

Deep UL2DL: Channel Knowledge Transfer from Uplink to Downlink

Mohammad Sadegh Safari and Vahid Pourahmadi

Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran

Abstract

Knowledge of the channel state information (CSI) at the transmitter side is one of the primary sources of information that can be used for efficient allocation of wireless resources. Obtaining Down-Link (DL) CSI in FDD systems from Up-Link (UL) CSI is not as straightforward as TDD systems, and so usually users feedback the DL-CSI to the transmitter. To remove the need for feedback (and thus having less signaling overhead), several methods have been studied to estimate DL-CSI from UL-CSI. In this paper, we propose a scheme to infer DL-CSI by observing UL-CSI in which we use two recent deep neural network structures: a) Convolutional Neural network and b) Generative Adversarial Networks. The proposed deep network structures are first learning a latent model of the environment from the training data. Then, the resulted latent model is used to predict the DL-CSI from the UL-CSI. We have simulated the proposed scheme and evaluated its performance in a few network settings.

Index Terms

Deep Learning, Uplink, Downlink, Channel Prediction, FDD Systems, Generative Adversarial Networks, Convolutional Neural Networks.

I. INTRODUCTION

One key feature of new generation of cellular networks is their efficient use of frequency band and energy. To achieve this goal, they use various techniques such as water-filling, appropriate precoding and beamforming. In most of these techniques, the Channel State Information (CSI) should be available at the transmitter side (CSIT). In TDD systems, Up-Link (UL) and Down-Link (DL) frequencies are equal, so we can use channel reciprocity and simply infer DL channel by observing UL channel. In Frequency Division Duplexing (FDD) systems, however,

DL channel and UL channel have different frequencies so we can't use channel reciprocity to infer DL channel. The most commonly used solution is that the user (receiver) first measures (estimates) the DL channel, and then sends its information back to the transmitter. This solution has two major disadvantages: delay and overhead. If the delay is greater than coherence time, the actual DL channel is different from what has been feedbacked by the user. In addition, in new generations of mobile networks the transmitter has a large number of antennas. For example, for a fourth-generation transmitter with 64 antennas, the need to learn DL channel (pilot transmission and feedback data) consumes large portion of the transmitter's traffic [1]. This overhead is very large and is one major challenges of LTE networks [1]. These challenges have such important affect on the network performance that despite some significant advantages of FDD system such as continuous transmission [2], in recent years, Time Division Duplexing (TDD) has attract more attentions.

To eliminate the need for the feedback (and so its associated overhead and delay), there are several studies that aim to infer DL channel by observing UL channel in FDD systems. DL-CSI estimation methods in [3]–[5] are based on the assumption that the difference between dominant angle of arrival (AOA) in UL and dominant angle of departure (AOD) in DL is small and directional properties of UL and DL are correlated. For example, [6] through a great deal of measurements shows that with probability of about 81 percent, this difference is smaller than 4.5 degrees. Therefore, by having dominant AOA in UL, the dominant AOD in DL can be obtained and used for purposes like beamforming.

References [7]–[10] are based on covariance matrix due to channel matrix slow variations. In [7] a transformation matrix is used to convert UL covariance matrix to DL covariance matrix. [9] is based on the concept of dictionary learning and it has two phases: training and exploitation. In training phase, they make a dictionary with corresponding DL and UL covariance matrices (by changing user location, they have constructed the dataset of different input and output pairs). In exploitation phase, by observing UL covariance matrix, DL covariance matrix is constructed by interpolation of stored dictionary with various methods.

In [11]–[13] by taking into account multipath structure of the channel, they extract paths of signal independent of frequency and hence can infer channel response in any desired frequency band. For example, in [11], the authors consider 4 parameters for every path (path attenuation, path length, an independent phase shift for modeling reflection and angle of arrival of path) and then they have tried to estimate these parameters using the UL-CSI. The resulted model, is then

used for prediction of the DL-CSI.

AOA-Based methods are often not usable in cases where accurate response of channel is required and often used only for beamforming. In path extraction-based methods, we can obtain the accurate channel response at any desired frequency, but they often based on assumptions that may not practically possible. For example, path attenuation is considered independent of frequency [11]. This assumption is only true if difference between DL and UL frequencies is small. In [13], to consider path attenuation dependency on frequency, limited feedback is used and verified that deriving DL channel, with the assumption of frequency independent path attenuation, is not very accurate. Covariance-based algorithms also depend on different environmental factors such as correlation between antennas. When antenna correlation is poor, [14] showed that it is not appropriate to use correlation based methods.

In recent years, artificial intelligence has revolutionized human life, so that some called it the fourth industrial revolution. One of the leading areas of artificial intelligence and machine learning is deep learning which has been very successful in many cases such as machine vision, speech recognition and object detection and in some cases even exceeded human performance [15]. Deep Learning have been also used in physical layer communications [16]–[21]. [16] considered communication system at physical layer as an autoencoder and designed an end-to-end system that optimizes transmitter and receiver simultaneously in one process. [17] used a variational GAN to capture stochastic model of channel and learns its probability density function (PDF). In [18] authors used an adversarial network to model the channel input-output conditional probability. In [20], a super resolution network cascaded with a denoising autoencoder used to estimate channel response based on some known pilots. CsiNet which introduced in [21], to perform limited CSI feedback in FDD systems, encodes channel response at one side (user) and decodes received feedback with a decoder in other side (base station).

Motivated by such applications, in this paper, we propose a method based on deep neural networks that predict the DL-CSI based on the past UL-CSI measurements. Use of the deep networks enables us to expand the search space of the environment propagation model (beyond the the current mathematical models) and therefore it can capture more on how UL-CSI should be transferred to DL-CSI.

In essence, the core idea of our scheme is that the way that the channel affects the transmitted signal (regardless that it is UL or DL) is related to the structure of the environment the signal is propagating in (e.g., the objects which are in the environment, the materials that they are

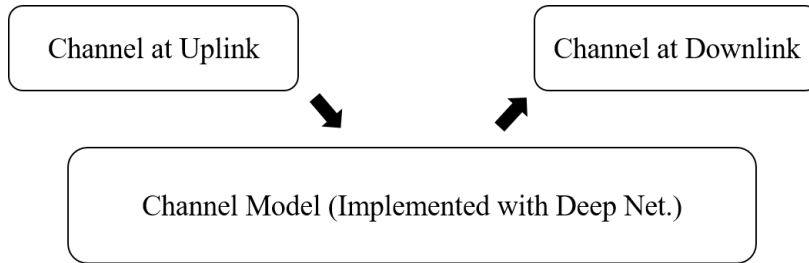


Fig. 1: UL to DL knowledge transfer procedure: extracting environment information from UL and transferring it to DL

made up, their shape). By knowing this fact that both DL and UL channels share the same propagating environment (assuming of course no sudden change in the environment), we use data-driven approach to extract environment information from UL channel response to a latent domain and then transfer derived environment information from the latent domain to the DL channel, Fig. 1.

To achieve this goal we use two types of Deep Networks: Convolutional Neural Networks (CNNs) and a specific type of Generative Adversarial Networks (GANs) called Boundary equilibrium GAN (BEGAN) which is based on Mean Square Error (MSE). For training and testing we use simulated data of Extended Vehicular-A (EVA) and Extended Typical Urban (ETU) models. Our results verifies the effectiveness of our schemes.

It worth also mentioning that if we want to fully characterize a DL channel of a multiple-input multiple-output (MIMO) system, we should characterize a 4-dimensional space, i.e., we should find out what is the channel effect (both on amplitude and phase of the signal) between 1) each transmit antenna and 2) each receive antenna, for 3) each of the subcarriers in our frequency range and 4) for each time slot. In most of the previous studies the prediction of DL matrix is investigated in terms of the MIMO channel matrix and their aim was not to determine the channel effect in time-frequency domain. In this work, instead of looking high level at the transmitter and receiver antennas and give one value for each pair, we focus on one transmit-receive antenna pair and predict DL channel over a block of time and frequency. The results can be further extended to the MIMO case but we will not discuss that in this paper.

The rest of this paper is organized as follows. In section II, we will describe CNNs and GANs as tools we used to predict DL channel. Section III provides detailed discussion on prediction problem. Section IV contains two approaches for solving DL prediction problem: direct approach

and generative approach. In section V, we explain implementation details of networks. Simulation results also are provided. Section VI draws a conclusion on this paper.

II. BACKGROUND

In following subsections we briefly discuss about two special types of deep neural networks which we used in this paper to predict DL channel.

A. Convolutional neural networks

One of the interesting neural network structures used widely in artificial intelligence (AI) community is Convolutional Neural Network (CNN). CNNs could have many hidden layers and usually they are one of the three type of convolutional layers, pooling layers and fully connected layers. CNN is a powerful tool specially in analyzing 2D data like images. It is mainly due to the structure of the convolutional layer which computes the output by convolving filter (kernel) weights with the input image (data). Value of each point in the output image is equal to cross-correlation of filter and corresponding area in the input image [22]. After applying convolution operation on the input data, an activation function will applied. Output of the activation function will pass through the next layer as the input.

B. Generative adversarial networks

Generative adversarial networks (GANs) are One of the most powerful generative models that captures data distribution [23]. They are based on game theory and consist of two networks: generator and discriminator which are trained simultaneously. Considering a noise vector z as the input, (typically with normal distribution) generator tries to create images similar to real ones while discriminator tries to distinguish generated images from real ones. Training of GANs is a two-player mini-max game. Generator tries to maximize error probability of discriminator (this means discriminator assigns high probability of being real to generated images) while discriminator tries to minimize probability of being real to generated images. GANs are hard to train and non-convergence or instability is their main problem [24].

Many different structures have been recently proposed for new implementation of GANs. In the original GAN, output of discriminator is a positive number between 0 and 1. This number represent the probability that the discriminator input image is in-fact a real image (not a generated image). Such discriminator is the most common type of discriminator network in GANs literature.

For first time in Energy-based GAN (EBGAN) [25], an autoencoder used as the discriminator. In EBGAN, the discriminator objective is to maximize reconstruction error of generated images while minimizing it for real ones. EBGAN generator's structure is similar to decoder part of discriminator. Using an autoencoder as discriminator makes training easier, faster and more stable. Boundary equilibrium GANs (BEGANs) [26] are improved version of EBGANs and use same structure but BEGANs aim to match autoencoder loss distribution instead of matching data distribution directly. To training such networks an equilibrium is

$$\gamma \times E[L(x)] = E[L(G(z))] , \quad (1)$$

where $L(x)$ and $L(G(z))$ are the autoencoder reconstruction loss when it gets a real image and a generated image, respectively. $E[\cdot]$ represents the expectation operation. In (1), $\gamma \in [0, 1]$ has an inverse relation with the diversity of the generated images, meaning that if γ is set to a larger number, the generator creates less diverse images.

In BEGAN, L_D and L_G denote the discriminator and generator loss or objective functions respectively, and they defined as

$$\begin{aligned} L_D &= L(x) - k_t \cdot L(G(z_D)) \\ L_G &= L(G(z_G)) \\ k_{t+1} &= k_t + \lambda_k (\gamma L(x) - L(G(z_G))) \end{aligned} , \quad (2)$$

where L_D is difference between the reconstruction loss of real images and the reconstruction loss of generated images which is scaled with parameter k_t . k_t introduced to maintain (1) and based on proportional control theory, [26] suggests that k_t should be updated using last equation in (2) and λ_k is its learning rate. Visual inspection is typically the only way to determine convergence in GANs but [26] also defines a convergence measure as

$$M_{Global} = L(x) + |\gamma L(x) - L(G(z_G))| . \quad (3)$$

In (3), smaller M_{Global} is desired as it means both small reconstruction loss for real images and maintaining (1).

III. PROBLEM DEFINITION AND FORMULATION

Suppose a base-station (transmitter) and a user (receiver) in the network. To increase the network spectral efficiency (by techniques like water-filling and beamforming in the case of multiple antennas), the base-station needs to know DL-CSI. When a user sends its data on UL

channel, for example if it uses OFDM method, it allocates some of its subcarriers and time-slots for pilots transmission. Using this pilots, the base-station can estimate UL-CSI, but assuming an FDD system UL and DL CSI reciprocity does not hold. So, one way to get that information is to first send pilot in downlink and after user estimate the DL-CSI, it sends DL-CSI over the feedback link. Such scheme leads to high over head in the system. Eliminating the feedback, we should find a way to derive DL-CSI using UL-CSI, which is only available information about the environment at the base-station.

To better describe the problem, consider a block of time-frequency between a pair of transmitter-receiver antennas, Fig. 2. Assuming a grid over this block, knowledge of the channel state information is equivalent to having information about what is the effect of channel (on both amplitude and phase of the transmitted signal) over each cell of the grid (i.e., we should know a complex value for each grid cell)

This block itself consists of two main portions of UL and DL: a) The first Δf_{UL} rows (subcarriers) of frame are assigned to UL and b) the next Δf_{DL} rows (subcarriers) are assigned to DL. Columns of our frame represent different time slots (how the channel effect changes over time). By considering this structure, we can say that the problem at hand is that we have the UL-CSI information over the Δf_{UL} subcarriers and Δt_{UL} time slots (part 1) and want to predict DL-CSI in Δf_{DL} subcarriers and Δt_{DL} next time slots (part 2). It worth mentioning that to make the model realistic and causal, we only use the past UL-CSI information for DL-CSI prediction (and not the UL-CSI that is measured at the same time-slots of DL-CSI).

To solve DL prediction problem, most of the previous studies are based on first considering a mathematical channel model for the environment. For example, the *multipath model* for channel modeling is defined as

$$h = \sum_{n=1}^N a_n e^{-j2\pi \frac{d_n}{\lambda_0} + j\phi_n}, \quad (4)$$

where h is the channel response over particular frequency of $\frac{1}{\lambda_0}$. In (4), it is assumed that there are N distinct paths in the environment where a_n is path attenuation and ϕ_n is a frequency-independent phase shift that captures reflection and attenuation of the signal along that path.

In a machine learning terminology, the common approach is that they first consider a parametric model for the environment and then use UL-CSI to estimate the parameters of the model. Having the resulted complete model, the DL-CSI can be predicted.

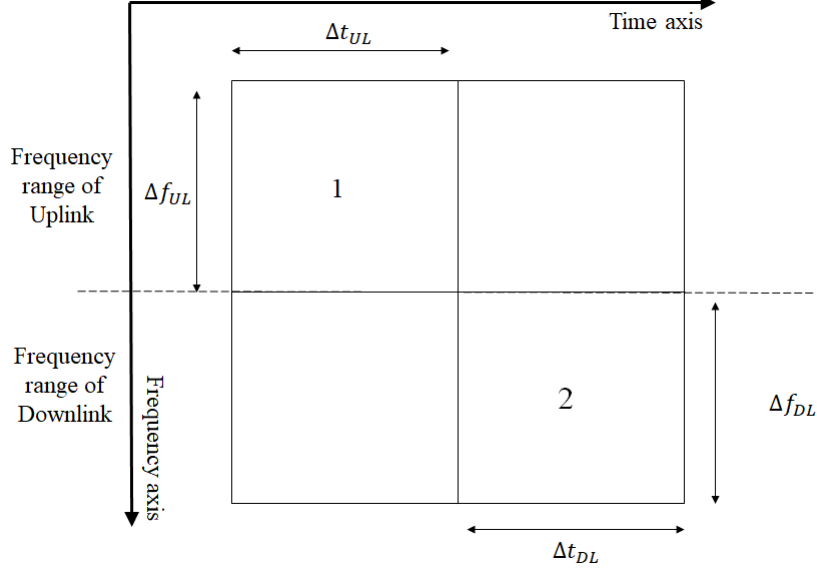


Fig. 2: UL-DL joint frame structure

Incorrect assumption about the parametric model and/or incorrect derivation of the parameters both lead to loss of some parts of UL-CSI information and consequently resulting low accuracy of DL-CSI prediction. Furthermore, when parametric models are estimated, often some simplifying assumptions should be considered that may not be true in some cases or even violated. For example, as mentioned in section I, in (4), a_n is assumed to be constant for UL and DL, but as some studies, [13], suggest that this assumption is not always correct.

To avoid forcing incorrect latent domain and problem of simplifying assumptions, in this work, we do not use any specific parametric model and instead use data-driven approaches to discover underlying structure of data without any prior model assumption. More details of the proposed scheme is presented in Section IV.

IV. PROPOSED SCHEME

In this section, we first explain how CSI information can be considered as an image, then we present the two approaches which are proposed for DL channel prediction.

A. CSI as an Image

Looking back at Fig. 2, CSI is a 2D complex matrix with size of $N_s \times N_t$, where N_s is the number of subcarriers and N_t is the number of time slots in the time-frequency block.

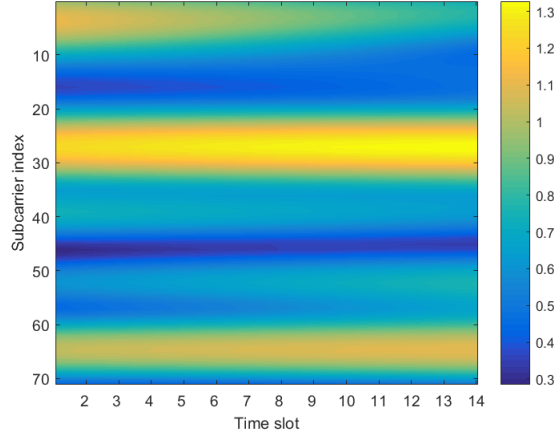


Fig. 3: CSI between a pair of antennas considered as an image

Recently many advanced techniques have been proposed for analyzing image data using neural networks. Image data are in fact 2D real matrices with one or more channels. For example the color images are 2D images with 3 channels (for red, green and blue components). To use image-based techniques in this work we have considered the 2D CSI matrix as an image. For example in Fig. 3, heat map of CSI absolute values, plotted for a sample FDD frame.

If we use only absolute values, phase information will be lost. So to solve this problem, we consider complex values CSI matrix as a real values matrix with 2 channels. There are two choices to put complex values as 2 channel of an image: put absolute values in first channel and phase values in second channel or put real values in first channel and imaginary values in second channel. We selected the second approach to prevent the problem of phase wrapping that may happen for the phase information.

So in the rest of this paper wherever that we use "image" term, it is refer to the CSI matrix that is considered as a $N_s \times N_t \times 2$ image with real values in the first channel and imaginary values in the second channel.

B. Up-Link to Down-Link Knowledge Transfer

In this paper, to extract environment information from UL and then transfer the derived knowledge back to DL domain, two approaches are introduced based on deep neural networks: direct approach and generative approach. In following subsections we explain each of this two approaches in details.

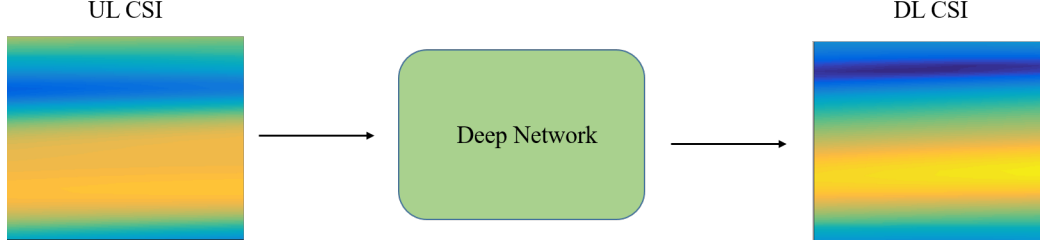


Fig. 4: Direct approach: Considering Fig.2, the network gets "part 1" of the image (UL-CSI) as input, and tries to predict "part 2" of the image (DL-CSI) as output. It is clear that we don't use information of other parts of image in this approach

1) **Direct Approach:** As mentioned before in UL to DL transfer, we have two steps: first encode environment information from UL-CSI to a latent domain, second transfer and decode derived latent domain model into DL-CSI. In direct approach we use a network to accomplish both of this two steps in a single process as one deep network (Fig. 4).

In this approach, we feed UL-CSI in Δf_{UL} subcarriers and past Δt_{UL} time slots (part 1 in Fig. 2) as the input and network tries to predict DL-CSI in Δf_{DL} subcarriers and next Δt_{DL} time slots (part 2 in Fig. 2) as the output.

As discussed in Section II, CNNs are one of the most successful tools in analyzing image data, so in this work we design an specific convolutional neural network to implement direct approach and using the designed model to predict DL CSI. The details of the network structure will be discussed in Section V.

2) **Generative Approach:** We still have similar desired input and output: considering Fig 2, *part 1* as input and *part 2* as output. The difference is that we do not directly learn the UL to DL relation, instead, we consider the whole time-frequency block (a matrix of size $N_s \times N_t$) as the image that we want to learn, i.e., we want to learn the joint distribution between different pixels of the complete CSI image. Knowing the joint distribution of the whole frame, we also know joint distribution of the UL section and DL section (Part 1 and part 2). Using UL-DL joint distribution, we can predict DL-CSI when we have the UL-CSI. It is clear that capturing such joint distribution is a very complicated task (considering the size of the CSI matrix).

As briefly described in Section II, researchers in the AI filed, recently proposed GAN structure as a very successful tool for estimating joint distribution of the input data, specially when we are dealing with images. After training a GAN with a set of images, it can generate images that

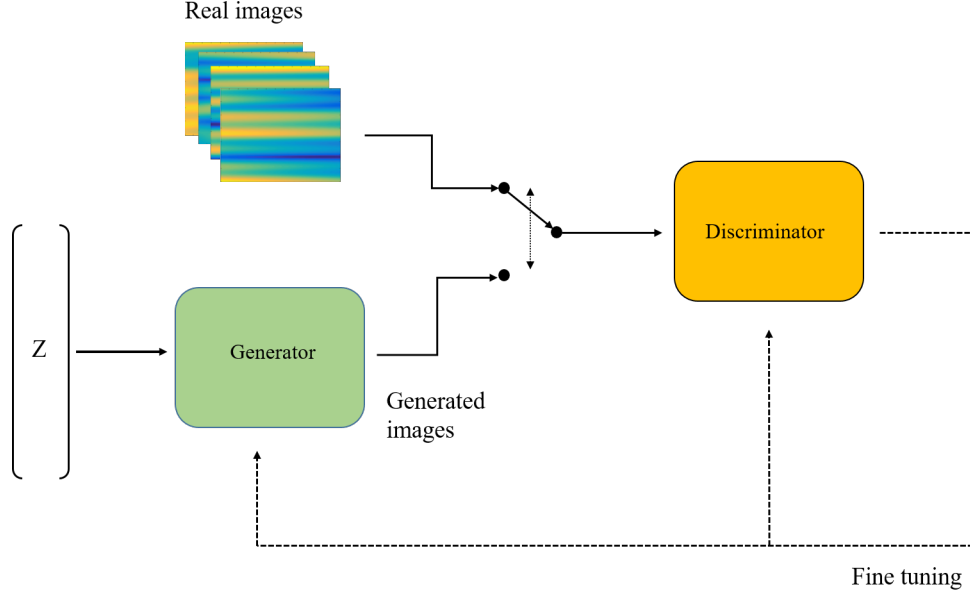


Fig. 5: Given a set of complete time-frequency CSI blocks, the generator learns how to create CSI images that are similar to the real CSI blocks

are very similar to the real images. *One interesting application of GANs* is in image completion, i.e., given a corrupted image (like when one part of the image is missing) the GAN tries to find the missing part. Several schemes have been proposed for image completion. The core idea of them is that first GAN tries to generate an image that resembles (w.r.t some metric) to the corrupted image and then use the generated image to predict the missing part.

Motivated by the success of GANs and image completion schemes (and since we are able to consider the CSI matrix as an image) we should be able to use similar network structure to find the joint distribution of the CSI and then use that model to predict DL-CSI from UL-CSI.

The steps of our proposed scheme can be summarized as:

- **Training Phase:** First train a GAN with CSI images of the whole time-frequency block. After complete learning, the generator network is capable of getting a random vector z as input and creating images which are very similar to real CSI images, Fig. 5.

It worth mentioning that from different types of GANs, we first selected the most common structure which called Deep Convolutional Generative Adversarial Network (DCGAN) [27]. Although DCGANs are able to produce similar images like our CSI images, during the completion phase, we didn't get desired results and MSE of prediction was relatively high.

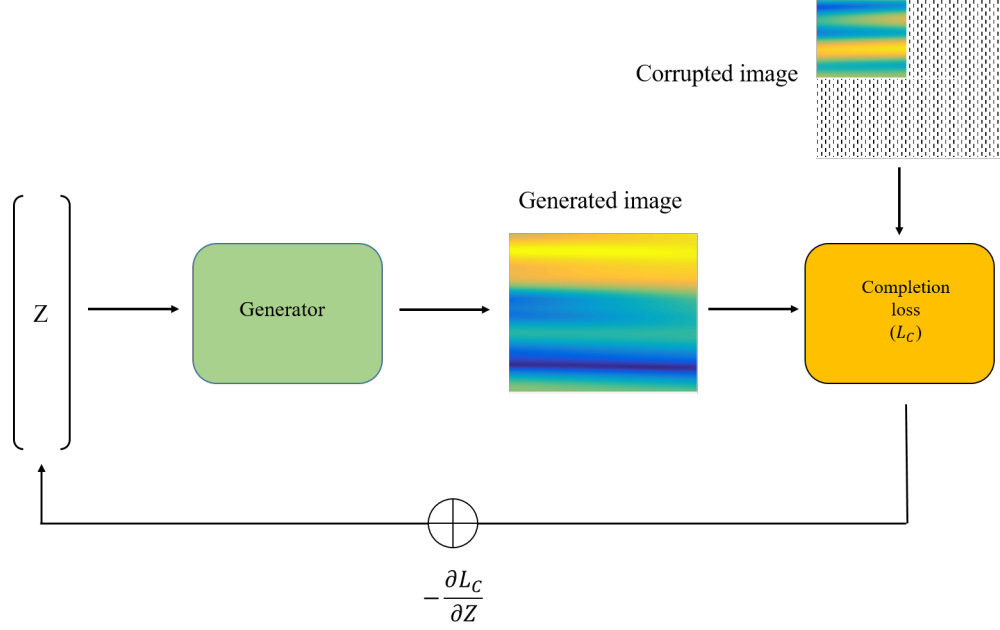


Fig. 6: Image completion procedure: Given a corrupted image (Only UL-CSI is known), the network generates appropriate image that can be used for DL-CSI prediction

Additionally, given a UL-CSI frame, with different initializations of the input vector z we got very different predictions of DL-CSI.

To solve this problem, in this work, we have used BEGAN (described in Section II-B). As discussed there BEGANs are designed based on the MSE error and have better convergence property.

- **Completion Phase:**

In this step we want to predict DL-CSI. The idea is that we consider *the time-frequency block that only has the UL-CSI as the corrupted image*, then we use different GAN-Based image completion algorithms to complete the missing part (DL-CSI). As we treat the prediction task as completing a corrupted image, we name the prediction phase, completion phase, Fig. 6.

More accurately, in completion phase, vector z (input of the generator) is initialized with a random state. Then we update z using gradient descent so that the generated image and the corrupted image are getting more similar (a loss function is reduced). After several iterations, the generated image is considered equal to complete real image and desired output (DL-CSI) will be derived. In this work we have used two different loss functions

and tried image completion using both methods.

a) **Contextual Loss**: Contextual loss is defined as distance between known part of image and its corresponding part in generated image. If we define mask as

$$\text{mask}[i, j] = \begin{cases} 1, & 0 \leq i \leq \Delta f_{UL}, 0 \leq j \leq \Delta t_{UL} \\ 0, & \text{otherwise} \end{cases}. \quad (5)$$

Then the contextual loss will be

$$\text{contextual loss} = \|\text{mask} \odot x - \text{mask} \odot G(z)\|, \quad (6)$$

where x is the image that we want to complete and $G(z)$ is the generated image.

b) **Contextual Loss + Perceptual Loss**: Reference [28] first used such loss function for image completion. If we use only contextual loss, final completed image may seem artificial (have different structure compared to real data) so one use discriminator loss of generated image as a new term in total loss and call that perceptual loss because it is a sense of being real. So in BEGAN

$$\text{perceptual loss} = D(G(z)), \quad (7)$$

and,

$$\text{total loss} = \text{contextual loss} + \lambda \times \text{perceptual loss}, \quad (8)$$

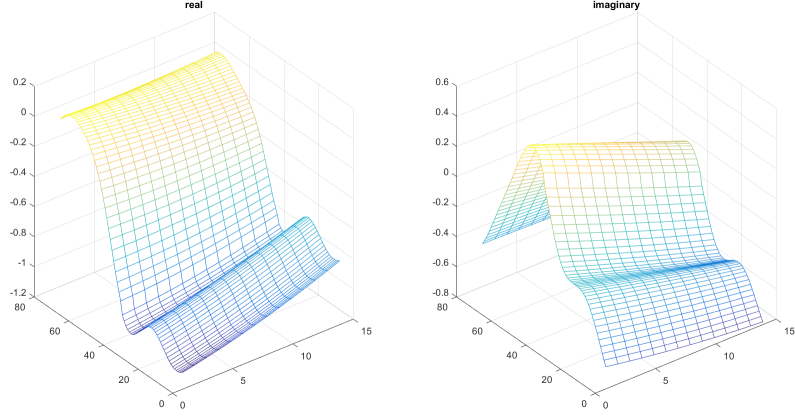
where λ is a hyper parameter to control how much emphasis is put on perceptual respect to contextual loss during gradient descent. Its default value in BEGAN is 0.01.

V. IMPLEMENTATION AND SIMULATION RESULTS

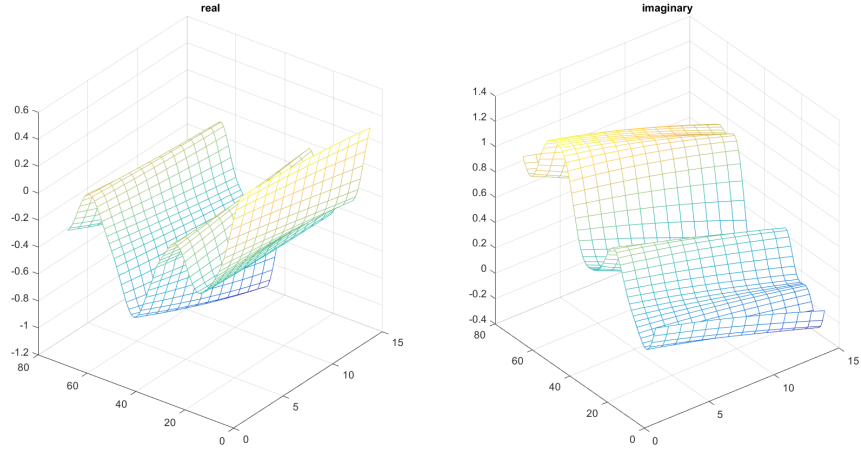
In the following, we discuss the details of the implementation and the simulation results. The source code of the implementation can be found at <https://github.com/safarisadegh/UL2DL>

A. Dataset Generation

To evaluate performance of our proposed schemes described, we have used Vienna LTE-A Downlink link level simulator [29] to simulate multipath fading channels. Two 3GPP fading models were simulated for a single-input single-output (SISO) channel: Extended Vehicular A (EVA) and Extended Typical Urban (ETU) and we used speed of 50 km/h to take into account Doppler effect. Our simulated time-frequency frames had size of 72×14 (72 subcarriers in 14 time slots equivalent to 6 resource blocks in a 1ms subframe) based on standard LTE FDD frame



(a) a sample EVA simulated frame



(b) a sample ETU simulated frame

Fig. 7: Samples of time-frequency CSI block: EVA (7a) and ETU (7b)

size. As for simulations, we select the first 36 subcarriers over the first 7 time slots as the UL channel, and the second 36 subcarriers over the second 7 time slots as the DL channel.

Number of simulated frames that were created independently was 40K (35K for train and 5K for test). Sample of simulated frames are shown in Fig. 7 while their real and imaginary parts are shown in two separate subfigures. For convenience in the rest of paper, we will show only absolute values of frames.

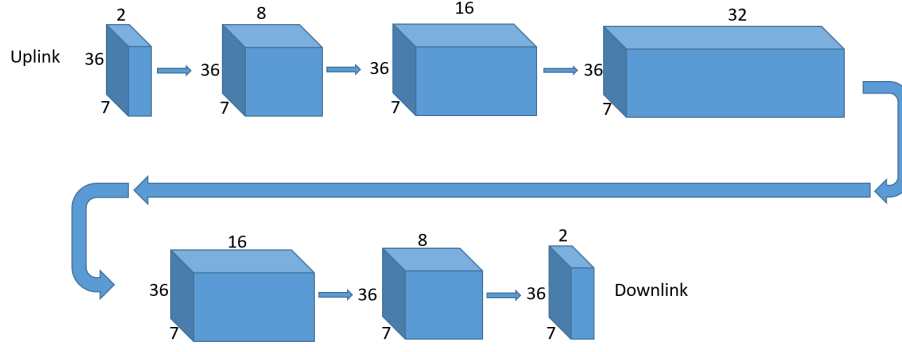


Fig. 8: CNN structure used in direct approach

B. Network Structure

In this section we explain networks structure that we have used for DL-CSI prediction for each of the direct and generative approaches.

1) *Direct Approach:* As mentioned before, CNN is used to implement direct approach. The designed network only contains convolutional layers (there is no pooling or fully connected layers) that results in lower training and testing complexity. We aimed to design our network as simple as possible thus it has only 5 hidden layers. It has totally about 12K learnable parameters that is very small compared to a few million parameters that have typical deep networks. tanh activation function is used in all layers of network except output layer (*lrelu* activation function is also tested but tanh resulted in better predictions). Network structure is shown in Fig. 8.

We used Xavier method [30] to initialize network parameters. For optimization, we used Adam optimizer [31]. Except first two layers which have symmetric padding, we used zero padding in other layers.

2) *Generative Approach:* To adapt network structure with our image size ($72 \times 14 \times 2$), we used main BEGAN [26] structure with some modifications. Network structure is shown in Fig. 9. CSI values also normalized to their maximum value and during training procedure a zero mean normal noise with decaying variance added to input values to improve network regularization. Other settings are similar to [26].

During the training phase we also faced the mode collapse problem. In mode collapse, generator generates one or limited set of images. It is a common problem in training GANs and as mentioned in [26] can be seen in BEGANs. Reference [26] suggests that decreasing the

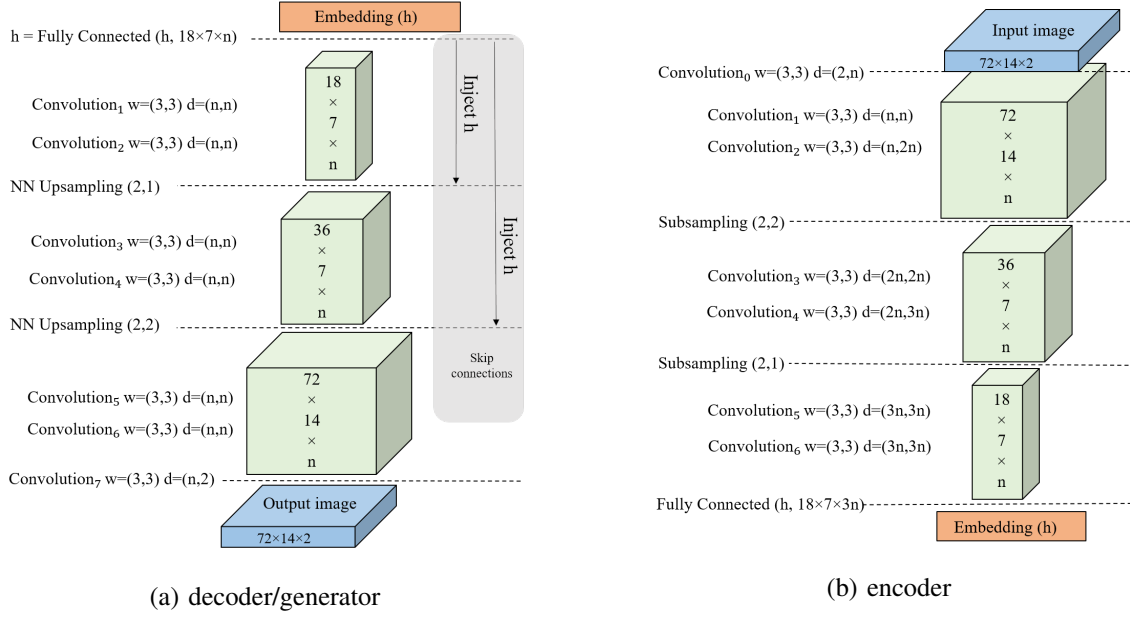


Fig. 9: BEGAN structure used in direct approach

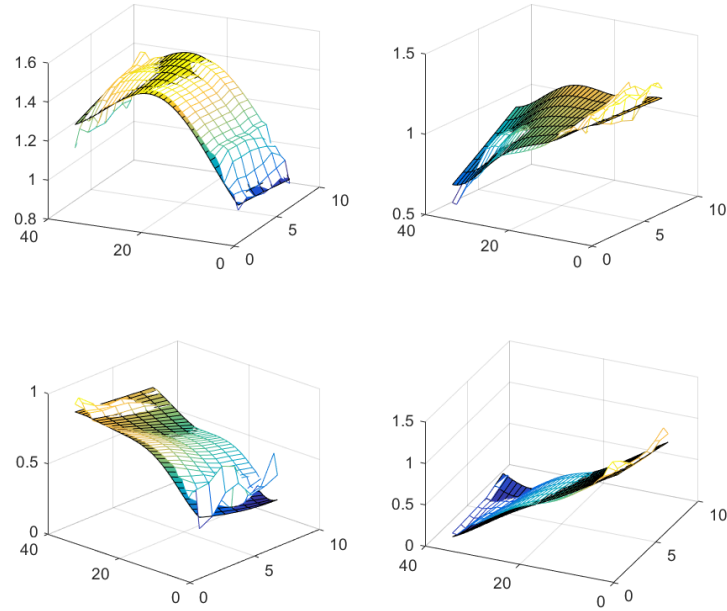
start learning rate can help the network recover from the mode collapse issue but in our case it doesn't help. The solution we used was weight sharing between generator and decoder part of discriminator.

C. Simulation results

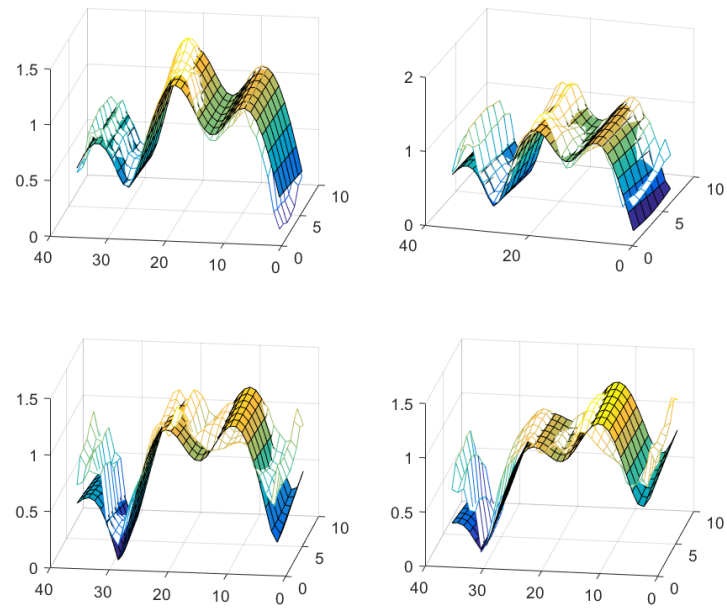
1) Direct Approach: After we trained our CNN on EVA and ETU datasets, we used trained CNN to predict DL-CSI. Some samples of the predicted DL-CSI and their corresponding ground truth (matrices of shapes 36×7) are shown in Fig. 10. The actual and predicted DL-CSI are depicted as surface (solid face colors) and meshgrid plots, respectively.

Despite good prediction quality, as can be seen in Fig 10, sometimes we have relatively high errors at the edge subcarriers, this is due to the structure of convolutional layers. One way to correct the edges is to train the network on larger images than 36×7 and then look at the middle 36×7 block.

To see the performance of the predicted DL-CSI, let's consider a pair of a transmitter (base-station) and a receiver (user). Also assume that the user and the signaling need to be simple so the user is not able to estimate the channel and feed it back to the server. In such settings if the base-station wants to send data to the user, it needs to pre-compensate the transmitted data.



(a) EVA model



(b) ETU model

Fig. 10: Sample of DL-CSi prediction using the direct approach (ground truths are depicted as a surface (solid face color) and predicted DL-CSi are depicted as a meshgrid plot)

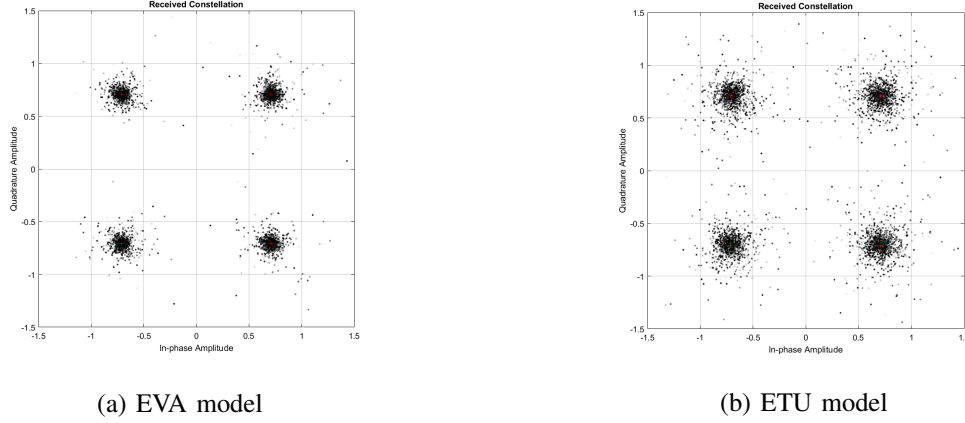


Fig. 11: Received constellations for 5K pre-compensated QPSK symbols using DL-CSI prediction (Direct approach)

To perform the pre-compensation, the base-station needs to know the DL-CSI but since there is no feedback, it should predict that. The procedure therefore will be that the transmitter first measures the UL-CSI and then using the proposed scheme, it predicts the DL channel state in the next time slot. Having a prediction of the DL channel it can pre-compensate the signal.

To examine such settings, we have simulated many channel realizations (both UL and DL channels). For each case, we assume that the base-station only knows the UL channel and it uses that information to predict DL channel. Assuming that the base-station wants to send QPSK modulated signal, before transmission it divides the signal by what it has *predicted for the DL channel*. The pre-compensated signal is then transmitted through the downlink (and thus will be multiplied by the *actual realization of the DL channel*). Therefore, if we have a good prediction of downlink they will cancel out each other. Constellations of the received symbols are shown in Fig. 11 for EVA and ETU channel models. As can be seen, the received constellations are well centered around the QPSK points verifying high prediction accuracy.

2) **Generative Approach:** We repeat the above studies to see the performance of the second proposed scheme.

First, we trained BEGAN on EVA and ETU datasets to generate images like complete frequency-time CSI block. Some BEGAN generated images are shown in Fig. 12. Note that these are images of size 72×14 (the whole time-frequency block) not just the DL-CSI.

During the completion phase, the trained BEGAN is used to determine a complete CSI image

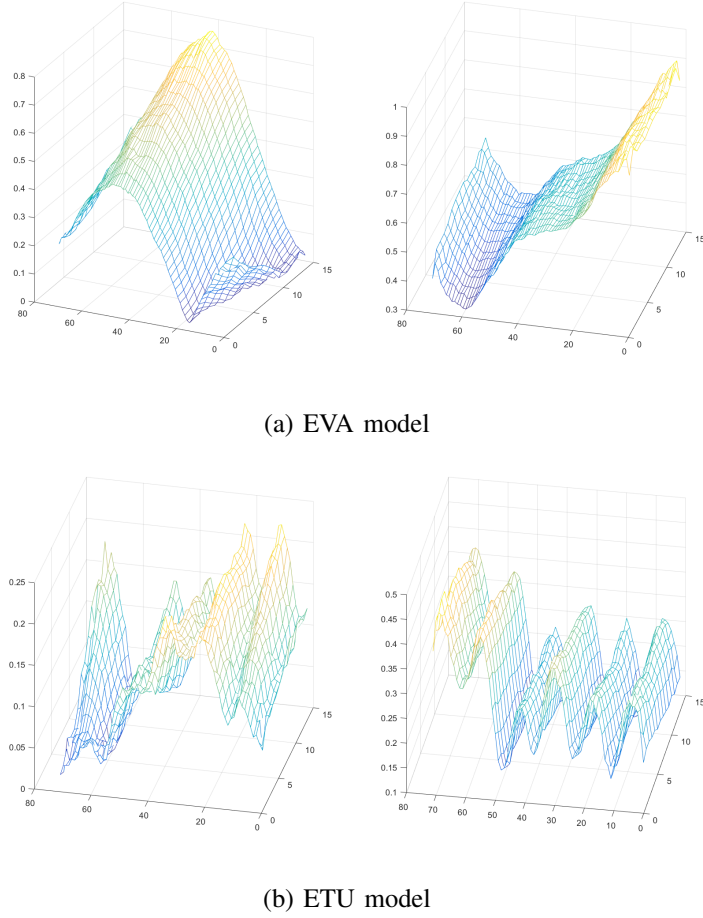
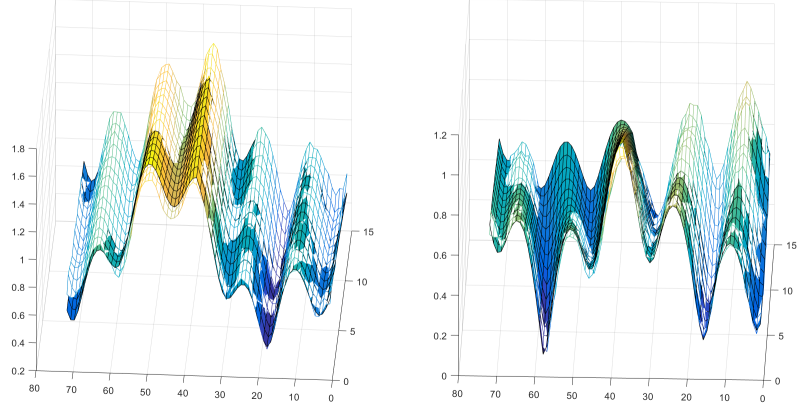


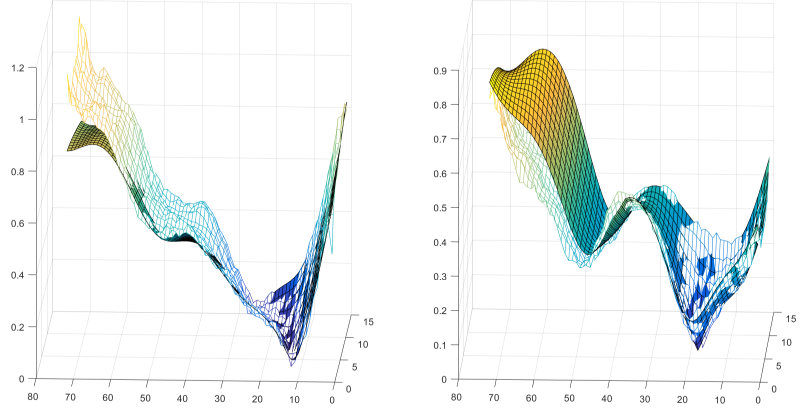
Fig. 12: Samples of the Generated Images

that matches the corrupted CSI images (UL-DCI is known and other parts are missing) in the known parts. The DL-CSI part of the resulted generated CSI image is then considered as the DL-CSI. Some completion examples using contextual loss are shown in Fig. 13 (different losses doesn't have notable visual difference so we do not include them here, the numerical result is reported at the end though). In Fig. 13, the actual time-frequency block of size 72×14 is depicted as a surface (solid face colors) and the generated image is shown as a meshgrid. The DL-CSI is the subcarriers 36 to 72 and time slots 7 to 14.

We have also tested the consistency of the DL-CSI image prediction, meaning that we fixed the UL-CSI part and then executed the image completion algorithm with different initialization of the z vector to produce the complete image. We remind that one of the main problems we faced in DCGAN structure was high difference between completed images (for a fixed corrupted image)



(a) ETU model



(b) EVA model

Fig. 13: Samples of image completion using the generative approach (ground truth vs generated CSI block)

for different initialization of z vector. As seen in Fig 14, this problem is solved by using BEGAN. The ground truth is shown as a surface and the generated images are depicted as meshgrids (there are five generated images but as they are very close they are not easily differentiable).

To see the performance of this approach, we followed the same procedure of the direct approach and simulated the constellation map of the received signal when we perform signal pre-compensation using the predicted DL-CSI. Resulted constellations are shown in Fig. 15 for EVA and ETU channel models.

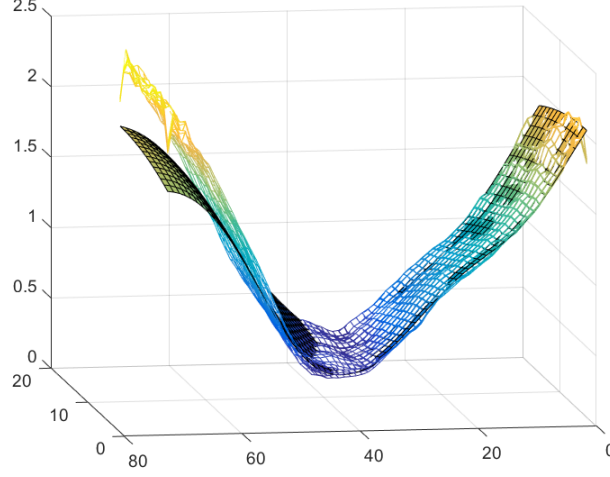
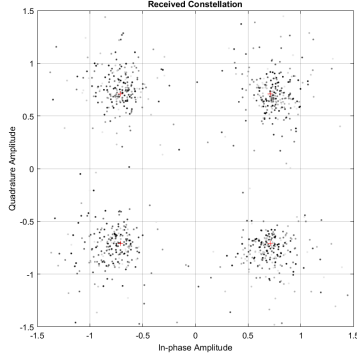
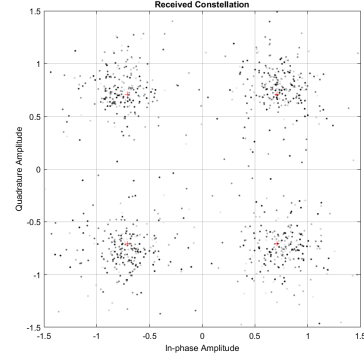


Fig. 14: Completion results with different initialization of the z vector: Completion is performed for 5 different initialized z vector. As shown they are almost same.



(a) ETU model



(b) EVA model

Fig. 15: Received constellations for 1K pre-compensated QPSK symbols using DL-CSI prediction (Generative approach)

3) **Comparable results:** As the last section, we present the comparative results between the performance of different proposed DL-CSI estimation methods. As for the comparison metric, we have used Normalized MSE (NMSE) which is defined:

$$NMSE = \frac{\|H - \hat{H}\|^2}{\|H\|^2}, \quad (9)$$

where H is the ground truth DL-CSI, and \hat{H} is the predicted values.

Dataset		Error	
CNN		BEGAN	
		Contextual	Contextual+Perceptual
EVA	0.0102	0.0638	0.0632
ETU	0.0376	0.0297	0.0308

TABLE I: NMSE of prediction for different methods on EVA and ETU datasets

As can be seen in Table I, performance of the CNN (direct approach) is better on EVA dataset compared to BEGAN (generative approach), while on ETU dataset BEGAN has relatively better results. In BEGAN, using contextual loss only has better result on ETU dataset while its performance is worse than using contextual+perceptual loss on EVA dataset (however on both datasets their performance is almost close).

VI. CONCLUSION

In this paper, we have proposed two data-driven approaches to predict DL-CSI from UL-CSI in FDD systems: direct approach and generative approach. Both of the proposed approaches try to use UL-CSI to determine a latent model that represents the environment propagation properties. The latent model is then used to predict DL-CSI. To determine the latent model, we have used Convolutional Neural Network (CNN) and Generative Adversarial Network (GAN) architecture for the direct and generative approaches, respectively. Our simulation results on EVA and ETU channel models demonstrate efficiency of both direct and generative approaches for UL2DL prediction.

REFERENCES

- [1] H. Ji, Y. Kim, J. Lee, E. Onggosanusi, Y. Nam, J. Zhang, B. Lee, and B. Shim, "Overview of full-dimension MIMO in LTE-advanced pro," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 176–184, 2017.
- [2] P. W. Chan, E. S. Lo, R. R. Wang, E. K. Au, V. K. Lau, R. S. Cheng, W. H. Mow, R. D. Murch, and K. B. Letaief, "The evolution path of 4G networks: FDD or TDD?," *IEEE Communications Magazine*, vol. 44, no. 12, pp. 42–50, 2006.
- [3] H. Almosa, R. Shafin, S. Mosleh, Z. Zhou, Y. Li, J. Zhang, and L. Liu, "Downlink channel estimation and precoding for FDD 3D Massive MIMO/FD-MIMO systems," in *Wireless and Optical Communication Conference (WOCC), 2017 26th*, pp. 1–4, IEEE, 2017.
- [4] N. Palleit and T. Weber, "Prediction of frequency selective SIMO channels," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2011 IEEE 22nd International Symposium on*, pp. 1428–1432, IEEE, 2011.

- [5] K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "Joint-directional properties of uplink and downlink channel in mobile communications," *Electronics Letters*, vol. 35, no. 16, pp. 1311–1312, 1999.
- [6] K. Hugl, K. Kalliola, and J. Laurila, "Spatial reciprocity of uplink and downlink radio channels in FDD systems," *Proc. COST 273 Technical Document TD (02)*, vol. 66, p. 7, 2002.
- [7] A. A. Esswie, M. El-Absi, O. A. Dobre, S. Ikki, and T. Kaiser, "A novel FDD massive MIMO system based on downlink spatial channel estimation without CSIT," in *Communications (ICC), 2017 IEEE International Conference on*, pp. 1–6, IEEE, 2017.
- [8] H. Xie, F. Gao, S. Jin, J. Fang, and Y.-C. Liang, "Channel estimation for TDD/FDD massive MIMO systems with channel covariance computing," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 4206–4218, 2018.
- [9] A. Decurninge, M. Guillaud, and D. T. Slock, "Channel covariance estimation in massive MIMO frequency division duplex systems," in *Globecom Workshops (GC Wkshps), 2015 IEEE*, pp. 1–6, IEEE, 2015.
- [10] T. Asté, P. Forster, L. Fety, and S. Mayrargue, "Downlink beamforming avoiding DOA estimation for cellular mobile communications," in *IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS SPEECH AND SIGNAL PROCESSING*, vol. 6, pp. VI–3313, INSTITUTE OF ELECTRICAL ENGINEERS INC (IEE), 1998.
- [11] D. Vasisht, S. Kumar, H. Rahul, and D. Katabi, "Eliminating channel feedback in next-generation cellular networks," in *Proceedings of the 2016 ACM SIGCOMM Conference*, pp. 398–411, ACM, 2016.
- [12] D. Hu and L. He, "Channel Estimation for FDD Massive MIMO OFDM Systems," in *Vehicular Technology Conference (VTC-Fall), 2017 IEEE 86th*, pp. 1–5, IEEE, 2017.
- [13] Y. Han, T.-H. Hsu, C.-K. Wen, K.-K. Wong, and S. Jin, "Efficient downlink channel reconstruction for FDD transmission systems," in *Wireless and Optical Communication Conference (WOCC), 2018 27th*, pp. 1–5, IEEE, 2018.
- [14] Y. Han, J. Ni, and G. Du, "The potential approaches to achieve channel reciprocity in FDD system with frequency correction algorithms," in *Communications and Networking in China (CHINACOM), 2010 5th International ICST Conference on*, pp. 1–5, IEEE, 2010.
- [15] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.
- [16] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, 2017.
- [17] T. J. O'Shea, T. Roy, and N. West, "Approximating the Void: Learning Stochastic Channel Models from Observation with Variational Generative Adversarial Networks," *arXiv preprint arXiv:1805.06350*, 2018.
- [18] T. J. O'Shea, T. Roy, N. West, and B. C. Hilburn, "Physical Layer Communications System Design Over-the-Air Using Adversarial Networks," *arXiv preprint arXiv:1803.03145*, 2018.
- [19] Z. Qin, H. Ye, G. Y. Li, and B.-H. F. Juang, "Deep learning in physical layer communications," *arXiv preprint arXiv:1807.11713*, 2018.
- [20] M. Soltani, A. Mirzaei, V. Pourahmadi, and H. Sheikhzadeh, "Deep Learning-Based Channel Estimation," *arXiv preprint arXiv:1810.05893*, 2018.
- [21] C.-K. Wen, W.-T. Shih, and S. Jin, "Deep Learning for Massive MIMO CSI Feedback," *IEEE Wireless Communications Letters*, 2018.
- [22] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," *arXiv preprint arXiv:1603.07285*, 2016.
- [23] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, pp. 2672–2680, 2014.
- [24] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training gans," in *Advances in Neural Information Processing Systems*, pp. 2234–2242, 2016.

- [25] J. Zhao, M. Mathieu, and Y. LeCun, “Energy-based generative adversarial network,” *arXiv preprint arXiv:1609.03126*, 2016.
- [26] D. Berthelot, T. Schumm, and L. Metz, “BEGAN: boundary equilibrium generative adversarial networks,” *arXiv preprint arXiv:1703.10717*, 2017.
- [27] A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” *arXiv preprint arXiv:1511.06434*, 2015.
- [28] R. A. Yeh, C. Chen, T.-Y. Lim, A. G. Schwing, M. Hasegawa-Johnson, and M. N. Do, “Semantic Image Inpainting with Deep Generative Models,” in *CVPR*, vol. 2, p. 4, 2017.
- [29] C. Mehlführer, J. C. Ikuno, M. Šimko, S. Schwarz, M. Wrulich, and M. Rupp, “The Vienna LTE simulators-Enabling reproducibility in wireless communications research,” *EURASIP Journal on Advances in Signal Processing*, vol. 2011, no. 1, p. 29, 2011.
- [30] X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pp. 249–256, 2010.
- [31] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.