RECONFIGURABLE-INTELLIGENT-SURFACE EMPOWERED WIRELESS COMMUNICATIONS: CHALLENGES AND OPPORTUNITIES

Xiaojun Yuan, Ying-Jun Angela Zhang, Yuanming Shi, Wenjing Yan, and Hang Liu

ABSTRACT

Reconfigurable intelligent surfaces (RISs) are regarded as a promising emerging hardware technology to improve the spectrum and energy efficiency of wireless networks by artificially reconfiguring the propagation environment of electromagnetic waves. Due to the unique advantages in enhancing wireless channel capacity, RISs have recently become a hot research topic. In this article, we focus on three fundamental physical-layer challenges for the incorporation of RISs into wireless networks, namely, channel state information acquisition, passive information transfer, and low-complexity robust system design. We summarize the state-of-the-art solutions and explore potential research directions. Furthermore, we discuss other promising research directions of RISs, including edge intelligence and physical-layer security.

INTRODUCTION

Wireless connectivity is a key enabler to support the evolution from "connected things" to "connected intelligence" in the future information society. In particular, ultra-high data rates, ultra-high reliability, ultra-low latency, and extremely massive connectivity are essential for data sensing, collection, transmission, and processing across future wireless networks. This is achieved by the key enabling wireless technologies, including massive multiple-input multiple-output (MIMO), millimeter wave (mmWave), ultra dense networks and AI-empowered wireless networks [1], complemented by new network functionalities such as edge computing, caching, learning and network slicing. The development of existing wireless systems is based on a basic principle that the radio environments are fixed exogenously and cannot be controlled. Combating the detrimental effects of the radio environments has been the main aim of the design of algorithms, protocols, and systems. Moving forward to unleash the full potential of future wireless networks, it is time to investigate the feasibility of building a controllable and programable radio environment, so that the radio environment itself becomes a degree of freedom for system optimization.

Reconfigurable intelligent surfaces (RISs), a.k.a., large intelligent metasurfaces (LIMs) [2] and intelligent reflecting surfaces (IRSs) [3], have been envisioned to reduce the energy consumption and improve the spectral efficiency of wireless networks by artificially reconfiguring the propagation environment of electromagnetic waves. Specifically, RISs are typically constructed by the planar (or even conformal) artificial metasurfaces consisting of many reflection amplitude/phase shifts, which are adjustable by a smart controller. The RIS-empowered smart radio environments are thus able to control the phases and/or amplitudes of the incident signals, thereby enhancing the desired signals and mitigating the interference signals. As such, RISs have a huge potential to revolutionize the design of wireless networks, particularly when combined and integrated together with other promising wireless technologies such as ultra-massive MIMO, terahertz communications, AI-empowered wireless networks, and edge intelligence. As such, the purpose of this article is to draw attention to and spur activities on this new research direction.

RIS-empowered smart radio is able to combat the unfavorable propagation conditions (such as deep fading) by manipulating the radio propagation environment in future-generation wireless communications. Millimetre wave and Terahertz communications have been envisioned as enabling technologies for 5G-and-beyond wireless communications. As the increase of the radio frequency, tens and even hundreds of antenna elements will be installed on base stations and even on portable devices. As such, massive MIMO with antenna arrays deployed at both base stations and device terminals is able to provide unprecedented capacity gains to meet the exponential increase in the demand of wireless data services. Meanwhile, however, the increase of the radio frequency weakens the diffraction and scattering effect, rendering electromagnetic waves prone to blockage by obstacles such as buildings in urban areas. As a result, it is difficult to ensure a universal coverage of wireless services in 5G-and-beyond wireless communications using conventional cellular techniques. The recent advancement of the RIS technique provides a revolutionarily new solution to tackle the problem by artificially controlling the propagation environment of electromagnetic waves. Typically, a RIS is composed of a large number of low-cost and energy-efficient reconfigurable reflecting elements that can reflect impinging electromagnetic waves with a controllable phase shift via the help of a

Xiaojun Yuan and Wenjing Yan are with the University of Electronic Science and Technology of China; Ying-Jun Angela Zhang and Hang Liu are with The Chinese University of Hong Kong; Yuanming Shi is with ShanghaiTech University.

smart controller. Through intelligent placement and reflect/passive beamforming, a RIS is able to provide an extra high-quality channel link to overcome the unfavorable propagation conditions of wireless communication systems.

There are undeniable performance advantages of RIS-empowered smart radio. First, RISs can be deployed almost everywhere to make use of electromagnetic waves that are otherwise dissipated in space. Reconfigurable electromagnetic materials can be used to coat objects in the environment, including but not limited to building facades, ceilings, furniture, and clothes, and so on. Second, RISs are environmentally friendly to meet the requirement of green communications. As the RISs are nearly passive, no additional energy is consumed by RIS-aided systems compared to conventional wireless systems. Third, RISs support full-duplex and full-band transmission, since they only reflect electromagnetic waves. In addition, RISs are cost-effective since they need neither analog-to-digital/digital-to-analog converters nor power amplifiers. Fourth, the power gain of a RIS follows the quadratic scaling law, in contrast to the linear power scaling law of a conventional active antenna array.

At the same time, however, the use of RISs poses a number of new challenges for the transceiver design of wireless communication systems. In particular, compared with the existing magazine articles [3–5], we mainly focus on three physical-layer problems, namely, channel state information (CSI) estimation, passive information transfer, and low-complexity robust system design. These are chosen because they pose fundamental and unique challenges for the design of RIS-aided wireless networks. More specifically, the CSI acquisition of a RIS-aided channel is very different from that of a conventional channel since the transmitter-RIS and RIS-receiver links are cascaded and difficult to measure separately due to the passive nature of the RIS. In addition, the RIS is required to communicate with the transceiver to coordinate data transmission, which gives rise to the necessity of passive information transfer. Given these new features of the physical-layer design, it is therefore imperative to understand the fundamental tradeoffs between system performance and computational cost, so as to fully exploit the potential of RIS-empowered wireless networks. We will also highlight some new research directions, including RIS-aided edge intelligence and RIS-aided physical-layer security, to encourage future research endeavors.

The remainder of this article is organized as follows. The following section discusses channel acquisition in RIS systems. We then discuss passive information transfer in RIS systems. Robust and low-complexity system design is then discussed. Following that we highlight some of the other research challenges. The article is concluded in the final section.

Channel State Information Acquisition in RIS-Aided Communication Systems
Problem Description

The acquisition of CSI is a fundamental problem to achieve the full potential of RIS-aided wireless networks. Recent studies show that the transceiv-
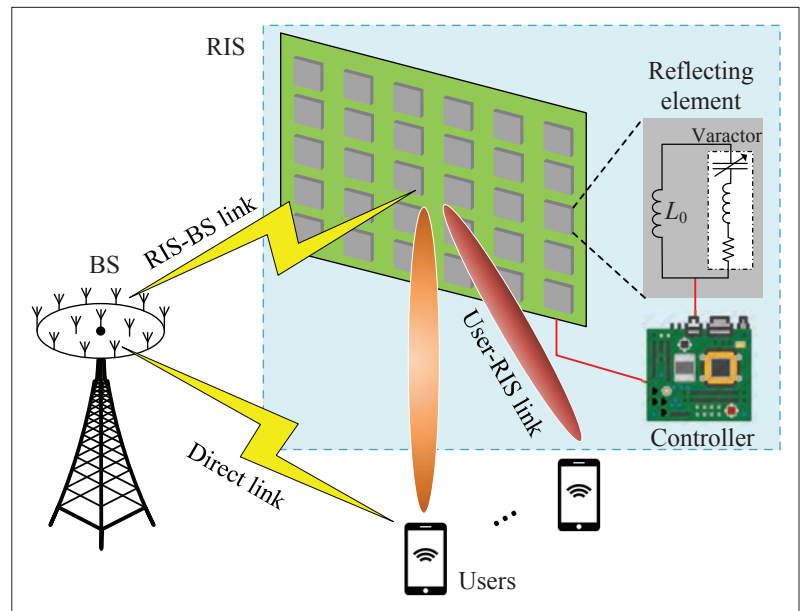

FIGURE 1. A RIS-assisted massive MIMO system.

er design of a RIS-aided system critically depends on the knowledge of CSI, for example, the joint active and passive beamforming design in [3] and the joint transmit power allocation and beamforming design in [6].

The CSI acquisition problem in a RIS-aided system, however, is quite different from that in traditional MIMO communication systems. To be specific, we consider a typical RIS-aided massive MIMO communication system, where a number of users communicate with a multi-antenna base station (BS) via the help of a RIS, as illustrated in Fig. 1. In such a system, besides the estimation of the direct channel link (i.e., the channel link between the BS and the users), two additional channel links, namely, the user-RIS channel link and the RIS-BS channel link, also need to be estimated. By turning off all the RIS elements, the direct channel can be obtained based on traditional channel estimation methods. Yet, the remaining problem of estimating the user-RIS and RIS-BS links is far more difficult since the RIS is expected to be a nearly passive device with very limited capability of transmitting, receiving, and processing radio frequency (RF) signals. This means that unlike conventional pilot-assisted channel estimation, one can neither rely on the RIS to estimate the user-RIS and RIS-BS channel links by processing the pilots from the users and the BS, nor rely on the RIS to transmit pilot signals to facilitate the channel estimation at the users and the BS. As such, CSI acquisition in RIS-aided systems gives rise to a cascaded channel estimation problem, that is, the problem of estimating the user-RIS and RIS-BS channel links upon the observation of their noisy product. We henceforth refer to the cascade of the user-RIS and the RIS-BS links as the cascaded channel. The cascaded channel estimation problem is in general a bilinear estimation problem, in contrast to the linear estimation problem for channel estimation in conventional communication systems. In addition, the size of a RIS is usually very large (e.g., in the scale of tens or hundreds). This implies that a large number of channel coefficients need to be estimated, which

Since the channel coefficients of a large antenna array at the RIS have strong correlation, these coefficients can be constructed based on the sampled channel information by utilizing compressive sensing tools. The channel links from the RIS to the users and to the BS can be obtained by assuming channel reciprocity.

imposes extra difficulty on the cascaded channel estimation problem.

STATE-OF-THE-ART SOLUTIONS

At present, the design of CSI acquisition in RIS-aided systems is still in its infancy. Initial attempts to solve this problem can be roughly divided into three categories.

Active-Channel-Sensor Based CSI Acquisition: This approach is based on the insertion of active channel sensors into the array of passive elements for sensing channel information [7]. Each active channel sensor is equipped with not only an RF phase shifter like a passive reflecting element for reflecting the incident electromagnetic (EM) wave, but also an additional baseband processing unit for the channel estimation. Correspondingly, the active sensors have two work modes, namely, the channel sensing mode (using the baseband unit for channel estimation) and the reflection mode (using the RF phase shifters for reflecting EM wave) [7]. During the channel sensing mode, the sensors receive the pilot signals from the users and the BS to estimate their corresponding channel links. Since the channel coefficients of a large antenna array at the RIS have strong correlation, these coefficients can be constructed based on the sampled channel information by utilizing compressive sensing tools. The channel links from the RIS to the users and to the BS can be obtained by assuming channel reciprocity.

The active-channel-sensor based approach, however, has some disadvantages. First, the active sensors require additional baseband processing units, which increases the hardware cost of the RIS. Second, the active sensors consume extra energy, which may pose a heavy burden on the RIS as a nearly passive device. Last but not least, the channel information obtained at the RIS needs to be uploaded to a control centre (usually located at the BS) for beamforming design and resource allocation, which gives rise to the need of information transfer for RIS. To address the above issues, we next describe two CSI acquisition approaches based on a RIS with all passive elements.

Channel-Decomposition Based CSI Acquisition: As discussed earlier, the cascaded channel estimation problem arising from RIS-aided systems is difficult to solve due to the multiplication of the two coefficient matrixes of the user-RIS and RIS-BS links. An idea to tackle this difficulty is to decompose the cascaded channel into a series of sub-channels that are easier to estimate. For example, one may decompose the cascaded channel into a series of rank-1 matrices with each corresponding to a RIS element. Each sub-channel can be estimated by turning on only one RIS element (and turning off all the other elements). Performing this procedure for each RIS element, the CSI of the whole cascaded channel can be obtained. This method has been adopted in [8] to estimate the channels in a RIS-aided multiple-input single-output system. A total pilot length of MNK is required for reliable CSI acquisition, where M and N are the numbers of the antennas/elements at the BS and at the RIS, respectively. This method can be extended to the multiuser cases. However, the required pilot length increases to NK (with K being the number of users), incurring a prohibitively high training overhead when K is large.

Another decomposition method for cascaded channel estimation is to estimate the channel by activating each user one by one, that is, the cascaded channel is decomposed into a series of single-input multiple-output channels seen by each user. By exploiting the fact that all the users share a common RIS-BS link, the authors in [9] proposed an efficient three-phase channel estimation method, which reduces the required pilot length to

$$K + N + \max\left(K-1, \frac{(K-1)N}{M} \right).$$

Structure-Learning Based CSI Acquisition: The cascaded channel of a RIS-aided system usually exhibits strong structural features, such as sparsity and low-rankness, which can be exploited to reduce the overhead of CSI acquisition. The deployment of large-scale antenna arrays enables the RIS and the BS to distinguish EM waves from different directions with high resolution, thereby yielding a sparse representation of the channel matrices in the angular domain (i.e., a large portion of angular channel coefficients are zeros). Likewise, the low-rankness of the cascaded channel matrices arises from the effect of a limited number of scattering paths in the user-RIS and RIS-BS propagation environment. Moreover, signal sparsity can be artificially introduced to assist the cascaded channel estimation by controlling the on/off states of the RIS elements. With this structural information, the estimation of a cascaded channel can be done by utilizing advanced signal processing tools, such as compressed sensing, sparse matrix factorization, and low-rank matrix recovery algorithms. For example, the authors in [2] proposed a two-stage algorithm to estimate the cascaded channel in the RIS-aided MIMO systems. The algorithm includes a sparse matrix factorization stage to estimate the information of the RIS-BS channel link, and a matrix completion stage to estimate the information of the user-RIS channel link. Besides channel sparsity, the authors in [10] also exploited the quasi-static property of the RIS-BS channel link (i.e., the channel components of RIS-BS channel vary much slower than those of the user-BS channel link and the user-RIS channel link since the RIS is usually deployed at a fixed location) to further reduce the training overhead, and proposed a matrix-calibration-based cascaded channel estimation method to solve the corresponding estimation problem. Table 1 gives a heuristic comparison of the minimum training lengths of the channel estimation algorithms in [2, 8–10] versus the number of RIS elements.[1]

RESEARCH CHALLENGES

Channel Modelling and Channel Acquisition: Channel modelling of RIS-aided MIMO systems has not yet been well understood. A conventional MIMO channel is usually assumed to be far-field, where EM waves impinge upon an antenna array nearly in parallel under the assumption that the radio source, the scatterers, and the receiver are located sufficiently apart from each other. However, the passive antenna array of a RIS, coated on a facade of a building or on the ceiling of a room, can have a size that is comparable with its distance from the base stations or the mobile devices. As such, it is necessary to take into account near-field propagation properties in modelling the

[1] Note that the minimum training lengths in [8] and [9] are calculated based on the theoretical formulas given earlier, while the figures of [2, 10], and the analytical lower bound in [10] are obtained based on simulations with the key parameter settings are listed in the notes of Table 1. We refer the interested readers to [2] and [10] for details.

BS-RIS-user channels. Other propagation properties, such as line-of-sight (LoS)/non-LoS/narrow-band/broadband, and so on, may also be radically different for RIS-aided MIMO systems compared to conventional MIMO systems. Therefore, new models are needed to characterize the propagation environment of a RIS-aided MIMO system more precisely.

The new models also encourage the use of new mathematical tools in channel acquisition. For example, matrix factorization and matrix completion techniques are involved in the cascaded channel estimation algorithm developed in [2]. Yet, the channel model employed in [2] is very primitive. We believe that with more realistic channel modelling, other advanced statistical signal processing techniques, such as tensor factorization and structured signal reconstruction, will find their roles in reliable acquisition of the CSI.

System Design Under CSI Uncertainty: The existing studies on the design of RIS-aided systems are mostly based on the assumption of perfect CSI, so that the phases of the reflecting elements of the RIS can be judiciously adjusted for performance enhancement. As mentioned previously, in practice, CSI acquisition of the cascaded channel links is a difficult problem due to limited signal processing capabilities of the RIS. As such, the design of RIS-aided systems including transceiver design and passive beamforming optimization at the RIS needs to be carried out under CSI uncertainty. It is worth mentioning that the outdated CSI problem, as a model of CSI uncertainty due to CSI acquisition delay, is more likely to occur in RIS-aided systems. The reason is that, besides the CSI of the direct link, a large amount of the channel coefficients of the user-RIS and RIS-BS links need to be delivered between the transceiver and the RIS, which causes additional delay. This inspires us to investigate the design of RIS-aided systems under outdated CSI.

In addition, it is known that in wireless communication systems, joint channel estimation and signal detection provides significant performance improvement over separate processing approaches. Thus, it is desirable to investigate the possibility of jointly estimating the user-RIS and RIS-BS channels and at the same time detecting the data from the users. This is a highly non-linear signal estimation problem that calls for urgent solutions.

PASSIVE INFORMATION TRANSFER OF RIS
WHY PASSIVE INFORMATION TRANSFER?

The existing studies mostly focused on the utilization of RISs to enhance the primary end-to-end communications by performing passive beamforming. However, in practice, RISs also need to transfer information. The potential sources of the RIS information are listed as follows.

Control Signaling of RIS: To coordinate with the transceiver, a RIS is required to report its state information in real time. For example, to synchronize with the transceiver for packet delivery, the RIS needs to acknowledge its current status by sending out control signals.

Maintenance of RIS: It is important to monitor the environmental conditions (such as temperature, humidity, pressure, and so on) of the RIS in real time to guarantee its normal operation. In addition, if some elements of the RIS are impaired,

Algorithm	No. of RIS elements				
	128	200	256	300	400
Method of [8]	12800	20000	25600	30000	40000
Method of [2]	245	324	438	468	981
Method of [9]	227	299	383	449	598
Method of [10]	53	55	58	65	83
Analytical lower bound in [10]	20	36	40	47	62

[a] We consider a RIS-aided uplink MIMO system where the number of antennas at the BS and the number of single antenna users are 200 and 100, respectively. The noise power is set to 0. For the method of [2], the rank of the user-RIS channel is 5. The minimum training lengths of respective algorithms are obtained when the channel MSEs are less than -60 dB. The active rate of the RIS elements required in [2] is set to 0.15.

[b] The unit used for the training lengths in the table is "channel use."

TABLE 1. Minimum training lengths of respective channel estimation algorithms versus the number of RIS elements.

such information needs to be reported to the control center.

Assistance of CSI Acquisition: As mentioned earlier, with active-channel-sensor based CSI acquisition, the channel is estimated at the RIS based on the received signals of the inserted active sensors. The CSI acquired at the RIS needs to be forwarded to the transmitter for beamforming design.

Green IoT: Hardware cost and energy consumption are fundamental bottlenecks to constrain the extensive implementation of IoT devices. The combination of IoT with backscatter communication is regarded as a promising solution to overcome these obstacles and achieve green IoT [11]. Compared with backscattering, RISs are able to reflect incident signals in a much more efficient way. This inspires the use of RISs to assist information transfer of IoT devices.

Based on the above discussions, information transfer at RISs is a critical problem for the development of the RIS technology. An immediate solution to this problem is to equip each RIS with a dedicated transmitter. However, this solution is not cost-effective and requires extra power consumption. Instead, it is more desirable to modulate the RIS data onto its reflected signals to achieve passive information transfer. We next briefly introduce the state-of-the-art research along this line.

STATE-OF-THE-ART SOLUTIONS

In [12], the RIS is regarded as an access point by assuming that the RIS is supported by a nearby radio frequency (RF) signal generator. To transmit RIS information, the RF signal generator emits an unmodulated carrier signal to the RIS, and the RIS modulates its information onto the reflected carrier signal. Simultaneously, the RIS is required to maximize the received signal-to-noise ratio (SNR) by manipulating the phase shifts of the RIS. However, the use of the dedicated RF signal generator is not cost-effective.

In [13], the authors proposed a joint passive beamforming and information transfer (PBIT) technique for the RIS-aided communication systems, which aims to simultaneously transmit the RIS information and enhance the primary communication quality. Furthermore, the authors proposed to adopt the spatial modulation on the index of
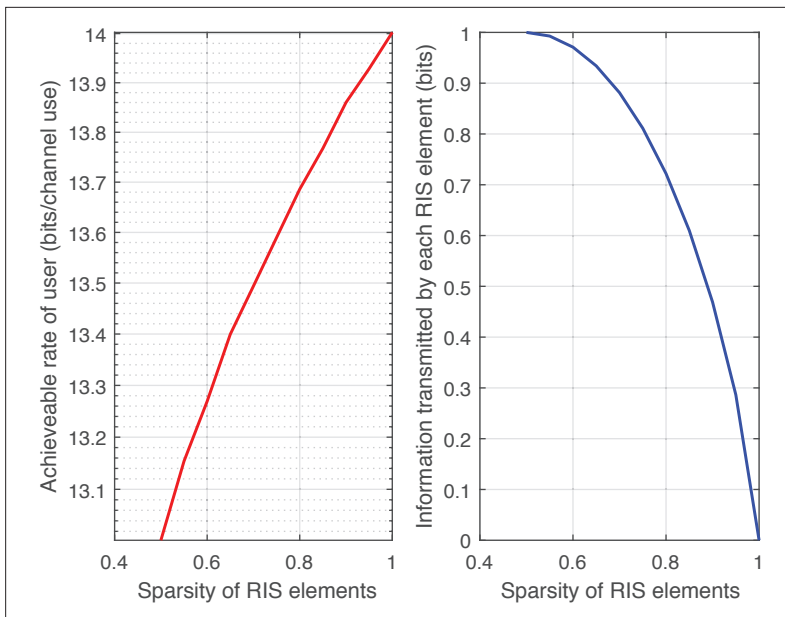
FIGURE 2. An illustration of the trade-off between the achievable rate of the primary system and the amount of information carried by each RIS element in a RIS-aided SIMO system, where the numbers of the receive antennas and the RIS elements are both set to 128. The phase shifts of the RIS elements are randomly taken as a complex number with unit modulus. All the channel coefficients are independently drawn from the standard circularly symmetric complex Gaussian distribution.

the RIS elements to transmit the RIS information in a completely passive manner. That is, the RIS information is transmitted by manipulating the on/off states of the RIS elements. Compared to [12], the PBIT scheme in [13] is more promising since it does not consume any extra time/frequency resource in information transfer.

Research Challenges

Passive information transfer of RIS, especially its joint design with passive beamforming, is an emerging research direction rich in open challenges as discussed below.

RIS Design: In the PBIT scheme, the RIS is required to enhance the primary communication and simultaneously deliver its private information. Then, how to balance these two functionalities becomes an essential problem for the design of RIS. A straightforward approach is to divide all the RIS elements into two groups, one for performing passive beamforming and the other for transferring information. One disadvantage of this approach is that the RIS elements used to transfer information do not contribute to enhancing the primary communication. To address this issue, the authors in [13] proposed to enable simultaneous passive beamforming and information transfer at each RIS element, where spatial modulation is applied to each RIS element for information delivery. The spatial modulation method randomly turns off a portion of RIS elements for information delivery, which generally compromises the capabilities of enhancing the primary communication at the RIS. A simple illustration of the trade-off between the achievable rate of the primary system and the amount of information carried by each RIS element in a RIS-aided single input multiple output (SIMO) system under the PBIT scheme [13] is given in Fig. 2. We see that,

with the sparsity of the RIS elements (i.e., the probability of each RIS element being turned on) increases from 0.5 to 1, the achievable rate of the primary system increases from about 13 bits to 14 bits, while the information carried by each RIS element deceases from 1 bit to 0. A deeper understanding of the trade-off between the passive information transfer capability and the passive beamforming gain calls for urgent investigation.

The passive beamforming design in the PBIT scheme generally involves stochastic optimization since the carried information introduces randomness on the RIS reflecting coefficients. Solving stochastic optimization problems is far more difficult than solving deterministic optimization problems involved in traditional beamforming design. In addition, various design criteria in terms of spectrum and power efficiency shall be considered in the problem formulations.

Joint Transceiver and RIS Design: We start with the transmitter side. In a RIS-assisted massive MIMO system, the active beamforming design at the transmitter needs to be optimized together with the passive beamforming at the RIS to achieve globally optimal system performance. In the PBIT scheme, joint active and passive beamforming design is particularly challenging due to the need for stochastic optimization caused by the randomness of RIS information.

The receiver of the PBIT scheme is required to retrieve the information from both the RIS and the transmitters. The signals from the transmitters and the RIS are multiplied together, resulting in bilinear signal detection problems. This inspires the development of new statistical inference techniques for bilinear models.

From an information theoretic perspective, the PBIT system can be modelled by a multiplicative multiple access channel. The capacity of such a channel is not well understood so far. The joint design of the channel coding and beamforming strategies at the transmitter and the RIS, together with the detection and decoding algorithms at the receiver, poses a highly challenging task worthy of further investigation.

Low-Complexity Robust System Design

The unprecedented capacity gain brought by RISs comes at a high computational cost. As discussed earlier, to significantly reduce the training overhead, Bayesian inference methods are needed to directly factorize the cascaded channel matrices, yielding much higher computational complexity than the traditional channel estimation methods (e.g., in [8]). Likewise, the joint optimization of RIS phase shifts and transceiver design leads to large-scale non-convex optimization problems, which are difficult to solve especially when the size of the RIS grows. This section discusses the trade-off between the system performance and the computational cost. In particular, we conjecture that due to the large number of RIS elements, the system performance is often robust against the imperfect phase tuning of individual elements.

Robustness Against Channel Estimation Errors

Figure 3 illustrates the robustness of the achievable data rate against the channel estimation error.[2] This figure shows that the loss of data rate is negligible when the normalized mean square

---

[2] For more simulation details, please refer to the demo code in https://github.com/liuhang1994/RIS_Rate_VS_Channel_Error.

error (MSE) is as large as –10 dB. Intuitively, this is because a large number of RIS elements offsets the need of accurate phase calibration. The relatively large error tolerance would allow us to put the channel estimation algorithm to an early stop when the MSE drops to an acceptable level. To see this, Fig. 4 depicts the convergence of the message passing algorithm for Bayesian matrix factorization in [10]. The figure shows that the algorithm achieves –15dB normalized MSE very quickly within the first 100 iterations. However, to further increase the accuracy to –20 dB, more than 400 iterations are needed. Considering the performance robustness in Fig. 3, we can safely stop the algorithm at the 100th iteration without waiting for the algorithm to fully converge.

The above observation naturally leads to the question: *How much computational cost shall we spend on cascaded channel estimation?* To answer this question, it is critical to seek the following fundamental understandings:

- How to characterize the relationship between the performance metrics (such as achievable data rate and outage probability) and CSI accuracy.
- How to analytically track the MSE evolution in the iterative Bayesian inference algorithms. For example, the authors in [10] borrowed the replica method from statistical physics to analyze the MSE evolution in the matrix factorization algorithm therein.

Moreover, the robustness of RIS-assisted systems against CSI error motivates new designs in LoS-dominate channels. As a LoS-dominate channel usually contains a dominant LoS component and multiple trivial non-LoS components, we only need to estimate the LoS path using its corresponding low-complexity estimation algorithms (e.g., by exploiting channel structural information such as low-rankness of the LoS component and/or user location information) to approach the performance with accurate CSI. In this case, the resultant data rate is robust against the information loss on the non-LoS channel components.

ROBUSTNESS AGAINST LOW-RESOLUTION PHASE SHIFTS

In practice, the RIS phase shifts can only take discrete values due to the finite quantization levels of the hardware, rendering the RIS optimization problem an intractable mixed-integer non-convex optimization problem. To reduce the hardware cost and control signaling overhead, recent studies have investigated the robustness of system performance with respect to low-resolution quantization. It has been shown that the capacity degradation is below 1 bit/s/Hz when the number of quantization bits is reduced from infinity to 2. Indeed, analytically characterizing the impact of low-resolution quantization on the overall system performance will be a challenging yet important research topic.

Low-resolution quantization allows the design of low-complexity optimization algorithms due to the limited number of candidate solutions. In addition to conventional integer programming algorithms, low-resolution quantization opens the possibility of designing reinforcement-learning based algorithms with the action space greatly reduced. For example, the authors in [14] recently proposed an actor-critic based mixed integer programming solver, where the actor network adopts a deep neural
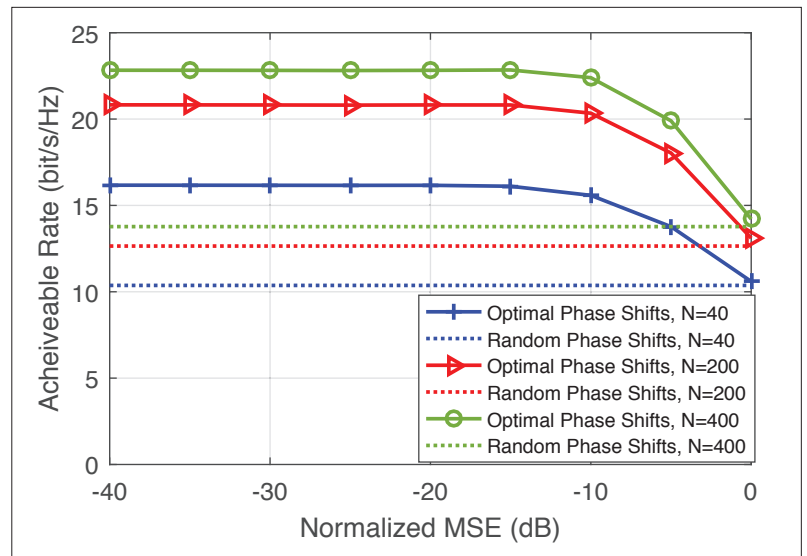

FIGURE 3. Achievable rate versus channel estimation error in a RIS assisted system. A BS with M =40 antennas communicates with a single-antenna user with the help of a RIS, where the number of RIS elements N ranges from 40 to 400. Solid lines represent the cases of optimal phase shifts, where the RIS phase shifts are optimized with respect to estimated channel information. Dotted lines represent the cases of random phase shifts, where the RIS phase shifts are randomly selected in $[0, 2\pi]^N$.
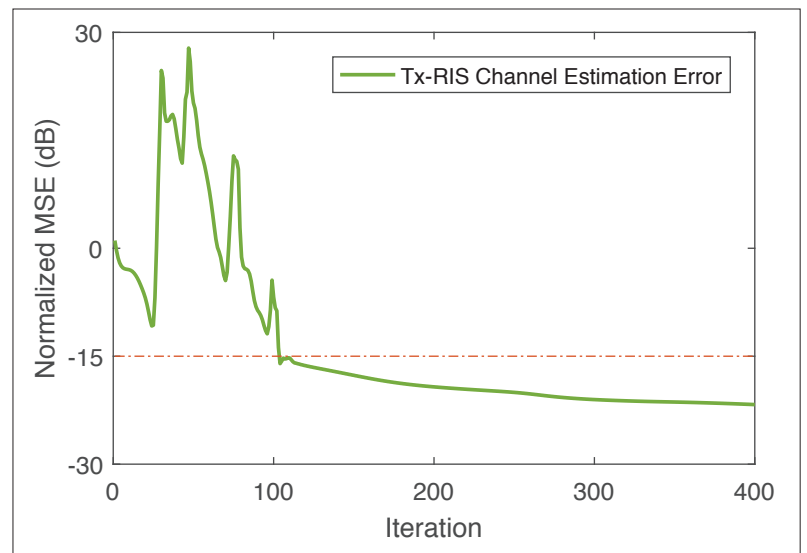

FIGURE 4. MSE vs. message passing iterations in the matrix calibration based cascaded channel estimation algorithm [10]. We simulate on a multiuser system with a 60-antenna base station, 20 single-antenna users, and a 4×4 RIS. The length of training pilots is set to 50. More detailed simulation settings can be found in [10, Fig. 8].}

network to learn the integer variables and the critic network uses mathematical optimization to tackle the continuous variables. The solver is shown to be very effective when the action space is small.

OTHER CHALLENGES OF RIS-AIDED WIRELESS COMMUNICATIONS

RIS-AIDED EDGE INTELLIGENCE

Edge intelligence, including edge caching, edge computing, and edge learning, is an advanced technology to relieve the data traffic of networks by utilizing the storage units at edge servers, to

It is a pressing challenge to develop new optimization techniques for solving these problems. Furthermore, since the delay and interference in the system may yield outdated cascaded channel acquisition, it becomes critical to characterize the performance of the RIS-aided physical layer security under outdated channel state information.

solve the computation latency problem of highly demanding devices via computation offloading, and to guarantee the privacy and security of big data analysis by adding computing and learning functionalities into radio access networks. However, the utilization of edge intelligence heavily relies on the inhabited network topology and the limited energy budget of edge devices. To design communication-efficient content delivery strategies for an edge caching network and data shuffling strategies for an edge computing system, the RIS provides a promising solution to improve the achievable degree-of-freedoms via coping with rank-deficient channels and mitigating the co-channel interference. This is achieved by improving the feasibility of the interference alignment conditions via actively controlling the network environments. The RIS can also boost the received signal power for over-the-air computation in the user scheduling procedure, thereby enabling low-latency global model aggregation in edge learning. This is achieved by intelligently tuning the phases of the incident EM waves and exploiting the waveform superposition property of a wireless multiple-access channel to adapt to the local model updates, thereby improving the scheduling policy for fast edge learning.

The typical deployments of the federated edge learning include single-server system, hierarchical system, and decentralized system, for which the RIS provides a promising way to minimize the training delay considering the heterogeneity in terms of data distributions, model sizes, communication and computation resources. This is achieved by the joint design of distributed learning and IRS-aided wireless networking techniques.

RIS-Aided Physical-Layer Security

The security of wireless networks is of crucial importance since wireless networks have been used increasingly for a wide variety of security-sensitive applications, including but not limited to banking, social networking, and environmental surveillance. Recently, there has been a rising interest in the development of secure data transmission based on physical properties of the wireless channel (hence the name physical-layer security). The use of RISs provides a radically new mechanism in which the propagation environment around the insecure nodes is manipulated to prevent potential information leakage and ensure network security. This is achieved by smartly transforming or recycling the existing signals via programming the wireless channel propagations to enhance the signals for the legitimate users and cancel out the signals to eavesdroppers. In particular, in emerging network architectures such as ad hoc networks and sensor networks, the use of conventional data encryption for network security is difficult to implement. As a result, there has been a rising interest in the development of secure data transmission based on physical properties of the wireless channel. The usage of RISs provides new mechanisms to manipulate the signals at the insecure nodes, so as to prevent potential information leakage and ensure network security.

Mathematically, the integration of RISs with physical-layer security brings about new communication models, and hence new optimization

problems. These problems are generally non-convex and difficult to solve. As such, it is a pressing challenge to develop new optimization techniques for solving these problems. Furthermore, since the delay and interference in the system may yield outdated cascaded channel acquisition, it becomes critical to characterize the performance of the RIS-aided physical layer security under outdated channel state information [15].

Conclusions

This article discussed the fundamental physical-layer issues related to the deployment of RISs in practical wireless networks, including CSI acquisition, passive information transfer, and low-complexity robust system design. For each of these issues, we explained the main challenges, discussed the state-of-the-art solutions, and pointed out the open research directions. Other RIS design problems, such as edge intelligence and physical-layer security, were briefly introduced. In conclusion, this article serves as a humble attempt to provide useful guidance and insightful inspiration to future research endeavors on RIS-aided wireless communication systems. We believe that the RIS technique, due to its superb advantages in enhancing wireless communications, will play an indispensable role in the 5G-and-beyond era.

References

[1] K. B. Letaief *et al.*, "The Roadmap to 6G: AI Empowered Wireless Networks," *IEEE Commun. Mag.*, vol. 57, no. 8, Aug. 2019, pp. 84–90.
[2] Z.-Q. He and X. Yuan, "Cascaded Channel Estimation for Large Intelligent Metasurface Assisted Massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, Feb. 2019, pp. 210–14.
[3] Q. Wu and R. Zhang, "Towards Smart and Reconfigurable Environment: Intelligent Reflecting Surface Aided Wireless Network," *IEEE Commun. Mag.*, vol. 58, no. 1, Jan. 2020, pp. 106–12.
[4] W. Tang *et al.*, "Wireless Communications with Programmable Metasurface: New Paradigms, Opportunities, and Challenges on Transceiver Design," *IEEE Wireless Commun.*, vol. 27, no. 2, Apr 2020, pp. 180–87.
[5] M. Di Renzo *et al.*, "Smart Radio Environments Empowered by AI Reconfigurable Meta-Surfaces: An Idea Whose Time Has Come," *EURASIP J. Wirel. Commun. Netw.*, vol. 2019, no. 1, May 2019, pp. 1–20.
[6] C. Huang *et al.*, "Reconfigurable Intelligent Surfaces for Energy Efficiency in Wireless Communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, Aug. 2019, pp. 4157–70.
[7] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling Large Intelligent Surfaces with Compressive Sensing and Deep Learning," Apr. 2019; available: https://arxiv.org/abs/1904.10136.
[8] D. Mishra and H. Johansson, "Channel Estimation and Low-Complexity Beamforming Design for Passive Intelligent Surface Assisted MISO Wireless Energy Transfer," *Proc. IEEE Int'l. Conf. Acoust. Speech Signal Process, (ICASSP)*, Brighton, UK, 2019, pp. 4659–63.
[9] Z. Wang, L. Liu, and S. Cui, "Channel Estimation for Intelligent Reflecting Surface Assisted Multiuser Communications: Framework, Algorithms, and Analysis," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, Oct. 2020, pp. 6607–20.
[10] H. Liu, X. Yuan, and Y.-J. A. Zhang, "Matrix-Calibration-Based Cascaded Channel Estimation for Reconfigurable Intelligent Surface Assisted Multiuser MIMO," *IEEE JSAC*, vol. 38, no. 11, Nov. 2020, pp. 2621–36.
[11] V. Liu *et al.*, "Ambient Backscatter: Wireless Communication out of Thin Air," *ACM SIGCOMM*, vol. 43, no. 4. ACM, Aug. 2013, pp. 39–50.

[12] E. Basar, "Transmission through Large Intelligent Surfaces: A New Frontier in Wireless Communications," *Proc. Eur. Conf. Netw. Commu. (EuCNC)*, Valencia, Spain, 2019, pp. 112–17.

[13] W. Yan, X. Yuan, and X. Kuai, "Passive Beamforming and Information Transfer via Large Intelligent Surface," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, Apr. 2020, pp. 533–37.

[14] L. Huang *et al.*, "Deep Reinforcement Learning for Online Computation Offloading in Wireless Powered Mobile-Edge Computing Networks," *IEEE Trans. Mobile Comput.*, vol. 19, no. 11, Nov. 2020, pp. 2581–93.

[15] K. N. Le, "Secrecy and End-to-End Analyses Employing Opportunistic Relays Under Outdated Channel State Information and Dual Correlated Rayleigh Fading," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, Nov. 2018, pp. 10 504–18.

BIOGRAPHIES

XIAOJUN YUAN [S'04, M'09, SM'15] (xjyuan@uestc.edu.cn) received his Ph.D. degree in electrical engineering from the City University of Hong Kong. He is now a professor with the University of Electronic Science and Technology of China.

YING-JUN ANGELA ZHANG [S'00, M'05, SM'10, F'20] (yjzhang@ie.cuhk.edu.hk) received her Ph.D. degree from The Hong Kong University of Science and Technology. She is now an associate professor with the Department of Information Engineering, The Chinese University of Hong Kong.

YUANMING SHI [S'13, M'15, SM'20] (shiym@shanghaitech.edu.cn) received his B.S. degree from Tsinghua University and the Ph.D. degree from The Hong Kong University of Science and Technology. He is currently a tenured associate professor at the School of Information Science and Technology, ShanghaiTech University.

WENJING YAN [S'20] (wjyan@std.uestc.edu.cn) received the B.S. degree from Chongqing University in 2018. She is currently working toward the M.S. degree at the University of Electronic Science and Technology of China.

HANG LIU [S'19] (lh117@ie.cuhk.edu.hk) received the B.Sc. degree in mathematics and information engineering from The Chinese University of Hong Kong, Hong Kong, in 2017, where he is currently pursuing the Ph.D. degree.