End-to-End MIMO Systems With Conditional Generative Adversarial Networks

Yifan Yao
Division of Electronic Engineering
Jeonbuk National University
Jeonju, Korea
yaoyifan1995@gmail.com

Juin Shin
Division of Electronic Engineering
Jeonbuk National University
Jeonju, Korea
juin0128@gmail.com

Xianglan Jin
Division of Electronic Engineering
Jeonbuk National University
Jeonju, Korea
jinxl77@jbnu.ac.kr

***Abstract*—Conventional autoencoder-based systems, comprising a neural-network encoder at the transmitter and a neural-network decoder at the receiver, often face limitations in realistic signal transmission due to its reliance on differentiable channel models. In response to this limitation, an end-to-end framework has emerged, employing a conditional generative adversarial network (CGAN) for channel learning. The CGAN not only learns to represent channel effects through feature extraction but also facilitates gradient back-propagation between the receiver and the transmitter, thereby enhancing the system's adaptability. This paper extends the CGAN-based end-to-end system from single input single output channels to multiple input multiple output (MIMO) scenarios. This CGAN-based MIMO system demonstrates promising performance comparable to the autoencoder communication systems, highlighting the potential of CGANs as effective alternatives for original channels.**

***Index Terms*—Autoencoder, deep learning, generative adversarial network (GAN), multiple input multiple output (MIMO)**

I. Introduction

Traditional communication systems generally involve multiple signal processing blocks in both the transmitter and receiver as illustrated in Fig. 1 (a). Although the technologies within this system are well-established, individual modules are designed and optimized with diverse assumptions and goals, challenging the determination of global optimal for the entire system.

In recent years, deep learning has significantly boosted conventional communication systems. Various deep neural networks (DNNs) have shown notable success in different tasks, such as channel estimation [1], channel decoding [2], [3] and multiple input multiple output (MIMO) detection [4]. Although these existing methods have brought flexible and efficient modules, it is not yet known whether independently optimized processing modules can lead to optimal end-to-end performance. Autoencoders were introduced into communications [5], thereby pioneering an end-to-end communication system learning approach. In the autoencoder-based communication system, the encoder and decoder networks serve as the transmitter and receiver, offering a unique method for joint optimization of modules as depicted in Fig. 1 (b). Remarkably,

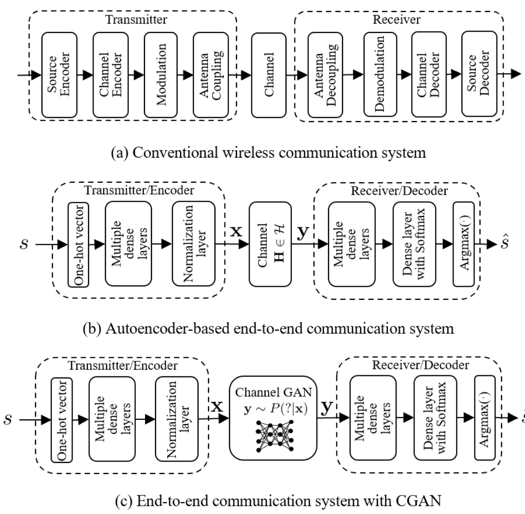

Fig. 1. Structure of MIMO communication systems.

its performance is comparable to traditional communications in both single input single output (SISO) [5] and MIMO scenarios [6], contributing to the widespread use of end-to-end communication systems.

However, the major practical limitation of this idea is its reliance on a differentiable channel model function, which is crucial for back-propagation in the training process [7]. A simplest instance of such a differentiable channel model is the additive white Gaussian noise channel. In practical applications, performance may need to be optimized for specific over-the-air channel effects or combinations of effects such as device responses, interferences, distortions, and other noises. While simplified analytical models can be employed in certain scenarios to describe or simulate specific phenomena, they often perform inadequately in capturing nonlinear effects in the real world, particularly when dealing with combinations of such effects [7]. This limitation arises due to the inherent complexity and degrees of freedom associated with these phenomena.

As indicated in Fig. 1 (c), conditional generative adversarial networks (CGANs) [8] are considered as an appealing solution capable of representing channel effects and acting as a bridge

This research has been supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT under Grant 2022R1F1A1063048.

to back-propagate gradients from the receiver neural network to the transmitter neural network [9], [10]. This work extends the application of CGANs from SISO communication systems [10] to MIMO scenarios, acknowledging the important role of MIMO systems in future communications. To the best of our knowledge, it is the first attempt to construct an end-to-end MIMO communication system by leveraging CGANs. Simulation results demonstrate that the CGAN is a promising channel-agnostic learning algorithm, exhibiting performance comparable to that of the autoencoder in MIMO communication systems [6].

II. Modeling MIMO Channel With CGAN

A traditional communication system consists of a transmitter and a receiver. The end-to-end communication system, utilizing autoencoders, incorporates the transmitter and receiver as distinct neural networks. However, the existence of an unknown channel presents a challenge to the backpropagation algorithm, impeding end-to-end learning. To tackle this issue, the CGAN is employed. The CGAN proves effectiveness in learning channel effects and facilitating the propagation of gradients between the transmitter and receiver. This section introduces the CGAN-based channel modeling algorithm.

A. Channel Model

In a MIMO system, the transmitter and receiver are equipped with N_t transmitting antennas and N_r transmitting antennas, respectively. During each time slot, the transmitter sends a symbol vector $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_{N_t}]^T \in \mathbb{C}^{N_t}$ to the receiver, subject to the transmission power constraint $\|\mathbf{x}\|^2 = 1$. The received signal $\mathbf{y} \in \mathbb{C}^{N_r}$ is influenced by the spatial interactions and multipath effects, and it can be expressed as a linear combination of the transmitted symbols as follows:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z} \tag{1}$$

where $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ represents the channel coefficient matrix, and $\mathbf{z}$ denotes the noise term at the receiver. Here, the elements of $\mathbf{z}$ represent i.i.d. circularly symmetric complex Gaussian random variables with zero mean and variance σ^2, i.e., $\mathbf{z} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{N_r})$. The transmit signal-to-noise ratio (SNR) is $1/\sigma^2$ in this work.

B. Channel Modeling with CGAN

The distinctive feature of generative adversarial networks (GANs) lies in its generative model structure, consisting of two key components: a generator and a discriminator. The generator is designed to model the target distribution, while the discriminator's role is to distinguish whether received samples come from the real distribution or the generator. During training, the real channel is substituted with a learned generator, enabling the back-propagation of gradients from the receiver through the learned channel to the transmitter. The discriminator is optimized to maximize its ability to differentiate between generated and real samples, while the generator is simultaneously trained to minimize the distance between generated and real samples.

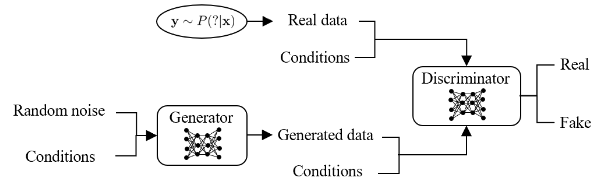

Fig. 2. Architecture of CGAN.

The CGAN represents a special example of the GAN, when both the generator and the discriminator are conditioned on some extra information $\mathbf{c}$. As depicted in Fig. 2, the conditional information is fed as an extra input into both the generator and the discriminator. The objective function for optimization related to the generator G and the discriminator D can be formulated as

$$\begin{aligned}\min_G \max_D \mathcal{L}(D, G) = E_{\mathbf{y}}[\log D(\mathbf{y}|\mathbf{c})] \\ + E_{\mathbf{n}}[\log(1 - D(G(\mathbf{n}|\mathbf{c})|\mathbf{c})]\end{aligned} \tag{2}$$

where $\mathbf{y}$ is the real received data, and $\mathbf{n}$ is a random noise distributed corresponding to some distribution p_n. $D(\mathbf{y}|\mathbf{c})$ is the output of the discriminator when the received data $\mathbf{y}$ in (1) is provided as input, and $D(G(\mathbf{n}|\mathbf{c})|\mathbf{c})$ denotes the output of the discriminator when the generated data $G(\mathbf{n}|\mathbf{c})$ comes as input, where $\mathbf{y}_G = G(\mathbf{n}|\mathbf{c})$ is the output of the generator. The transmitted signal and the received signal for the pilot symbol are selected as the conditional information in the SISO case [9], [10]. In this MIMO scenario, the condition $\mathbf{c}$ still includes the transmitted signal, denoted as $\mathbf{x}$, and the received pilot data, denoted as $\mathbf{y}_{\text{pilot}}$, but differs slightly from the case of the SISO system. Specifically, the received pilot data $\mathbf{y}_{\text{pilot}}$ involves the received signals when the transmitter sequentially sends the N_t pilot signals, employing only one antenna among N_t transmitting antennas at a time. In this end-to-end MIMO system, the CGAN is employed to learn the channel distribution by applying both the transmitted signals ($\mathbf{x}$) and the received pilot signals ($\mathbf{y}_{\text{pilot}}$) as depicted in Fig. 3.

III. End-to-End System Based on CGAN

With the help of the CGAN, the independent transmitter neural network and receiver neural network are connected. The connection facilitates the propagation of gradients to the transmitter during its training, leading to the updating of transmitter parameters. Consequently, an end-to-end communication system based on the CGAN is established. The details of the system and training process are outlined in this section.

A. System Description

The end-to-end CGAN-based system is illustrated in Fig. 1 (c). A message $s \in \mathcal{S} = \{1, \cdots, M^{N_t}\}$ is to be transmitted from the transmitter with N_t antennas. The transmitter encodes the message into a one-hot vector $\mathbf{s}_{oh}$ of length M^{N_t}. The end-to-end algorithm treats signal detection as an M^{N_t}-class classification problem, producing a probability vector $\hat{\mathbf{s}}_{oh}$ with

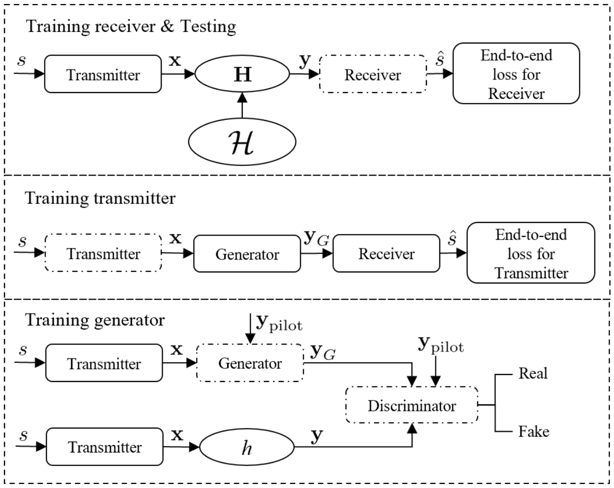

Fig. 3. Training and testing process

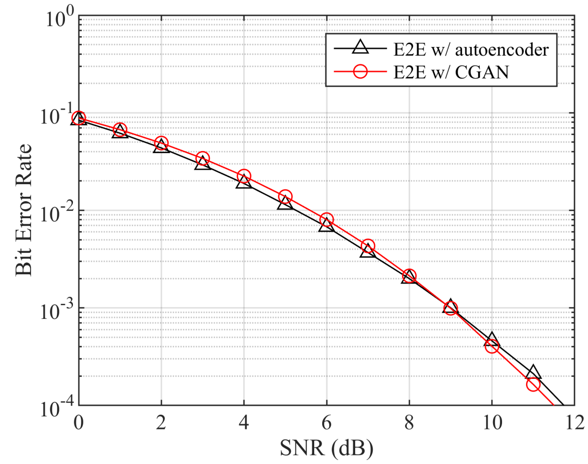

Fig. 4. BER comparsion with $N_t = N_r = 2$ and $M = 2$.

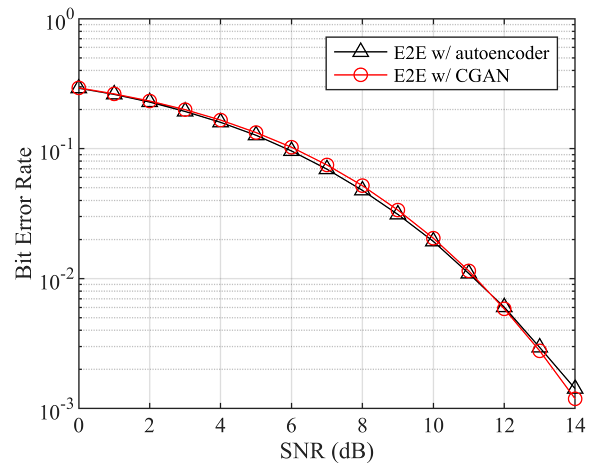

Fig. 5. BER comparsion with $N_t = N_r = 2$ and $M = 4$.

M^{N_t} potential categories. The performance is evaluated using cross-entropy, defined as follows:

$$\mathcal{L} = \sum_{m=1}^{M^{N_t}} -s_{oh}[m] \log(\hat{s}_{oh}[m]) \tag{3}$$

where $s_{oh}[m]$ and $\hat{s}_{oh}[m]$ denote the m-th element of $\mathbf{s}_{oh}$ and $\hat{\mathbf{s}}_{oh}$, respectively.

B. Training and Testing Process

The training and testing procedures are shown in Fig. 3. The instantaneous CSI is randomly sampled from an extensive channel set $\mathcal{H}$. At each iteration, one module (dash-dot box) is trained, while the parameters of the remaining modules (solid box) are fixed. This iterative process alternately trains the receiver, transmitter, and the channel generator.

During receiver training, the purpose is to reconstruct the received signal $\mathbf{y}$ to determine the transmitted message s. The optimization of the loss function in (3) aims to minimize the difference between s and the estimated value $\hat{s}$. The feasibility of direct back-propagation of gradients arises from the differentiable nature of the receiver as a neural network.

In the training of the transmitter neural network, supervision of the encoder's output, $\mathbf{x}$, by the true value s becomes challenging if gradient propagation is hindered due to a non-differentiable channel. To enable gradient propagation from the receiver to the transmitter during transmitter training, a differentiable generator substitutes the real channel and is trained to emulate the real channel's behavior.

As mentioned in II-B, the discriminator and the generator are trained simultaneously during the generator training process. Real data $\mathbf{y}$ for the CGAN training is obtained from the output of the real channel in (1), passing through the transmitter, while fake data $\mathbf{y}_G$ is generated by the generator. The objective function (2) guides the parameter updates in the CGAN training.

IV. Simulation and Results

Simulation details and results are presented in this section. Our method is implemented in Tensorflow 2.7, and all simulations are conducted under Rayleigh fading channels. During the training stage, the models are trained using $100,000$ randomly generated messages and their corresponding received signals $\mathbf{y}$ following the system structure in (1) considering a given channel condition $\mathbf{H}$ at a fixed SNR of 10 dB. Adam optimizer is used with the learning of 0.001. During the testing phase, the evaluation is conducted on a dataset of $10,000$ samples under the same channel condition $\mathbf{H}$. The process is conducted for multiple Rayleigh-distributed fading channels. We compare the CGAN-based end-to-end system with the autoencoder-based research [6] for MIMO systems in term of bit error rates (BERs) across various SNRs. This work provides comparison results for four scenarios, as shown from Fig. 4 to Fig. 7.

In the case of $N_t = N_r = 2$ and $M = 2$ as depicted in Fig. 4, the performance of the end-to-end MIMO system

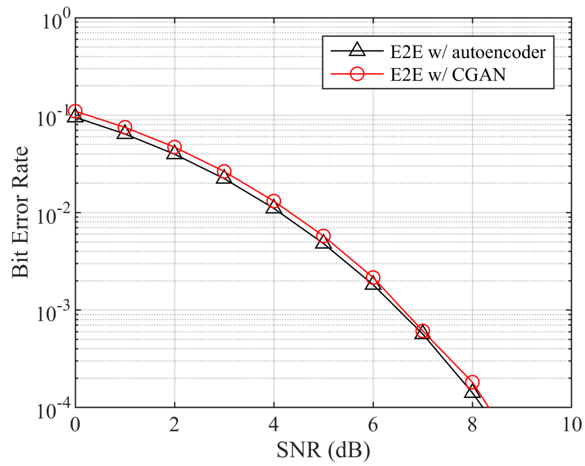

Fig. 6. BER comparsion with $N_t = N_r = 4$ and $M = 2$.

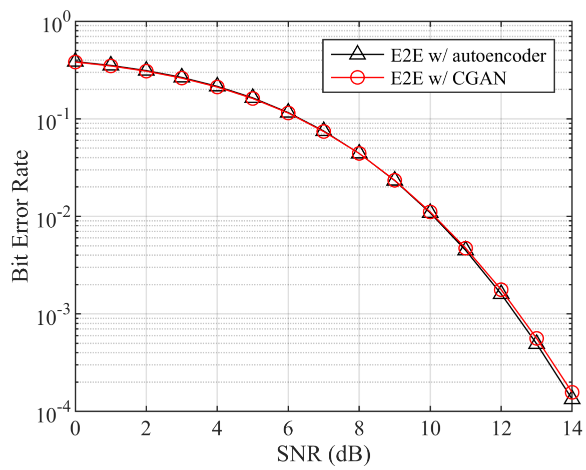

Fig. 7. BER comparsion with $N_t = N_r = 4$ and $M = 4$.

closely resembles that of autoencoder-based approach [6] in terms of BER. Fig. 5 demonstrates a comparable trend when $M = 4$, providing further validation of the effectiveness of the CGAN method.

Subsequently, with an increase in the number of antennas to $N_t = N_r = 4$ in Figs 6 and 7, performance results for both $M = 2$ and $M = 4$ demonstrate that the CGAN-based MIMO system can achieve comparable superior performance compared to the autoencoder-based approach.

V. Conclusion

In summary, the challenges encountered by autoencoder-based end-to-end communication systems, particularly in the context of gradient propagation issues for non-differentiable channel models, have prompted the exploration of alternative solutions. In response to this, the CGAN has been introduced as a viable solution. Building upon this concept, we extend the CGAN-based system from SISO systems to MIMO scenarios. The presented end-to-end approach demonstrates its effectiveness in the Rayleigh-fading MIMO channel environment, exhibiting performance comparable to the autoencoder-based approach. Simulation results support the notion that the CGAN serves as a promising alternative to conventional MIMO channel models. In light of these compelling findings, our future work will be dedicated to extending the application of the CGAN to practical non-differentiable channel conditions, thereby enhancing the application of the method in real channel scenarios.

References

[1] H. He, C. -K. Wen, S. Jin and G. Y. Li, “Deep learning-based channel estimation for beamspace mmWave massive MIMO systems,” *IEEE Wireless Commun. Letters*, vol. 7, no. 5, pp. 852-855, Oct. 2018.

[2] E. Nachmani, Y. Be’ery and D. Burshtein, “Learning to decode linear codes using deep learning,” *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, USA, 2016, pp. 341-346.

[3] T. Gruber, S. Cammerer, J. Hoydis and S. t. Brink, “On deep learning-based channel decoding,” *2017 51st Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, USA, 2017, pp. 1-6.

[4] N. Samuel, T. Diskin and A. Wiesel, “Deep MIMO detection,” *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Sapporo, Japan, 2017, pp. 1-5.

[5] T. O’Shea and J. Hoydis, “An introduction to deep learning for the physical layer,” *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 563-575, Dec. 2017.

[6] T. O’Shea, T. Erpek, and T. C. Clancy, “Deep learning-based MIMO communications,” 2017, *arXiv:1707.07980*. [Online]. Available: https://arxiv.org/abs/1707.07980

[7] T. J. O’Shea, T. Roy and N. West, “Approximating the void: learning stochastic channel models from observation with variational generative adversarial networks,” *2019 International Conference on Computing, Networking and Communications (ICNC)*, Honolulu, HI, USA, 2019, pp. 681-686.

[8] M. Mirza, and S. Osindero, “Conditional generative adversarial nets,” *arXiv preprint arXiv:1411.1784*, 2014.

[9] H. Ye, G. Y. Li, B. -H. F. Juang and K. Sivanesan, “Channel agnostic end-to-end learning based communication systems with conditional GAN,” *2018 IEEE Globecom Workshops (GC Wkshps)*, Abu Dhabi, United Arab Emirates, 2018, pp. 1-5.

[10] H. Ye, L. Liang, G. Y. Li and B. -H. Juang, “Deep learning-based end-to-end wireless communication systems with conditional GANs as unknown channels,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3133-3143, May 2020.