Transfer Generative Adversarial Networks (T-GAN)-based Terahertz Temporal Channel Modeling and Generating

Zhengdong Hu, Yuanbo Li, and Chong Han

Shanghai Jiao Tong University, Email: {huzhengdong, yuanbo.li, chong.han}@sjtu.edu.cn

ABSTRACT

Terahertz (THz) communications are a promising technology for 6G and beyond wireless systems, offering ultra-broad bandwidth and thus data rates of Terabit-per-second (Tbps). However, accurate channel modeling and characterization are fundamental for the design of THz communications. Relying on channel measurements, traditional statistical channel modeling methods suffer from low accuracy due to the assumed certain distributions and empirical parameters. Moreover, acquiring extensive channel measurement is time-consuming and expensive in the THz band. To address these challenges, a transfer generative adversarial network (T-GAN) based modeling method is proposed in the THz band, which exploits the advantage of GAN in modeling the complex distribution. Moreover, the transfer learning technique is introduced in T-GAN, which transfers the knowledge stored in a pre-trained model based on simulated data, to a new model based on a small amount of measured data. The simulation data is generated by the standard channel model from 3rd generation partnerships project (3GPP), which contains the knowledge that can be transferred to reduce the demand of measurement data and improve the accuracy of T-GAN. Experimental results reveal that the distribution of power delay profiles (PDPs) generated by the proposed T-GAN method shows good agreement with measurement. Moreover, T-GAN achieves good performance in channel modeling, with 9 dB improved root-mean-square error (RMSE) and higher Structure Similarity Index Measure (SSIM), compared with traditional 3GPP method.

CCS CONCEPTS

• **Hardware → Signal processing systems**.

KEYWORDS

Terahertz, channel modeling, transfer learning, generative adversarial networks.

1 INTRODUCTION

As the number of interconnected devices continues to grow exponentially, the sixth generation (6G) is expected to achieve intelligent connections of everything, anywhere, anytime [2], with data rates reaching Terabit-per-second (Tbps). To meet the demand, Terahertz (THz) communications have emerged as a vital technology of 6G systems, due to the ultra-broad bandwidth ranging from tens of GHz to hundreds of GHz [3]. The THz band is a promising solution to address the spectrum scarcity and capacity limitations of current wireless systems, and enable new applications, such as wireless cognition, localization/positioning, integrated sensing and communication.

To design reliable THz wireless systems, one fundamental challenge lies in developing an accurate channel model to portray the propagation phenomena. Due to the high frequencies, new characteristics occur in the THz band, such as frequency-, distance- and environment-dependent molecular absorption and rough-surface diffusely scattering. Due to these attributes, channel modeling needs to be properly conducted to capture the THz characteristics. However, traditional statistical channel modeling methods suffer from the problem of low accuracy with the assumed certain distributions and empirical parameters. For example, a geometric based stochastic channel model (GSCM) assumes that the positions of scatters follow certain statistical distributions, such as the uniform distribution within a circle around the transmitters and receivers [8]. However, the positions of

scatters are hard to characterize by certain statistical distributions, making the GSCM not accurate for utilization in the THz band. Moreover, it is time-consuming and costly to acquire extensive channel measurement for THz channel modeling. To this end, an accurate channel modeling method with limited measurement data for the THz band is needed.

Recently, deep learning (DL) is popular and widely applied in wireless communications. Among different kinds of DL methods, the generative adversarial network (GAN) has the advantage of modeling complex distribution accurately without any statistical assumptions, based on which GAN can be utilized to develop channel models. The authors in [7] train GAN to approximate the probability distribution functions (PDFs) of stochastic channel response. In [11], GAN is applied to generate channel samples close to the distribution of original channel samples. The researchers in [6] model the channel with GAN through channel input-output measurements. In [10], a model-driven GAN-based channel modeling method is developed in intelligent reflecting surface (IRS) aided communication system. All of the aforementioned works train the GAN network with a large number of synthetic channel samples, generated by the conventional channel models. This causes mismatch between the formulated GAN models and the models developed based on measurement data. Therefore, incorporating measurement data into the training and testing of the GAN network is necessary to overcome this limitation. Nevertheless, it is impractical to leverage a large amount of measurement data in designing the GAN.

In this paper, a transfer GAN (T-GAN)-based THz temporal channel modeling method is proposed, which models the channel by learning the distribution of power delay profile (PDP). PDP indicates the dispersion of power over the time delay, which can well characterize the channel in the temporal domain. Moreover, to tackle the challenge of limited channel measurement in the THz band, the transfer learning technique is introduced in T-GAN, which transfers the knowledge stored in a pre-trained model based on synthetic data from 3GPP standardized channel simulation to a new model based on a small amount of measured data [4, 9]. This can alleviate the demand of large amount of measurement data for training and improve the accuracy of T-GAN, since the synthetic data can serve as a good compensation for the initialization of GAN network. Furthermore, the performance of T-GAN in modeling the channel distribution is validated by real measurement [5].

The contributions of this paper are listed as follows.

- **We propose a T-GAN based THz channel modeling method**, in which a GAN is designed to capture the distribution of PDPs of the THz channel, by training on the dataset of PDP samples.

- **To tackle the challenge of limited measurement data for THz channel modeling, transfer learning is further exploited by T-GAN**, which reduces the size requirement of training dataset, and enhances the performance of T-GAN, through transferring the knowledge stored in a pre-trained model based on 3GPP simulated data to a new model based on a small amount of measured data.

The rest of the sections are organized as follows. Sec. 2 details the proposed T-GAN based channel modeling method. Sec. 3 demonstrates the performance of the proposed T-GAN method. The paper is concluded in Sec. 4.

Notation: a is a scalar. $\mathbf{a}$ denotes a vector. $\mathbf{A}$ represents a matrix. $\mathbb{E}\{\cdot\}$ describes the expectation. ∇ denotes the gradient operation. $\|\cdot\|$ represent the L2 norm. $\mathbf{I}_N$ defines an N dimensional identity matrix. $\mathcal{N}$ denotes the normal distribution.

2 TRANSFER GAN (T-GAN) BASED CHANNEL MODELING

In this section, the channel modeling problem is first formulated into a channel distribution learning problem. Then, the proposed GAN in T-GAN method is elaborated. Finally, T-GAN is presented.

2.1 Problem Formulation

The THz channel can be represented as

$$h(\tau) = \sum_{l=0}^{L-1} \alpha_l e^{j\phi_l} \delta(\tau - \tau_l), \tag{1}$$

where τ_l denotes the delay of the lth multi-path component (MPC), L denotes the number of MPCs, α_l refers to the path gain and ϕ_l represents the phase of the corresponding MPC. Moreover, PDP is an important feature to characterize the channel in the temporal domain, which represents the received power with respect to the delay in a multi-path channel. It can be expressed as

$$P(\tau) = |h(\tau)|^2, \tag{2}$$

Then, the channel modeling problem is exploited by learning the distribution of PDPs denoted by p_r, which is difficult to be analytically represented. Instead, the distribution p_r can be implicitly captured by generating fake PDP samples with distribution p_g, such that the generated distribution p_g of PDPs can match the actual distribution p_r.

2.2 Proposed GAN

The GAN network is proposed to learn the distribution of PDPs, with the framework depicted in Fig 1. The GAN consists of two sub-networks, namely, generator and discriminator. The generator is aimed at generating fake samples $G(z)$
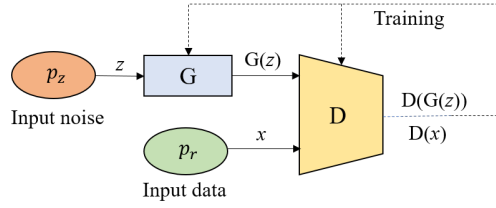
Figure 1: Framework of GAN.

to fool the discriminator, while the discriminator tries to distinguish between real samples x and fake samples $G(z)$.

The input to the generator is a noise vector z with dimension $n_z = 100$, which is sampled from the probability density function $\mathcal{N}(0, \sigma^2 \mathbf{I}_{n_z})$. The generator consists of five dense layers, and the numbers of neurons in the dense layers are 128, 128, 128, 128, 401, respectively. It is worth noting that the size of the output layer is equal to the size of PDP. The activation function of the first four dense layers is LeakyReLU function, which can speed up the convergence and avoid the gradient vanishing problem. The formula of the LeakyReLU function is expressed as

$$f(x) = \begin{cases} x, & \text{if } x \geq 0 \\ \alpha x, & \text{if } x < 0 \end{cases}, \tag{3}$$

where α is the slope coefficient when the value of neuron x is negative. In addition to the LeakyReLU function, a Sigmoid function is utilized in the last layer, which maps the output to the range of $[0, 1]$. The Sigmoid function is defined as

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{4}$$

After going through the dense layers and activation functions in the generator, the input noise vectors are transformed into the generated samples. Then, the generated samples $G(z)$ together with real samples x are passed to the discriminator.

The discriminator is designed to distinguish between generated samples and real samples. The numbers of neurons for the five dense layers in the discriminator are 512, 256, 128, 64, 1, respectively. The activation function chosen for the first 4 layers is the LeakyReLU function as introduced before, while the final output of the discriminator $D(x)$ and $D(G(z))$ are activated linearly.

Then, the generator and discriminator are trained in an adversarial manner, which can be considered as a two-player zero-sum minimax game. Specifically, the training objective can be represented by

$$\min_G \max_D \mathbb{E}_{x \sim p_r}[D(x)] + \mathbb{E}_{z \sim p_z}[(1 - D(G(z)))]$$
$$+ \lambda \mathbb{E}_{\tilde{x}}[(\|\nabla_{\tilde{x}} D(\tilde{x})\| - 1)^2)], \tag{5}$$

where p_r and p_z represent the distributions of real channels and noise vector, respectively. The generator minimizes the
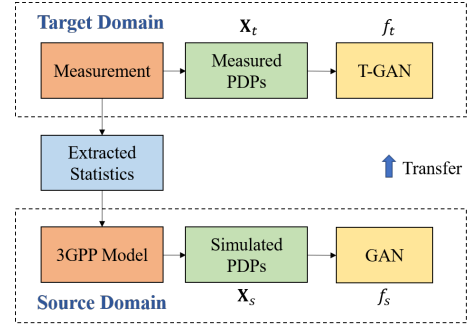

Figure 2: Framework of T-GAN.

probability $(1 - D(G(z))$ that the generated sample is detected as fake by the discriminator, while the discriminator maximizes this probability. Therefore, the generator and discriminator compete against each other with the opposite objectives in the training process. Through the adversarial training, the Nash equilibrium can be achieved, such that the generator and discriminator cannot improve their objectives by changing only their own network. Moreover, the last term in (5) is the gradient penalty to enforce Lipschitz constraint such that the gradient of the GAN network is upper-bounded by a maximum value. This term is essential to mitigate the gradient vanishing or exploding problem in the training of GAN network. Specifically, the symbol $\tilde{x}$ is the sampled point along the straight lines between the paired points of x and $G(z)$, and the parameter λ is the penalty coefficient.

2.3 Proposed T-GAN

The framework for the proposed T-GAN is depicted in Fig. 2, in which the transfer learning is conducted to transfer the knowledge from the source domain to the target domain. The source domain is defined on the PDPs with large size simulated by 3GPP model [1], while the target domain is based on the measured PDPs of small size. Moreover, the symbols $\mathbf{X_s}$ and $\mathbf{X_t}$ denotes the feature spaces of source domain and target domain, respectively. The two feature spaces are closely related, since the 3GPP model are implemented with the extracted statics from measurement including delay spread, angular spread and so on. Therefore, there exists shared knowledge between the two domains, which can be transferred.

In both of the source domain and the target domain, a task is defined to capture the distribution of the PDPs. Moreover, the symbols f_s and f_t denote the predictive functions of the two domains to fulfil the task, which can be learned by the proposed GAN network through the training process. The GAN network can be well trained in the source domain, with a large number of simulated PDPs. However, the size of measured PDPs is quite small for the training of GAN in the
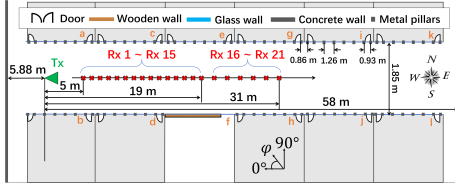
Figure 3: Measurement layout in the indoor corridor scenario [5].

Figure 4: Loss of the generator and discriminator in the GAN network.

target domain, which can cause the difficulty of converging or the over-fitting problem. Then, the solution is to transfer the knowledge learned in the pre-trained GAN network in the source domain to the target domain, which yields the T-GAN network.

The method of fine-tuning [9] is adopted for the transfer learning process. The T-GAN is initialized with the weights of the GAN pre-trained on the simulated PDPs, and is then fine-tuned on the measured PDPs. It is worth noting that the generator and discriminator in the GAN are both transferred, which can yield the better performance in generating high quality samples and fast convergence, compared with transferring only the generator or the discriminator [9].

With transfer learning, the performance of T-GAN can be largely enhanced. Specifically, the channel statistics extracted for 3GPP method are captured by the proposed GAN trained on simulated PDPs, which are further transferred to T-GAN. Moreover, T-GAN can learn the features of PDPs that are not characterized by 3GPP method, directly from measurement, which further improves the performance of T-GAN in modeling the distribution of PDPs.

3 EXPERIMENT AND PERFORMANCE EVALUATION

In this section, the experiment settings are elaborated. Moreover, the performance of the T-GAN are evaluated by comparing the generated distribution of PDPs with measurement.

3.1 Dataset and Setup

The dataset is collected from the measurement campaign in [5]. which is conducted in an indoor corridor scenario at 306-321 GHz with 400 ns maximum delay, as depicted in Fig. 3. With the measurement data, the PDPs can be extracted to characterize the channel in the 21 receiver points. Since the sample frequency interval is relatively small, as 2.5 MHz, the measured PDPs are very long, including 6001 sample points, which results in extraordinary computation and time consumption to train the GANs. To address this problem, we only use the measured channel transfer functions in the frequency band from 314 to 315 GHz, based on which the PDPs can be shorten to 401 sample points.
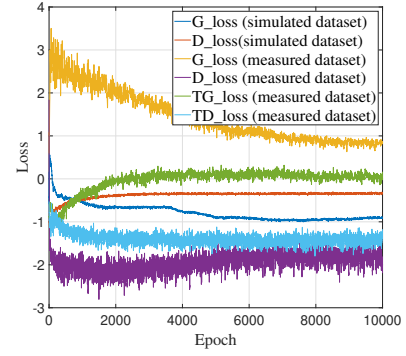
The PDPs of the 21 measured channels make up the measured dataset. In addition to the measured dataset, the dataset of simulated PDPs can be generated by 3GPP model with the extracted statistics from the measurement, which consists of 10000 channels. Compared to the measured dataset, the simulated dataset has larger data size with the channel statistics embedded. Moreover, the PDPs in two datasets are normalized into the range of [0, 1] by the min-max normalization method.

The training procedure of the GAN network is explained in detail as follows. Firstly, the input noise vector z of size 100 is generated by the multivariate normal distribution, which can provide the capabilities to transform into the desired distribution. The gradient penalty parameter λ in (5) is set as 10, which works well in the training process. Moreover, the stochastic gradient descent (SGD) optimizer is applied for the generator network, and the adaptive moment estimation (Adam) optimizer is chosen for the discriminator network. In addition, the learning rates of the two optimizers are both set as 0.0002 to stabilize the training.

All the experimental results are implemented on a PC with AMD Ryzen Threadripper 3990X @ 2.19 GHz and four Nvidia GeForce RTX 3090 Ti GPUs. In addition, the training of GAN network is carried out in the Pytorch framework.

3.2 Convergence

The proposed GAN is first trained on the simulated dataset, and is then fine-tuned on the measured dataset with transfer learning to develop the T-GAN. The numbers of epochs for training the proposed GAN and T-GAN are both set as 10000. A epoch is defined as a complete training cycle through the training dataset, in which the generator and discriminator are iteratively trained for once. To demonstrate the benefits of transfer learning, the GAN is also trained on the measured dataset without transfer learning for comparison. The loss
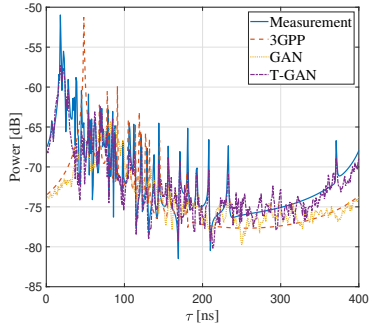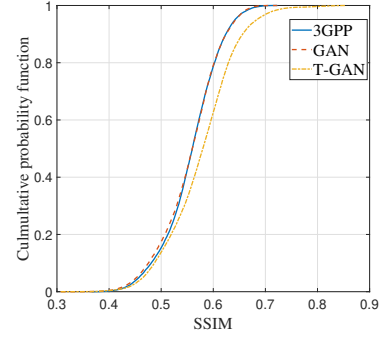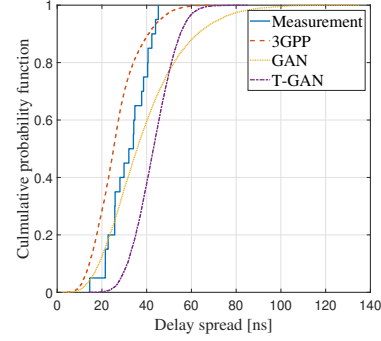
(a) Samples of PDP.
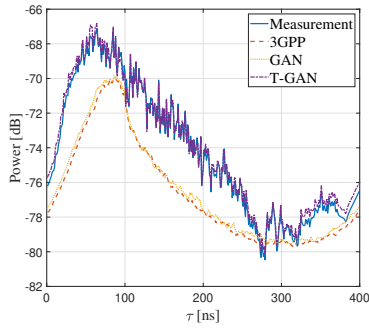
(b) Average PDP.

Figure 5: Plot of PDPs generated by measurement, 3GPP, the proposed GAN and T-GAN.

Figure 6: SSIM of PDP for 3GPP, the proposed GAN and T-GAN.

Figure 7: Delay spread for 3GPP, the proposed GAN and T-GAN.

of generator denoted by G_loss and loss of discriminator denoted by D_loss are shown in the Fig. 4, in which the TG_loss and TD_loss correspond to the losses for T-GAN. For the simulated dataset, it is clear that the generator and discriminator reach the equilibrium in the end. For the measured dataset, the loss of T-GAN is close to the loss for the simulated dataset except for some small fluctuations. The fluctuations are due to the small size of the measured dataset. By comparison, the training is not stable for the GAN network without transfer leaning. There is large fluctuation in the discriminator loss, and the absolute values of G_loss and D_loss are quite large compared to the losses for the simulated dataset. The comparison demonstrates the benefits of the transfer learning in the training of GAN network, which enables T-GAN to converge with a small training dataset. Moreover, it takes 4000 epochs for T-GAN to converge, compared to 6000 epochs for GAN trained on the simulated dataset. From these results, it is clear that the transfer learning technique can improve the convergence rate of T-GAN, and reduce the training overhead with the knowledge from the pre-trained model.

3.3 Power Delay Profile

In the experiment, the samples of PDP from measurement, 3GPP method, the proposed GAN and T-GAN are compared as in Fig. 5(a). It is clear that the PDPs are similar to each other, which proves that the proposed GAN and T-GAN can learn the features of PDPs. Moreover, it is observed that PDP of measurement is more complex than PDP of 3GPP method. There are more peaks and fluctuations in the temporal domain. This shows that 3GPP cannot well capture the channel effects embedded in PDP. Comparing PDPs generated by the proposed GAN and T-GAN, the PDP generated by T-GAN is close to measurement, while the PDP generated by the proposed GAN is similar to the 3GPP approach. This is reasonable, since the T-GAN can capture the features of PDP from measurement through transfer learning, while the propose GAN can only learn the features of the simulated PDPs by 3GPP method.

In addition, the average PDPs for these method are plotted in Fig. 5(b). It is clear that T-GAN shows good agreement with measurement, while 3GPP and GAN have large deviations

from measurement. The deviations can be measured by root-mean-square error (RMSE), calculated as

$$\text{RMSE} = \sqrt{\frac{1}{N_\tau} \sum (P_m(i\Delta\tau) - P_g(i\Delta\tau))^2}, \quad (6)$$

where N_τ denotes the number of sampling points in PDP, i indexes temporal sample points of PDPs, N_τ represents the number of sampling points and $\Delta\tau$ is the sampling interval. Moreover, $P_m(i\Delta\tau)$ and $P_g(i\Delta\tau)$ are the average power in the ith sample point of measured PDPs and generated PDPs, respectively. The results of RMSE for 3GPP, the proposed GAN and T-GAN are 4.29 dB, 4.12 dB and -4.82 dB, respectively. The T-GAN improves the performance of RMSE by about 9 dB, compared with other methods, which demonstrates that the T-GAN outperforms the other methods in terms of modeling the average power of PDP. This is attributed to the powerful capability of GAN in modeling the complex distribution, and the benefits of transfer learning in better utilizing the small measurement dataset.

Moreover, to measure the similarity quantitatively, Structure Similarity Index Measure (SSIM) is introduced, which is widely applied to evaluate the quality and similarity of images. The range of SSIM is from 0 to 1, and the value of SSIM is larger when the similarity between images is higher. The PDPs generated by 3GPP method, the proposed GAN and T-GAN are compared with measurement. The cumulative probability functions (CDFs) of SSIM for these method are shown in Fig. 6. It can be observed that the proposed T-GAN can achieve higher SSIM values compared with other methods. More than 40% of SSIM values are higher than 0.6 for T-GAN, compared to only 20% for 3GPP and the proposed GAN. This further demonstrates the better performance of T-GAN in modeling the PDPs.

3.4 Delay Spread

Delay spread characterizes the power dispersion of multipath components in the temporal domain, which can be calculated as the second central moment of PDPs, by

$$\bar{\tau} = \frac{\sum_{i=0}^{N_\tau} i\Delta\tau P(i\Delta\tau)\Delta\tau}{\sum_{i=0}^{N_\tau} P(i\Delta\tau)\Delta\tau},$$
$$\tau_{rms} = \sqrt{\frac{\sum_{i=0}^{N_\tau} (i\Delta\tau - \bar{\tau})^2 P(i\Delta\tau)\Delta\tau}{\sum_{i=0}^{N_\tau} P(i\Delta\tau)\Delta\tau}}, \quad (7)$$

where $\bar{\tau}$ denotes the mean delay weighted by the power, τ_{rms} refers to the root-mean-square (RMS) delay spread, and $P(i\Delta\tau)$ are the power in the ith sample point of PDPs.

Then, the CDF plot of delay spread for measurement, 3GPP, the proposed GAN and T-GAN is depicted in Fig. 7. It can be observed that the CDFs of delay spread for 3GPP, the proposed GAN and T-GAN match the measurement well.

4 CONCLUSION

In this paper, we proposed a T-GAN based THz temporal channel modeling method, which can capture the distribution of PDPs for the THz channel. Moreover, the transfer learning is exploited in T-GAN to reduce the size requirement of training dataset and enhance the performance of T-GAN, through transferring the knowledge stored in the pre-trained GAN on the simulated dataset by 3GPP method to the target T-GAN trained on limited measurement. Finally, we validate the performance of T-GAN with measurement. T-GAN can generate the PDPs that have good agreement with measurement. Compared with conventional methods, T-GAN has better performance in modeling the distribution of PDPs, with 9 dB improved RMSE and higher SSIM. More than 40% of SSIM values are higher than 0.6 for T-GAN, compared with only 20% for 3GPP and the proposed GAN.

5 ACKNOWLEDGEMENT

REFERENCES

[1] 3GPP. 2018. *Study on Channel Model for Frequencies From 0.5 to 100 GHz (Release 15)*. TR 38.901.
[2] Ian F Akyildiz, Chong Han, Zhifeng Hu, Shuai Nie, and Josep Miquel Jornet. 2022. Terahertz band communication: An old problem revisited and research directions for the next decade (invited paper). *IEEE Trans. Commun.* 70, 6 (2022), 4250–4285.
[3] Z. Chen et al. 2021. Terahertz Wireless Communications for 2030 and Beyond: A Cutting-Edge Frontier. *IEEE Commun. Mag.* 59, 11 (2021), 66–72.
[4] Nguyen Van Huynh and Geoffrey Ye Li. 2022. Transfer Learning for Signal Detection in Wireless Networks. *IEEE Wirel. Commun. Lett.* (2022), 1–1.
[5] Yuanbo Li, Yiqin Wang, Yi Chen, Ziming Yu, and Chong Han. 2022. Channel Measurement and Analysis in an Indoor Corridor Scenario at 300 GHz. In *Proc. IEEE Int. Conf. Commun.* 2888–2893.
[6] Tribhuvanesh Orekondy, Arash Behboodi, and Joseph B. Soriaga. 2022. MIMO-GAN: Generative MIMO Channel Modeling. In *Proc. IEEE Int. Conf. Commun.* 5322–5328.
[7] Timothy J. O'Shea, Tamoghna Roy, and Nathan West. 2019. Approximating the Void: Learning Stochastic Channel Models from Observation with Variational Generative Adversarial Networks. In *Proc. Int. Conf. Comput., Netw. Commun.* 681–686.
[8] P. Petrus, J.H. Reed, and T.S. Rappaport. 2002. Geometrical-based statistical macrocell channel model for mobile environments. *IEEE Trans. Commun.* 50, 3 (2002), 495–502.
[9] Yaxing Wang, Chenshen Wu, Luis Herranz, Joost van de Weijer, Abel Gonzalez-Garcia, and Bogdan Raducanu. 2018. Transferring gans: generating images from limited data. In *ECCV*.
[10] Yi Wei, Ming-Min Zhao, and Min-Jian Zhao. 2022. Channel Distribution Learning: Model-Driven GAN-Based Channel Modeling for IRS-Aided Wireless Communication. *IEEE Trans. Commun.* 70, 7 (2022), 4482–4497.
[11] Han Xiao, Wenqiang Tian, Wendong Liu, and Jia Shen. 2022. Channel-GAN: Deep Learning-Based Channel Modeling and Generating. *IEEE Wirel. Commun. Lett.* 11, 3 (2022), 650–654.