One-Bit mmWave MIMO Channel Estimation Using Deep Generative Networks

Akash Doshi\textsuperscript{c*} and Jeffrey G. Andrews\textsuperscript{c}, Fellow, IEEE

Abstract—As future wireless systems trend towards higher carrier frequencies and large antenna arrays, receivers with one-bit analog-to-digital converters (ADCs) are being explored owing to their reduced power consumption. However, the combination of large antenna arrays and one-bit ADCs makes channel estimation challenging. In this letter, we formulate channel estimation from a limited number of one-bit quantized pilot measurements as an inverse problem and reconstruct the channel by optimizing the input vector of a pre-trained deep generative model with the objective of maximizing a novel correlation-based loss function. We observe that deep generative priors adapted to the underlying channel model significantly outperform Bernoulli-Gaussian Approximate Message Passing (BG-GAMP), while a single generative model that uses a conditional input to distinguish between Line-of-Sight (LOS) and Non-Line-of-Sight (NLOS) channel realizations outperforms BG-GAMP on LOS channels and achieves comparable performance on NLOS channels in terms of the normalized channel reconstruction error.

Index Terms—Deep generative models, low resolution receivers, mmWave MIMO channel estimation.

I. INTRODUCTION

Channel estimation (CE) in 6G and beyond will be performed at increasingly higher carrier frequencies, leading to an increase in the dimensionality and complexity of the problem due to the associated increase in antenna array sizes at the base station (BS) and user (UE). Conventional sub-6 GHz CE techniques such as least squares (LS) and minimum mean squared error (MMSE) estimators require full rank pilot measurements to recover the channel, hence will not scale to this squared error (MMSE) estimators require full rank pilot measurements at the base station (BS) and user (UE). Conventional sub-6 GHz CE techniques such as least squares (LS) and minimum mean squared error (MMSE) estimators require full rank pilot measurements to recover the channel, hence will not scale to this large antenna arrays and one-bit ADCs makes channel estimation challenging. In this letter, we formulate channel estimation from a limited number of one-bit quantized pilot measurements as an inverse problem and reconstruct the channel by optimizing the input vector of a pre-trained deep generative model with the objective of maximizing a novel correlation-based loss function. We observe that deep generative priors adapted to the underlying channel model significantly outperform Bernoulli-Gaussian Approximate Message Passing (BG-GAMP), while a single generative model that uses a conditional input to distinguish between Line-of-Sight (LOS) and Non-Line-of-Sight (NLOS) channel realizations outperforms BG-GAMP on LOS channels and achieves comparable performance on NLOS channels in terms of the normalized channel reconstruction error.

The transmission of a fixed training signal. Utilizing one-bit ADCs at the receiver makes channel estimation even more challenging due to their severe non-linearity [8]. In this letter, we describe an unsupervised learning technique to perform channel estimation from a small number of one-bit quantized pilot measurements using deep generative priors. This technique was introduced in [9] to perform full-resolution channel estimation from a limited number of pilots. However, the prior work in [9] assumed an antenna spacing of \(\lambda_c/10\) (where \(\lambda_c = c/f_c\) and \(f_c\) is the carrier frequency) in place of the conventional \(\lambda_c/2\) to generate high spatial correlation in channel realizations and also trained a deep generative model and evaluated its performance only on channels with a strong LOS component. While we addressed the aforementioned limitations in the context of full-resolution channel estimation in [10], in this letter, we will additionally extend the framework to heavily quantized channel estimation. To be more specific, we will perform channel estimation from compressive one-bit pilot measurements by optimizing the input vector to a pre-trained deep generative model with the objective of maximizing the correlation between the generator output and pilot measurements.

II. SYSTEM MODEL

Consider a single user setup with a transmitter and receiver having \(N_t\) and \(N_r\) antennas respectively. We want to estimate the downlink (DL) narrowband mmWave MIMO channel matrix \(H \in \mathbb{C}^{N_r \times N_t}\) from one-bit quantized received pilot signals \(Y\). Denote the hybrid precoder by \(F \in \mathbb{C}^{N_t \times N_p}\), and the hybrid combiner by \(W \in \mathbb{C}^{N_r \times N_p}\), with \(N_p\) being the number of data streams that can be transmitted. Consequently, the transmitted pilot symbols \(S \in \mathbb{C}^{N_r \times N_p}\) are received as

\[
Y = Q_1(W^H F H F^H S + W^H N),
\]

where each element of \(N \in \mathbb{C}^{N_r \times N_p}\) is an independent and identically distributed (i.i.d.) complex Gaussian random variable with mean 0 and variance \(\sigma^2\). The operator \(Q_1\) represents one-bit quantization and mathematically is given by

\[
Q_1(\cdot) = \text{sign}(\text{Re}(\cdot)) + j \cdot \text{sign}(\text{Im}(\cdot)).
\]

Note that while \(Y \in \{\pm 1\}^{N_r \times N_p}\), we want to recover a “full-resolution” estimate of \(H\). Implicitly assumed in (1) is a block fading model over \(\geq N_p\) pilot symbols, i.e., a new i.i.d. channel realization \(H\) is chosen almost every \(N_p\) time slots. We assume a fully connected phase shifting network [1], and constrain the angles realized by the phase shifters to quantized sets [11] given by

\[
A = \left\{ 0, \frac{2\pi}{2^N}, \ldots, \frac{(2^N - 1)2\pi}{2^N} \right\},
\]

The authors are with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712 USA (e-mail: akashdoshi@utexas.edu; jandrews@ece.utexas.edu).

Digital Object Identifier 10.1109/LWC.2023.3283926

Manuscript received 16 May 2023; accepted 5 June 2023. Date of publication 7 June 2023; date of current version 8 September 2023. This work was supported in part by NVIDIA; in part by Qualcomm Innovation Fellowship (QIF); and in part by NSF under Grant CNS-2148141 and Grant CCF-2008710. The associate editor coordinating the review of this article and approving it for publication was G. Zheng. (Corresponding author: Akash Doshi.)

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
where \( N_Q \) is the number of quantization bits. We assume \( N_{bit,t} \) and \( N_{bit,r} \) phase shift quantization bits at the transmitter and receiver respectively, with the quantization sets denoted by \( A_t \) and \( A_r \) respectively. This implies \( [F]_{i,j} = \frac{1}{\sqrt{N_t}} e^{j \psi_{i,j}} \) and \( [W]_{i,j} = \frac{1}{\sqrt{N_r}} e^{j \phi_{i,j}} \) where \( \psi_{i,j} \in A_t \) and \( \phi_{i,j} \in A_r \). Vectorizing (1) and utilizing the Kronecker product identity \( ABC = (C^T \otimes A)B \), we obtain

\[
y = Q_1((S^T F^T \otimes W^H)H + (I_{N_p} \otimes W^H)u), \tag{4}
\]

where \( y \in \{\pm 1\}^{N_N \times N_1 \times 1} \), \( H \in \mathbb{C}^{N_N \times N_1 \times 1} \) and \( u \in \mathbb{C}^{N_N \times N_1 \times 1} \). Since the received signal is 1-bit quantized, channel estimation in the noiseless, full-rank \((N_N N_p = N_t N_r)\) setting is also an ill-posed inverse problem. The technique presented in this letter will provide for channel estimation from noisy one-bit pilot measurements with \( N_p N_q < N_t N_r \).

### III. Quantized Generative Channel Estimation

Deep generative models \( G \) are feed-forward neural networks (NN) that take as input a low dimensional vector \( z \in \mathbb{R}^d \) and output high dimensional matrices \( G(z) \in \mathbb{R}^{c \times l \times w} \), where \( c, l \) and \( w \) refer to the number of channels, length and width of an image outputted by \( G \) and \( d \ll c l w \). Such a model can be trained via Generative Adversarial Networks [12] to take a i.i.d. Gaussian vector \( z \) as input and produce samples from complicated distributions, such as human faces [13].

In [9], we developed an algorithm – Generative Channel Estimation (GCE) – that utilized compressed sensing using deep generative models [14] to perform MIMO channel estimation. We trained \( G \) to output channel realizations \( H \) from a given distribution, and then utilized \( G \) to recover \( H \) from compressive pilot measurements \( y \). However, we experimentally demonstrated that a reduced \( \lambda_c \) antenna spacing of \( \lambda_c / 10 \) in the antenna arrays at the transmitter and receiver was key to training \( G \) successfully. We attributed this to the high spatial correlation generated in channel realizations by such an antenna spacing, making it easier for \( G \) to learn the underlying channel distribution.

In this letter, we will utilize the following key insight to output channel realizations with the conventional and realistic \( \lambda_c / 2 \) antenna spacing: beamspace representation of mmWave MIMO channels have high spatial correlation due to clustering in the angular domain. To be precise, assuming uniformly spaced linear arrays at the transmitter and receiver, the array response matrices are given by the unitary DFT matrices \( A_T \in \mathbb{C}^{N_t \times N_t} \) and \( A_R \in \mathbb{C}^{N_r \times N_r} \) respectively. Then, we can represent \( H \) as

\[
H = A_R H_v A_T^H
\]

\[
H = ((A_T^H)^T \otimes A_R) H_v. \tag{5}
\]

We will train \( G \) to output samples of \( H_v \), i.e., \( \mathbb{P}_G \) converges to \( \mathbb{P}_{H_v} \) as GAN training progresses.

Subsequently, in order to adapt GCE to quantized channel estimation, we propose a new empirical optimization objective, drawing inspiration from the loss function proposed in [15] for robust one-bit recovery using deep generative networks. Given a trained generator \( G \) and pilot measurements \( y \) as defined in (4), we will solve the following optimization problem

\[
z^* = \arg \max_{z \in \mathbb{R}^d} \sum_{i=1}^{N_p N_q} \Re(y[i]) \Re((A_{sp}[i], G(z)) + \sum_{i=1}^{N_p N_q} \Im(y[i]) \Im((A_{sp}[i], G(z)) \tag{6}
\]

where \( A_{sp} = (A_T^H FS)T \otimes W^H A_R \). This heuristically designed loss function attempts to maximize the correlation between \( y \) (which is constrained to a vector with entries \( \pm 1 \) for one-bit quantization) and \( A_{sp} G(z) \). The summation in (6) should be interpreted as the sum over the real and imaginary parts, separately.

\[
\sum_{i=1}^{N_p N_q} \Re(y[i]) \Re((A_{sp}[i], G(z)) + \sum_{i=1}^{N_p N_q} \Im(y[i]) \Im((A_{sp}[i], G(z)) \tag{7}
\]

The beamspace channel estimate is then given by \( H_{v,est} = G(z^*) \). The performance metric used to assess the quality of \( H_{v,est} \) is the normalized mean square error (NMSE), defined as

\[
NMSE = \mathbb{E} \left[ \frac{||H_v - \kappa H_{v,est}||^2}{||H_v||^2} \right], \tag{8}
\]

where \( \kappa = \text{argmin} ||H_v - \kappa H_{v,est}||^2 \) for a given \( H_v \) and \( H_{v,est} \).

In order to train \( G \), we utilize two different GAN architectures - (i) Wasserstein GAN with Gradient Penalty (WGAN-GP) and (ii) Conditional Wasserstein GAN (CWGAN). These have both been presented in [10] in the context of narrow-band beamspace MIMO channel generation. Due to space constraints, we will not be presenting them in this letter.

### IV. RESULTS & DISCUSSION

#### A. Data Generation & Preprocessing

Channel realizations have been generated using the 5G Toolbox in MATLAB in accordance with the 3GPP specifications TR 38.901 [16], consisting of an equal number of realizations of all categories of CDL channels, i.e., CDL-A,B,C (which are NLOS) and CDL-D,E (which are LOS). The channel simulation parameters are summarized in Table I. We assume a narrowband block fading model in this letter.

The generator output \( G(z) \) and discriminator input are of size \((2, N_t, N_r)\), where the first dimension allows us to stack the real and imaginary parts. Based on empirical
evidence that a GAN is unable to learn mean-shifted distributions [17], it is important to normalize the data used to train a GAN. Given a beamspace channel realization \( \mathbf{H}_v \), \( \mu_{i,j} = \mathbb{E}[\mathbf{H}_v[i,j]] \), \( \text{Re}(\sigma_{i,j}) = \mathbb{E}[(\text{Re}((\mathbf{H}_v[i,j] - \mu_{i,j}))^2)]^{0.5} \) and \( \text{Im}(\sigma_{i,j}) = \mathbb{E}[(\text{Im}((\mathbf{H}_v[i,j] - \mu_{i,j}))^2)]^{0.5} \), we normalize the matrix element-wise as

\[
\text{Re}(\mathbf{H}_v[i,j]) \leftarrow \frac{\text{Re}(\mathbf{H}_v[i,j] - \mu_{i,j})}{\sigma_{i,j}} \\
\text{Im}(\mathbf{H}_v[i,j]) \leftarrow \frac{\text{Im}(\mathbf{H}_v[i,j] - \mu_{i,j})}{\sigma_{i,j}}
\]

(9)

In lieu of (9), the operations \( \mathbf{G}(\cdot) \) and \( \mathbf{D}(\cdot, \chi) \) will implicitly be used to denote \( \text{SN}^{-1}(\mathbf{G}(\cdot)) \) and \( \text{SN}(\cdot, \chi) \) respectively throughout this letter without exception. Here \( \text{SN}(\cdot, \chi) \) denotes the operation of stacking the real and imaginary part followed by normalization using \( \{\mu_{i,j}, \sigma_{i,j}\}_{i,j} \) and then \( \text{SN}^{-1}(\cdot) \) corresponds to unnormalization followed by unstacking to generate a complex-valued output.

**B. Neural Network Architectures & Training Hyperparameters**

The generator and discriminator employed in the Wasserstein GAN are Deep Convolutional NNs. While the discriminator architecture was adopted from [18], the generator was fine-tuned to improve its ability to learn the underlying probability distribution. The generator \( \mathbf{G} \) takes an input \( z \in \mathbb{R}^d \), passes it through a dense layer with output size \( 128N_tN_r/16 \), and reshapes it to an output size of \( (N_t/4, N_r/4, 128) \). This latent representation is then passed through \( k = 2 \) layers, each consisting of the following units: \( 2 \times 2 \) upsampling, 2D Convolution with a kernel size of \( 4 \) and Batch Normalization. All BatchNorm2D layers have momentum = 0.8 [18] and Conv2D layers have bias = False. We utilize \( d = 65 \) for \( z \in \mathbb{R}^d \) (refer [10, Appendix B] for an empirical justification).

In order to extend the generator and critic architectures to the conditional setting, we employ an Embedding(2, 10) layer in both. This layer learns a 10-dimensional embedding for \( \chi = 0 \) and \( \chi = 1 \). Subsequently, \( \mathbf{G} \) passes this embedding through Linear(10, \( N_tN_r/16 \)) and Reshape(1, \( N_t/4, N_r/4 \)) before concatenating it to \( \text{Linear}(\varphi) \) of size \( (127, N_t/4, N_r/4) \).

For performing QGCE, we utilize an Adam [19] optimizer with a step size \( \eta = 0.1 \) and iteration count 500. We will utilize \( N_s = 16, N_p = 25, N_{\text{bit}} = 6 \) and \( N_{\text{bit},e} = 2 \). Note that \( N_sN_p < N_tN_r \), hence the sensing matrix \( \mathbf{A}_\text{sp} \) is not full rank, and channel estimation is an ill-posed inverse problem even in the absence of quantization. The computational complexity of the algorithm is \( O(N_tN_r^2) \) [9].

**C. Baselines**

1) **BG-GAMP**: We utilize the Generalized Approximate Message Passing (GAMP) algorithm proposed in [3], [4] as a compressed sensing baseline for 1-bit quantized channel estimation. Specifically, [4] models the angular domain coefficients of the signal to be recovered - in this case, the beamspace channel - as a Bernoulli-Gaussian (BG) mixture random variable and uses AMP to compute approximately the MMSE estimates of the channel coefficients. We also tune the sparsity hyperparameter in accordance with the approximate beamspace sparsity of each channel model. To be precise, [4] defines the channel sparsity rate \( 1 - \lambda_0 \) as the ratio of the number of non-zero elements in \( \mathbf{H}_v \) and \( N_tN_r \). Based on the beamspace CDL channel representations, we use the following estimates for \( (1 - \lambda_0)N_tN_r \) while implementing BG-GAMP: CDL-A (20), CDL-B,C (50), CDL-D,E (5).

2) **GCE**: Given noisy un-quantized pilot measurements \( y \) and \( \lambda_{\text{reg}} = 0.001 \), GCE [10] recovers the channel estimate \( \mathbf{H}_{v,\text{est}} = \mathbf{G}(z^*) \), where \( z^* \) is given by

\[
z^* = \arg \min_{z \in \mathbb{R}^d} || y - \mathbf{A}_{\text{sp}}\mathbf{G}(z) ||_2^2 + \lambda_{\text{reg}} || z ||_2^2.
\]

**D. Results**

In accordance with the WGAN-GP model presented in [10, Sec. IV-A], we design a separate generator for each of the five CDL channel models by training a WGAN-GP (\( 1_{\text{GP}} = 1 \)) for 60,000 training iterations. We also design a single conditional generative model by training a CWGAN (\( 1_{\text{GP}} = 0 \)) as outlined in [10, Sec. IV-B] for 100,000 training iterations. In both cases, we extract the final trained generator \( \mathbf{G} \) and perform QGCE at varying SNR to plot NMSE vs SNR, as shown in Fig. 1.

Clearly, the individually trained WGAN-GP outperforms BG-GAMP, by 0.5 dB in CDL-B and C, 1 dB in CDL-A and 5 dB for CDL-E and D. We also observe that across CDL channel models, the performance of QGCE is consistently \( B < C < A < E < D \). This is in agreement with the decreasing number of rays/clusters and the increasing magnitude of the LOS component in \( \mathbf{H}_v \) as we go from left to right (refer [16, Table 7.7.1] for the precise channel profiles). It is important to note that the NMSEs obtained using QGCE, going as low as \( -7.7 \) dB for CDL-D and CDL-E, have been obtained using only a fraction \( -\alpha = N_pN_s/N_tN_r = 0.4 \) of the pilot symbols that would have been required for full-rank channel estimation in the absence of quantization. For comparison, one-bit quantized channel estimation algorithms in [4] and [6] use \( \alpha = 8 \) to perform GAMP-based CE. Aside from the excessive training overhead, such methods also implicitly assume that the channel does not change over a large number of pilot symbols, rendering them inapplicable in the presence of UE mobility.

As expected, we see a degradation in NMSE compared to full-resolution GCE, ranging from \( \sim 8 \) dB for CDL-A and D to 2 dB for CDL-B at an SNR of 15 dB. However, it is interesting to observe that QGCE outperforms GCE at SNR \( \leq 0 \) dB for CDL-D and E. This suggests that the QGCE optimization in (6) is more robust to noise than the GCE optimization in (11), and a weighted objective combining (6) and (11) could be used to improve the performance of full-resolution channel estimation at low SNRs.

On switching to a single CWGAN model, we observe that the gain in NMSE over BG-GAMP is reduced. The LOS channel models CDL-D and E still outperform BG-GAMP, but the

---

1Code available at https://github.com/akashsdoshi96/obq-gan-mimo-ce.
NLOS channel models only achieve performance competitive with BG-GAMP. The performance degradation in CWGAN as compared to the individually trained WGAN-GP generative models can be attributed to the usage of a simple LOS/NLOS label to distinguish between the different channel modalities as well as the inability of the convolutional architecture of the generator $G$ to learn the NLOS channel models that have “richer” beamspace representations.

At the same time, it should be noted that the BG-GAMP baseline was adapted to each CDL channel model by a careful tuning of the sparsity rate for CDL A-E. Such sparsity rates cannot be obtained in practice from pilot measurements alone. We initially considered the usage of EM-BG-GAMP as described in [4], where the Expectation Maximization – EM – step would be responsible for the automated tuning of the sparsity and noise variance estimates, however we were unable to obtain any reasonable NMSE for one-bit quantized pilot measurements, even with a higher value of $N_p$. A possible reason for this could be that [4] only tested channels with a small number of multi-path clusters ($\leq 4$) in their geometric channel model representation, while the CDL channel models contain up to 23 clusters.

To compute the NMSE in (8), note that we utilize a seemingly genie-aided scaling factor $\kappa$, since both the one-bit quantized pilot measurements $y$ as well as the correlation-based optimization objective in (6) do not provide for optimal scaling of the reconstructed channel. In order to verify that the channel estimate $H_{est} = A_R G(z^*) A_H^T$ from WGAN-GP based QGCE is in fact “better” than the estimate obtained from BG-GAMP, we perform a simple achievable rate computation.

---

**Fig. 1.** NMSE vs SNR for WGAN-GP GCE & QGCE and CWGAN QGCE. For reference, BG-GAMP baselines have also been plotted.

**Fig. 2.** Spectral Efficiency vs SNR for WGAN-GP QGCE. For reference, BG-GAMP and Perfect CSI baselines have also been plotted.
Utilizing the SVD of \( H_{\text{est}} = U \Sigma V^H \), we set the optimal precoding and combining vector as the first column of \( V \) and \( U \) respectively. The spectral efficiency is then given by

\[
C(\text{SNR}; H) = \log_2 (1 + \text{SNR} \|U[1]^H HV[1]\|^2)
\]  

(12)

The spectral efficiency is plotted as a function of SNR in Fig. 2. One can observe the clear correspondence with the NMSE in Fig. 1. For e.g., \( C(\text{SNR}) \) for CDL-D and E almost matches the Perfect CSI curve, unlike CDL-A, B and C, since CDL-D and E achieve NMSEs as low as -8 dB. Similarly, \( C(\text{SNR}) \) for CDL-D and C has the least improvement over BG-GAMP, which is again consistent with the small \(~0.5\) dB improvement in NMSE over BG-GAMP in Fig. 1.

V. CONCLUSION AND FUTURE DIRECTIONS

Channel estimation in mmWave MIMO using one-bit quantized pilot measurements typically requires a large number of pilot measurements \( (N_p > N_t N_r / N_s) \) in order to recover a channel estimate with low NMSE. In this letter, we demonstrate how a deep generative prior \( G \), trained using Wasserstein GAN, can be used to perform channel estimation from a limited number of pilot measurements \( (N_p < N_t N_r / N_s) \) by optimizing the input vector \( z \) to a deep generative model \( G \) with the objective of maximizing the correlation between the quantized pilot measurements \( y \) and the estimated transmit signal \( A_{\text{opt}} G(z) \). Our results indicate that a carefully tuned generative prior significantly outperforms state-of-the-art baselines such as BG-GAMP, while a single conditional generative model outperforms BG-GAMP on LOS channel models and achieves competitive results on NLOS channel models.

A key shortcoming of our approach is the need for clean channel realizations to train the WGAN. While techniques such as Ambient GAN [20] can be used to train WGAN from noisy un-quantized pilot measurements, the usage of one-bit ADCs destroys the invertibility of the function mapping the probability density \( p_{H_v}(H_v) \) to \( p_y(y) \), rendering Ambient GAN inapplicable. Hence training a GAN using noisy quantized pilot measurements should be investigated. Additionally, instead of a simple LOS/NLOS label, we could train \( N \) GANs – for example \( N = 3 \) for low, medium and high levels of beamspace sparsity – and then learn a classifier that will indicate which generative model to utilize.

ACKNOWLEDGMENT

Code and data will be made publicly available at https://github.com/akashsdoshi96/obq-gan-mimo-ce.

REFERENCES

[1] R. Méndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, and R. W. Heath, “Hybrid MIMO architectures for millimeter wave communications: Phase switches or switches?” IEEE Access, vol. 4, pp. 247–267, 2016.

[2] P. Sun, Z. Wang, and P. Schniter, “Joint channel-estimation and equalization of single-carrier systems via bilinear AMP,” IEEE Trans. Signal Process., vol. 66, no. 10, pp. 2772–2785, May 2018.

[3] J. Mo, P. Schniter, N. G. Prelcic, and R. W. Heath, “Channel estimation in millimeter wave MIMO systems with one-bit quantization,” in Proc. Asilomar SSC, Nov. 2014, pp. 957–961.

[4] J. Mo, P. Schniter, and R. W. Heath, “Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs,” IEEE Trans. Signal Process., vol. 66, no. 5, pp. 1141–1154, Mar. 2018.

[5] P. Sun, Z. Wang, R. W. Heather, and P. Schniter, “Joint channel-estimation/decoding with frequency-selective channels and few-bit ADCs,” IEEE Trans. Signal Process., vol. 67, no. 4, pp. 899–914, Feb. 2019.

[6] N. J. Myers, K. N. Tran, and R. W. Heath, “Low-rank mmWave MIMO channel estimation in one-bit receivers,” in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), May 2020, pp. 5005–5009.

[7] D. H. N. Nguyen, “Neural network-optimized channel estimator and training signal design for MIMO systems with few-bit ADCs,” IEEE Signal Process. Lett., vol. 27, pp. 1370–1374, 2020.

[8] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, “One-bit massive MIMO: Channel estimation and high-order modulations,” in Proc. IEEE Int. Conf. Commun. Workshop (ICCW), Jun. 2015, pp. 1304–1309.

[9] E. Talevi, A. Doshi, A. Jalal, A. Dimakis, and J. G. Andrews, “High dimensional channel estimation using deep generative networks,” IEEE J. Sel. Areas Commun., vol. 39, no. 1, pp. 18–30, Jan. 2021.

[10] A. Doshi, M. Gupta, and J. G. Andrews, “Over-the-air design of GAN training for mmWave MIMO channel estimation,” IEEE J. Sel. Areas Inf. Theory, vol. 3, no. 3, pp. 557–573, Sep. 2022.

[11] K. Venugopal, A. Alkhateeb, N. G. Prelcic, and R. W. Heath, “Channel estimation for hybrid architecture-based wideband millimeter wave systems,” IEEE J. Sel. Areas Commun., vol. 35, no. 9, pp. 1996–2009, Sep. 2017.

[12] I. Goodfellow et al., “Generative adversarial nets,” in Proc. Adv. NIPS, Dec. 2014, pp. 2672–2680.

[13] A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” in Proc. ICLR, Nov. 2015, pp. 1–16.

[14] A. Bora, A. Jalal, E. Price, and A. G. Dimakis, “Compressed sensing using generative models,” in Proc. Int. Conf. Mach. Learn. (ICML), Aug. 2017, pp. 537–546.

[15] S. Qiu, X. Wei, and Z. Qiu, “Robust one-bit recovery via ReLU generative networks: Improved statistical rates and global landscape analysis,” in Proc. NeurIPS Deep Inverse Workshop, Dec. 2019, pp. 1–10.

[16] “Study on channel model for frequencies from 0.5 to 100 GHz, version 16.1.0,” 3GPP, Sophia Antipolis, France, Rep. (TR) 38.901, Jan. 2020.

[17] A. Srivastava, L. Valkov, C. Russell, M. U. Gutmann, and C. Sutton, “VeeGAN: Reducing mode collapse in GANs using implicit variational learning,” in Proc. Adv. NIPS, Dec. 2017, pp. 3308–3318.

[18] M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein generative adversarial networks,” in Proc. Int. Conf. Mach. Learn. (ICML), Dec. 2017, pp. 214–223.

[19] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in Proc. ICLR, Dec. 2014, pp. 1–15.

[20] A. Bora, E. Price, and A. G. Dimakis, “AmbientGAN: Generative models from lossy measurements,” in Proc. ICLR, Feb. 2018, pp. 1–22.