Adaptive Transmission for Reconfigurable Intelligent Surface-Assisted OFDM Wireless Communications

Shaoe Lin, Beixiong Zheng, *Member, IEEE*, George C. Alexandropoulos, *Senior Member, IEEE*, Miaowen Wen, *Senior Member, IEEE*, Fangjiong Chen, *Member, IEEE*, and Shahid Mumtaz, *Senior Member, IEEE*

***Abstract*— Reconfigurable intelligent surfaces (RISs) have recently emerged as an innovative technology for improving the coverage, throughput, and energy/spectrum efficiency of future wireless communications. In this paper, we propose a new transmission protocol for wideband RIS-assisted single-input multiple-output (SIMO) orthogonal frequency division multiplexing (OFDM) communication systems, where each transmission frame is divided into multiple sub-frames to execute channel estimation simultaneously with passive beamforming. As the training symbols are discretely distributed over multiple sub-frames, the channel state information (CSI) associated with RIS cannot be estimated at once. As such, we propose a new channel estimation method to progressively estimate the associated CSI over consecutive sub-frames, based on which the passive beamforming at the RIS is fine-tuned to improve the achievable rate for data transmission. In particular, during the channel training, the RIS plays two roles of embedding training reflection states for progressive channel estimation and performing passive beamforming for data transmission on the data tones. Based on the estimated CSI in each sub-frame, we formulate an optimization problem to maximize the average achievable rate by designing the passive beamforming at the RIS, which needs to balance the received signal power over different sub-carriers and different receive antennas. As the formulated problem is non-convex and thus difficult to solve optimally, we propose two efficient algorithms to find high-quality solutions. Simulation results validate the effectiveness of the proposed channel estimation and beamforming optimization methods. It is shown that the proposed joint channel estimation and passive beamforming scheme is able to drastically improve the average achievable rate and reduce the delay for data transmission as compared to existing schemes.**

***Index Terms*— Reconfigurable intelligent surface (RIS), transmission protocol, channel estimation, beamforming optimization, rate maximization, OFDM.**

Manuscript received December 21, 2019; revised April 30, 2020; accepted May 9, 2020. Date of publication July 3, 2020; date of current version October 16, 2020. This work was supported in part by the Pearl River Nova Program of Guangzhou under Grant 201806010171, in part by the Natural Science Foundation of Guangdong Province under Grant 2018B030306005, in part by the National Natural Science Foundation of China under Grant 61871190, in part by the Open Research Fund of the National Mobile Communications Research Laboratory, Southeast University, under Grant 2020D03, in part by the Guangdong Provincial Key Laboratory of Short-Range Wireless Detection and Communication under Grant 2017B030314003, in part by the Fundamental Research Funds for the Central Universities under Grant 2019SJ02, and in part by the Project funded by China Postdoctoral Science Foundation under Grant 2018M640781 and Grant 2019T120731. *(Corresponding author: Miaowen Wen.)*

Shaoe Lin and Fangjiong Chen are with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China (e-mail: eeshe.lin@mail.scut.edu.cn; eefjchen@scut.edu.cn).

Beixiong Zheng was with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China. He is now with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117583 (e-mail: elezbe@nus.edu.sg).

George C. Alexandropoulos is with the Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Panepistimiopolis Ilissia, 15784 Athens, Greece (e-mail: alexandg@di.uoa.gr).

Miaowen Wen is with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China, and also with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: eemwwen@scut.edu.cn).

Shahid Mumtaz is with the Instituto de Telecomunicações, Campus Universitário de Santiago, University of Aveiro, 3810-193 Aveiro, Portugal (e-mail: smumtaz@av.it.pt).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2020.3007038

I. Introduction

In THE last decade, various wireless communication technologies such as millimeter wave (mmWave) communications and massive multiple-input multiple-output (MIMO) systems have been extensively investigated to meet the requirements for higher data rate, enhanced energy efficiency, low-latency, ultra-reliability, etc. of the fifth-generation (5G) wireless communication system [1]. However, none of them can support all 5G requirements and applications individually. Moreover, most of them require costly hardware and suffer from increasingly higher complexity and energy consumption, thus severely hindering their extensive practical deployment [2], [3]. For example, massive MIMO systems operating at mmWave frequency bands require costly and energy consuming radio frequency (RF) chains as well as sophisticated signal processing capability. Thanks to recent advances in reconfigurable surfaces [4], [5], reconfigurable intelligent surfaces (RISs) (a.k.a. intelligent reflecting surfaces) have emerged as an innovative technology for improving the coverage, throughput, and energy/spectrum efficiency of future wireless networks [6]–[16]. Specifically, RISs are planar surfaces consisting of a large number of low-cost unit cell elements, each of which is able to independently adjust the amplitude and/or phase shift of the reflected signal, thus reconfiguring the wireless propagation environment. Compared to existing techniques such as amplify-and-forward (AF) relays, RISs work in a full-duplex mode without incurring self-interference and thermal noise, and yet possess substantially reduced hardware cost and energy consumption due to the nearly passive components [17]–[19].

To characterize the theoretical performance upper bound of RIS-assisted wireless communication systems, perfect channel state information (CSI) was assumed to be available at the access point (AP), e.g., in [20] and [21]. However, as fully passive RISs lack signal processing and transmitting/receiving capabilities as well as involve a large number of unit cell elements, channel acquisition/estimation in RIS-assisted systems is one of the most challenging issues, which has spurred rapidly growing interests [22]–[29]. Existing methods for acquiring the CSI associated with the RIS can be classified into two categories. Methods of the first category consider that the RIS is equipped with receive RF chains, which enable the RIS to explicitly estimate the CSI of individual channels associated with its elements similarly as in the conventional wireless communication without RIS. However, these methods complicate the hardware design of the RIS as well as incur prohibitive hardware cost and energy consumption due to the large number of receive RF chains at the RIS. To alleviate this issue, a cost-effective RIS hardware architecture with minimal active elements was proposed in [24] and [25].

On the other hand, methods of the second category propose to estimate the CSI of the AP-RIS-user cascaded channels at the AP/user, by sending training symbols at the user/AP and adjusting reflection states of the RIS according to a pre-designed training reflection pattern [26]–[29]. In this manner, costly receive RF chains can be completely removed from the RIS for energy saving and cost reduction. More specifically, a channel estimation scheme using the simple ON/OFF-based training reflection pattern was proposed in [26] and [27], where numerous training symbols are needed to estimate the cascaded channel associated with each element at the RIS, thus incurring a prohibitive channel training overhead due to the massive number of RIS elements. To reduce the channel training overhead, an RIS-elements grouping strategy was proposed in [27] and [28], where the RIS is split into multiple disjoint sub-surfaces with RIS elements in identical sub-surfaces sharing a common reflection coefficient so that only aggregated cascaded channels for sub-surfaces need to be estimated. Based on RIS-elements grouping, a channel estimation method using the discrete Fourier transform (DFT)-based training reflection pattern was first presented in [28], which significantly improves the channel estimation accuracy by leveraging the large aperture gain of RIS. Moreover, a new DFT-Hadamard-based training reflection pattern with discrete phase shifts was developed in [29] to minimize the mean square error (MSE) of channel estimation. More recently, some preliminary works (e.g., [30], [31]) have appeared for channel estimation in RIS-assisted multi-user systems.

In existing RIS channel estimation works, the associated CSI is estimated at one time and then used for designing the passive beamforming at the RIS to maximize the achievable rate for data transmission. A transmission protocol with continuous channel training was proposed in [27]–[29], where the channel estimation and passive beamforming are executed successively. However, this protocol may not be suitable for delay-sensitive applications and/or short-packet transmissions, since estimating all required CSI in an all-at-once manner will cause a long delay for data transmission. Note that in mobile environments where the channel coherence time may be not long enough or varying, such delay will have a negative impact on the system performance and short-packet transmission is needed. For internet of things (IoT) networks, short-packet transmission is preferred as well due to the massive access and limited device power. To reduce the delay for data transmission, we propose a new transmission protocol in this paper to execute channel estimation and passive beamforming simultaneously. As illustrated in Fig. 1, we consider an RIS-assisted single-input multiple-output (SIMO) orthogonal frequency division multiplexing (OFDM) wireless communication system under frequency-selective fading channels, where an RIS is deployed to coordinate the uplink transmission from a single-antenna user to a multi-antenna AP. The main contributions of this paper are summarized as follows:

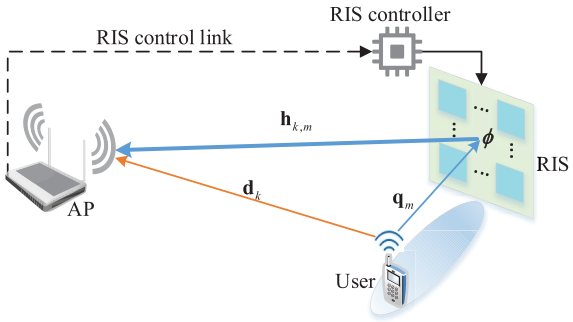


Fig. 1. The considered wideband RIS-assisted SIMO OFDM wireless communication system in the uplink.

- We propose a novel transmission protocol with quasi-periodic training symbols to execute channel estimation and passive beamforming simultaneously, where each transmission frame is divided into multiple sub-frames and only a small number of training symbols are allocated in each sub-frame. This protocol is adaptive in two folds: 1) based on RIS-elements grouping, it enables a flexible tradeoff between channel estimation overhead and passive beamforming performance for supporting short-packet transmissions; 2) by varying the time interval of adjacent training symbols, it provides a fundamental tradeoff between data transmission delay and effective transmission rate for supporting delay-sensitive transmissions.
- Based on RIS-elements grouping, we propose a new channel estimation method to progressively estimate the CSI associated with each sub-surface over consecutive sub-frames, by designing a binary-reflection (i.e., 0 or π) training pattern. Moreover, unlike existing schemes that perform channel training without enhancing data transmission, our proposed scheme improves the achievable rate for data transmission on the data tones by performing passive beamfoming at the RIS. In this manner, the RIS plays two roles during the channel training: 1) it embeds training reflection states for progressive channel estimation; 2) it performs passive beamforming for data transmission on the data tones.
- Based on the estimated CSI in each sub-frame, we formulate an optimization problem to maximize the average

achievable rate by designing the passive beamforming at the RIS, which needs to balance the received signal power over different sub-carriers and different receive antennas. This problem, however, is non-convex and thus difficult to solve. As such, we propose an efficient algorithm based on the semidefinite relaxation technique to sub-optimally solve it. To reduce the complexity for solving the formulated problem, we further propose a low-complexity algorithm to find a high-quality solution, by exploiting the strongest channel tap in the time domain.

- Simulation results validate the effectiveness of the proposed channel estimation and beamforming optimization methods, and show the superiority of the proposed protocol over the existing one. It is shown that the proposed protocol is able to drastically improve the average achievable rate and reduce the delay for data transmission as compared to the existing one. Moreover, the superiority of the proposed protocol over the existing one becomes more significant as the training time gets longer, since both the estimated CSI and the passive beamforming performance are progressively improved during the channel training.

The rest of this paper is organized as follows. In Section II, we introduce the system model and propose a new transmission protocol. In Section III, we propose the progressive channel estimation method. In Section IV, we formulate an optimization problem for designing the passive beamforming based on the estimated CSI in each sub-frame. Simulation results are presented in Section V, and Section VI concludes the paper.

Notation: Upper and lower case boldface letters denote matrices and column vectors, respectively. Upper-case calligraphic letters (e.g., $\mathcal{A}$) denote discrete and finite sets. $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^{-1}$ stand for transpose, Hermitian transpose, and matrix inversion operations, respectively. $\mathbb{C}^{a\times b}$ denotes the space of $a \times b$ complex-valued matrices. $\mathrm{diag}(\mathbf{x})$ returns a diagonal matrix whose diagonal elements are included in $\mathbf{x}$. $|\cdot|$ denotes the absolute value if applied to a complex number or the cardinality if applied to a set. $\|\cdot\|$ denotes the ℓ_2-norm, and $\mathrm{tr}(\cdot)$ returns the trace of the matrix. $\odot$ denotes the Hadamard product, and $\mathbb{E}\{\cdot\}$ stands for expectation. $\arg\max$ returns the argument that achieves the maximum function value, and $\mathrm{rank}\,(\cdot)$ returns the rank of the matrix. $\mathbf{S} \succeq 0$ implies that $\mathbf{S}$ is positive semidefinite. $\mathbf{I}_a$, $\mathbf{1}_{a\times b}$, and $\mathbf{0}_{a\times b}$ denote an identity matrix of sizes $a\times a$, an all-one matrix of sizes $a\times b$, and an all-zero matrix of sizes $a\times b$, respectively. $\angle(\mathbf{x})$ denotes a vector with each element being the phase of the corresponding element in $\mathbf{x}$. The relative complement of $\mathcal{A}$ in $\mathcal{B}$ is denoted by $\mathcal{B} \setminus \mathcal{A}$, and the union of $\mathcal{A}$ and $\mathcal{B}$ is denoted by $\mathcal{A} \cup \mathcal{B}$. $[\mathbf{X}]_{\iota,\jmath}$ denotes the $(\iota,\jmath)$-th entry of matrix $\mathbf{X}$, and $[\mathbf{x}]_\iota$ denotes the ι-th entry of vector $\mathbf{x}$. The distribution of a circularly symmetric complex Gaussian (CSCG) random vector with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is denoted by $\mathcal{N}_c(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and $\sim$ stands for "distributed as."

II. System Model and Transmission Protocol

As illustrated in Fig. 1, we consider an RIS-assisted OFDM system in the uplink, where an RIS composed of M_0 unit cell elements is deployed to coordinate the transmission from a single-antenna user to an AP equipped with K receive antennas indexed by the set $\mathcal{K} \triangleq \{1, 2, \ldots, K\}$. Note that the number of unit cell elements at the RIS is practically large, which implies that massive training overhead is required for estimating the cascaded channel associated with each element. By assuming that multiple adjacent elements form a sub-surface to share a common reflection coefficient, the channel estimation overhead and the complexity of passive beamforming design can be significantly reduced [27], [28]. Accordingly, the RIS is assumed to comprise M sub-surfaces indexed by the set $\mathcal{M} \triangleq \{1, 2, \ldots, M\}$, each of which consists of $\eta = M_0/M$ adjacent elements. For notational convenience, we assume that η is a positive integer that we refer to as grouping size. The RIS is connected to a smart controller, which is responsible for reconfiguring the phase shifts of its elements [32] and exchanging information with the AP through a feedback link to realize its passive beamforming functionality. We assume that the feedback link between the AP and RIS controller is a wired link such as optical fiber. Note that the signal processing and information feedback delay can be neglected when the AP has sufficient signal processing capability and the wired connection is applied as the feedback link. As such, the optimized passive beamforming vector (that describes the reflection amplitudes and/or phase shifts of the M sub-surfaces) can be fed back to the RIS without any error or delay. Moreover, quasi-static frequency-selective fading channels are considered for the user-AP, user-RIS, and RIS-AP individual links, which remain approximately constant within the transmission frame of our interest.

A. System Model

Let N denote the number of sub-carriers in each OFDM symbol. For simplicity, we consider the case where the user has no knowledge of the CSI. As such, the total transmit power of the user P_t can be assumed to be equally allocated over the N sub-carriers. We assume that the baseband equivalent channels of the user-AP, user-RIS, and RIS-AP links have L_d, L_1, and L_2 taps in the impulse response, respectively. The baseband equivalent channel of the cascaded user-RIS-AP link, which is a convolution of the user-RIS channel with the RIS-AP channel weighted by RIS reflection coefficients, has $L_r = L_1 + L_2 - 1$ taps in the impulse response. Therefore, the baseband equivalent channel from the user to AP has $L = \max\{L_r, L_d\}$ taps in the impulse response.

At the user side, each OFDM symbol denoted by $\mathbf{x} \triangleq [X_0\ X_1\ \ldots\ X_{N-1}]^T$ is first transformed into the time domain via an N-point inverse discrete Fourier transform (IDFT), and then appended by a cyclic prefix (CP) of length L_{cp} with $L_{cp} \geq L - 1$ to avoid inter-symbol interference. Let $\mathbf{d}_k \triangleq [D_{k,0}\ D_{k,1}\ \ldots\ D_{k,N-1}]^T$ for $k \in \mathcal{K}$ denote the frequency response of the direct user-AP channel associated with the k-th receive antenna. Let $\mathbf{q}_m \in \mathbb{C}^{N\times 1}$ and $\mathbf{h}_{k,m} \in \mathbb{C}^{N\times 1}$ for $m \in \mathcal{M}$ and $k \in \mathcal{K}$ denote the frequency response of the user-RIS channel associated with the m-th sub-surface, and the frequency response of the RIS-AP channel associated with the m-th sub-surface and the k-th receive antenna, respectively. At the AP side, after removing the CP and performing

the N-point DFT, the baseband received signal at the k-th ($k = 1, 2, \ldots, K$) receive antenna in the frequency domain is given by

$$\mathbf{y}_k = \mathbf{X}\left(\sum_{m=1}^{M} \phi_m \mathbf{q}_m \odot \mathbf{h}_{k,m} + \mathbf{d}_k\right) + \mathbf{v}_k \tag{1}$$

where $\mathbf{X} = \text{diag}(\mathbf{x})$ is the diagonal matrix of the transmitted OFDM symbol $\mathbf{x}$, $\mathbf{v}_k \triangleq [V_{k,0}\ V_{k,1}\ \ldots\ V_{k,N-1}]^T \sim \mathcal{N}_c(\mathbf{0}, \sigma^2\mathbf{I}_N)$ denotes the additive white Gaussian noise (AWGN) vector at the k-th receive antenna with noise power of σ^2, and $\phi_m = e^{-j\varphi_m}$ denotes the common reflection coefficient of the m-th sub-surface with phase shift $\varphi_m \in (0, 2\pi]$. It is worth pointing out that as each sub-surface is fabricated with the maximum reflection amplitude (i.e., unit amplitude) to maximize the RIS reflection power, no amplitude control circuits are needed for RIS elements, thus simplifying the hardware design of RIS and saving energy consumption. In this way, both channel estimation and passive beamforming are realized by adjusting the phase shifts of the M sub-surfaces only, i.e., φ_m, $m = 1, 2, \ldots, M$.

Let $\mathbf{g}_{k,m} \triangleq [G_{k,m,0}\ G_{k,m,1}\ \ldots\ G_{k,m,N-1}]^T = \mathbf{q}_m \odot \mathbf{h}_{k,m}$ denote the frequency response of the cascaded user-RIS-AP channel associated with the k-th receive antenna and the m-th sub-surface without the effect of phase shift. Thus, (1) is rewritten as

$$\mathbf{y}_k = \mathbf{X}\left(\sum_{m=1}^{M} \phi_m \mathbf{g}_{k,m} + \mathbf{d}_k\right) + \mathbf{v}_k, \quad k = 1, 2, \ldots, K \tag{2}$$

where the explicit knowledge of $\mathbf{q}_m$ and $\mathbf{h}_{k,m}$ is no longer needed. By stacking $\mathbf{g}_{k,m}$ with $m = 1, 2, \ldots, M$ into $\mathbf{G}_k = [\mathbf{g}_{k,1}\ \mathbf{g}_{k,2}\ \ldots\ \mathbf{g}_{k,M}]$, we can rewrite (2) compactly as

$$\mathbf{y}_k = \mathbf{X}\left(\mathbf{G}_k\boldsymbol{\phi} + \mathbf{d}_k\right) + \mathbf{v}_k, \quad k = 1, 2, \ldots, K \tag{3}$$

where $\boldsymbol{\phi} \triangleq [\phi_1\ \phi_2\ \ldots\ \phi_M]^T$ denotes the passive beamforming vector describing the phase shifts of the M sub-surfaces. Therefore, the frequency response of the equivalent channel from the RIS-serviced user to the k-th receive antenna at the AP is given by

$$\mathbf{r}_k \triangleq [R_{k,0}\ R_{k,1}\ \ldots\ R_{k,N-1}]^T \tag{4}$$

$$= \mathbf{G}_k\boldsymbol{\phi} + \mathbf{d}_k, \quad k = 1, 2, \ldots, K. \tag{5}$$

According to (5), the knowledge of $\mathbf{G}_k$ and $\mathbf{d}_k$ with $k = 1, 2, \ldots, K$ is required for designing the passive beamforming vector $\boldsymbol{\phi}$ to maximize the achievable rate. We next present a new transmission protocol with quasi-periodic training symbols to execute channel estimation and passive beamforming simultaneously. It is worth noting that the proposed framework can be readily extended to the multi-user case by assigning disjoint pilot tones to different users and performing the same channel estimation procedure for each user [31].

B. Proposed Adaptive Transmission Protocol (ATP)

As illustrated in Fig. 2, each transmission frame of T_f OFDM symbol durations is divided into M consecutive sub-frames, except for the first sub-frame, each of which consists of one training OFDM symbol followed by the data OFDM symbols, while the first sub-frame contains one additional training OFDM symbol. Let Δ_t denote the time interval between two adjacent training OFDM symbols except for the first one, which is normalized to the duration of one OFDM symbol. Thus, the training OFDM symbols are indexed in increasing order by

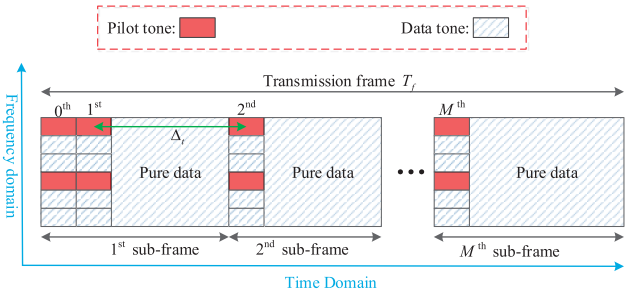


Fig. 2. An illustration of the proposed adaptive transmission protocol.

$$\mathcal{T} \triangleq \{0\} \cup \{1, \Delta_t + 1, 2\Delta_t + 1, \ldots, (M-1)\Delta_t + 1\} \tag{6}$$

where Δ_t is an integer satisfying $\Delta_t M < T_f$ or $1 \le \Delta_t < \frac{T_f}{M}$.

Since the $M + 1$ training OFDM symbols are distributed over M consecutive sub-frames, the AP cannot estimate in total KM cascaded channels (i.e., $\mathbf{g}_{k,m}$, $\forall k \in \mathcal{K}$, $\forall m \in \mathcal{M}$) at once. As such, we propose a new channel estimation scheme to progressively estimate the cascaded channel associated with each sub-surface. Specifically, in the i-th ($i = 1, 2, \ldots, M$) sub-frame, by adjusting the RIS reflection state as $\boldsymbol{\phi}^{(i)}$, the AP obtains the channel frequency responses (CFRs) of $\mathbf{g}_{k,i-1}$ with $k = 1, 2, \ldots, K$ based on the i-th training OFDM symbol sent by the user and the estimated CSI in the previous sub-frame, and then updates the estimated CSI set to $\{\mathbf{g}_{k,1}, \mathbf{g}_{k,2}, \ldots, \mathbf{g}_{k,i-1}, \sum_{m=i}^{M} \mathbf{g}_{k,m}\}$. The details of the proposed progressive channel estimation are described later in Section III. On the other hand, the newly estimated CSI provides one more degree of freedom for designing the passive beamforming vector, which should be refined accordingly. Based on the estimated CSI in the i-th sub-frame, the passive beamforming vector $\boldsymbol{\theta}^{(i)} = [\theta_1, \theta_2, \ldots, \theta_{i-1}, \bar{\theta}_i \mathbf{1}_{1\times(M-i+1)}]^T$ is optimized for maximizing the average achievable rate subject to the unit-modulus reflection constraint, which can be stated mathematically as

$$\text{(P1):}\ \max_{\boldsymbol{\theta}^{(i)}}\ C\left(\boldsymbol{\theta}^{(i)}\right) = \frac{1}{N + L_{cp}} \sum_{n=0}^{N-1} \log_2\left(1 + \frac{\hat{W}_n^{(i)}\left(\boldsymbol{\theta}^{(i)}\right) P_t}{N\Gamma\sigma^2}\right) \tag{7}$$

$$\text{s.t.}\ \ \left|\bar{\theta}_i\right| = 1, \quad |\theta_m| = 1,\ \forall m = 1, 2, \ldots, i-1 \tag{8}$$

where $\hat{W}_n^{(i)}\left(\boldsymbol{\theta}^{(i)}\right)$ is the channel gain of the n-th sub-carrier which varies with the passive beamforming vector $\boldsymbol{\theta}^{(i)}$, $\Gamma \ge 1$ represents the achievable rate gap due to a practical modulation and coding scheme, θ_m is the reflection coefficient for the m-th sub-surface, and $\bar{\theta}_i$ is the common reflection coefficient for the remaining sub-surfaces indexed by $\mathcal{I} \triangleq \{i, i+1, \ldots, M\}$. The average achievable rate in (7) is measured in bits per second per Hertz (bps/Hz). Note that to achieve the optimal solution

of problem (P1), the passive beamforming vector $\boldsymbol{\theta}^{(i)}$ needs to balance $\hat{W}_n^{(i)}\left(\boldsymbol{\theta}^{(i)}\right)$ over different receive antennas at the AP and different sub-carriers, since the RIS reflection is not frequency-selective.[1] It can be verified that problem (P1) is non-convex and thus difficult to solve directly. As such, we present two efficient algorithms to sub-optimally solve it in Section IV. Apparently, both the estimated CSI and the passive beamforming vector are updated progressively on a sub-frame basis.

After solving problem (P1), the optimized beamforming vector $\boldsymbol{\theta}^{(i)}$ is fed back to the RIS controller via the feedback link. Accordingly, the RIS controller adjusts the RIS reflection state as $\boldsymbol{\theta}^{(i)}$ to achieve the desired signal reflection for data transmission. The above channel estimation process is repeated periodically to progressively estimate all required CSI, while the RIS performs progressive passive beamforming adaptation throughout the frame to improve the system performance in terms of achievable rate. It is worth pointing out that this protocol is adaptive in two folds:

- Given the channel training time or the number of training OFDM symbols, the number of sub-surfaces M can be adjusted to provide a flexible system tradeoff between the channel estimation overhead and the passive beamforming performance. For short channel coherent time, the value of M is envisioned to be small due to the strictly limited channel training time.
- Given the number of sub-frames M, the time interval Δ_t can be adjusted to get a fundamental tradeoff between the data transmission delay and effective transmission rate within one frame. For long channel coherent time, the value of Δ_t is envisioned to be small to allow a much longer duration of passive beamforming that is optimized based on the full CSI, thus increasing the average achievable rate for the whole frame.

C. Consecutive Transmission Protocol (CTP)

In this subsection, we briefly introduce the existing transmission protocol using consecutive training symbols, which serves as a benchmark scheme in this paper. Specifically, each transmission frame composed of T_f OFDM symbol durations is divided into two sub-frames: the first sub-frame consists of $M+1$ consecutive training OFDM symbols while the second sub-frame consists of multiple consecutive data OFDM symbols in the remaining part of the frame. The $M+1$ consecutive training OFDM symbols in conjunction with a predesigned RIS training reflection pattern, e.g., a DFT-based reflection pattern [28] or an ON/OFF-based reflection pattern [27], allow the AP to *fully* estimate all the required CSI of $\mathbf{G}_k$ and $\mathbf{d}_k$, $\forall k \in \mathcal{K}$, at once in the first sub-frame. With the estimated CSI of each $\mathbf{G}_k$ and $\mathbf{d}_k$, the passive beamforming vector is designed to maximize the average achievable rate for the second sub-frame subject to the unit-modulus reflection constraint. Then, the optimized passive beamforming vector is fed back to the RIS controller via the feedback link. According to the feedback information, the RIS controller adjusts the phase shift of each of the RIS elements to achieve the desired signal reflection for data transmission during the second sub-frame.

While this protocol appears similar to the proposed ATP with $\Delta_t = 1$, which also consists of $M+1$ consecutive training OFDM symbols followed by consecutive data OFDM symbols in the remaining part of the frame, they differ in terms of the functionality of the RIS. During the transmission of the $M+1$ consecutive training OFDM symbols, the RIS under this protocol assists in channel extraction only, while the one under the proposed ATP with $\Delta_t = 1$ not only assists in channel extraction but also enhances the channel gains on the data tones via progressive passive beamforming adaptation. We will compare their performance in the simulations to draw some important insights into our proposed protocol design.

III. Progressive Channel Estimation/Extraction

Although the channel estimation method using the DFT-based training reflection pattern achieves superior performance in [28], it cannot be applied for the proposed ATP since it requires the joint processing over multiple training symbols in an all-at-once manner. Note that the ON/OFF-based channel estimation method proposed in [27] can be applied for the proposed ATP, where the direct channel is estimated with the whole RIS switched OFF and the cascaded channel associated with each sub-surface is estimated with only one out of M sub-surfaces switched ON sequentially. However, on one hand, the large aperture gain of the RIS is lost in the ON/OFF-based channel estimation scheme, thus degrading the channel estimation accuracy; on the other hand, the channel gains on the data tones in each training OFDM symbol have not been improved. To overcome these issues, we propose a novel channel estimation scheme in this section for the proposed ATP, which extracts all the required CSI of $\mathbf{g}_{k,m}$, $\forall k \in \mathcal{K}$, $\forall m \in \mathcal{M}$ in a progressive manner and integrates training reflection states with progressive passive beamforming vectors.

As shown in Fig. 3, N_p pilots are inserted in each training OFDM symbol as the pilot tones indexed by

$$\mathcal{P} = \{0, \Delta_p, \ldots, (N_p - 1)\Delta_p\} \tag{9}$$

where $\Delta_p = \lfloor N/N_p \rfloor$ is the frequency spacing of adjacent pilots. Note that this channel estimation scheme can be readily extended to the multi-user case by assigning disjoint pilot tones to different users and performing the same channel estimation procedure to each user. Let $\mathcal{D} = \mathcal{N} \setminus \mathcal{P}$ denote the index set of $N_d = N - N_p$ data tones in each training OFDM symbol, where $\mathcal{N} \triangleq \{0, 1, \ldots, N-1\}$ denotes the index set of the N sub-carriers. With the known pilot sequence $\mathbf{s}_p$, the CFRs on the pilot tones $\mathcal{P}$ at the k-th ($k = 1, 2, \ldots, K$) receive antenna can be estimated as

$$\hat{\mathbf{r}}_{k,\mathcal{P}} = \mathbf{S}_p^{-1}\mathbf{y}_{k,\mathcal{P}} = \mathbf{r}_{k,\mathcal{P}} + \mathbf{S}_p^{-1}\mathbf{v}_{k,\mathcal{P}} \tag{10}$$

where $\mathbf{S}_p = \operatorname{diag}(\mathbf{s}_p)$ is the diagonal matrix of the pilot sequence $\mathbf{s}_p$; $\mathbf{y}_{k,\mathcal{P}}$ denotes the received signals on the pilot tones $\mathcal{P}$ at the k-th receive antenna; $\mathbf{r}_{k,\mathcal{P}}$ and $\mathbf{v}_{k,\mathcal{P}}$ respectively denote the CFRs and AWGNs on the pilot tones $\mathcal{P}$ associated with the k-th receive antenna. Based on the

[1] The same phase shift will be applied to the signals on all sub-carriers.

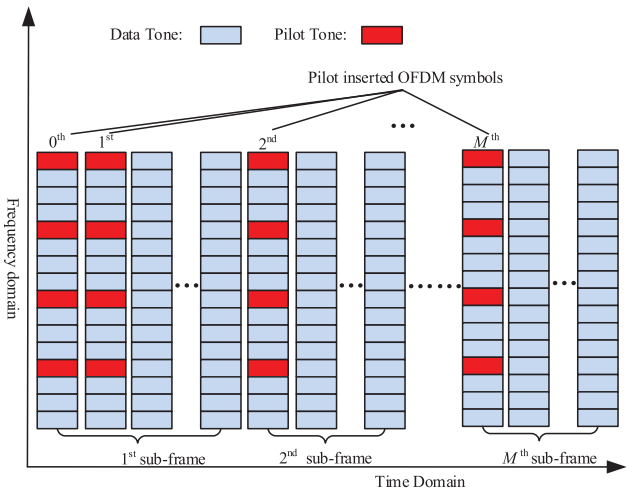


Fig. 3. An illustration of the comb-type pilot structure within the proposed transmission frame.

estimated CFRs on the pilot tones $\mathcal{P}$, the CFRs on the data tones $\mathcal{D}$ can be acquired by adopting time/frequency domain interpolation, e.g., the trigonometric interpolation whose optimality for frequency-selective fading channels has been proved in [33]. Specifically, according to the trigonometric interpolation, the combined CFR vector of the direct and reflected links associated with the k-th receive antenna is estimated as

$$\hat{\mathbf{r}}_k = \frac{1}{\sqrt{N_p}}\mathbf{F}_N \left[\left[\tilde{\mathbf{r}}_{k,\mathcal{P}}^T\right]_{1:L} \mathbf{0}_{1\times(N-L)}\right]^T \tag{11}$$

$$= \mathbf{r}_k + \bar{\mathbf{v}}_k = \mathbf{G}_k\boldsymbol{\phi} + \mathbf{d}_k + \bar{\mathbf{v}}_k \tag{12}$$

where $\mathbf{F}_N$ is an $N\times N$ DFT matrix, $\tilde{\mathbf{r}}_{k,\mathcal{P}}$ denotes the N_p-point IDFT of $\hat{\mathbf{r}}_{k,\mathcal{P}}$ in (10), $\bar{\mathbf{v}}_k$ is the corresponding noise vector, and the last equality holds according to (5).

According to (10) and (11), only the combined CFR vector $\mathbf{r}_k$ can be estimated based on the pilot sequence $\mathbf{s}_p$. From (12), we observe that the CSI of $\mathbf{G}_k$ and $\mathbf{d}_k$ can be extracted by leveraging the RIS reflection state $\boldsymbol{\phi}$. For ease of exposition, we hereafter consider the noiseless case to disclose the channel extraction process for $\mathbf{G}_k$ and $\mathbf{d}_k$, while practical channel estimation corrupted by different noise levels will be taken into account in the simulations. As each receive antenna can perform the channel estimation and extraction in parallel, we hereafter drop the receive antenna index k in this section for notational simplicity without causing confusion. Denote the training reflection state at the RIS for the duration of the i-th training OFDM symbol as $\boldsymbol{\phi}^{(i)} \triangleq \left[e^{-j\varphi_1^{(i)}}\ e^{-j\varphi_2^{(i)}}\ \ldots\ e^{-j\varphi_M^{(i)}}\right]^T$. Accordingly, the estimate of the combined CFR vector based on the i-th received training OFDM symbol at each receive antenna can be expressed as

$$\mathbf{r}^{(i)} = \mathbf{G}\boldsymbol{\phi}^{(i)} + \mathbf{d}, \quad i = 0, 1, \ldots, M. \tag{13}$$

1) Initialization Process: By setting the initial training reflection states $\boldsymbol{\phi}^{(0)} = \mathbf{z}^{(0)} = \mathbf{1}_{M\times 1}$ and $\boldsymbol{\phi}^{(1)} = \mathbf{z}^{(1)} = -\mathbf{1}_{M\times 1}$, i.e., $\varphi_m^{(0)} = 0$ and $\varphi_m^{(1)} = \pi$, $\forall\ m = 1, 2, \ldots, M$, we can extract the CFR vector of the direct user-AP link as

$$\mathbf{d} = (\mathbf{r}^{(0)} + \mathbf{r}^{(1)})/2 \tag{14}$$

and the unseparated CFR vector of the cascaded user-RIS-AP link as

$$\bar{\mathbf{g}}_1 = \sum_{m=1}^{M} \mathbf{g}_m = (\mathbf{r}^{(0)} - \mathbf{r}^{(1)})/2. \tag{15}$$

Let $\bar{\mathbf{g}}_i = \sum_{m=i}^{M} \mathbf{g}_m$ with $i = 1, 2, \ldots, M$ denote the unseparated CFR vector of the cascaded user-RIS-AP link associated with sub-surfaces indexed by $\mathcal{I}$. Furthermore, we denote the partially estimated CSI by the end of the i-th training OFDM symbol duration in a set form as

$$\mathcal{G}^i = \{\mathbf{d}, \mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{i-1}, \bar{\mathbf{g}}_i\} \tag{16}$$

which has i resolvable CFR vectors for the cascaded user-RIS-AP link. After initialization, with the estimated CSI $\mathcal{G}^1 = \{\mathbf{d}, \bar{\mathbf{g}}_1\}$, the M sub-surfaces share a common phase shift, which is optimized for maximizing the average achievable rate in the first sub-frame.

2) Update Process: In order to extract $\mathbf{g}_m$ for $m = 1, 2, \ldots, M$ from $\bar{\mathbf{g}}_1$, we design the i-th training reflection state at the RIS for the duration of the i-th training OFDM symbol as $\mathbf{z}^{(i)} \triangleq \left[z_1^{(i)}\ z_2^{(i)}\ \ldots\ z_M^{(i)}\right]^T$ with each entry given by

$$z_m^{(i)} = \begin{cases} e^{-j\pi} = -1, m = i-1 & m = 1, 2, \ldots, M \\ e^{-j0} = +1, \text{otherwise}, & i = 2, 3, \ldots, M. \end{cases} \tag{17}$$

On the other hand, unlike existing channel estimation schemes that perform channel training without enhancing channel gain through passive beamforming, our proposed progressive channel estimation scheme not only improves the channel estimation accuracy by leveraging the large aperture gain of the RIS but also boosts the transmission rate of the data tones by performing passive beamfoming during the channel training. In this manner, the RIS plays two roles: not only to embed the training reflection pattern for progressive channel estimation, but also to enhance the communication link via progressive passive beamforming adaptation. Specifically, the reflection state $\boldsymbol{\phi}^{(i)}$ at the RIS for the duration of the i-th training OFDM symbol is given by

$$\boldsymbol{\phi}^{(i)} = \text{diag}(\mathbf{z}^{(i)})\boldsymbol{\theta}^{(i-1)}, \quad i = 2, 3, \ldots M \tag{18}$$

where $\boldsymbol{\theta}^{(i-1)}$ denotes the passive beamforming vector based on the estimated CSI $\mathcal{G}^{i-1} = \{\mathbf{d}, \mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{i-2}, \bar{\mathbf{g}}_{i-1}\}$. Since $\mathcal{G}^{i-1}$ only has $(i-1)$ resolvable CFR vectors as the CSI of the cascaded user-RIS-AP channel, $\boldsymbol{\theta}^{(i-1)}$ thus only has $(i-1)$ design degrees of freedom, which is given by

$$\boldsymbol{\theta}^{(i-1)} = [\theta_1, \theta_2, \ldots, \theta_{i-2}, \bar{\theta}_{i-1}\mathbf{1}_{1\times(M-i+2)}]^T \tag{19}$$

and optimized for achieving coherent channel combination of the extracted CFR vectors in $\mathcal{G}^{i-1}$, with the details given in Section IV. Substituting (17) and (19) into (18), we have

$$\boldsymbol{\phi}^{(i)} = [\theta_1, \theta_2, \ldots, \theta_{i-2}, -\bar{\theta}_{i-1}, \bar{\theta}_{i-1}\mathbf{1}_{1\times(M-i+1)}]^T. \tag{20}$$

Then, (13) can be rewritten as

$$\mathbf{r}^{(i)} = \sum_{m=1}^{i-2} \theta_m \mathbf{g}_m + \bar{\theta}_{i-1}\left(-\mathbf{g}_{i-1} + \bar{\mathbf{g}}_i\right) + \mathbf{d}. \tag{21}$$

TABLE I
AN ILLUSTRATION OF CHANNEL EXTRACTION AND RIS REFLECTION STATE IN PROGRESSIVE CHANNEL ESTIMATION

	RIS reflection state	Channel estimation	Chanel extraction	Beamforming vector
1st sub-frame	$\boldsymbol{\phi}^{(0)}=\mathbf{z}^{(0)}=\mathbf{1}_{4\times1}$ $\boldsymbol{\phi}^{(1)}=\mathbf{z}^{(1)}=-\mathbf{1}_{4\times1}$	$\mathbf{r}^{(0)}=\bar{\mathbf{g}}_1+\mathbf{d}$ $\mathbf{r}^{(1)}=-\bar{\mathbf{g}}_1+\mathbf{d}$	$\mathbf{d}=(\mathbf{r}^{(0)}+\mathbf{r}^{(1)})/2$ $\bar{\mathbf{g}}_1=(\mathbf{r}^{(0)}-\mathbf{r}^{(1)})/2$	$\boldsymbol{\theta}^{(1)}=\bar{\theta}_1\mathbf{1}_{4\times1}$
2nd sub-frame	$\mathbf{z}^{(2)}=[-1,1,1,1]^T$ $\boldsymbol{\phi}^{(2)}=\text{diag}(\mathbf{z}^{(2)})\boldsymbol{\theta}^{(1)}$	$\mathbf{r}^{(2)}=-\bar{\theta}_1\mathbf{g}_1+\bar{\theta}_1\bar{\mathbf{g}}_2+\mathbf{d}$	$\mathbf{g}_1=\frac{\bar{\mathbf{g}}_1}{2}-\frac{\mathbf{r}^{(2)}-\mathbf{d}}{2\bar{\theta}_1}$ $\bar{\mathbf{g}}_2=\bar{\mathbf{g}}_1-\mathbf{g}_1$	$\boldsymbol{\theta}^{(2)}=[\theta_1\ \bar{\theta}_2\mathbf{1}_{3\times1}]$
3rd sub-frame	$\mathbf{z}^{(3)}=[1,-1,1,1]^T$ $\boldsymbol{\phi}^{(3)}=\text{diag}(\mathbf{z}^{(3)})\boldsymbol{\theta}^{(2)}$	$\mathbf{r}^{(3)}=\theta_1\mathbf{g}_1-\bar{\theta}_2\mathbf{g}_2+\bar{\theta}_2\bar{\mathbf{g}}_3+\mathbf{d}$	$\mathbf{g}_2=\frac{\bar{\mathbf{g}}_2}{2}-\frac{\mathbf{r}^{(3)}-\mathbf{d}-\theta_1\mathbf{g}_1}{2\bar{\theta}_2}$ $\bar{\mathbf{g}}_3=\bar{\mathbf{g}}_2-\mathbf{g}_2$	$\boldsymbol{\theta}^{(3)}=[\theta_1\ \theta_2\ \bar{\theta}_3\mathbf{1}_{2\times1}]$
4th sub-frame	$\mathbf{z}^{(4)}=[1,1-1,1]^T$ $\boldsymbol{\phi}^{(4)}=\text{diag}(\mathbf{z}^{(4)})\boldsymbol{\theta}^{(3)}$	$\mathbf{r}^{(4)}=\sum_{m=1}^{2}\theta_m\mathbf{g}_m-\bar{\theta}_3\mathbf{g}_3+\bar{\theta}_3\mathbf{g}_4+\mathbf{d}$	$\mathbf{g}_3=\frac{\bar{\mathbf{g}}_3}{2}-\frac{\mathbf{r}^{(4)}-\mathbf{d}-\sum_{m=1}^{2}\theta_m\mathbf{g}_m}{2\bar{\theta}_3}$ $\mathbf{g}_4=\bar{\mathbf{g}}_4=\bar{\mathbf{g}}_3-\mathbf{g}_3$	$\boldsymbol{\theta}^{(4)}=[\theta_1\ \theta_2\ \theta_3\ \theta_4]$

After stripping the extracted CFR vectors in $\mathcal{G}^{i-1}$ off, each receive antenna can extract a new CFR vector from (21) as

$$\mathbf{g}_{i-1}=\bar{\mathbf{g}}_{i-1}/2-\left(\mathbf{r}^{(i)}-\mathbf{d}-\sum_{m=1}^{i-2}\theta_m\mathbf{g}_m\right)/2\bar{\theta}_{i-1} \quad (22)$$

and the remaining unseparated CFR vector of the cascaded user-RIS-AP link as

$$\bar{\mathbf{g}}_i=\bar{\mathbf{g}}_{i-1}-\mathbf{g}_{i-1} \quad (23)$$

which is similar to the concept of successive interference cancellation (SIC). Thanks to the newly extracted CFR vector $\mathbf{g}_{i-1}$, the degrees of freedom for designing passive beamforming vector increase. As a result, the AP updates the passive beamforming vector to achieve coherent channel combination of the extracted CFR vectors in $\mathcal{G}^i$ as

$$\boldsymbol{\theta}^{(i)}=[\theta_1,\theta_2,\ldots,\theta_{i-1},\bar{\theta}_i\mathbf{1}_{1\times(M-i+1)}]^T \quad (24)$$

which is in agreement with (19). An example with $M=4$ is given in Table I to illustrate the process of channel estimation and extraction as well as RIS reflection state in the proposed progressive channel estimation scheme. The whole procedure of the progressive channel estimation and beamforming optimization is summarized in Algorithm 1.

Algorithm 1 Channel Extraction & RIS Reflection State in Progressive Channel Estimation

Initialization Process:

1: Estimate the combined CFR vectors $\mathbf{r}^{(0)}$ and $\mathbf{r}^{(1)}$ based on (10) and (11) with training reflection states $\boldsymbol{\phi}^{(0)}=\mathbf{z}^{(0)}$ and $\boldsymbol{\phi}^{(1)}=\mathbf{z}^{(1)}$ at the RIS for the duration of the 0-th and 1-st training OFDM symbols, respectively;
2: Extract the CFR vectors of $\mathbf{d}$ and $\bar{\mathbf{g}}_1$ according to (14) and (15);
3: Store the initial set of the extracted CFR vectors $\mathcal{G}^1$;
4: Optimize the passive beamforming vector $\boldsymbol{\theta}^{(1)}=\bar{\theta}_1\mathbf{1}_{M\times1}$ and feed it back to the RIS controller;
5: The RIS performs passive beamforming according to $\boldsymbol{\theta}^{(1)}$;

Update Process:

6: **for** $i=2$ to M **do**
7: Estimate the combined CFR vector $\mathbf{r}^{(i)}$ based on the i-th received training OFDM symbol associated with RIS reflection state $\boldsymbol{\phi}^{(i)}$ given in (20), according to (10) and (11);
8: Extract $\mathbf{g}_{i-1}$ and $\bar{\mathbf{g}}_i$ according to (22) and (23), respectively;
9: Update the set of the estimated CSI as (16);
10: Optimize the passive beamforming vector $\boldsymbol{\theta}^{(i)}$ in form of (24) and feed it back to the RIS controller;
11: The RIS performs passive beamforming according to $\boldsymbol{\theta}^{(i)}$;
12: **end for**

IV. PROGRESSIVE BEAMFORMING OPTIMIZATION

In this section, our objective is to optimize the passive beamforming vector for maximizing the average achievable rate in each sub-frame based on the partially estimated CSI, which is updated progressively on a sub-frame basis. Specifically, let $\mathcal{G}_k^i \triangleq \left\{\hat{\mathbf{d}}_k,\hat{\mathbf{g}}_{k,1},\hat{\mathbf{g}}_{k,2},\ldots,\hat{\mathbf{g}}_{k,i-1},\hat{\bar{\mathbf{g}}}_{k,i}\right\}$ denote the set of the partially estimated CSI at the k-th ($k=1,2,\ldots,K$) receive antenna in the i-th ($i=1,2,\ldots,M$) sub-frame, where $\hat{\bar{\mathbf{g}}}_{k,i} \triangleq \left[\hat{\bar{G}}_{k,i,0}\ \hat{\bar{G}}_{k,i,1}\ \ldots\ \hat{\bar{G}}_{k,i,N-1}\right]^T$ denotes the unseparated CFR vector for the cascaded user-RIS-AP link associated with sub-surfaces indexed by $\mathcal{I}$. After extracting the CFR vectors in $\mathcal{G}_k^i$ at each receive antenna and applying the maximal-ratio combination (MRC), the channel gain of the n-th ($n=0,1,\ldots,N-1$) sub-carrier is given by

$$\hat{W}_n^{(i)}\left(\boldsymbol{\theta}^{(i)}\right)=\sum_{k=1}^{K}\left|\sum_{m=1}^{i-1}\theta_m\hat{G}_{k,m,n}+\bar{\theta}_i\hat{\bar{G}}_{k,i,n}+\hat{D}_{k,n}\right|^2 \quad (25)$$

which varies with the passive beamforming vector $\boldsymbol{\theta}^{(i)}$. Based on (25), we formulate an optimization problem for maximizing the average achievable rate $C\left(\boldsymbol{\theta}^{(i)}\right)$ in (7) by designing the passive beamforming vector $\boldsymbol{\theta}^{(i)}$ as problem (P1) in Section II-B. However, problem (P1) is difficult to be solved optimally since it is non-concave with respect to $\boldsymbol{\theta}^{(i)}$ and the unit-modulus reflection constraint in (8) is non-convex. Towards maximizing $C\left(\boldsymbol{\theta}^{(i)}\right)$, we turn to maximize the upper bound of $C\left(\boldsymbol{\theta}^{(i)}\right)$, which is given by

$$C\left(\boldsymbol{\theta}^{(i)}\right) \leq \frac{N}{N+L_{cp}} \log_2\left(1+\frac{1}{N}\sum_{n=0}^{N-1}\frac{\hat{W}_n^{(i)}\left(\boldsymbol{\theta}^{(i)}\right)P_t}{N\Gamma\sigma^2}\right). \quad (26)$$

Specifically, we reformulate the following optimization problem:

$$\text{(P2):} \quad \max_{\boldsymbol{\theta}^{(i)}} \sum_{n=0}^{N-1}\sum_{k=1}^{K}\left|\sum_{m=1}^{i-1}\theta_m\hat{G}_{k,m,n}+\bar{\theta}_i\hat{\bar{G}}_{k,i,n}+\hat{D}_{k,n}\right|^2 \quad (27)$$

$$\text{s.t.} \quad \left|\bar{\theta}_i\right|=1, \quad |\theta_m|=1, \ \forall m=1,2,\ldots,i-1 \quad (28)$$

where irrelevant terms have been omitted for brevity. Note that the optimal value of problem (P2) is the maximization of the sum channel gain at the AP. In the following, we solve problem (P2) sub-optimally by using either the semidefinite relaxation method or the strongest tap maximization method, which enjoy near-optimal performance and low complexity, respectively.

A. Semidefinite Relaxation (SDR) Method

Let $\mathbf{a}_{k,i,n}^H \triangleq \left[\hat{G}_{k,1,n}\ \hat{G}_{k,2,n}\ \ldots\ \hat{G}_{k,i-1,n}\ \hat{\bar{G}}_{k,i,n}\right]$ denote the estimated CFR vector for the cascaded user-RIS-AP link on the n-th sub-carrier at the k-th receive antenna. Thus, the objective function of (27) in problem (P2) can be rewritten as

$$\sum_{n=0}^{N-1}\sum_{k=1}^{K}\left|\mathbf{a}_{k,i,n}^H\mathbf{v}_i+\hat{D}_{k,n}\right|^2 \quad (29)$$

where $\mathbf{v}_i \triangleq [\theta_1,\theta_2,\ldots,\theta_{i-1},\bar{\theta}_i]^T$. By denoting

$$\mathbf{A}_i = \sum_{n=0}^{N-1}\sum_{k=1}^{K}\mathbf{a}_{k,i,n}\mathbf{a}_{k,i,n}^H \quad (30)$$

$$\mathbf{u}_i = \sum_{n=0}^{N-1}\sum_{k=1}^{K}\mathbf{a}_{k,i,n}\hat{D}_{k,n} \quad (31)$$

$$\bar{\hat{D}} = \sum_{n=0}^{N-1}\sum_{k=1}^{K}\left|\hat{D}_{k,n}\right|^2 \quad (32)$$

problem (P2) is equivalent to

$$\text{(P3):} \quad \max_{\mathbf{v}_i} \mathbf{v}_i^H\mathbf{A}_i\mathbf{v}_i+\mathbf{v}_i^H\mathbf{u}_i+\mathbf{u}_i^H\mathbf{v}_i+\bar{\hat{D}} \quad (33)$$

$$\text{s.t.} \quad \left|[\mathbf{v}_i]_m\right|=1, \quad \forall m=1,2,\ldots,i. \quad (34)$$

It can be readily recognized that problem (P3) is a non-convex quadratically constrained quadratic program (QCQP) problem, which can be reformulated as a homogeneous QCQP problem [34]. Specifically, by defining

$$\tilde{\mathbf{A}}_i = \begin{bmatrix}\mathbf{A}_i & \mathbf{u}_i\\ \mathbf{u}_i^H & 0\end{bmatrix}, \quad \tilde{\mathbf{v}}_i = \begin{bmatrix}\mathbf{v}_i\\ t\end{bmatrix} \quad (35)$$

with t as an auxiliary variable, we transform problem (P3) into the following problem:

$$\text{(P4):} \quad \max_{\boldsymbol{\Phi}_i} \operatorname{tr}\left(\tilde{\mathbf{A}}_i\boldsymbol{\Phi}_i\right)+\bar{\hat{D}} \quad (36)$$

$$\text{s.t.} \quad [\boldsymbol{\Phi}_i]_{m,m}=1, \quad \forall m=1,2,\ldots,i+1 \quad (37)$$

$$\boldsymbol{\Phi}_i \succeq 0 \quad (38)$$

where $\boldsymbol{\Phi}_i = \tilde{\mathbf{v}}_i\tilde{\mathbf{v}}_i^H$ with $\operatorname{rank}(\boldsymbol{\Phi}_i)=1$. However, problem (P4) is still non-convex due to the rank-one constraint. As such, we use the SDR technique to relax the rank-one constraint and further transform problem (P4) into a convex semidefinite program (SDP) problem, which can be solved optimally by using existing convex optimization solvers such as CVX [35]. It is worth pointing out that the optimal objective value of problem (P4) serves as an upper bound on that of problem (P2). On the other hand, the optimal solution $\boldsymbol{\Phi}_i^*$ to problem (P4) may not be a rank-one solution. Therefore, we apply the Gaussian randomization method to construct a rank-one solution, as follows.

$$\tilde{\mathbf{v}}_i^* = \begin{cases}\mathbf{U}\boldsymbol{\Lambda}^{1/2}\mathbf{1}_{(i+1)\times 1}, & \operatorname{rank}(\boldsymbol{\Phi}_i^*)=1\\ \mathbf{U}\boldsymbol{\Lambda}^{1/2}\mathbf{u}, & \operatorname{rank}(\boldsymbol{\Phi}_i^*)\neq 1\end{cases} \quad (39)$$

where $\boldsymbol{\Phi}_i^* = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$ is the eigenvalue decomposition of $\boldsymbol{\Phi}_i^*$ and $\mathbf{u}\sim\mathcal{N}_c(\mathbf{0},\mathbf{I}_{i+1})$ is a random vector. Finally, we obtain a high-quality sub-optimal solution to problem (P2), i.e.,

$$\mathbf{v}_i^* = e^{j\angle([\tilde{\mathbf{v}}_i^*]_{1:i}/[\tilde{\mathbf{v}}_i^*]_{i+1})}. \quad (40)$$

The complexity for solving the SDP problem in (P4) is $\mathcal{O}((i+1)^{4.5})$ [36], which could be practically high for large i.

B. Strongest Tap Maximization (STM) Method

To reduce the computational complexity of solving problem (P2), we propose in this subsection a low-complexity method to solve it sub-optimally by exploiting the strongest channel tap in the time domain. Notice that the objective function in (27) is the total energy of the received signals computed in the frequency domain. According to the Plancherel theorem, signal energy can be computed by summing power-per-sample across time as well. Let $\tilde{\mathbf{d}}_k \triangleq \left[\hat{d}_{k,0}\ \hat{d}_{k,1}\ \ldots\ \hat{d}_{k,N-1}\right]^T$, $\tilde{\mathbf{g}}_{k,m} \triangleq [\hat{g}_{k,m,0}\ \hat{g}_{k,m,1}\ \cdots\ \hat{g}_{k,m,N-1}]^T$, and $\tilde{\bar{\mathbf{g}}}_{k,i} \triangleq \left[\hat{\bar{g}}_{k,i,0}\ \hat{\bar{g}}_{k,i,1}\ \ldots\ \hat{\bar{g}}_{k,i,N-1}\right]^T$ denote the N-point IDFT of $\hat{\mathbf{d}}_k$, $\hat{\mathbf{g}}_{k,m}$, and $\hat{\bar{\mathbf{g}}}_{k,i}$, respectively, $k=1,2,\ldots,K$, $m=1,2,\ldots,i-1$, and $i=1,2,\ldots,M$. Moreover, let $[\tilde{\mathbf{d}}_k]_{L_d+1:N}=\mathbf{0}_{1\times(N-L_d)}$, $[\tilde{\mathbf{g}}_{k,m}]_{L_r+1:N}=\mathbf{0}_{1\times(N-L_r)}$, and $[\tilde{\bar{\mathbf{g}}}_{k,i}]_{L_r+1:N}=\mathbf{0}_{1\times(N-L_r)}$. Thus, problem (P2) is equivalent to the optimization problem given below

$$\text{(P5):} \quad \max_{\boldsymbol{\theta}^{(i)}} \sum_{l=0}^{L-1}\sum_{k=1}^{K}\left|\sum_{m=1}^{i-1}\theta_m\hat{g}_{k,m,l}+\bar{\theta}_i\hat{\bar{g}}_{k,i,l}+\hat{d}_{k,l}\right|^2 \quad (41)$$

$$\text{s.t.} \quad \left|\bar{\theta}_i\right|=1, \quad |\theta_m|=1, \ \forall m=1,2,\ldots,i-1. \quad (42)$$

Note that calculating the objective function in (41) is more efficient than calculating that in (27) since we have $L \leq L_{cp}+1 \ll N$ in practical wireless communication systems. To reduce the computational complexity of solving problem (P5), we turn to find the strongest channel tap with respect to the receive antenna index k and the tap index l, i.e.,

$$\langle \breve{k}, \breve{l} \rangle = \arg \max_{\substack{l\in\{0,\ldots,L-1\}\\ k\in\{1,\ldots,K\}}} \left| \sum_{m=1}^{i-1} |\hat{g}_{k,m,l}| + |\hat{\bar{g}}_{k,i,l}| + \left|\hat{d}_{k,l}\right| \right|^2. \quad (43)$$

Then, we obtain a sub-optimal solution to problem (P2) as follows:

$$\theta_m^* = e^{-j(\angle \hat{g}_{\breve{k},m,\breve{l}} - \angle \hat{d}_{\breve{k},\breve{l}})}, \quad m = 1,2,\ldots,i-1 \quad (44)$$

$$\theta_i^* = e^{-j(\angle \hat{\bar{g}}_{\breve{k},i,\breve{l}} - \angle \hat{d}_{\breve{k},\breve{l}})} \quad (45)$$

which maximizes the energy of the strongest channel tap. The complexity for searching the optimal tap index and the optimal receive antenna index is $\mathcal{O}(KL)$. If there is a strong line-of-sight path in the wireless environment, the strongest channel tap is most likely to be the first tap. As such, we can simply set $\breve{l}=0$ in (43)-(45), and only search for the optimal receive antenna index $\breve{k}$ in (43) with a significantly reduced complexity of $\mathcal{O}(K)$.

Compared to the SDR method, the STM solution enjoys a much lower complexity, yet only suffers from a small performance loss, as will be shown in the simulations. The effectiveness of the proposed STM method can be explained by the fact that the energy of the strongest time-domain channel tap will diffuse to all the sub-carriers in the frequency domain. In particular, the closed-form results in (44) and (45) are the optimal solutions to problems (P1) and (P2) when $K=1$ and $L=1$.

V. Simulation Results and Discussions

In this section, we provide some simulation results to demonstrate the effectiveness of our proposed ATP, which includes the progressive channel estimation scheme and beamforming optimization methods. Transmission frame structures under both the CTP and the proposed ATP are considered for the RIS-assisted OFDM-SIMO system. Moreover, as the RIS is practically deployed in the wireless system to enhance the communication link, the AP does not need to be equipped with too many antennas to achieve good performance, thus reducing the implementation cost. As such, the number of antennas equipped at the AP is set as $K=2$ in our simulations. Each OFDM symbol consists of $N=64$ sub-carriers and is appended by a CP of length $L_{cp}=8$. For the training OFDM symbol, the pilot overhead ratio is 12.5% ($N_p=8$) unless otherwise specified. The Zadoff-Chu sequence [37] is employed as the pilot sequence, i.e., $\mathbf{s}_p = \left[1, e^{j\omega\pi/N_p}, e^{j4\omega\pi/N_p}, \ldots, e^{j(N_p-1)^2\omega\pi/N_p}\right]^T$ with ω being an integer relatively prime to N_p and set to be 3 in our simulations. The frequency-selective fading channel is considered for all the involved links, which is modeled by an exponentially decaying power delay profile with a root-mean-square delay spread, where each tap is generated according to Rayleigh distribution with the spread power decaying factor $\alpha=4$. Since the RIS is practically deployed at the network edge to serve the edge user nearby, the RIS-AP and user-AP channels are expected to have relatively large multipath delay spread due to the long distance, and their numbers of taps in the impulse response are set as $L_2=5$ and $L_d=6$, respectively, whereas the number of taps of the user-RIS channel is set as $L_1=2$ due to the relatively short distance between the nearby edge user and the RIS. Taking the direct user-AP link as a baseline, the average received signal-to-noise ratio (SNR) per sub-carrier is defined as $\mathbb{E}\{|D_{k,n}|^2\}/\mathbb{E}\{|V_{k,n}|^2\} = \mathbb{E}\{|D_{k,n}|^2\}/\sigma^2$. Moreover, the average gain (energy) ratio of the cascaded user-RIS-AP link associated with each RIS element to the direct user-AP link is set to $\epsilon=0.02$ in the simulations. The performance gap of the practical modulation and coding scheme is $\Gamma=9$ dB, the transmission frame totally consists of $T_f=20$ OFDM symbols, the total number of unit cell elements at the RIS is $M_0=100$, and the number of randomizations for the SDR method is 100 in our setting. Other parameters will be specified later to examine their effects on the system performance. In the following simulations, all the results are obtained by averaging more than 100 independent channel realizations.

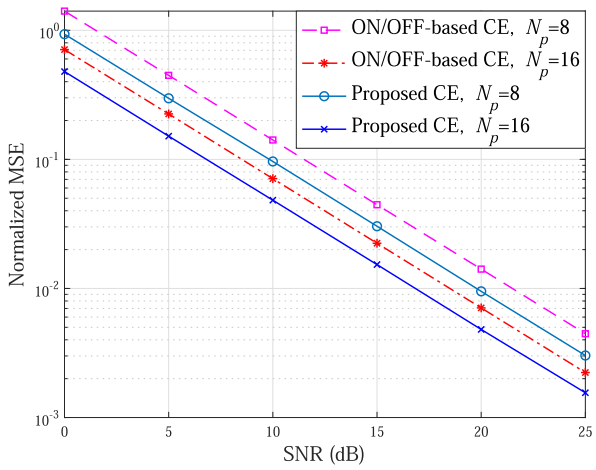


Fig. 4. Normalized MSE versus SNR under different channel estimation schemes with $M=10$.

A. Performance of Proposed Channel Estimation Scheme

In Fig. 4, we examine the performance of different channel estimation schemes in terms of normalized MSE with $M=10$. As can be seen, under the same number of pilot tones N_p, the proposed progressive channel estimation scheme outperforms the ON/OFF-based counterpart [27]. This can be explained by the fact that our proposed channel estimation scheme adopts the full-reflection training pattern to fully achieve the large aperture gain of RIS for channel training. However, for the scheme using the ON/OFF-based training reflection pattern, the large aperture gain of RIS is lost since only one of its sub-surfaces is switched ON at each time, thus degrading the channel estimation accuracy. In addition, by doubling the number of pilot tones N_p, our proposed channel estimation scheme achieves about 3 dB power gain.

B. Performance of Proposed Optimization Methods

In Fig. 5, we compare the achievable rates of different schemes versus the SNR with $M=10$. First, we can observe that all RIS-assisted schemes significantly outperform the one

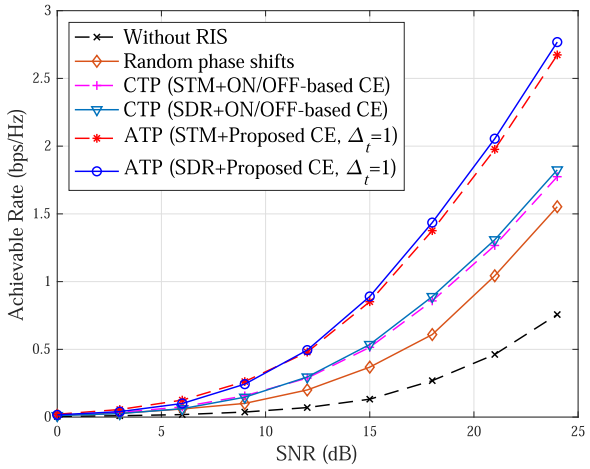


Fig. 5. Achievable rate versus SNR with $M = 10$.

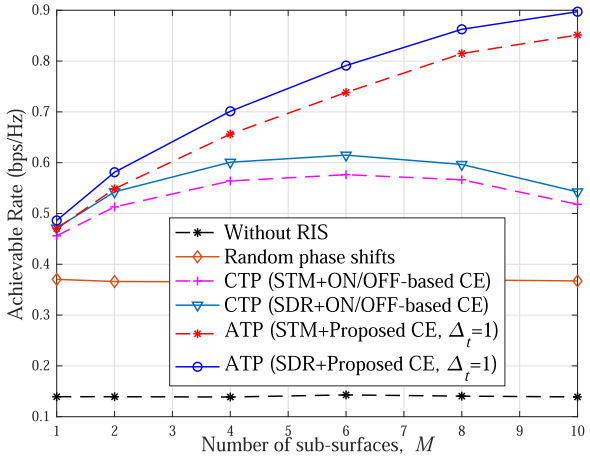


Fig. 6. Achievable rate versus number of sub-surfaces M with SNR equal to 15 dB.

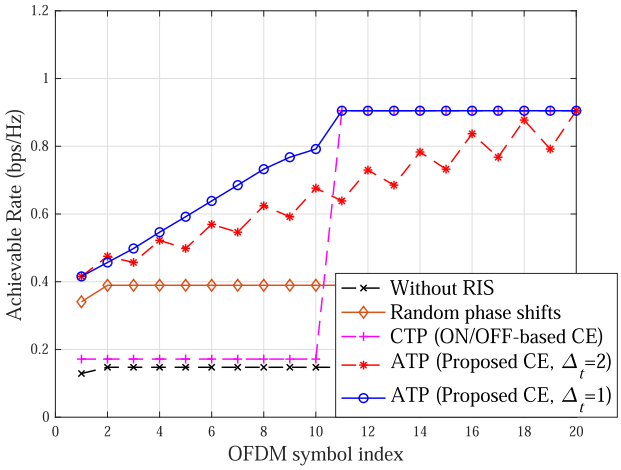


Fig. 7. Achievable rate versus OFDM symbol index with $M = 10$, SNR $= 15$ dB, and STM method for both the CTP and the ATP.

without RIS. Second, it can be observed that the scheme with random phase shifts at the RIS performs worse than the other two schemes with optimal phase shifts obtained by the two proposed optimization methods based on the estimated CSI. Using the same channel estimation method, the two proposed optimization methods achieve comparable performance, while the STM and SDR methods show slightly superior performance to each other at low and high SNRs, respectively. On the other hand, one can observe that the proposed ATP with $\Delta_t = 1$ achieves much better performance than the CTP counterpart. This can be attributed to two aspects of the existing CTP: 1) the CSI obtained by the ON/OFF-based channel estimation method is not accurate enough for achieving high passive beamforming gain; 2) the RIS reflection during the channel training is not optimized to improve the channel gain on data tones. In contrast, our proposed ATP fully reaps the large aperture gain of the RIS by adopting full-reflection of the IRS at all time, and optimizes the RIS reflection during the transmission of training symbols to achieve coherent channel combination on the data tones.

C. Performance of Proposed Transmission Protocol

In Fig. 6, we compare the achievable rates of different schemes versus the number of sub-surfaces M with SNR equal to 15 dB. First, it can be observed that the rate performance of the scheme with random phase shifts at the RIS is invariant with respect to the number of sub-surfaces. This is expected since this scheme only has the aperture gain but no passive beamforming gain over the scheme without RIS, which is independent of the grouping manner. This implies that the aperture gain depends on the number of RIS elements only. Second, we can observe that both the CTP and the proposed ATP show superior performance as compared to the scheme with random phase shifts at the RIS. The achievable rate of the CTP first increases and then decreases with the increase of the number of sub-surfaces M or the training time equivalently, i.e., $M+1$. As pointed out in [28], for the CTP, with too little training the estimated CSI is not accurate enough for achieving high passive beamforming gain, while too much training reduces the effective data transmission time, both reducing the achievable rate. In contrast, the achievable rate of the proposed ATP keeps increasing as the number of sub-surfaces increases. This can be understood since both the estimated CSI and the passive beamforming performance are progressively improved during the channel training. On one hand, by increasing the number of sub-surfaces (training time), the accuracy of estimated CSI can be further improved, i.e., more channels associated with the RIS can be resolved for achieving higher passive beamforming gain. On the other hand, more refined CSI is achieved without sacrificing the achievable rate, since the RIS performs adaptive passive beamforming to enhance the channnel gain on the data tones during the channel training.

In Fig. 7, we compare the achievable rates of different schemes versus the OFDM symbol index with $M = 10$, SNR $= 15$ dB, and the STM method. First, it is observed that unlike the CTP, the achievable rates of the proposed ATP with different Δ_t values increase progressively before acquiring the full CSI. Given a target transmission rate, the proposed ATP requires a much shorter time than the CTP. This is because the proposed scheme progressively resolves the CSI of the cascaded user-RIS-AP channels over consecutive subframes and accordingly refines the passive beamforming to improve the achievable rate, which thus doesn't need to wait until the end of channel training. However, the CTP scheme has to wait until the end of channel training for passive beamforming,

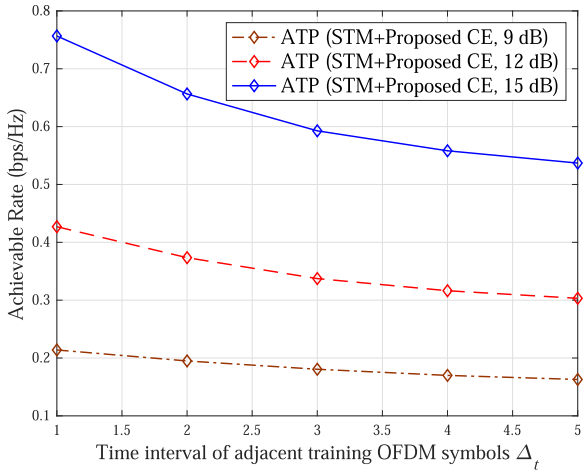


Fig. 8. Effect of time interval of adjacent training OFDM symbols Δ_t on the achievable rate of proposed ATP with $M = 10$.

so it encounters a very low rate during the channel training, resulting in a longer delay. Moreover, it can be observed that before acquiring the full CSI, the achievable rate of the ATP with $\Delta_t = 1$ increases steadily, while with $\Delta_t = 2$, the achievable rate fluctuates with the OFDM symbol index. This can be explained by the fact that the increased design degrees of freedom for passive beamforming improves the achievable rate, while the reduced number of data tones due to pilot insertion reduces the achievable rate. On the other hand, it is clear from Fig. 7 that given the same target transmission rate, the ATP with a larger Δ_t has a longer delay. The reasons can be found in Fig. 8, which plots the average achievable rate of the proposed ATP versus the time interval of adjacent training OFDM symbols Δ_t with $M = 10$. It is shown in Fig. 8 that the average achievable rate of the proposed ATP monotonically decreases with the time interval of adjacent training OFDM symbols Δ_t. In addition, as shown in Fig. 7, for the scheme without RIS and the one with random phase shifts at the RIS, their achievable rates increase after acquiring the CSI of the direct user-AP channel.

VI. Conclusion

In this paper, we have proposed an adaptive transmission protocol in conjunction with a progressive channel estimation method for the wideband RIS-assisted OFDM system to execute channel estimation and passive beamforming simultaneously. In particular, the RIS reflection was designed for embedding the training reflection pattern and enhancing data transmission on the data tones simultaneously during the channel training. Next, we formulated an optimization problem to maximize the average achievable rate by designing the passive beamforming at the RIS based on the estimated CSI in each sub-frame subject to the unit-modulus reflection constraint. The SDR and STM methods were proposed to obtain high-quality sub-optimal solutions for the formulated problem. Simulation results have demonstrated the superiority of the proposed protocol over the existing one. It was shown that the proposed progressive channel estimation and passive beamforming schemes were able to drastically improve the average achievable rate and reduce the delay for data transmission as compared to existing schemes.

References

[1] F. Boccardi, R. W. Heath, Jr., A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[2] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, p. 501–513, Apr. 2016.

[3] H. Q. Ngo, E. Larsson, and T. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, p. 1436–1449, Apr. 2013.

[4] F. Liu *et al.*, "Programmable metasurfaces: State of the art and prospects," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Florence, Italy, 2018, pp. 1–5.

[5] T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, "Coding metamaterials, digital metamaterials and programmable metamaterials," *Light, Sci. Appl.*, vol. 3, no. 10, p. e218, Oct. 2014.

[6] G. C. Alexandropoulos, G. Lerosey, M. Debbah, and M. Fink, "Reconfigurable intelligent surfaces and metamaterials: The potential of wave propagation control for 6G wireless communications," *IEEE ComSoc TCCN Newslett.*, vol. 6, no. 1, Jun. 2020. [Online]. Available: https://arxiv.org/abs/2006.11136

[7] M. Di Renzo *et al.*, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and road ahead," 2020, *arXiv:2004.09352*. [Online]. Available: http://arxiv.org/abs/2004.09352

[8] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A new wireless communication paradigm through software-controlled metasurfaces," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 162–169, Sep. 2018.

[9] X. Tan, Z. Sun, D. Koutsonikolas, and J. M. Jornet, "Enabling indoor mobile millimeter-wave networks based on smart reflect-arrays," in *Proc. IEEE Conf. Comput. Commun.*, Hawaii, HI, USA, Apr. 2018, pp. 270–278.

[10] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116753–116773, 2019.

[11] M. D. Renzo *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, May 2019, Art. no. 129.

[12] R. Karasik, O. Simeone, M. Di Renzo, and S. Shamai, "Beyond max-SNR: Joint encoding for reconfigurable intelligent surfaces," 2019, *arXiv:1911.09443*. [Online]. Available: http://arxiv.org/abs/1911.09443

[13] W. Tang *et al.*, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," 2019, *arXiv:1911.05326*. [Online]. Available: http://arxiv.org/abs/1911.05326

[14] N. S. Perović, M. Di Renzo, and M. F. Flanagan, "Channel capacity optimization using reconfigurable intelligent surfaces in indoor mmWave environments," 2019, *arXiv:1910.14310*. [Online]. Available: http://arxiv.org/abs/1910.14310

[15] C. Huang *et al.*, "Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends," 2019, *arXiv:1911.12296*. [Online]. Available: http://arxiv.org/abs/1911.12296

[16] M. Di Renzo and J. Song, "Reflection probability in wireless networks with metasurface-coated environmental objects: An approach based on random spatial processes," 2019, *arXiv:1901.01046*. [Online]. Available: http://arxiv.org/abs/1901.01046

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

[18] C. Huang, G. C. Alexandropoulos, A. Zappone, M. Debbah, and C. Yuen, "Energy efficient multi-user MISO communication using low resolution large intelligent surfaces," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Abu Dhabi, UAE, Dec. 2018, pp. 1–6.

[19] E. Bjornson, O. Ozdogan, and E. G. Larsson, "Intelligent reflecting surface versus decode-and-forward: How large surfaces are needed to beat relaying?" *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 244–248, Feb. 2020.

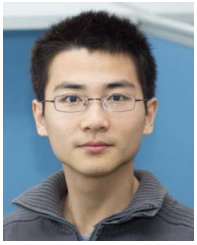
[20] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[21] B. Zheng, Q. Wu, and R. Zhang, "Intelligent reflecting surface-assisted multiple access with user pairing: NOMA or OMA?" *IEEE Commun. Lett.*, vol. 24, no. 4, pp. 753–757, Apr. 2020.

[22] C. Huang, G. C. Alexandropoulos, C. Yuen, and M. Debbah, "Indoor signal focusing with deep learning designed reconfigurable intelligent surfaces," in *Proc. IEEE 20th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Cannes, France, Jul. 2019, pp. 1–5.

[23] Z.-Q. He and X. Yuan, "Cascaded channel estimation for large intelligent metasurface assisted massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 210–214, Feb. 2020.
[24] G. C. Alexandropoulos and E. Vlachos, "A hardware architecture for reconfigurable intelligent surfaces with minimal active elements for explicit channel estimation," in *Proc. ICASSP - IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Barcelona, Spain, May 2020, p. 9175.
[25] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling large intelligent surfaces with compressive sensing and deep learning," 2019, *arXiv:1904.10136*. [Online]. Available: http://arxiv.org/abs/1904.10136
[26] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Brighton, U.K., May 2019, pp. 4659–4663.
[27] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization," *IEEE Trans. Commun.*, Early Access, Mar. 17, 2020, doi: 10.1109/TCOMM.2020.2981458.
[28] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 518–522, Apr. 2020.
[29] C. You, B. Zheng, and R. Zhang, "Intelligent reflecting surface with discrete phase shifts: Channel estimation and passive beamforming," 2019, *arXiv:1911.03916*. [Online]. Available: http://arxiv.org/abs/1911.03916
[30] L. Wei, C. Huang, G. C. Alexandropoulos, and C. Yuen, "Parallel factor decomposition channel estimation in RIS-assisted multi-user MISO communication," in *Proc. IEEE 11th Sensor Array Multichannel Signal Process. Workshop (SAM)*, Hangzhou, China, Jun. 2020, pp. 1–5.
[31] B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface assisted multi-user OFDMA: Channel estimation and training design," 2020, *arXiv:2003.00648*. [Online]. Available: http://arxiv.org/abs/2003.00648
[32] H. Yang *et al.*, "Design of resistor-loaded reflectarray elements for both amplitude and phase control," *IEEE Antennas Wireless Propag. Lett.*, vol. 16, pp. 1159–1162, 2017.
[33] Y. Mostofi, "ICI mitigation and time synchronization for pilot-aided OFDM mobile systems," Ph.D. dissertation, Dept. Elect. Eng., Stanford Univ., Stanford, CA, USA, 2003.
[34] A. M.-C. So, J. Zhang, and Y. Ye, "On approximating complex quadratic optimization problems via semidefinite programming relaxations," *Math. Program.*, vol. 110, no. 1, pp. 93–110, Mar. 2007.
[35] M. Grant and S. Boyd. (Jan. 2020). *CVX: MATLAB Software for Disciplined Convex Programming*. [Online]. Available: http://cvxr.com/cvx
[36] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, p. 20–34, May 2010.
[37] D. Chu, "Polyphase codes with good periodic correlation properties (Corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 4, pp. 531–532, Jul. 1972.

Shaoe Lin received the B.S. and M.S. degrees from the South China University of Technology, Guangzhou, China, in 2013 and 2016, respectively, where she is currently pursuing the Ph.D. degree. Since 2019, she has been a Visiting Student Research Collaborator with the University of Waterloo, Waterloo, ON, Canada. Her current research interests include MIMO systems, OFDM, and index modulation.



Beixiong Zheng (Member, IEEE) received the B.S. and Ph.D. degrees from the South China University of Technology, Guangzhou, China, in 2013 and 2018, respectively.

From 2015 to 2016, he was a Visiting Student Research Collaborator with Columbia University, New York, NY, USA. He is currently a Research Fellow with the Department of Electrical and Computer Engineering, National University of Singapore. His current research interests include intelligent reflecting surface (IRS), index modulation (IM), and non-orthogonal multiple access (NOMA). He was a recipient of the Best Paper Award from the IEEE International Conference on Computing, Networking, and Communications in 2016, the Best Ph.D. Thesis Award from China Education Society of Electronics in 2018, and the Outstanding Reviewer of Physical Communication in 2019.

George C. Alexandropoulos (Senior Member, IEEE) was born in Athens, Greece, in 1980. He received the Engineering Diploma degree in computer engineering and informatics, the M.A.Sc. degree (Hons.) in signal processing and communications, and the Ph.D. degree in wireless communications from the University of Patras (UoP), Rio-Patras, Greece, in 2003, 2005, and 2010, respectively. From 2001 to 2014, he has held research positions at various Greek universities and research institutes, such as UoP, University of Peloponnese, Technical University of Crete, National Center for Scientific Research Demokritos, National Observatory of Athens, and the Athens Information Technology Center for Research and Education, where he technically managed several national, European, and international Research and Development projects, and lectured mathematics and computer engineering courses. From 2014 to January 2019, he was a Senior Research Engineer with the Mathematical and Algorithmic Sciences Laboratory, Paris Research Center, Huawei Technologies France SASU delivering RAN solutions for 5G NR and beyond, and mainly for full duplex and massive MIMO, hybrid A/D beamforming management, and reconfigurable metasurfaces. He is currently an Assistant Professor for communication systems and signal processing with the Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Greece. His research and development activities span the general areas of algorithmic design, optimization, and performance analysis for wireless communication networks with emphasis on multiantenna systems, transceiver hardware architectures, high-frequency communications, and distributed machine learning algorithms.

Dr. Alexandropoulos is a Senior Member of the IEEE Communications and Signal Processing Societies and a Professional Engineer of the Technical Chamber of Greece. He currently serves as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE COMMUNICATIONS LETTERS, and *Elsevier Computer Networks*. He has received scholarships for his postgraduate and Ph.D. studies, a student travel grant for the 2010 IEEE Global Telecommunications Conference, the Best Ph.D. Thesis Award 2010 by a Greek University in the fields of informatics and telecommunications, and the IEEE Communications Society Best Young Professional in Industry Award 2018.

Miaowen Wen (Senior Member, IEEE) received the Ph.D. degree from Peking University, Beijing, China, in 2014.

From 2012 to 2013, he was a Visiting Student Research Collaborator with Princeton University, Princeton, NJ, USA. He is currently an Associate Professor with the South China University of Technology, Guangzhou, China, and a Hong Kong Scholar with The University of Hong Kong, Hong Kong. He has published a Springer book entitled *Index Modulation for 5G Wireless Communications* and more than 90 journal articles. His research interests include a variety of topics in the areas of wireless and molecular communications. He was a recipient of four Best Paper Awards from the IEEE ITST'12, the IEEE ITSC'14, the IEEE ICNC'16, and the IEEE ICCT'2019. He was recognized as an Exemplary Reviewer of the IEEE COMMUNICATIONS LETTERS in 2017 and the IEEE TRANSACTIONS ON COMMUNICATIONS in 2019. He has served as an Editor for IEEE ACCESS and the *EURASIP Journal on Wireless Communications and Networking*. He was a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Special Issue on Spatial Modulation for Emerging Wireless Systems) and the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (Special Issue on Index Modulation for Future Wireless Networks: A Signal Processing Perspective). He is serving as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE COMMUNICATIONS LETTERS, and the *Physical Communication* (Elsevier), and a Guest Editor for the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (Special Issue on Advanced Signal Processing for Local and Private 5G Networks).

Fangjiong Chen (Member, IEEE) received the B.S. degree in electronics and information technology from Zhejiang University, Hangzhou, China, in 1997, and the Ph.D. degree in communication and information engineering from the South China University of Technology, Guangzhou, China, in 2002.

He was with the School of Electronics and Information Engineering, South China University of Technology. From 2002 to 2005, he was a Lecturer, and from 2005 to 2011, he was an Associate Professor with the South China University of Technology. He is currently a full-time Professor with the School of Electronics and Information Engineering, South China University of Technology. He is the Vice Director of the Engineering Center for Short-Range Wireless Communication and Network, Ministry of Education. His research interests include signal detection and estimation, array signal processing, and wireless communication.

Prof. Chen received the National Science Fund for Outstanding Young Scientists in 2013. He was elected to the New Century Excellent Talent Program of MOE, China, in 2012.

Shahid Mumtaz (Senior Member, IEEE) received the master's degree in electrical and electronic engineering from the Blekinge Institute of Technology, Sweden, in 2006, and the Ph.D. degree in in electrical and electronic engineering from the University of Aveiro, Portugal, in 2011, respectively.

He is currently an ACM Distinguished Speaker. He is the author of four technical books, 12 book chapters, more than 180 technical articles (more than 130 Journal/transaction, more than 70 conferences, two IEEE Best Paper Awards in the area of mobile communications. He was a recipient of the IEEE ComSoC Young Researcher Award in 2020. He is the Founder and EiC of the IET *Journal of Quantum Communication*. He is the Vice-Chair: Europe/Africa Region-IEEE ComSoc: Green Communications and Computing Society and the IEEE standard on P1932.1: Standard for Licensed/Unlicensed Spectrum Interoperability in Wireless Mobile Networks. This standard resulted from his novel idea on WiFi in a Licensed band. He is also a Senior 5G Consultant at Huawei, Sweden. He has more than 12 years of wireless industry/academic experience.