Report on 2021 IEEE Signal Processing Cup

Team name: MBNRG_IUST Team ID: 26435
Alireza Navi(undergraduate, alirezanavi1999@gmail.com)
Ali Amiri hazaveh(undergraduate, ali.amiri.h.1410@gmail.com)
Fatemeh Zamanian(postgraduate, zamanian73@gmail.com)
Erfan Fatemi(undergraduate, erfanpro2012@gmail.com)
Mohammad Razavizadeh(Supervisor, smrazavi@gmail.com)

***Abstract*—This report presents the outcomes of the activities carried out by our team regarding the Signal Processing Cup (SP Cup) competition 2021. The objective of this project is finding the optimum configuration for an intelligent reflecting surface to maximize the individual rate of all of the 50 users in the network. We present the adopted ideas and solution methods as well as the final results obtained from our simulations and numerical experiments. The report consists of two sections. In the first section, the channel estimation method is explained and the second part discusses the optimization problem for IRS coefficients and its solution.**

***Index Terms*—Intelligent Reconfigurable Surface(IRS), OFDM, least squares(LS), sparsity, successive convex approximation(SCA)**

I. Introduction

THE first thing that needs to be clarified to make the rest of the report understandable, is the system model and problem formulation. a simple illustration of the system is depicted in Fig.1.

The transmission protocol is an uplink transmission since the computations must be done at the BS side and the coefficients will be sent to the IRS using a backhaul link. based on the description file since OFDM is implemented for the transmission, the received signal-for each user-in frequency domain is given by

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{h}_{\boldsymbol{\theta}} + \boldsymbol{V} \tag{1}$$

Where $\boldsymbol{y} \in \mathbb{C}^{500\times1}$ is the received signal, $\boldsymbol{X} \in \mathbb{C}^{500\times500} = \text{diag}(\boldsymbol{x})$ is a diagonal matrix of the transmitted OFDM pilot $\boldsymbol{x}$, $\boldsymbol{V} \sim \mathcal{CN}(0, \sigma_n^2 \boldsymbol{I}_{500})$ is the 500*1 AWGN vector and $\boldsymbol{h}_{\boldsymbol{\theta}} \in \mathbb{C}^{500\times1}$ is the equivalent channel between the transmitter and the receiver, that based on [1] is modeled as

$$\boldsymbol{h}_{\boldsymbol{\theta}} = \sum_{n=1}^{4096} \boldsymbol{q}_n \theta_n \odot \boldsymbol{b}_n + \boldsymbol{d} \tag{2}$$

Where $\odot$ is the hadamard product, $n \in \mathcal{N} \triangleq \{1, \cdots, 4096\}$ is the IRS element index, θ_n is the coefficient of the n_{th} IRS element that can be either $+1$ or -1, $\boldsymbol{d} \in \mathbb{C}^{500\times1}$ is the channel frequency response(CFR) of the user→BS direct link and $\boldsymbol{b}_n \in \mathbb{C}^{500\times1}$ and $\boldsymbol{q}_n \in \mathbb{C}^{500\times1}$ are CFRs of user→IRS link and IRS→link respectively. Since IRS is a passive element and there is no observation at the IRS, it is not possible to estimate $\boldsymbol{b}_n$ and $\boldsymbol{q}_n$ separately. Because of that by denoting $\boldsymbol{g}_n \triangleq \boldsymbol{q}_n \odot \boldsymbol{b}_n$ and $\boldsymbol{G} \triangleq [\boldsymbol{g}_1, \cdots, \boldsymbol{g}_{4096}]$, (2) can be written as

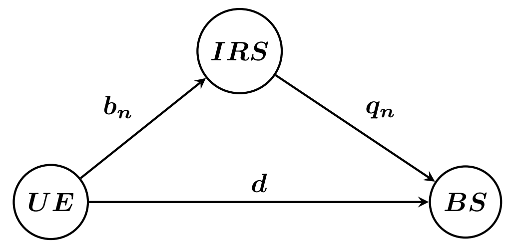

Fig. 1. a simple schematic of the purposed system model

$$\boldsymbol{y} = \boldsymbol{X}(\boldsymbol{G}\boldsymbol{\theta} + \boldsymbol{d}) + \boldsymbol{V} \tag{3}$$

Where $\boldsymbol{\theta} \triangleq [\theta_1, \cdots, \theta_{4096}]^T$ is the IRS coefficient vector. From (1) and (3) it is obvious that $\boldsymbol{h}_{\boldsymbol{\theta}} = \boldsymbol{G}\boldsymbol{\theta} + \boldsymbol{d}$ and $\boldsymbol{G} \in \mathbb{C}^{500\times4096}$. So, for the purpose of channel estimation, it is needed to estimate $\boldsymbol{G}$ and $\boldsymbol{d}$.

II. Channel estimation

There are different ways to perform channel estimation in IRS-assisted OFDM systems. Authors in [2] and [3] have deployed a compressive sensing method by placing few active elements among the passive IRS elements and estimate $\boldsymbol{b}_n$ and $\boldsymbol{q}_n$ independently, since there are no active elements in our setup, this method is not doable for our case. Also, there is no statistical information about the channel parameters, so the bayesian approches implemented in [4] and [5] is not doable either. So the first way that comes to mind is to perform a least square(LS) channel estimation. It is proved in [6] that for a proper channel estimation of $\boldsymbol{G}$ and $\boldsymbol{d}$, $N+1$ pilots shall be transmitted under different IRS configurations where N is the number of IRS elements. since we have only 4096 transmissions it is challenge to estimate $\boldsymbol{d}$ and $\boldsymbol{G}$ properly. assuming the direct link, $\boldsymbol{d}$, is totally blocked($\boldsymbol{d} = \boldsymbol{0}_{500\times1}$), the LS estimation for the i_{th} IRS configuration would be as follows

$$\boldsymbol{X}^{-1}\boldsymbol{y}^{(i)} = \boldsymbol{G}\boldsymbol{\theta}^{(i)} + \boldsymbol{X}^{-1}\boldsymbol{V}, \qquad i = 1, \cdots, 4096 \tag{4}$$

therefore by defining $\boldsymbol{Y} \triangleq [\boldsymbol{y}^{(1)}, \cdots, \boldsymbol{y}^{(4096)}]$ and $\bar{\boldsymbol{\theta}} \triangleq [\boldsymbol{\theta}^{(1)}, \cdots, \boldsymbol{\theta}^{(4096)}]$, (4) is written in a matrix form as

$$\boldsymbol{X}^{-1}\boldsymbol{Y} = \boldsymbol{G}\bar{\boldsymbol{\theta}} + \boldsymbol{X}^{-1}\boldsymbol{V} \tag{5}$$

so,

$$\boldsymbol{X}^{-1}\boldsymbol{Y}\bar{\boldsymbol{\theta}}^{-1} = \boldsymbol{G} + \boldsymbol{X}^{-1}\boldsymbol{V}\bar{\boldsymbol{\theta}}^{-1} = \hat{\boldsymbol{G}}_{\boldsymbol{LS}} \tag{6}$$

This is the first step for the channel estimation, But obviously it needs to be enhanced much more. The next step is to make use of the OFDM property [1]. As mentioned in the description file, the Channel Impulse Response(CIR) has a dimension of $\mathbf{20} \times \mathbf{1}$ so G,if transformed into time domain, will have a matrix $\boldsymbol{g}$ of size $\mathbf{20} \times \mathbf{4096}$. On the other hand, by transforming $\hat{\boldsymbol{G}}_{\boldsymbol{LS}}$ into time domain, using $\mathbf{500}-$point IDFT, we will obtain a $\hat{\boldsymbol{g}}_{\boldsymbol{LS}}$ matrix of size $\mathbf{500} \times \mathbf{4096}$. by comparing $\boldsymbol{g}$ and $\hat{\boldsymbol{g}}_{\boldsymbol{LS}}$ we can realize that the last $\mathbf{480}$ rows of $\hat{\boldsymbol{g}}_{\boldsymbol{LS}}$ are the unpleasant results of the noise. So, by zero forcing those rows and transforming the result back into the frequency domain we will have a much better estimation of $\boldsymbol{G}$.

$$\hat{\boldsymbol{g}}_{\boldsymbol{LS}} = \boldsymbol{F}_{\mathbf{500}}^{\boldsymbol{H}}\hat{\boldsymbol{G}}_{\boldsymbol{LS}} \tag{7}$$

$$\tilde{\boldsymbol{g}} = [\hat{\boldsymbol{g}}_{\boldsymbol{LS}}]_{\mathbf{1:20\times 4096}} \tag{8}$$

$$\tilde{\boldsymbol{G}} = \boldsymbol{F}_{\mathbf{500}}[\tilde{\boldsymbol{g}}^{\boldsymbol{T}}, \mathbf{0}_{\mathbf{4096\times 480}}]^{\boldsymbol{T}} \tag{9}$$

Where $\boldsymbol{F}_{\boldsymbol{N}}$ and $\boldsymbol{F}_{\boldsymbol{N}}^{\boldsymbol{H}}$ are $\boldsymbol{N}-$point DFT and IDFT matrix respectively. The next thing that needs to be done is a better assumption for the direct link channel ,$\boldsymbol{d}$. authors in [1] rewrote (3) as

$$\boldsymbol{y} = \boldsymbol{X}\bar{\boldsymbol{G}}\bar{\boldsymbol{\phi}} + \boldsymbol{V} \tag{10}$$

Where $\bar{\boldsymbol{G}} = [\boldsymbol{d}, \boldsymbol{G}]$ and $\bar{\boldsymbol{\phi}} = [\mathbf{1}_{\mathbf{4096\times 1}}, \bar{\boldsymbol{\theta}}^{\boldsymbol{T}}]^{\boldsymbol{T}}$ which can be interpreted as having an IRS of $\mathbf{4097}$ elements with the first element being set to $\mathbf{1}$ during all the transmissions and the direct link channel,$\boldsymbol{d}$, being ignored. Since the given pilot matrix in dataset-which is the same as $\bar{\boldsymbol{\theta}}$ in this report- is produced using HADAMARD function in MATLAB, it already has an element(column $\mathbf{1}$) being set to $\mathbf{1}$ during all the transmissions. So we can somehow see the system as a $\mathbf{4095}-$element IRS and direct link inside. Another assumption is to set the estimated channel responding to element $\mathbf{1}$ as the direct link channel,$\boldsymbol{d}$. By testing this direct link assumption (DLA) on the first dataset, it shows that it is a better than the first assumption we had($\boldsymbol{d} = \mathbf{0}_{\mathbf{500\times 1}}$).

Now it is time for the last idea that aims to improve the accuracy of channel estimation, which is using convolutional neural network(CNN), specifically Denoising convolutional neural network(DnCnn). Although this algorithm adds a lot of complexity, it enhances the performance of the Nlos users noticeably. This CNN is mainly used in image processing problems when the goal is to remove noise from images. Since an image is nothing more than three matrices and the nature of noise is the same for both an image and a wireless system, then why should the cure be any different? This method is vastly explained in [7]. In short it is composed of $\mathbf{20}$ convolutional layers followed by ReLUs and batch-normalization layers. After estimating the channel regardless of how it is estimated the real and imaginary part of the channel are separated and are given to the DnCnn and at the output of the network, they are reconnected again. Authors in [5][8] have used this method to further improve their channel estimation. A simple diagram of this method is shown in Fig.2. As for the last point, it needs to be stated that the absolute values of phaseshifts of $\boldsymbol{\theta_i s}$ does

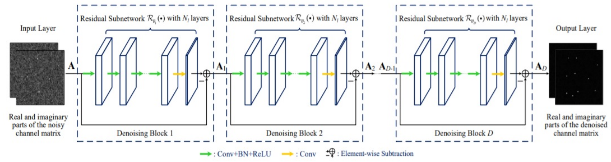

Fig. 2. figure rights belong to authors of [5].

not effect the estimation as long as the difference between the possible states of each element is set to $\mathbf{180°}$. So the absolute values are set to $\mathbf{0}$ and $\mathbf{180}$ to work with real values of $\boldsymbol{\theta}$ instead of complex ones.

III. Reflection Optimization

The rate of each user is given by

$$\boldsymbol{R} = \frac{\boldsymbol{B}}{\boldsymbol{K}+\boldsymbol{M}-\mathbf{1}}\sum_{\boldsymbol{v}=\mathbf{0}}^{\boldsymbol{K}-\mathbf{1}}\log_2\left(\mathbf{1} + \frac{\boldsymbol{P}|\boldsymbol{h}_{\boldsymbol{\theta}}[\boldsymbol{v}]|^2}{\boldsymbol{B}\boldsymbol{N}_{\mathbf{0}}}\right) \tag{11}$$

Therefore the optimization problem would be as

$$(\boldsymbol{P1}):\ \max_{\boldsymbol{\theta}}\sum_{\boldsymbol{v}=\mathbf{0}}^{\boldsymbol{K}-\mathbf{1}}\log_2\left(\mathbf{1} + \frac{\boldsymbol{P}|\boldsymbol{h}_{\boldsymbol{\theta}}[\boldsymbol{v}]|^2}{\boldsymbol{B}\boldsymbol{N}_{\mathbf{0}}}\right) \tag{12}$$

$$\boldsymbol{s.t.}\quad |\boldsymbol{\theta}_{\boldsymbol{i}}| = \mathbf{1} \quad \forall \boldsymbol{i} = \mathbf{1}, \cdots, \mathbf{4096} \tag{13}$$

$$\angle\boldsymbol{\theta}_{\boldsymbol{i}} = \{\mathbf{0}, \boldsymbol{\pi}\} \quad \forall \boldsymbol{i} = \mathbf{1}, \cdots, \mathbf{4096} \tag{14}$$

Where $\boldsymbol{B} = \mathbf{10}\boldsymbol{Mhz}$ is the channel bandwidth,$\boldsymbol{N0}$ is the noise power spectral density,$\boldsymbol{P}$ is the power od transmitted signal over each subcarrier,$\boldsymbol{K} = \mathbf{500}$ is the number of subcarriers and $\boldsymbol{M} = \mathbf{20}$ is the nuber of channel taps. Since solving an optimization problem with discrete variables is really challenging, at first we ignore the second constraint (14) and solve the problem for continuous values of $\boldsymbol{\theta_i s}$ and then approximate the results to its nearest acceptable value. This method is also used in [9]. But still there is a problem, (12) is non-convex over $\boldsymbol{\theta}$, meaning that it is still hard to solve. A method is purposed in [10] to change (12) into convex for and solve it using successive convex approximation (SCA). This method has to much complexity when the IRS is equipped with a large number of elements which in our case it is. Also the algorithm needs to be used 50 times, once for each user, So we need to implement an algorithm with less complexity. By using Jensen inequality, an upper bound for (11) is obtained as

$$\boldsymbol{R} \leq \frac{\boldsymbol{B}}{\boldsymbol{K}+\boldsymbol{M}-\mathbf{1}}\log_2\left(\mathbf{1} + \sum_{\boldsymbol{v}=\mathbf{0}}^{\boldsymbol{K}-\mathbf{1}}\frac{\boldsymbol{P}|\boldsymbol{h}_{\boldsymbol{\theta}}[\boldsymbol{v}]|^2}{\boldsymbol{B}\boldsymbol{N}_{\mathbf{0}}}\right) \tag{15}$$

Now, instead of optimizing the rate we will try to optimize the upper bound of rate which in [1] it is discussed that it will have a close to optimal performance but with the advantage of being much less complex. Since the upper bound (15) is a single $\mathbf{log}$ function the new optimization problem is written as

$$(\boldsymbol{P2}):\ \max_{\boldsymbol{\theta}}\sum_{\boldsymbol{v}=\mathbf{0}}^{\boldsymbol{K}-\mathbf{1}}|\boldsymbol{h}_{\boldsymbol{\theta}}[\boldsymbol{v}]|^2 \tag{16}$$

$$\boldsymbol{s.t.}\quad |\boldsymbol{\theta}_{\boldsymbol{i}}| = \mathbf{1} \quad \forall \boldsymbol{i} = \mathbf{1}, \cdots, \mathbf{4096} \tag{17}$$

This problem can be solved using Semi-definite Relaxation (SDR), but it will have a complexity in the order of $\mathcal{O}\left((N+1)^6\right)$. So another method should be approached[1]. First the objective of (P2) is reformed in time domain as

$$\sum_{l=0}^{L-1} |\tilde{g}_{(l,:)}\theta + \tilde{g}_{(l,1)}|^2 \tag{18}$$

Where $\tilde{g}_{(l,:)}$ is the l_{th} tap of CIR (7) and $\tilde{g}_{(l,1)}$ is the assumption made for the direct link in II. Now if we find the tap with the strongest CIR as

$$\hat{l} = \underset{l\in\{0,\cdots,19\}}{\arg\max} \left| \sum_{n=1}^{N} |\tilde{g}_{(l,n)}| + |\tilde{g}_{(l,1)}| \right|^2 \tag{19}$$

Finally the refelection coefficiant of each element is determined using the strongest tap in the CIR as

$$\angle\theta_n = -\angle\tilde{g}_{(\hat{l},n)} + \angle\tilde{g}_{(\hat{l},1)} \tag{20}$$

And the discrete reflection coefficient, θ_n, is defined as

$$\theta_n = \begin{cases} 1 & \text{if} \quad real(e^{j\angle\theta_n}) > 0 \\ -1 & \quad O.W \end{cases} \tag{21}$$

References

[1] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced ofdm: Channel estimation and reflection optimization," *arXiv preprint arXiv:1909.03272*, 2019.

[2] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling large intelligent surfaces with compressive sensing and deep learning," 2019.

[3] S. Liu, Z. Gao, J. Zhang, M. D. Renzo, and M.-S. Alouini, "Deep denoising neural network assisted compressive channel estimation for mmwave intelligent reflecting surfaces," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 9223–9228, 2020.

[4] N. K. Kundu and M. R. McKay, "A deep learning-based channel estimation approach for miso communications with large intelligent surfaces," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, 2020, pp. 1–6.

[5] C. Liu, X. Liu, D. W. K. Ng, and J. Yuan, "Deep residual learning for channel estimation in intelligent reflecting surface-assisted multi-user communications," 2021.

[6] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6607–6620, 2020.

[7] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, p. 3142–3155, Jul 2017. [Online]. Available: http://dx.doi.org/10.1109/TIP.2017.2662206

[8] M. Soltani, V. Pourahmadi, A. Mirzaei, and H. Sheikhzadeh, "Deep learning-based channel estimation," *IEEE Communications Letters*, vol. 23, no. 4, pp. 652–655, 2019.

[9] X. Pei, H. Yin, L. Tan, L. Cao, Z. Li, K. Wang, K. Zhang, and E. Björnson, "Ris-aided wireless communications: Prototyping, adaptive beamforming, and indoor/outdoor field trials," 2021.

[10] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets ofdm: Protocol design and rate maximization," 2019.