batch}}}\left(u_i - \sum_{i=1}^{N_{\text{batch}}} u_i\right)^2 + \epsilon_0}} + \delta, \tag{16}$$

where γ and δ are parameters to be learned, u_i is the i-th output of the ReLU layer. The receiver neural network operates similarly to the constellation mapping module. It applies the transformation $\mathbf{r}_3 = \mathrm{g}(\mathbf{r}_2)$ to generate the output signal $\mathbf{r}_3$ of the receiver neural network. The function of the receiver neural network can be expressed as

$$\mathrm{g}(\mathbf{r}_2) = \mathrm{Norm}\left[\sigma\left(W_6\left[\mathrm{Norm}\left[\sigma\left(W_5\mathbf{r}_2 + b_5\right)\right]\right] + b_6\right)\right]. \tag{17}$$

The cross-entropy loss is used in the training of the SC-FDMA Autoencoder. First, the Softmax function is operated on the output of autoencoder $\mathbf{r}_3$ to generate the probability distribution of the output belonging to each input symbol:

$$p_{i,j} = \frac{\exp\left[R_i(j)\right]}{\sum_{j=1}^{M}\exp\left[R_i(j)\right]}, \tag{18}$$

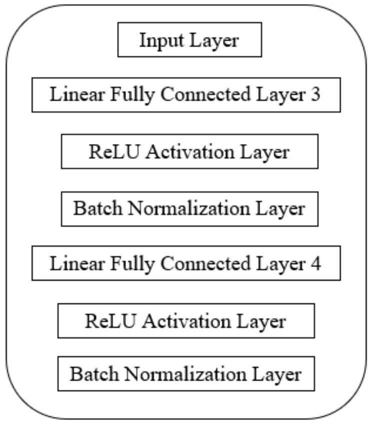
where $p_{i,j}$ is the prediction probability that the real category of the i-th transmitted signal is equal to j. Then, the loss and the parameters of the autoencoder are calculated. The loss function can be expressed as

$$l = -\frac{1}{N_{\text{batch}}}\sum_{i=1}^{N_{\text{batch}}}\sum_{j=1}^{M} y_{i,j}\log\left(p_{i,j}\right), \tag{19}$$

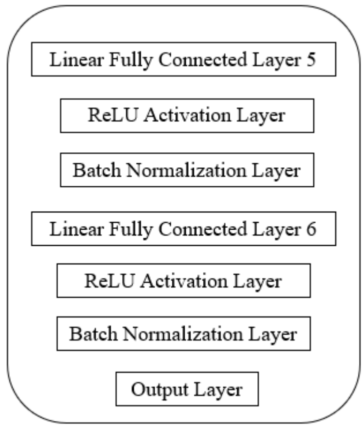
where R_i is the i-th row of $\mathbf{r}_3$; $R_i(j)$ is the value of the j-th column of R_i; $y_{i,j}$ is a symbol-function. If the real category of the i-th transmitted signal is equal to j, $y_{i,j}$ takes 1; otherwise, it takes 0.

IV. SIMULATIONS AND DISCUSSIONS

In this section, the simulation results of the proposed method are provided. In the simulation, the mapping pattern of SC-FDMA is localized mapping; the size of frequency-domain data is 32, and the number of subcarriers is 256.



(a) Constellation Mapping Module



(b) Receiver Neural Network

FIGURE 4. **The structure of constellation mapping module and receiver neural network of SC-FDMA autoencoder.**

Besides, the size of the training set is 51200; the batch size is 256; the initial learning rate is 0.001, and it decreases by 20% every 5 epochs. The training is stopped when the loss does not drop for 10 epochs.

In the simulations, for convenience, the output signal of PA is calculated by the parameters and the behavioral model of PA. The behavioral model of the PA performs odd-order Taylor series expansion of the Saleh model, and it has the form [24]:

$$x_{\text{out}} = \alpha_1 x_{\text{in}} + \alpha_3 x_{\text{in}}{}^3 + \cdots + a_n x_{\text{in}}{}^n, \tag{20}$$

where x_{in} and x_{out} are the PA input and output magnitude, respectively; α_n is the n-th coefficient of the odd-order Taylor series, and it can be expressed as [25]:

$$a_1 = 10^{\frac{G}{20}}, \tag{21}$$

$$a_n = -\frac{2^{\frac{n-1}{2}}}{\begin{pmatrix} n \\ \frac{n+1}{2} \end{pmatrix}} 10^{\frac{-(n-1)IP_n + nG}{20}}, \tag{22}$$

where G is the power gain in dB, and IP_n is the output n-order intercept point in dBw. For a practical PA module ZRL-2400LN, G is 28 dB, and IP_3 is 15 dBw. Substituting these values into (20) and (21), we have $\alpha_1 = 17.8$ and $\alpha_3 = -118.6$. For ZRL-2400LN+, another PA module with more severe nonlinearity, G is 28 dB, and IP_3 is 12 dBw. Based on this mode, we have $\alpha_1 = 17.8$, $\alpha_3 = -666.7$.

First, the SER and the transmit power of the proposed method are compared with those of the PBO method with 64-square QAM (64-SQAM), which is a square shaped constellation. In Fig. 5, the peak transmit power is decreased to the compression point of 1 dB to make the PA operate at a linear region. When the four curves have an identical SER of 10^{-2}, the proposed method achieved a gain of 3.9 dB in the noise power compared with PBO for ZRL-2400LN, and the gain increases to 4.6 dB for ZRL-2400LN+. Therefore, it can be seen that the proposed method can achieve more performance improvement for the PA under more serious nonlinearity. The ratio of the average power of PA to the saturated power of PA is shown in Fig. 6, it can be found that the proposed method enables the PA to operate at a higher power than the PBO method, and the power improvement is greater for ZRL-2400LN+, which has a more serious nonlinearity. This is because the PBO method reduces more transmit power to make a more nonlinear PA operate at a linear region. Therefore, the gain of the proposed method should be greater when the nonlinearity of PA is more serious. Besides, Fig. 5 and Fig. 6 also demonstrate that the reduction of SER by the proposed method is attributed to the optimization of transmit power.

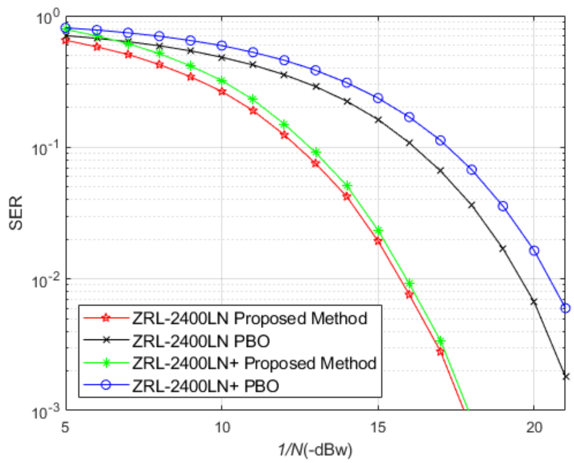


FIGURE 5. The SER of the proposed method and the PBO method for different PA.

Next, the influence of the modulation order M on the performance is analyzed for ZRL-2400LN. The designed constellation shaping under $M = 4, 16, 64$ is shown in Fig. 7, and the SER of the proposed method and the PBO method under $M = 4, 16, 64$ is shown in Fig. 8. When the different curves have an identical SER of 10^{-2}, the proposed method has a gain of 2 dB in noise power when $M = 4$, but the gain becomes 4.7 dB and 4.3 dB when $M = 16$ and $M = 64$. Compared with the PBO method, the SER improvement of the proposed method when $M = 4$ is lower than that when $M = 16, 64$. This is because the PAPR of the proposed constellation and the quadrature phase shift keying (QPSK) is equal to 1 for $M = 4$. As shown in Table. 1, for $M = 16, 64$, the PAPR of the proposed constellation is lower than 16-SQAM and 64-SQAM, which makes the PBO for 16-SQAM and 64-SQAM reduce more power than PBO for QPSK.

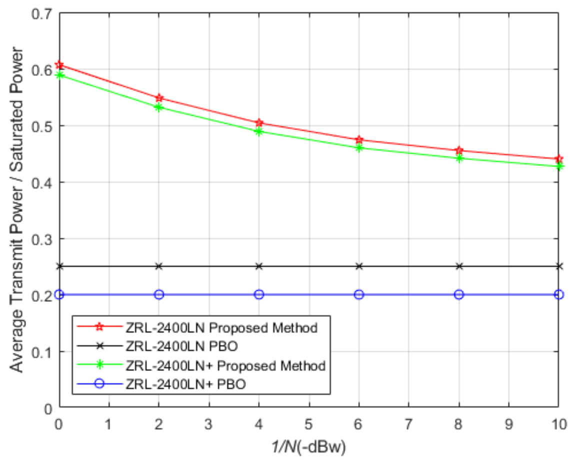


FIGURE 6. The average transmit power/saturated power of the proposed method and the PBO method.

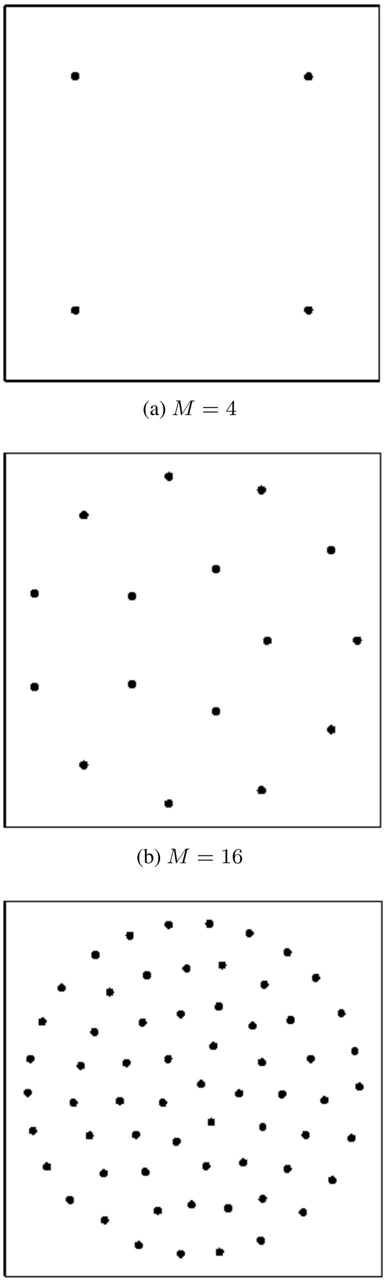
TABLE 1. The PAPR of the proposed method and SQAM under $M = 4, 16, 64$.

Constellation, M	PAPR
Proposed Method, $M = 4$	1
SQAM ,$M = 4$(QPSK), $M = 4$	1
Proposed Method, $M = 16$	1.33
SQAM, $M = 16$	1.80
Proposed Method, $M = 64$	1.65
SQAM, $M = 64$	2.33

Furthermore, the proposed method is compared with the DPD method of M-SQAM and polygon constellation [26] with traversed power. In the comparison, the transmit power P is traversed from 0 to the saturated power P_{sat} at an interval of 10^{-3} to find the transmit power that leads to the lowest SER. The optimized power of the comparison method can be expressed as

$$P_{\text{opt}} = \underset{0<P\leq P_{\text{sat}}}{\operatorname{argmin}} \frac{1}{M} \sum_{m=1}^{M} \Pr\left(\mathbf{f}_m \notin \mathbf{Z}_m \mid \sqrt{P}\cdot \mathbf{d}_m\right), \tag{23}$$

where $\mathbf{d}_m$ is the m-th signal of the M possible transmitted signals with normalized power; $\mathbf{f}_m$ is the received signal when $\mathbf{d}_m$ is transmitted; $\mathbf{Z}_m$ is the decision region of the transmitted signal $\mathbf{d}_m$. The SER of the proposed method and the DPD method of 64-SQAM and 64-order polygon constellation with traversed power is shown in Fig. 9.



(a) $M = 4$

(b) $M = 16$

(c) $M = 64$

FIGURE 7. **Constellations of proposed method under $M = 4, 16, 64$.**

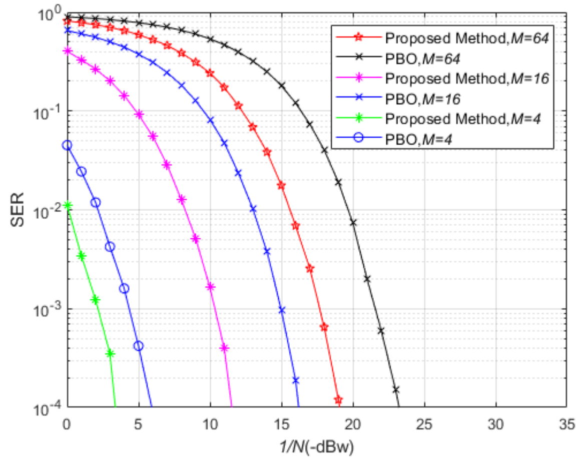


FIGURE 8. **SER of proposed method and PBO under $M = 6, 16, 64$.**

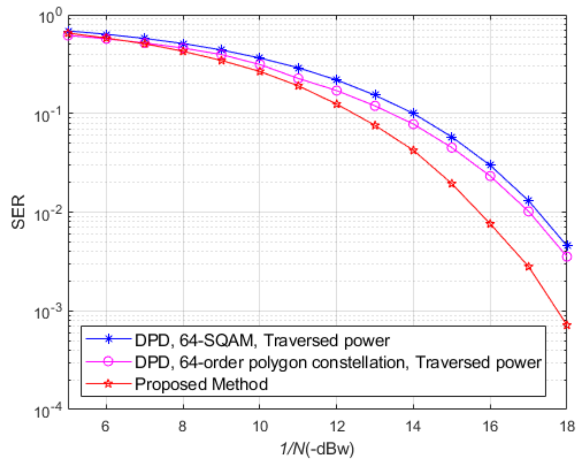


FIGURE 9. **SER of the proposed method and the DPD method with traversed power.**

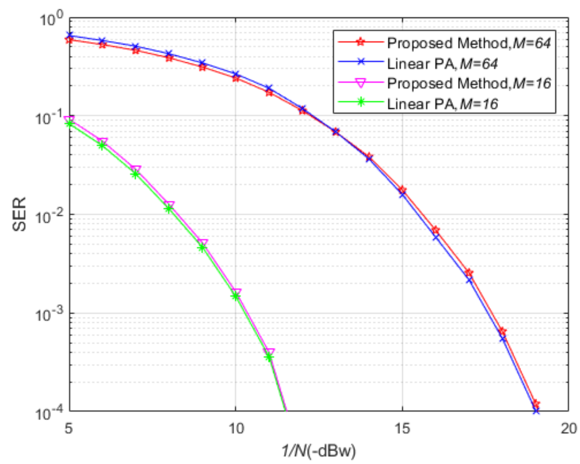


FIGURE 10. **The SER of the linear PA and the proposed method.**

It can be found that although the proposed method and the DPD method of 64-SQAM and 64-order polygon constellation have an identical SER of 10^{-2}, the DPD method operates at the traversed power that minimizes the SER. Compared with the DPD method of 64-order polygon constellation, the proposed method achieves a gain of about 1.5 dB in noise power, and the DPD method of 64-SQAM achieves a gain of about 1.3 dB. For these given constellations, the DPD method obtains a lower SER performance than the proposed method at any transmit power, which demonstrates the benefits of the joint optimization of constellation shape and transmit power.

In Fig. 10, the SER of the proposed method is compared with that of the linear PA with the same modulation and PA output power under $M = 16, 64$. It can be seen from

Fig. 10 that while the proposed method and the linear PA have an identical SER of 10^{-2}, the proposed method is 0.1 dB worse than the linear PA in noise power when $M = 16$, and 0.15 dB worse than linear PA in noise power when $M = 64$. The SER lower bound of nonlinear PA cannot be even lower than that of linear PA, and the SER of nonlinear PA with the proposed method is close to the SER of linear PA with the same modulation. Thus, the SER of the proposed method is close to the SER lower bound of nonlinear PA.

V. CONCLUSION

In this paper, a system design for a nonlinearity-constrained system is proposed. By training the constellation mapping module, the power control unit, and the receiver neural network of the SC-FDMA Autoencoder, the constellation shape, transmit power, and the demodulation scheme can be jointly optimized to minimize the SER for the given channel and the PA in the SC-FDMA system. This can increase the transmit power and make the PA operate in a region of severe nonlinearity. Based on this, the high-power transmission performance can be improved for the scenarios such as satellite communication and communication systems with severe interference. The simulation results indicate that the proposed method outperforms conventional methods in terms of SER and transmit power, and the SER of the proposed method is close to the SER lower bound of nonlinear PA.

REFERENCES

[1] C.-P. Liang, J.-H. Jong, W. E. Stark, and J. R. East, ''Nonlinear amplifier effects in communications systems,'' *IEEE Trans. Microw. Theory Techn.*, vol. 47, no. 8, pp. 1461–1466, Aug. 1999.

[2] L. E. Larson, ''Radio frequency integrated circuit technology for low-power wireless communications,'' *IEEE Pers. Commun.*, vol. 5, no. 3, pp. 11–19, Jun. 1998.

[3] A. I. Perez-Neira, M. A. Vazquez, M. R. B. Shankar, S. Maleki, and S. Chatzinotas, ''Signal processing for high-throughput satellites: Challenges in new interference-limited scenarios,'' *IEEE Signal Process. Mag.*, vol. 36, no. 4, pp. 112–131, Jul. 2019.

[4] M. Faulkner, ''Amplifier linearization using RF feedback and feedforward techniques,'' *IEEE Trans. Veh. Technol.*, vol. 47, no. 1, pp. 209–215, Feb. 1998.

[5] J. K. Cavers, ''Adaptation behavior of a feedforward amplifier linearizer,'' *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 31–40, Feb. 1995.

[6] E. E. Eid, F. M. Ghannouchi, and F. Beauregard, ''Optimal feedforward linearization system design,'' *Microw. J.*, vol. 38, no. 11, pp. 78–83, Nov. 1995.

[7] A. Katz, J. Wood, and D. Chokola, ''The evolution of PA linearization: From classic feedforward and feedback through analog and digital predistortion,'' *IEEE Microw. Mag.*, vol. 17, no. 2, pp. 32–40, Feb. 2016.

[8] J. L. Dawson, *Feedback Linearization RF Power Modeling*. Stanford, CA, USA: Stanford Univ., 2003.

[9] J. Kim, C. Park, J. Moon, and B. Kim, ''Analysis of adaptive digital feedback linearization techniques,'' *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, no. 2, pp. 345–354, Feb. 2010.

[10] J. Kim and K. Konstantinou, ''Digital predistortion of wideband signals based on power amplifier model with memory,'' *Electron. Lett.*, vol. 37, no. 23, pp. 1417–1418, Nov. 2001.

[11] 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.

[12] G. Karam and H. Sari, ''A data predistortion technique with memory for QAM radio systems,'' *IEEE Trans. Commun.*, vol. 39, no. 2, pp. 336–344, Feb. 1991.

[13] S. Chen, ''An efficient predistorter design for compensating nonlinear memory high power amplifiers,'' *IEEE Trans. Broadcast.*, vol. 57, no. 4, pp. 856–865, Dec. 2011.

[14] Z. Alina and O. Amrani, ''On digital post-distortion techniques,'' *IEEE Trans. Signal Process.*, vol. 64, no. 3, pp. 603–614, Feb. 2016.

[15] X. Lyu, W. Feng, and N. Ge, ''Deep neural network-based symbol detection for highly dynamic channels,'' in *Proc. IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–6.

[16] L. Zou, M. Jiang, C. Zhao, Y. He, D. Zhu, and Q. Huang, ''BLDNet: Robust learning-based detection for high-order QAM with nonlinear distortion,'' in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Aug. 2020, pp. 262–266.

[17] R. T. Jones, T. A. Eriksson, M. P. Yankov, and D. Zibar, ''Deep learning of geometric constellation shaping including fiber nonlinearities,'' in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2018, pp. 1–3.

[18] C. Liu, W. Feng, Y. Chen, C.-X. Wang, and N. Ge, ''Optimal beamforming for hybrid satellite terrestrial networks with nonlinear PA and imperfect CSIT,'' *IEEE Wireless Commun. Lett.*, vol. 9, no. 3, pp. 276–280, Mar. 2020.

[19] 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.

[20] 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.

[21] D. P. Kingma and J. Ba, ''Adam: A method for stochastic optimization,'' 2014, *arXiv:1412.6980*.

[22] Y. LeCun, Y. Bengio, and G. Hinton, ''Deep learning,'' *Nature*, vol. 521, pp. 436–444, Feb. 2015.

[23] S. Ioffe and C. Szegedy, ''Batch normalization: Accelerating deep network training by reducing internal covariate shift,'' in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.

[24] H. Lai and Y. Bar-Ness, ''A new predistorter design for nonlinear power amplifiers using the minimum distortion power polynomial model (MDP-PM),'' in *Proc. IEEE 54th Veh. Technol. Conf. VTC Fall*, Oct. 2001, pp. 2216–2220.

[25] X. Li, C. M. Liu, Y. Xu, and F. Li, ''Obtaining polynomial coefficients from intercept points of RF power amplifiers,'' *Electron. Lett.*, vol. 48, no. 19, pp. 1238–1240, Sep. 2012.

[26] L. Rugini, ''Symbol error probability of hexagonal QAM,'' *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1523–1526, Aug. 2016.

ZHIYUAN LIU was born in Hefei, China. He received the B.S. degree in electronic and information engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2020. He is currently pursuing the M.S. degree in signal and information processing with the Department of Electronics, Peking University, Beijing. His research interests include wireless communication systems and signal processing techniques.

MENG MA received the B.S. and Ph.D. degrees in electrical engineering from Peking University, Beijing, China, in 2001 and 2007, respectively. From 2007 to 2011, he was an Assistant Professor with the School of Electronics Engineering and Computer Science, Peking University, where he has been an Associate Professor, since 2011. From 2009 to 2010, he was a Visiting Scholar with the Commonwealth Scientific and Industrial Research Organization, Sydney, Australia. His research interests include signal processing in wireless communication systems, interference cancellation technique, and information theory.