ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-7605-5/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICASSP39728.2021.9413598

META-LEARNING FOR 6G COMMUNICATION NETWORKS WITH RECONFIGURABLE INTELLIGENT SURFACES

Minchae Jung[1] *and Walid Saad*[2]

[1]Department of Electronic Engineering, Soonchunhyang University, Asan, South Korea
[2]Wireless@VT, Bradley Department of Electrical and Computer Engineering, Virginia Tech, VA, USA

ABSTRACT

Channel acquisition is one of the main challenges in a reconfigurable intelligent surface (RIS) system due to the passive nature of an RIS. In order to accurately estimate RIS channels, a large number of pilot symbols are required, which could yield a severe performance degradation in terms of the spectral efficiency (SE). In this paper, a practical channel acquisition and passive beamforming technique is proposed using a limited number of pilot symbols in an RIS-assisted cellular network. In particular, the proposed technique relies on a meta-learning framework. To address practical RIS challenges, the problem of maximizing the instantaneous SE is formulated and a novel approach to solve this optimization problem is developed. The proposed algorithm enables an RIS to select the optimal phase shift matrix without the need for perfect channel state information. Also, the trained parameter resulting from the proposed meta-learning algorithm can be guaranteed to converge to an optimal solution. Simulation results show a comparable performance to an exhaustive search method with a few training symbols which validates the advantages of meta-learning for an RIS system.

***Index Terms*—** RIS, meta-learning, spectral efficiency.

1. INTRODUCTION

The concept of a reconfigurable intelligent surface (RIS) is rapidly emerging topic to support the demand for upcoming Internet of Things (IoT) services without using additional radio resources [1–12]. Given that future man-made structures, such as buildings and walls, are expected to be electromagnetically active, these structures can be exploited to provide wireless connectivity to future 6G services [13, 14], via the emerging concept of a RIS. An RIS is composed of a large number of passive reflecting elements, each of which controls the incident radio frequency (RF) signals impinging on it, in terms of frequency, amplitude, and phase, and reflect it to a destination, without additional signal processing. In an RIS system, channel acquisition is one of the main challenges given that RISs encompass reflective and passive elements and cannot estimate accurate RIS channels.

In [15] and [16], under a bounded channel state information (CSI) error, joint active and passive beamformers are designed based on the worst case optimization approach. In [2] and [17], the authors analyze the asymptotic performance of an RIS system under realistic channel estimation errors. The works in [18] and [19] considered a statistical CSI error distribution and designed the passive shift matrix. Hence, in order to accurately estimate the RIS-related channels, a large number of pilot symbols, i.e., at least the number of elements on RIS, are required. However, this assumption is impractical for a limited channel coherence time and will result in severe performance degradation in terms of the spectral efficiency (SE).

The main contribution of this paper is to develop a novel meta-learning based channel acquisition and beamforming technique that can achieve an optimal performance with a few training symbols. Meta-learning [20] provides a way to automate the selection of an initialization of model parameters (i.e., inductive bias) and allows a quick adaptation to a new task with limited training examples. Based on this characteristic of meta-learning, an inductive bias is trained at a base station (BS) based on the previously received pilot signals so that an inductive bias can be located as close as possible to the optimally trained models. Then, this updated inductive bias allows a quick adaptation for a newly observed RIS channel with a small number of training steps and a limited number of pilots. In particular, the BS sends different control signals to an RIS controller for different uplink pilot symbols and estimates the received downlink signal power for pilot symbols. This received signal power can be defined as a reward and the BS performs the meta-training to update the inductive bias based on this reward. For a new RIS channel, the BS performs a meta-test for new downlink pilot symbols based on the updated inductive bias and eventually estimates an effective RIS channel within a few training steps. Given this estimated RIS channel, we design a passive beamformer that achieves a performance that is comparable to conventional techniques with improved spectral efficiency.

This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant NRF- 2016R1A6A3A11936259 and in part by the U.S. National Science Foundation under Grants IIS-1633363 and CNS-2030215.

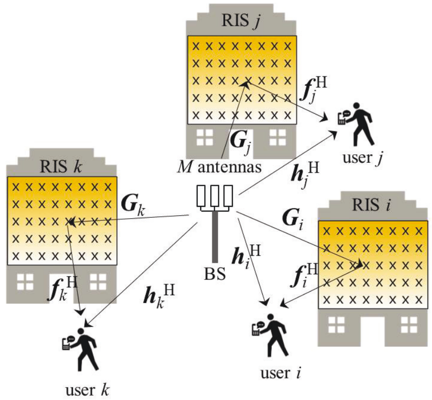

Fig. 1. Illustrative system model of the considered RIS-based downlink system.

Simulation results show that, using a limited number of pilot symbols, the proposed meta-learning algorithm can achieve the optimal performance that can be obtained by the exhaustive search method.

The rest of this paper is organized as follows. The system model is presented in Section 2. In Section 3, we formulate the proposed optimization problem. The meta-learning algorithm is proposed in Section 3.1 and its convergence is proved in Section 3.2. In Section 4, we provide simulation results. Finally, conclusions are drawn in Section 5.

2. SYSTEM MODEL

Consider the downlink of an RIS-based wireless network that consists of a single BS with M antennas, K RISs with N reflecting elements, and K single-antenna users. Each user receives the downlink signal directly from the BS and also receives, simultaneously, the reflected signal via an RIS, as shown in Fig. 1. In our system model, each user is connected to different RISs depending on its location (i.e., one user per one RIS) and the BS allocates orthogonal resources among users using an appropriate resource allocation scheme. Let $\boldsymbol{h}_k \in \mathbb{C}^{M\times 1}$, $\boldsymbol{G}_k \in \mathbb{C}^{N\times M}$, and $\boldsymbol{f}_k \in \mathbb{C}^{N\times 1}$ be, respectively, the wireless channels between the BS and user k, between the BS and RIS k, and between RIS k and user k. Then, the received signal at user k will be:

$$y_k = \sqrt{P}\left(\boldsymbol{h}_k^{\mathrm{H}} + \boldsymbol{f}_k^{\mathrm{H}}\boldsymbol{\Phi}_k\boldsymbol{G}_k\right)\boldsymbol{w}_k x_k + n_k, \qquad (1)$$

where P is the downlink transmit power at the BS, $\boldsymbol{\Phi}_k$ is the phase shift matrix adopted at RIS k, $\boldsymbol{w}_k \in \mathbb{C}^{M\times 1}$ is the transmit beamforming vector for user k, and x_k is downlink transmit symbol for user k with noise term $n_k \sim \mathcal{CN}(0, N_0)$. The phase shift matrix $\boldsymbol{\Phi}_k$ is given by:

$$\boldsymbol{\Phi}_k = \mathrm{diag}\left(a_1\phi_1, a_2\phi_2, \cdots, a_N\phi_N\right), \qquad (2)$$

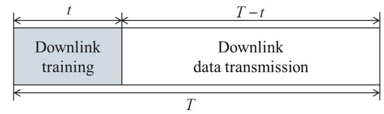

Fig. 2. Illustrative downlink frame structure with a pilot training period t and a data transmission period $T-t$.

where $a_n \in [0,1]$ and $\phi_n = e^{j\theta_n}$ where $\theta_n \in [0, 2\pi)$. For analytical simplicity, we assume that each reflecting element is designed to maximize the reflected signal, i.e., $a_n = 1$, $\forall n$. For practical RIS implementation, the BS will send a control signal with limited bits to the RIS controller and each RIS selects a phase shift matrix from the feasible set of discrete phase shifts based on the received control signal. We use a uniform discrete Fourier transform (DFT)-based codebook. Then, the i-th phase shift matrix for user k is obtained as follows:

$$\boldsymbol{\Phi}_k^i = \mathrm{diag}\left(1, e^{j\frac{2\pi}{2^b}i}, e^{j\frac{2\pi}{2^b}2i}, \cdots, e^{j\frac{2\pi}{2^b}(N-1)i}\right), \qquad (3)$$

where $i = 0, 1, \ldots, 2^b - 1$ and b is the number of RIS control bits.

From (1), the received signal-to-noise ratio (SNR) at user k is obtained as follows:

$$\gamma_k = \frac{P}{N_0}\left|\left(\boldsymbol{h}_k^{\mathrm{H}} + \boldsymbol{f}_k^{\mathrm{H}}\boldsymbol{\Phi}_k\boldsymbol{G}_k\right)\boldsymbol{w}_k\right|^2. \qquad (4)$$

Note that the BS transmits downlink signals after sending the pilot signals within the channel coherence time T. We consider the downlink frame structure shown in Fig. 2, where the coherence time T is divided into a period t for pilot training and a period $T-t$ for data transmission. For notational simplicity, we assume that the BS transmits t pilot symbols during t period. Each user estimates the CSI for the effective channel $\boldsymbol{h}_k^{\mathrm{H}} + \boldsymbol{f}_k^{\mathrm{H}}\boldsymbol{\Phi}_k\boldsymbol{G}_k$ during the pilot training period and sends back the estimated CSI to the BS. Then, the BS can have knowledge about the effective channels $\boldsymbol{h}_k^{\mathrm{H}} + \boldsymbol{f}_k^{\mathrm{H}}\boldsymbol{\Phi}_k\boldsymbol{G}_k$ $\forall k$, and can design $\boldsymbol{w}_k$ and $\boldsymbol{\Phi}_k$ based on this information. Considering the downlink frame structure shown in Fig. 2, the instantaneous spectral efficiency (SE) at user k can be obtained using (4) as follows:

$$R_k = \left(1 - \frac{t}{T}\right)\log\left(1 + \gamma_k\right). \qquad (5)$$

3. META-LEARNING FRAMEWORK FOR PRACTICAL RIS SYSTEM

Given the received SNR from (4), it is well known that the maximum SNR can be achieved by using a maximum ratio transmission (MRT) precoder such that [10, 11]:

$$\boldsymbol{w}_k = \frac{\boldsymbol{h}_k + \boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{\Phi}_k^{\mathrm{H}}\boldsymbol{f}_k}{\left\|\boldsymbol{h}_k + \boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{\Phi}_k^{\mathrm{H}}\boldsymbol{f}_k\right\|}. \qquad (6)$$

We then formulate the following optimization problem whose goal is to maximize the instantaneous SE at user k with respect to $\mathbf{\Phi}_k$:

$$\max_{\mathbf{\Phi}_k} \left(1-\frac{t}{T}\right)\log\left(1+\frac{P}{N_0}\left\|\boldsymbol{h}_k+\boldsymbol{G}_k^{\mathrm{H}}\mathbf{\Phi}_k^{\mathrm{H}}\boldsymbol{f}_k\right\|^2\right), \quad (7)$$

$$\text{s.t. } \mathbf{\Phi}_k \in \mathcal{F}, \quad (7a)$$

where $\mathcal{F}=\left\{\mathbf{\Phi}_k^i\right\}_{i=0}^{2^b-1}$. The optimal performance can be obtained by an exhaustive search method for all possible $\mathbf{\Phi}_k \in \mathcal{F}$. The exhaustive search method compares the instantaneous SE, R_k, for all possible $\mathbf{\Phi}_k \in \mathcal{F}$ and selects the optimal $\mathbf{\Phi}_k$ that maximizes R_k. In order to calculate R_k, the acquisition of accurate CSI for the RIS-related channels (i.e., $\boldsymbol{h}_k$, $\boldsymbol{G}_k$, and $\boldsymbol{f}_k$) is required. However, it is practically difficult due to the passive nature of RIS and the limited pilot training period. Therefore, we propose a novel meta-learning based channel acquisition and beamforming technique.

3.1. Proposed Meta-Learning Algorithm

In the proposed meta-learning algorithm, a total of t pilot symbols are divided into t_{L} symbols for meta-training and t_{T} symbols for meta-testing, and different phase shift matrices are used for each pilot symbol duration. Then, user k can measure the values of y_k and γ_k for t pilot symbols and send this information back to the BS. The meta-training and meta-testing datasets are defined, respectively, as $\mathcal{D}^{\mathrm{L}}=\left\{\mathcal{D}_k^{\mathrm{L}}\right\}_{k=1}^{K}$ and $\mathcal{D}^{\mathrm{T}}=\left\{\mathcal{D}_k^{\mathrm{T}}\right\}_{k=1}^{K}$, where $\mathcal{D}_k^{\mathrm{L}}=\{(y_k^p,\gamma_k^p)\}_{p=1}^{t_{\mathrm{L}}}$ and $\mathcal{D}_k^{\mathrm{T}}=\{(y_k^p,\gamma_k^p)\}_{p=t_{\mathrm{L}}+1}^{t}$. Here, (y_k^p,γ_k^p) is the p-th pilot-related data pair. Hence, the BS has knowledge about the datasets $\mathcal{D}^{\mathrm{L}}$ and $\mathcal{D}^{\mathrm{T}}$, and will use $\mathcal{D}^{\mathrm{L}}$ for meta-training and also use $\mathcal{D}^{\mathrm{T}}$ for meta-testing. An inductive bias is trained based on a dataset $\mathcal{D}^{\mathrm{L}}$ at the BS so that an inductive bias can be as close as possible to the optimally trained model. Then, this updated inductive bias allows a quick adaptation for a new channel with a small number of data from the dataset $\mathcal{D}^{\mathrm{T}}$. The proposed meta-learning algorithm is summarized in Algorithm 1. In Algorithm 1, we define the standard cross-entropy loss function as:

$$L\left(\bar{\boldsymbol{\theta}}\right)=-\sum\nolimits_{i=0}^{2^b-1} p\left(\mathbf{\Phi}^i|y,\gamma,\bar{\boldsymbol{\theta}}\right). \quad (8)$$

In fact, the expectation-maximization (EM) method is used as a standard tool to train the meta-learning model based on the meta-training dataset as follows:

$$p\left(\mathbf{\Phi}^i|y,\gamma,\bar{\boldsymbol{\theta}}\right)=\mathrm{E}_{\boldsymbol{\theta}^i}\left[p\left(\mathbf{\Phi}^i|y,\gamma,\boldsymbol{\theta}^i,\bar{\boldsymbol{\theta}}\right)\right]. \quad (9)$$

From (8) and (9), the cross-entropy loss function is obtained as:

$$L\left(\bar{\boldsymbol{\theta}}\right)=\lim\nolimits_{G\to\infty}\frac{1}{G}\sum\nolimits_{g=1}^{G} L_g\left(\bar{\boldsymbol{\theta}}\right), \quad (10)$$

where $L_g\left(\bar{\boldsymbol{\theta}}\right)=-\sum_i p\left(\mathbf{\Phi}^i|y,\gamma,\boldsymbol{\theta}^i,\bar{\boldsymbol{\theta}}\right)$ for the g-th realization of the dataset $\mathcal{D}^{\mathrm{L}}$. However, the computational complexity required to calculate $L\left(\bar{\boldsymbol{\theta}}\right)$ in (10) will be high and (9) may not be obtained, accurately, in practical fading channels. Hence, in the proposed meta-learning algorithm, we use the following practical cross-entropy loss function:

Algorithm 1 Proposed Meta-Learning Algorithm

1: **Initialize** inductive bias $\boldsymbol{\theta}$, $\boldsymbol{\theta}^0=\boldsymbol{\theta}$, and $i=0$.
2: **For each time slot**:
3: **For meta-training data** $\mathcal{D}_k^{\mathrm{L}}\in\mathcal{D}^{\mathrm{L}}$:
4: Calculate

$$\boldsymbol{\theta}^{i+1}=\boldsymbol{\theta}^i-\alpha\nabla_{\boldsymbol{\theta}^i}L\left(\boldsymbol{\theta}^i\right),$$
$$i\leftarrow i+1.$$

5: **End for**
6: Set $\bar{\boldsymbol{\theta}}=\boldsymbol{\theta}^i$.
7: **For meta-testing data** $\mathcal{D}_k^{\mathrm{T}}\in\mathcal{D}^{\mathrm{T}}$:
8: Calculate

$$\bar{\boldsymbol{\theta}}\leftarrow\bar{\boldsymbol{\theta}}-\beta\nabla_{\bar{\boldsymbol{\theta}}}L\left(\bar{\boldsymbol{\theta}}\right).$$

9: **End for**
10: Select $\mathbf{\Phi}_k\in\mathcal{F},\forall k$.
11: **End for**

$$L\left(\bar{\boldsymbol{\theta}}\right)=\frac{1}{G}\sum\nolimits_{g=1}^{G} L_g\left(\bar{\boldsymbol{\theta}}\right), \quad (11)$$

where G is a finite value. Given this cross-entropy loss function, we will prove the convergence of the proposed meta-learning algorithm.

3.2. Convergence Proof

Given a standard cross-entropy loss function, $L\left(\boldsymbol{\theta}\right)$ is strongly convex and twice differentiable such that: $l\boldsymbol{I}\preceq\nabla^2 L\left(\boldsymbol{\theta}\right)\preceq u\boldsymbol{I}$, where l and u are positive constants. From Step 4 in Algorithm 1, we have

$$\boldsymbol{\theta}^{i+1}-\boldsymbol{\theta}^i=-\alpha\nabla L\left(\boldsymbol{\theta}^i\right). \quad (12)$$

By using the upperbound of the second derivative of $L\left(\boldsymbol{\theta}\right)$, we have:

$$\begin{aligned} L\left(\boldsymbol{\theta}^{i+1}\right) &\le L\left(\boldsymbol{\theta}^i\right)+\left(\boldsymbol{\theta}^{i+1}-\boldsymbol{\theta}^i\right)^{\mathrm{T}}\nabla L\left(\boldsymbol{\theta}^i\right)+\frac{u}{2}\left\|\boldsymbol{\theta}^{i+1}-\boldsymbol{\theta}^i\right\|^2 \\ &= L\left(\boldsymbol{\theta}^i\right)-\alpha\left(1-\frac{u\alpha}{2}\right)\left\|\nabla L\left(\boldsymbol{\theta}^i\right)\right\|^2. \quad (13)\end{aligned}$$

From (11) and (13), the proposed cross-entropy loss function satisfies the following inequality:

$$\sum L_g\left(\boldsymbol{\theta}^{i+1}\right)\le\sum L_g\left(\boldsymbol{\theta}^i\right)-\frac{\alpha}{G}\left(1-\frac{u\alpha}{2}\right)\sum\left\|\nabla L_g\left(\boldsymbol{\theta}^i\right)\right\|^2.$$

Since $L\left(\boldsymbol{\theta}\right)$ is strongly convex function, we have

$$\begin{aligned}\left\|\nabla L_g\left(\boldsymbol{\theta}^i\right)\right\|^2 &= \left\|\nabla L_g\left(\boldsymbol{\theta}^i\right)-\nabla L_g\left(\boldsymbol{\theta}^*\right)\right\|^2 \\ &\ge l\left(\nabla L_g\left(\boldsymbol{\theta}^i\right)-\nabla L_g\left(\boldsymbol{\theta}^*\right)\right)^{\mathrm{T}}\left(\boldsymbol{\theta}^i-\boldsymbol{\theta}^*\right) \\ &\ge l\left(L_g\left(\boldsymbol{\theta}^i\right)-L_g\left(\boldsymbol{\theta}^*\right)\right), \quad (14)\end{aligned}$$

where $\boldsymbol{\theta}^*$ is the optimal solution that results in $\nabla L_g\left(\boldsymbol{\theta}^*\right)=0$. Hence, we have the following inequality when $u\alpha/2<1$:

$$\sum L_g\left(\boldsymbol{\theta}^{i+1}\right)\le\sum L_g\left(\boldsymbol{\theta}^i\right)-\mu\left(\sum L_g\left(\boldsymbol{\theta}^i\right)-L_g\left(\boldsymbol{\theta}^*\right)\right), \quad (15)$$

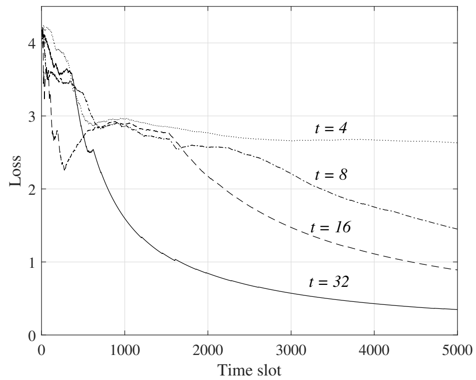

Fig. 3. Convergence of the loss function resulting from the proposed meta-learning algorithm.

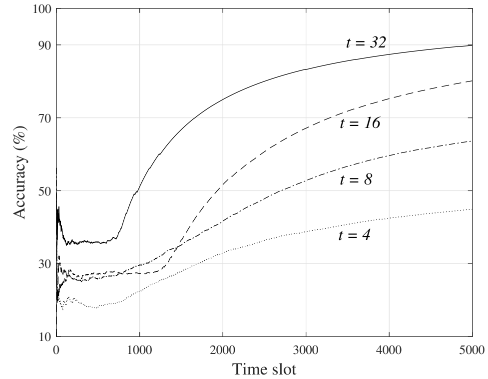

Fig. 4. The training and validation process of the proposed meta-learning algorithm.

where $\mu = \frac{l\alpha}{G}\left(1 - \frac{u\alpha}{2}\right)$. If $G \to \infty$, μ goes to zero and we then have

$$\sum L_g\left(\boldsymbol{\theta}^{i+1}\right) \leq \sum L_g\left(\boldsymbol{\theta}^{i}\right). \tag{16}$$

We can observe from (16) that $\sum L_g\left(\boldsymbol{\theta}^{i}\right)$ is a decreasing function with respect to i when $G \to \infty$. Since $\sum L_g\left(\boldsymbol{\theta}^{i}\right) \geq 0$, Algorithm 1 converges to the optimal model when $G \to \infty$, which completes the convergence proof of the EM method in (9). Next, we prove the convergence of Algorithm 1 with a finite G. By adding $-\sum L_g\left(\boldsymbol{\theta}^{*}\right)$ to both sides of (15), we have the following inequalities:

$$\sum L_g\left(\boldsymbol{\theta}^{i+1}\right) - L_g\left(\boldsymbol{\theta}^{*}\right) \leq \sum \left(L_g\left(\boldsymbol{\theta}^{i}\right) - L_g\left(\boldsymbol{\theta}^{*}\right)\right)(1-\mu),$$

$$\sum L_g\left(\boldsymbol{\theta}^{i}\right) - L_g\left(\boldsymbol{\theta}^{*}\right) \leq \sum \left(L_g\left(\boldsymbol{\theta}^{0}\right) - L_g\left(\boldsymbol{\theta}^{*}\right)\right)(1-\mu)^{i}.$$

If $0 < \mu < 1$, $L_g\left(\boldsymbol{\theta}^{i}\right)$ converges to its optimal value $L_g\left(\boldsymbol{\theta}^{*}\right)$ as i increases. Hence, the proposed meta-learning will converge to the optimal model as i increases, when we set α and G to satisfy the condition $0 < \mu < 1$. Note that the EM method is generally used as a standard tool to solve the problem of the maximum likelihood estimation in (9) and (10). Since (10) can be obtained when $G \to \infty$, the computation of the EM method is generally of high complexity. However, Algorithm 1 provides a general convergence rate with an arbitrary G. If we set α to satisfy the condition $0 < l\alpha\left(1 - \frac{u\alpha}{2}\right) < 1$, then the proposed meta-learning will always converge to the optimal model regardless of G.

4. SIMULATION RESULTS

For our simulations, we consider the following system parameters: $T = 100$, $P/N_0 = 1$, $M = 4$, $N = 100$, $K = 1$, $t_{\mathrm{L}} = t_{\mathrm{T}} = t/2$ and $b = 6$. The wireless channels, $\boldsymbol{h}_k$, $\boldsymbol{G}_k$, and $\boldsymbol{f}_k$, are generated by spatially correlated Rician fading as in [1, 2]. The Rician factors correspond to the BS-RIS link, RIS-user link, and BS-user link are, respectively, set to 10^4, 0, and 0. We use a deep reinforcement learning (RL) to seek a solution that maximizes the instantaneous SE. The states consist of y_k^p and we define γ_k as a reward.

In Fig. 3, we analyze the convergence of the proposed meta-learning algorithm, in terms of the training loss. As proved in Section 3.2, the training loss decreases as time slot increases (i.e., i increases). Moreover, we can observe from that the training loss with $t = 32$ converges faster than the others, due to accurately trained inductive bias $\bar{\boldsymbol{\theta}}$.

Fig. 4 shows the accuracy of the proposed meta-learning algorithm. We define the accuracy as the ratio between γ_k values resulting from the proposed algorithm and the exhaustive search method. The exhaustive search method compares γ_k for all possible $\boldsymbol{\Phi}_k \in \mathcal{F}$ and therefore, it requires $2^b = 64$ pilot symbols. Hence, the exhaustive search method can only use $\frac{T-2^b}{T} = 36$ % of radio resource for data transmission. However, when $t = 32$, our algorithm can use $\frac{T-2^b}{T} = 68$ % of radio resource for data transmission with 90 % accuracy compared to the optimal performance. Furthermore, when $t = 16$, our algorithm can use 84 % of radio resource for data transmission with 80 % accuracy compared to the optimal performance. Therefore, our algorithm can achieve better performance than the exhaustive search method in terms of the instantaneous SE and average SE.

5. CONCLUSION

In this paper, we have investigated the problem of practical RIS system such as channel acquisition and phase shift matrix design. We have proposed a novel framework to optimize instantaneous SE and formulated the problem of maximizing the instantaneous SE, that enables an RIS to select the optimal phase shift matrix without the need of the perfect CSI. We have shown that the proposed meta-learning algorithm can achieve the performance of optimal learning model and also shown the convergence of the proposed algorithm. Furthermore, the problem of RIS channel acquision was solved using our meta-learning algorithm to reduce the pilot overhead and enhance the instantaneous SE.

6. REFERENCES

[1] M. Jung, W. Saad, Y. Jang, G. Kong, and S. Choi, "Reliability analysis of large intelligent surfaces (LISs): Rate distribution and outage probability," *IEEE Wireless Communications Letters*, vol. 8, no. 6, pp. 1662–1666, Dec. 2019.

[2] M. Jung, W. Saad, Y. Jang, G. Kong, and S. Choi, "Performance analysis of large intelligent surfaces (LISs): Asymptotic data rate and channel hardening effects," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 2052–2065, Mar. 2020.

[3] C. Chaccour, M. N. Soorki, W. Saad, M. Bennis, and P. Popovski, "Risk-based optimization of virtual reality over terahertz reconfigurable intelligent surfaces," in *IEEE International Conference on Communications (ICC)*, Dublin, Ireland, Jun. 2020.

[4] T. J. Cui, S. Liu, and L. Zhang, "Information metamaterials and metasurfaces," *Journal of Materials Chemistry C*, vol. 5, pp. 3644–3668, 2017.

[5] E. Basar, M. D. Renzo, J. Rosny, M. Debbah, M. S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116753–117773, Sep. 2019.

[6] L. Dai, B. Wang, M. Wang, X. Yang, J. Tan, S. Bi, S. Xu, F. Yang, Z. Chen, M. D. Renzo, C. Chae, and L. Hanzo, "Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, pp. 45913–45923, 2020.

[7] W. Tang, J. Y. Dai, M. Chen, X. Li, Q. Cheng, S. Jin, K. Wong, and T. J. Cui, "Programmable metasurface-based RF chain-free 8PSK wireless transmitter," *Electronics Letters*, vol. 55, no. 7, pp. 417–420, 2019.

[8] W. Tang, X. Li, J. Y. Dai, S. Jin, Y. Zeng, Q. Cheng, and T. J. Cui, "Wireless communications with programmable metasurface: Transceiver design and experimental results," *available online: arxiv.org/abs/1811.08119*, Nov. 2018.

[9] Y. Han, W. Tang, S. Jin, C. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical CSI," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8238–8242, Aug. 2019.

[10] C. Huang, G. Alexandropoulos, A. Zaponne, M. Debbah, and C. Yuen, "Energy efficient multi-user MISO communication using low resolution large intelligent surfaces," in *IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, UAE, Dec. 2018.

[11] M. Jung, W. Saad, M. Debbah, and C. S. Hong, "On the optimality of active reconfigurable intelligent surfaces (RISs): Passive beamforming, modulation, and resource allocation," *IEEE Transactions on Wireless Communications*, to appear, 2021.

[12] G. Lee, M. Jung, A. T. Z. Kasgari, W. Saad, and M. Bennis, "Deep reinforcement learning for energy-efficient networking with reconfigurable intelligent surfaces," in *IEEE International Conference on Communications (ICC)*, Dublin, Ireland, Jun. 2020.

[13] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, May 2020.

[14] M. Mozaffari, A. T. Z. Kasgari, W. Saad, M. Bennis, and M. Debbah, "Beyond 5G with UAVs: Foundations of a 3D wireless cellular network," *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 357–372, Jan. 2019.

[15] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, "Robust and secure wireless communications via intelligent reflecting surfaces," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2637–2652, Nov. 2020.

[16] G. Zhou, C. Pan, H. Ren, K. Wang, M. Di Renzo, and A. Nallanathan, "Robust beamforming design for intelligent reflecting surface aided MISO communication systems," *IEEE Wireless Communications Letters*, vol. 9, no. 10, pp. 1658–1662, Jun. 2020.

[17] M. Jung, W. Saad, and G. Kong, "Large intelligent surfaces (LISs): Uplink spectral efficiency and pilot training," *IEEE Transactions on Commun.*, to appear, 2021.

[18] C. You, B. Zheng, and R. Zhang, "Channel estimation and passive beamforming for intelligent reflecting surface: Discrete phase shift and progressive refinement," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2604–2620, Nov. 2020.

[19] H. Guo, Y. Liang, J. Chen, and E. G. Larsson, "Weighted sumrate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3064–3076, May 2020.

[20] S. Thrun, "Lifeling learning algorithms," in *Learning to Learn*, Berlin, Germany: Springer, 1998, pp. 181–209.