Configuring an Intelligent Reflecting Surface for Wireless Communications: the hUMAns at RISk approach

Juan Carlos Ruiz Sicilia, Javier Giménez de la Cuesta, Gonzalo J. Anaya López, F. Javier López-Martínez
{jcrs, javierg, gjal, fjlopezm}@ic.uma.es
Dpto. de Ingeniería de Comunicaciones. Universidad de Málaga, Málaga, España 29071.

***Abstract*—We present our solution for the IEEE Signal Processing Cup 2021 challenge. The goal is to maximize the weighted average achievable rate for a set of $N_{\mathrm{u}} = 50$ users assisted by an intelligent reflecting surface (IRS). We combined a modified least-squares smoothing estimator, several variations over the strongest-tap maximization (STM) algorithm, and a final fine-tuning optimization that explored solutions in the vicinity of the STM solutions. Simulations show that our IRS configurations yield an estimated weighted achievable rate slightly over 117 Mbps.**

I. Introduction

The increasing number of devices and connections by the year of 2022 [1] combined with increasingly demanding requirements on the emerging future wireless networks (5-th generation (5G) [2] or 6-th generation (6G) [3]) have focused the attention on energy consumption [4]. Massive multiple-input multiple-output (MIMO) has been established as the key technology for improving spectral efficiency in these emerging networks [5]. Hence, finding energy-efficient hardware solutions to enhance the networks performance has become one of the crucial tasks for the upcoming new generations.

Among all the proposed technologies, the intelligent reflecting surfaces (IRSs) [6–8] have been proposed to achieve the above described goals. An IRS is a (nearly) passive two-dimensional meta-surface composed of elements that can be controlled with integrated electronic to modify their reflection coefficient [9]. This technology is energy-efficient thanks to the possibility of re-radiating the incoming signal without using any power amplifier. Instead, this requires to design properly the phase shifts applied by each reflective element, in order to constructively combine all the reflected signals.

This motivated the topic of the eighth edition of the IEEE Signal Processing Cup [10]: control a wireless propagation environment using an IRS deployed at a fixed location with the goal of assisting the communication between a base station and a set of $N_{\mathrm{u}} = 50$ users. The IEEE Signal Processing Cup is a yearly competition organized by the IEEE Signal Processing society. In this competition, undergraduate students form teams to work on real-life challenges. Each team should include one faculty member as an advisor, at most one graduate student as a mentor, and up to ten undergraduate students.

In this article, we describe the problem proposed in the competition, and the solution proposal made by our team, *hUMAns at RISk*.

Notation: Throughout the document, lower-case bold letters denote vectors, and overline lower-case bold letters denote frequency-domain vectors; $\mathbf{1}$ is the all-one vector; $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ denote the transpose and Hermitian transpose operations, respectively; $\odot$ denotes the Hadamard product; $\|\cdot\|_1$ is the L_1 norm.

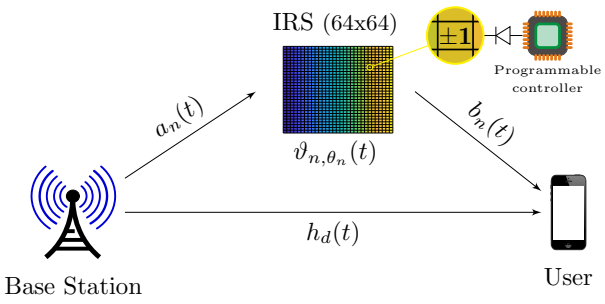

Fig. 1. System model with an IRS of 64x64 elements. The controller configures each element of the IRS in one of two states: $\theta_n = \pm 1$.

II. System model

The IRS consists of $N = 4096$ elements, distributed as a 64×64 array, where each of which can adjust its phase selecting from two possible states whose phase difference is equal to π. Hence, a total of 2^{4096} possible configurations can be chosen. Transmission is designed using orthogonal frequency division multiplexing (OFDM), so that data for each user are allocated on a set of $K = 500$ subcarriers spread over a bandwidth $B = 10$ MHz. Assuming that a cyclic extension longer than the channel impulse response (CIR) is used, i.e., $M = 20$ channel taps for the equivalent CIR, a frequency domain equivalent model can be used to describe the system response, as

$$\overline{\boldsymbol{z}} = \overline{\boldsymbol{h}}_{\boldsymbol{\theta}} \odot \overline{\boldsymbol{x}} + \overline{\boldsymbol{w}}, \tag{1}$$

where all frequency-domain vectors in (1) are of length K, $\overline{\boldsymbol{z}}$ is the received signal, $\overline{\boldsymbol{h}}_{\boldsymbol{\theta}}$ is the IRS-dependent channel frequency response, and $\overline{\boldsymbol{w}}$ represents the noise.

The end-to-end channel frequency response, whose expression in the time-domain is $h_{\boldsymbol{\theta}}(t)$, can be defined as follows:

$$h_{\boldsymbol{\theta}}(t) = h_d(t) + \sum_{n=1}^{N} \left(b_n * \vartheta_{n;\theta_n} * a_n\right)(t). \tag{2}$$

This expression results from the combination of the multiple channels over which the transmitted signal, $x(t)$, is propagated from the base station (BS) until it reaches the legitimate user, see Fig. 1. The signal from the BS is propagated over a linear time-invariant (LTI) channel with impulse response $a_n(t)$ to the IRS. The IRS is reconfigurable with an external stimulus, θ_n, being its behaviour defined by an LTI filter

with the impulse response defined by $\vartheta_{n;\theta_n}(t)$. Each element of the IRS modifies the phase of the impinging signal and propagates it over an LTI channel with impulse response $b_n(t)$. In addition, there are paths that are not propagated over the IRS, therefore out of control, that constitute the uncontrollable channel, $h_d(t)$.

The per-user signal-to-noise ratio (SNR) is given by $P|\bar{h}_{\boldsymbol{\theta}}^{(u)}[\nu]|^2/BN_0$, where $\bar{h}_{\boldsymbol{\theta}}^{(u)}[\nu]$ is the IRS-dependent channel frequency response for each user $u \in \{1, ..., N_\mathrm{u}\}$. For each of these users, u, the achievable data rate that we aim to maximize can be computed as

$$R_\mathrm{u} = \frac{B}{K+M-1} \sum_{\nu=0}^{K-1} \log_2 \left(1 + \frac{P\left|\bar{h}_{\boldsymbol{\theta}}^{(u)}[\nu]\right|^2}{BN_0} \right), \tag{3}$$

where $P = 1$ W is the transmit power and N_0 is the noise power spectral density. Note that this expression is a summation over the K subcarriers with a cyclic prefix compensation $K + M - 1$. In order to compensate the lower rates in the non-line-of-sight (NLOS) users, the final achievable data rate is computed as a weighted average rate:

$$\overline{R} = \frac{1}{N_\mathrm{u}} \sum_{u=1}^{N_\mathrm{u}} \alpha_u R_u, \tag{4}$$

where $\alpha_u = \{1, 2\}$ for LOS/NLOS (Line-of-sight/non-line-of-sight) users, respectively.

The goal of the competition is to design an algorithm that selects a good IRS configuration based on measurements made at the user location [10]. To this end there are two available data sets. The first one, *dataset1*, contains the received signal during a training phase using a known pilot sequence over $4N$ configurations of the IRS focused on a single user. The second one, *dataset2*, contains the received signals for the 50 users for which N possible IRS configurations are used during a training phase using a given pilot sequence.

Besides other tasks such as estimating the noise power density, and determining (if possible) which users are LOS/NLOS, the key problems to be solved are: (*i*) channel estimation, and (*ii*) IRS optimization.

III. Algorithms

A. Noise and channel estimation

In order to maximize the individual rates for each user according to (3), i.e., finding the optimal IRS configuration that maximizes such rates, we need to know first the noise power density, and also estimate the $N+1$ channels involved in the communication.

First, noise power is estimated by taking advantage of the extended data set available for a toy user (i.e., *dataset1*). Since there are $4N$ IRS configurations available, and some of them are repeated, combining two of the received signals under a constant lay-out (IRS configuration and assuming static channels) is enough to obtain a good noise estimate:

$$N_0 = \mathrm{var}\left\{ \frac{\overline{\boldsymbol{z}}_\mathrm{i} - \overline{\boldsymbol{z}}_\mathrm{j}}{\sqrt{2}} \right\} \tag{5}$$

which is computed for each pair $\{i, j\}$ of identical configurations in *dataset1*.

Next, the CIR needs to be estimated. This involves estimating the composite channels between the transmitter and receiver through each of the N IRS elements, plus the direct channel between the transmitter and the receiver. Since in the available data set for the N_u users (i.e., *dataset2*) the pilot transmission consisted of only N different configurations, this is not enough for directly using least squares (LS) estimation, and a different approach had to be taken to estimate the direct channel.

Formally, the direct channel can be assimilated as an additional IRS element whose configuration does not change. Then, the channel impulse response can be expressed as follows:

$$\boldsymbol{h}_{\boldsymbol{\theta}} = \boldsymbol{h}_d + \boldsymbol{V}^T \boldsymbol{\omega}_{\boldsymbol{\theta}} = \begin{bmatrix} \boldsymbol{h}_d & \boldsymbol{V}^T \end{bmatrix} \begin{bmatrix} 1 \\ \boldsymbol{\omega}_{\boldsymbol{\theta}} \end{bmatrix} \tag{6}$$

where the $N \times M$ matrix $\boldsymbol{V}$ includes the contribution of the transmitter-IRS and the IRS-receiver channels, and $\boldsymbol{\omega}_{\boldsymbol{\theta}}$ is an $N \times 1$ vector including the phase response for each element. Note that both $\boldsymbol{h}_{\boldsymbol{\theta}}$ and $\boldsymbol{h}_d$ are vectors of size M, i.e., the number of taps for the CIR. Hence, the received signal for the i-th configuration used during the training phase can be expressed as

$$\overline{\boldsymbol{z}}_i = \boldsymbol{X}\boldsymbol{F}\boldsymbol{h}_{\boldsymbol{\theta}_i} = \boldsymbol{X}\boldsymbol{F} \begin{bmatrix} \boldsymbol{h}_d & \boldsymbol{V}^T \end{bmatrix} \begin{bmatrix} 1 \\ \boldsymbol{\omega}_{\boldsymbol{\theta}i} \end{bmatrix} \tag{7}$$

where $\mathbf{X}$ is the $K \times K$ diagonal matrix with $[\mathbf{X}]_{k,k}$ being the symbol transmitted in the k-th carrier; and $\mathbf{F}$ is the $K \times M$ discrete Fourier transform (DFT) matrix of K points, and where we neglected the noise term for the sake of compactness.

Now, since the configuration matrix is a Hadamard matrix, then all configurations are orthogonal. Then, it can be proved that the product of the first row, which is the configuration of the first element in the IRS, with the other rows results in a row vector where all its elements but the first one are zero.

$$\boldsymbol{P}\,\mathbf{1} = \begin{bmatrix} N & 0 & \dots & 0 \end{bmatrix} \tag{8}$$

where $\mathbf{P}$ is the $N \times N$ pilot matrix and $\mathbf{1}$ has dimension N, which corresponds to the configuration of the first element. This property can be applied to the matrix $\mathbf{Z} = [\overline{\mathbf{z}}_1 \dots \overline{\mathbf{z}}_N]$ in the same fashion

$$\boldsymbol{Z}\,\mathbf{1} = \boldsymbol{X}\boldsymbol{F} \begin{bmatrix} \boldsymbol{h}_d & \boldsymbol{V}^T \end{bmatrix} \begin{bmatrix} \mathbf{1}^T \\ \boldsymbol{P} \end{bmatrix} \mathbf{1} = \boldsymbol{X}\boldsymbol{F}(N\boldsymbol{h}_d + N\boldsymbol{v}_1) \tag{9}$$

Unfortunately, the configuration of the first element collides with the one considered for the direct channel. To overcome this, we assume that the contribution of the direct channel is considerably greater than the one of the first element. Hence, the impulse response of the former can be estimated through (9). This estimation has proved to be good enough when evaluated in dataset1, where, as there are $4N$ pilot transmissions, an LS estimation could be performed for the N IRS channels and the direct channel.

For the other N-1 IRS channels, leaving out the first one, an LS estimation was performed, which was followed by a smoothing procedure over the pilots dimension in order to further reduce the noise in the estimation. As for the channel of the first element, we considered it to be equal to the one of

an adjacent element, namely the 65-th element. Once again, this proved to be suitable when evaluated with dataset1[1].

B. STM

After obtaining an accurate channel estimation, the next step is to optimize the IRS for achieving the higher average rate possible. Therefore, the goal is to find the $\boldsymbol{\omega_\theta}$ matrix that maximizes (3), which can be re-expressed as:

$$R_{\mathrm{u}} = \frac{B}{K+M-1} \sum_{\nu=0}^{K-1} \log_2 \left(1 + \frac{P\left|\boldsymbol{f}_\nu^{\mathrm{H}} \boldsymbol{h}_d + \boldsymbol{f}_\nu^{\mathrm{H}} \boldsymbol{V}^{\mathrm{T}} \boldsymbol{\omega_\theta}\right|^2}{BN_0}\right), \quad (10)$$

where $\boldsymbol{f}_\nu$ is the ν-th row of the DFT matrix F. Note that $\boldsymbol{h}_d$ and $\boldsymbol{V}$ now correspond to the channel estimates obtained in the previous subsection, although we decided to keep this notation for the sake of simplicity.

An upper bound for the achievable rate in (10) is achieved in the unrealistic case in which the state of the elements would be selected to get all the subcarriers aligned in phase at the receiver. In this case the rate would be:

$$R_{\mathrm{u}}^{\mathrm{UB}} = \frac{B}{K+M-1} \sum_{\nu=0}^{K-1} \log_2 \left(1 + \frac{P\left(\left|\boldsymbol{f}_\nu^{\mathrm{H}} \boldsymbol{h}_d\right| + \left\|\boldsymbol{f}_\nu^{\mathrm{H}} \boldsymbol{V}^{\mathrm{T}}\right\|_1\right)^2}{BN_0}\right), \quad (11)$$

The first approach to optimize the IRS has been done using the strongest-tap maximization (STM) [11]. As explained in [12], the intuition behind this technique is that it is easier to optimize the equivalent channel in time rather than the channel in frequency considering $M \ll K$. In order to select the tap to be maximized, we first calculate which one could be potentially larger. This is implemented by:

$$l^{\mathrm{opt}} = \underset{l=0\ldots M-1}{\operatorname{argmax}} \left| |h_d[l]| + \sum_{m=1}^{M} |v_m[l]| \right|^2 \quad (12)$$

For maximizing the l^{opt}-th tap, we need to shift the phase from the channel of each IRS element so that all channels are in phase at the receiver. As the direct channel is the uncontrollable element, the channel from all the elements must be aligned to it:

$$\varphi_n = -\angle v_n[l^{\mathrm{opt}}] + \angle h_d[l^{\mathrm{opt}}] \quad (13)$$

This expression must be adapted to our particular set-up, as it is only possible to select between two phase shifts. Therefore, the final state is:

$$\theta_n = \begin{cases} 1, & \text{if } -\pi/2 < \varphi_n < <\pi/2 \\ -1, & \text{otherwise} \end{cases} \quad (14)$$

One of the main disadvantages of this approach is that in NLOS channels there is no dominant tap. Hence, the expected performance in this type of scenario would be worse. Moreover, the reduced number of states of our particular problem is a handicap for achieving the maximum rate that potentially the STM would obtain for a particular tap. This might cause that for some users, it is possible to obtain a higher rate *by maximizing a different tap instead* of the one obtained in (12). Hence, we implemented a variation of STM that aims at maximizing not the largest, but the k_l-th largest channel tap. The configuration that obtains the higher rate for that user would be the output of the algorithm. We refer to this algorithm as STM-k, and includes STM as a special case.

We also observed that the equivalent channel for some line-of-sight (LOS) users, (e.g., user 9), after applying STM-k has two main taps larger than the rest. When these two dominant taps are not correctly aligned in phase, it could lead to a very frequency-selective channel (i.e., as in a two-ray channel). A possible solution for this type of users is to modify the expression from (13) to consider both paths:

$$\varphi_n = -\angle v_n[l^{opt}] + f(\angle h_d[l_1^{opt}, \angle h_d[l_2^{opt}], ..\angle h_d[l_{K_l}^{opt}]]) \quad (15)$$

where $f(\cdot)$ is a linear function that combines the phases of the most important taps. We refer to this last algorithm as generalized STM-k. Because of the strong phase quantization at the IRS, this function has been developed heuristically, selecting for each result the function with the best performance. For example, for the aforementioned user 9, we used the function:

$$f(\angle h_d[l_1^{opt}], \angle h_d[l_2^{opt}]) = \angle h_d[l_1^{opt}] + \angle h_d[l_2^{opt}] \quad (16)$$

C. Fine tuning

As a final step, we aim to improve the solution for the IRS configuration $\boldsymbol{\omega_\theta}$ obtained from the generalized STM-k solution. For this purpose, we try to optimize the $\boldsymbol{\omega_\theta}$ matrix by searching locally new solutions that improve the previous one. In this way, it is intended to simplify the problem, by reducing the dimension of the matrix, and, therefore, to accelerate the optimization process.

We use the *patternsearch* numerical optimization method in MATLAB [13] to obtain the maximization of the objective function (10). The possible values of the IRS matrix are restricted in the proposed problem statement, but in order to find a solution we proceed as in [14]. We relax the constraint so that $-1 \leq \theta_n \leq 1$ and then discretize the solution to fit the original constraints.

The local solution is based on a clustering method in which we iterate over $m \times n$ sub-matrices of $\boldsymbol{\omega_\theta}$ finding combinations that improve the global metric in (4) by simply changing some values in that sub-matrix and leaving the rest of the matrix $\boldsymbol{\omega_\theta}$ unchanged. In addition, we introduce a 50 % overlap over the different clusters to improve the result. Although different combinations of m and n have been tried, the best solution has been found for $m = n = 8$.

Finally, to consider the local convergence of the optimization method, *patternsearch*, it is decided to run the program several times and retain the best solution of the iterations by changing the initial generalized STM-k solution. Specifically, we tested multiple variations over the generalized STM-k with different linear combinations of the strongest taps that improve the rate on certain users.

IV. Numerical results

In this section we study the performance of the two main tasks of the challenge proposed in the SP Cup: the channel estimation and the IRS optimization. First, for the analysis of the channel estimation performance, we considered the one using the $4N$ pilot signal from the *dataset1* as ground truth.

[1]In a first step, we neglected the direct link and performed an initial estimation for the channels. This was used to determine whether the IRS structure was square (i.e. 64×64 or rectangular), by calculating the spatial correlation matrix for the IRS elements.

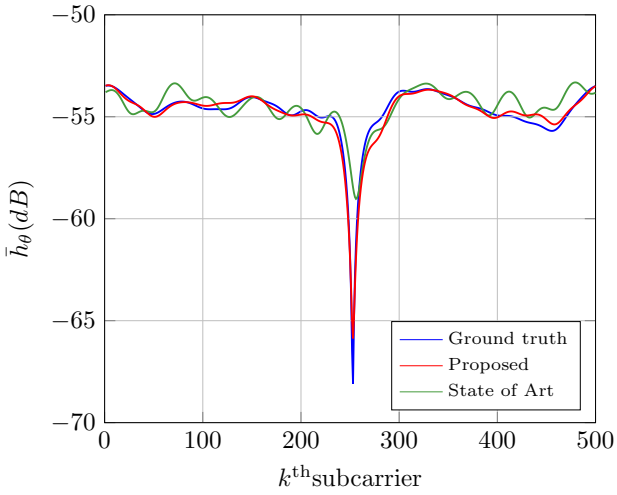

Fig. 2. Comparison of the channel estimation for a given RIS configuration between the ground truth, the stat-of-art algorithm and the proposed estimator.

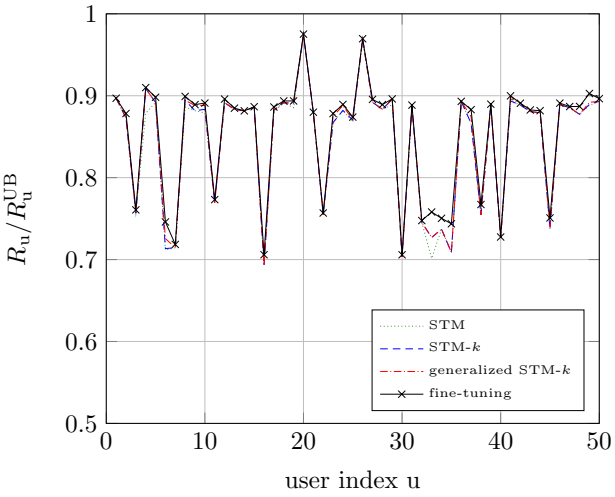

Fig. 3. Achievable rates per-user normalized to the upper-bound rate, for the STM, STM-k, generalized STM-k and the fine-tuning algorithms.

In Fig. 2, we show a comparison between such ground truth reference and the estimation obtained with the proposed algorithm. Moreover, the channel estimation obtained using the algorithm proposed in [12] is also used as a reference, which is referred to as the state-of-art estimator. It is possible to visually observe that our implementation outperforms the estimator initially proposed.

In the IRS optimization, the figure of merit for this challenge has been to achieve the highest averaged binary rate following the expression (3), but considering the double of the rate for the NLOS users.

In Fig. 3, we represent the normalized achievable rate obtained when using each of the previously described algorithms. We see that the STM alone does not provide the best possible rate, and that the different variations over the STM algorithm provide reasonable rate improvements for specific users. We also see that the fine-tuning algorithm also allows for a final rate boost over some users, especially for those far from the achievable upper bound rate.

In Fig. 3, it is possible to distinguish between two level of performances. Indeed, the users with the lowest normalized rate are those which are NLOS. The reason for this is that the STM, as it tries to improve just a single tap of $\boldsymbol{h_\theta}$, has a poor performance for this type of channels. This behaviour is explained in detail in [12].

V. Conclusion

We presented the set of algorithms that we implemented for the IEEE Signal Processing Cup 2021 challenge. A combination of modified versions of several state-of-the-art algorithms has allowed us to improve the performance obtained with the baseline STM method. The estimated average achievable rate with this solution is $\overline{R} \approx 117.5879$ Mbps. This solution achieved the ninth place out of a total of 30 final submissions, where the first place (University of Science and Technology, Kraków, Poland), obtained an average rate only a 0.6% higher than ours.

Acknowledgements

Thanks to Pablo Mateos Ruiz, Lorena Pérez Aguilar, Francisco Jurado Romero for their contribution in the early stages of this contest. Thanks to Rafael Larrosa Jiménez for his valuable support for setting up the simulations using PICASSO supercomputer.

References

[1] Forecast, GMDT, "Cisco visual networking index: global mobile data traffic forecast update, 2017–2022," *Update*, vol. 2017, p. 2022, 2019.

[2] S. Chen and J. Zhao, "The requirements, challenges, and technologies for 5g of terrestrial mobile telecommunication," *IEEE communications magazine*, vol. 52, no. 5, pp. 36–43, 2014.

[3] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Prospective multiple antenna technologies for beyond 5g," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1637–1660, 2020.

[4] N. Alliance, "5g white paper," *Next generation mobile networks, white paper*, vol. 1, 2015.

[5] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE transactions on wireless communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[6] O. Tsilipakos, A. C. Tasolamprou, A. Pitilakis, F. Liu, X. Wang, M. S. Mirmoosa, D. C. Tzarouchis, S. Abadal, H. Taghvaee, C. Liaskos *et al.*, "Toward intelligent metasurfaces: The progress from globally tunable metasurfaces to software-defined metasurfaces with an embedded network of controllers," *Advanced Optical Materials*, vol. 8, no. 17, p. 2000783, 2020.

[7] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019.

[8] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4157–4170, 2019.

[9] T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, "Coding metamaterials, digital metamaterials and programmable metamaterials," *Light: Science & Applications*, vol. 3, no. 10, pp. e218–e218, 2014.

[10] E. Björnson. (2021, May) IEEE Signal Processing Cup 2021. [Online]. Available: https://github.com/emilbjornson/SP_Cup_2021

[11] B. Zheng and R. Zhang, "Intelligent Reflecting Surface-Enhanced OFDM: Channel Estimation and Reflection Optimization," *IEEE Wirel. Commun. Lett.*, vol. 9, no. 4, pp. 518–522, 2020.

[12] E. Björnson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti, and E. de Carvalho, "Reconfigurable Intelligent Surfaces: A Signal Processing Perspective With Wireless Applications," 2021.

[13] I. The MathWorks. Find minimum of function using pattern search - MATLAB. [Online]. Available: https://www.mathworks.com/help/gads/patternsearch.html

[14] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, "Massive MIMO in real propagation environments: Do all antennas contribute equally?" *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 3917–3928, 2015.