Towards Smart and Reconfigurable Environment: Intelligent Reflecting Surface Aided Wireless Network

Qingqing Wu and Rui Zhang

ABSTRACT

IRS is a new and revolutionizing technology that is able to significantly improve the performance of wireless communication networks, by smartly reconfiguring the wireless propagation environment with the use of massive low-cost passive reflecting elements integrated on a planar surface. Specifically, different elements of an IRS can independently reflect the incident signal by controlling its amplitude and/or phase and thereby collaboratively achieve fine-grained 3D passive beamforming for directional signal enhancement or nulling. In this article, we first provide an overview of the IRS technology, including its main applications in wireless communication, competitive advantages over existing technologies, hardware architecture as well as the corresponding new signal model. We then address the key challenges in designing and implementing the new IRS-aided hybrid (with both active and passive components) wireless network, as compared to the traditional network comprising active components only. Finally, numerical results are provided to show the great performance enhancement with the use of IRS in typical wireless networks.

INTRODUCTION

The targeted 1000-fold network capacity increase and ubiquitous wireless connectivity for at least 100 billion devices by the forthcoming fifth-generation (5G) wireless network have been largely achieved, thanks to the various key enabling technologies such as ultra-dense network (UDN), massive multiple-input multiple-output (MIMO), millimeter wave (mmWave) communication, and so on [1]. However, the required high complexity and hardware cost as well as increased energy consumption are still crucial issues that remain unsolved. For instance, densely deploying base stations (BSs) or access points (APs) in a UDN not only entails increased hardware expenditure and maintenance cost, but also aggravates the network interference issue. In addition, extending massive MIMO from sub-6 GHz to mmWave frequency bands generally requires more complex signal processing as well as more costly and energy consuming hardware (e.g., radio frequency (RF) chains). Therefore, research on finding innovative, spectral and energy efficient, and yet cost-effective solutions for future/beyond-5G wireless networks is still imperative [2].

Recently, intelligent reflecting surface (IRS) has been proposed as a promising new technology for reconfiguring the wireless propagation environment via software-controlled reflection [3–6]. Specifically, IRS is a planar surface comprising a large number of low-cost passive reflecting elements, each being able to induce an amplitude and/or phase change to the incident signal independently, thereby collaboratively achieving fine-grained three-dimensional (3D) reflect beamforming. In a sharp contrast to the existing wireless link adaptation techniques at the transmitter/receiver, IRS proactively modifies the wireless channel between them via highly controllable and intelligent signal reflection. This thus provides a new degree of freedom (DoF) to further enhance the wireless communication performance and paves the way to realize a smart and programmable wireless environment. Since IRS eliminates the use of transmit RF chains and operates only in short range, it can be densely deployed with scalable cost and low energy consumption, yet without the need of sophisticated interference management among passive IRSs. Furthermore, IRSs can be fabricated to conform to an arbitrarily shaped surfaces to cater to different application scenarios, while the underlying communication modeling and problem need further investigation.

Figure 1 illustrates several typical applications of IRS-aided wireless networks. The first application considers a user located in a dead zone where the direct link between it and its serving BS is severely blocked by an obstacle. In this case, deploying an IRS that has clear links with the BS and user helps bypass the obstacle via intelligent signal reflection, thereby creating a virtual line-of-sight (LoS) link between them. This is particularly useful for the coverage extension in mmWave communications that are highly vulnerable to indoor blockage. The second application shows the use of IRS for improving the physical layer security. When the link distance from the BS to the eavesdropper is smaller than that to the legitimate user (e.g., user 1), or the eavesdropper lies in the same direction as the legitimate user (e.g., user 2), the achievable secrecy communi-
cation rates are highly limited (even by employing transmit beamforming at the BS in the latter case). However, if an IRS is deployed in the vicinity of the eavesdropper, the reflected signal by IRS can be tuned to cancel out the (non-IRS-reflected) signal from the BS at the eavesdropper, thus effectively reducing the information leakage. In the third application, for a cell-edge user that suffers both high signal attenuation from its serving BS and severe co-channel interference from a neighboring BS, an IRS can be deployed at the cell edge to help not only improve the desired signal power but also suppress the interference by properly designing its reflect beamforming, thus creating a “signal hotspot” as well as “interference-free zone” in its vicinity. The fourth application illustrates the use of IRS for enabling massive device-to-device (D2D) communications where the IRS acts as a signal reflection hub to support simultaneous low-power D2D transmissions via interference mitigation. The last application shows the use of IRS for realizing simultaneous wireless information and power transfer (SWIPT) to miscellaneous devices in an Internet-of-Things (IoT) network [7], where the large aperture of IRS is leveraged to compensate the significant power loss over long distance via passive beamforming to nearby IoT devices to improve the efficiency of wireless power transfer to them.

Next, we highlight the main differences as well as competitive advantages of IRS as compared to other related technologies, namely, active relay, backscatter communication, and active surface based massive MIMO [8]. First, compared to active wireless relay that assists in source-destination communication by signal regeneration and retransmission, IRS does not use any active transmit module (e.g., power amplifier) but only reflects the received signal as a passive array. Besides, active relay usually operates in half-duplex mode and is thus less spectrum efficient than IRS, which operates in full-duplex mode. Although full-duplex relay is also implementable, it requires effective self-interference cancellation techniques that are costly to implement. Second, different from the traditional backscatter communication such as the radio frequency identification (RFID) tag that communicates with the reader by modulating its reflected signal sent from the reader, IRS is mainly used to facilitate the existing communication link instead of sending any information of its own. As such, the reader in backscatter communication needs to implement self-interference cancellation at its receiver to decode the tag’s message [9]. By contrast, in IRS-aided communication, both the direct-path and reflect-path signals may carry the same useful information and thus can be coherently added at the receiver to improve the signal strength for decoding. Third, IRS is also different from the active surface based massive MIMO [8] due to their different array architectures (passive versus active) and operating mechanisms (reflect versus transmit).

Despite its many benefits, the IRS-aided wireless network constitutes both active (BS, AP, user terminal) and passive (IRS) components, thus differing significantly from the traditional network comprising active components only. This thus motivates this article to provide an overview of IRS, including its signal model, hardware architecture, passive beamforming design, channel acquisition, node deployment, and so on. In particular, the main challenges and their potential solutions for designing and implementing IRS-aided wireless networks are highlighted to inspire future research. Numerical results are also provided to validate the effectiveness of IRS in some typical wireless applications.

Figure 1. Typical IRS applications in wireless network.

IRS proactively modifies the wireless channel between them via highly controllable and intelligent signal reflection. This thus provides a new degree of freedom (DoF) to further enhance the wireless communication performance and paves the way to realizing a smart and programmable wireless environment.
In practice, field-programmable gate array (FPGA) can be implemented as the controller, which also acts as a gateway to communicate and coordinate with other network components (e.g., BSs, APs, and user terminals) through separate wireless links for low-rate information exchange with them.

Figure 2. Architecture of IRS.

**Signal Model and Hardware Architecture**

In this section, we first provide the general signal model for IRS’s reflection, and then discuss the IRS’s hardware implementation and resultant constraints on the design of IRS reflection coefficients in practice.

**Signal Model**

As shown in Fig. 1, the composite channel from the BS to the user through each element of the IRS is a concatenation of three components, namely, the BS-IRS link, IRS’s reflection, and the IRS-user link. Such a composite channel behaves different from the conventional point-to-point direct channel. Specifically, each element of the IRS receives the superposed multi-path signals from the transmitter, and then scatters the combined signal with adjustable amplitude and/or phase as if from a single point source, thus leading to a “multiplicative” channel model.

Mathematically, the reflected signal by the nth element of the IRS, denoted by $y_n$, is given by multiplying the corresponding incident signal, denoted by $x_n$, with a complex reflection coefficient, that is, $y_n = (\beta_n e^{j\theta_n})x_n$, $n = 1, \ldots, N$, where $\beta_n \in [0, 1]$ and $\theta_n \in [0, 2\pi]$ specify the reflection coefficient and control the reflected signal’s amplitude (or attenuation due to passive reflection) and phase shift, respectively, and $N$ denotes the total number of reflecting elements at the IRS. By smartly adjusting the reflection coefficients, the IRS can spatially control the reflected signal to achieve different purposes. For example, to maximize the received power of the user in a dead zone in Fig. 1, all elements of the IRS should set their reflection amplitude to the maximum value of one for maximum signal reflection; whereas to achieve signal/interference cancelation in Fig. 1, the reflection amplitude of IRS elements may not necessarily be equal to the maximum value, and can be set different over the elements. In practice, other factors such as elements’ mutual coupling, noise and hardware imperfections also need to be considered in the modeling, and their impact on the performance of IRS is still an ongoing research topic.

**Hardware Architecture**

The hardware implementation of IRS is based on the concept of “metasurface”, which is made of two-dimensional (2D) metamaterial that is digitally controllable [10]. Specifically, the metasurface is a planar array consisting of a large number of so-called meta-atoms with electrical thickness in the order of the subwavelength of the operating frequency of interest [11]. By properly designing the elements, including geometrical shape (e.g., square or split-ring), size/dimension, orientation, arrangement, and so on, their individual signal response (reflection amplitude and phase shift) can be modified accordingly. In wireless communication applications, the reflection coefficient of each element should be tunable to cater to dynamic wireless channels arising from the user mobility, thus requiring reconfigurability in real time. This can be achieved by leveraging electronic devices such as positive-intrinsic-negative (PIN) diodes, field-effect transistors (FETs), or microelectromechanical system (MEMS) switches.

As shown in Fig. 2, a typical architecture of IRS consists of three layers and a smart controller. In the outer layer, a large number of metallic patches (elements) are printed on a dielectric substrate to directly interact with incident signals. Behind this layer, a copper plate is used to avoid the signal energy leakage. Last, the inner layer is a control circuit board that is responsible for adjusting the reflection amplitude/phase shift of each element, triggered by a smart controller attached to the IRS. In practice, field-programmable gate array (FPGA) can be implemented as the controller, which also acts as a gateway to communicate and coordinate with other network components (e.g., BSs, APs, and user terminals) through separate wireless links for low-rate information exchange with them.

One example of an individual element’s structure is also shown in Fig. 2, where a PIN diode is
embedded in each element. By controlling its biasing voltage via a direct-current (DC) feeding line, the PIN diode can be switched between “On” and “Off” states as shown in the equivalent circuits, thereby generating a phase-shift difference of $\pi$ in rad [10]. As such, different phase shifts of IRS’s elements can be realized independently via setting the corresponding biasing voltages by the smart controller. On the other hand, to effectively control the reflection amplitude, variable resistor load can be applied in the element design [12]. For example, by changing the values of resistors in each element, different portions of the incident signal’s energy are dissipated, thus achieving controllable reflection amplitude in [0, 1]. In practice, it is desirable to have independent control of the amplitude and phase shift at each element, for which the above circuits need to be efficiently integrated [12].

**Discrete Amplitude and Phase-shift Model**

While continuously tuning the reflection amplitude and phase shift of each of the IRS’s elements is certainly advantageous for communication applications, it is costly to implement in practice because manufacturing such high-precision elements requires sophisticated design and expensive hardware, which may not be a scalable solution as the number of elements becomes very large. For example, to enable 16 levels of phase shift as shown in Fig. 2, four PIN diodes need to be integrated to each element. This not only makes the element design very challenging due to the limited element size, but also requires more controlling pins from the IRS controller to excite a large number of PIN diodes. As such, for practical IRSs that usually have massive elements, it is more cost-effective to implement only discrete amplitude/phase-shift levels requiring a small number of control bits for each element, for example, 1-bit for two-level (reflecting or absorbing) amplitude control, and/or two-level (0 or $\pi$) phase-shift control [10, 13].

**Main Design Challenges**

Besides the hardware, we present in this section other main challenges in designing and implementing IRS-aided wireless networks from the signal processing and communication perspective, including passive beamforming design, IRS channel acquisition, and IRS deployment.

**Passive Beamforming Design**

One challenge of designing the passive beamforming for IRS in practice lies in the discrete amplitude and phase-shift levels of each element. Instead of using exhaustive search, a practical approach is to first relax such constraints and solve the problem with continuous amplitude/phase-shift values, then quantize the obtained solutions to their nearest values in the discrete sets.

On the other hand, the passive reflect beamforming of IRS in general needs to be jointly designed with the transmit beamforming of other active components in the network such as BS/AP so as to optimize the network performance. For instance, when the BS-user direct link is severely blocked, the transmit beamforming of the BS ought to point toward the IRS to maximize its signal reflection for serving the user. In contrast, when the signal attenuation of the BS-user link is comparable to that of the IRS-reflected link, the transmit beamforming of the BS should be properly designed to strike a balance between the user’s and IRS’s directions. In the above cases, the reflection amplitude of all elements of the IRS should be set to the maximum value to achieve maximum signal reflection, while the phase shifts need to be tuned based on all channels such that the reflected signal by the IRS can be added constructively at the receiver with the signal directly from the BS.

For the general multiuser setup, an IRS-aided system benefits from not only the reflect beamforming for the desired signal but also the suppression of co-channel interference. For example, the user closer to the IRS can tolerate more interference from a neighboring BS, because the IRS’s reflect beamforming can be designed such that the interference reflected by the IRS adds destructively with that directly from the interfering BS to maximally cancel it at the user’s receiver. This in turn provides more flexibility for designing the transmit beamforming at the neighboring BS for serving other users outside the IRS’s covered region. Despite the above benefits, the active and passive beamforming designs are in general closely coupled and their joint design usually leads to complicated optimization problems that are hard to be solved optimally and efficiently. To reduce such high complexity, alternating optimization can be applied to obtain suboptimal solutions, by iteratively optimizing one of the transmit and reflect beamforming with the other being fixed, until the convergence is reached [4]. Furthermore, wireless networks generally operate in wideband channels with frequency selectivity. While active BSs can use digital processing in frequency domain such as digital beamforming or hybrid digital/analog beamforming to deal with the frequency-selective channel variation [14], it is practically difficult to implement such advanced signal processing for the passive IRS. As a result, the reflection coefficients of IRS needs to balance the channels at different frequency sub-bands, which further complicates the joint active and passive beamforming optimization.

Some interesting results have been reported in this new direction recently [3–5, 13]. Prior works [3, 4] revealed that in an IRS-aided single-user system, the received power increases asymptotically in the order of $N^2$ as the number of reflecting elements, $N$, goes to infinity. In other words, every doubling of $N$ achieves about 6 dB power gain in the large-$N$ regime. The fundamental reason behind such a “square law” of $N$ is that the IRS not only achieves a power gain of $N$ by reflect beamforming (similarly to the transmit beamforming with $N$ active antennas in massive MIMO [14]), but also captures another power gain of $N$ due to its large aperture for collecting...
Simulation setup.

The received signal energy from the BS (which is not available in massive MIMO). Moreover, compared to the ideal case with continuous phase shifts, it was shown in [13] that by using IRS with b-bit uniformly quantized phase shifts, the same asymptotic power scaling law of order $N^2$ can be achieved, while only a constant proportional power loss as a function of $b$ is incurred, which is insignificant as compared to $N^2$ for large $N$ and thus can be ignored as $N$ goes to infinity.

**IRS Channel Acquisition**

The various performance gains brought by the passive beamforming of IRS in general require the accurate knowledge of the channels between the IRS and the involved BSs and users. Note that by turning the IRS into the absorbing mode, the channel state information (CSI) of BS-user links without the IRS can be obtained by applying the conventional channel estimation methods [14]. Depending on whether receive RF chains are equipped with the elements at the IRS or not, the acquisition of CSI between the IRS and BSs/users can be classified into the following two categories.

First, although transmit RF chains are removed from the IRS for cost reduction and energy saving, each of its elements can be equipped with a low-power receive RF chain to enable the sensing capability for channel estimation. As such, the channels from the BSs/users to the IRS can be estimated at the IRS based on their training signals. To reduce the number of receive RF chains at the IRS, the sub-array technique can be applied where each sub-array consists of a cluster of neighboring elements arranged vertically and/or horizontally and each cluster is equipped with one receive RF chain for channel estimation. Accordingly, the reflection coefficients of all elements in each sub-array can be set to be either the same or different by applying proper interpolation over adjacent sub-arrays.

In contrast, when receive RF chains are not installed at the IRS, it is infeasible for the IRS to estimate the channels with BSs/users directly. However, a viable approach for this challenging case may be that, instead of estimating the IRS-BS/user channels separately, we estimate their concatenated channel with some known IRS reflection patterns (e.g., by turning on/off some of its elements). Alternatively, we can design the reflection coefficients for IRS’s passive beamforming based on the feedback from the BSs/users pertaining their received signals that are reflected by the IRS, thus without the need of explicitly estimating the IRS-BS/user channels. For example, the IRS can quickly sweep its reflect beamforming coefficients in a pre-designed codebook and the best beam is then selected based on the BS/user’s feedback. To reduce the complexity and time overhead of real-time training, historical data can be exploited. For example, for mmWave communication, due to the channel sparsity [14], the IRS-user channels are usually correlated in space and thus the IRS beamforming coefficients for users in nearby locations are similar and vary spatially like a smooth function. This can be utilized to obtain the IRS beamforming coefficients for a new user by interpolating those at its nearby locations obtained in the past.

**IRS Deployment**

How to judiciously deploy IRSs in a hybrid wireless network comprising both active BSs and passive IRSs to optimize its performance is another crucial problem to solve. Generally speaking, this problem should have different considerations as compared to that of deploying active BSs/relays in the traditional wireless network. As IRSs are deployed for local coverage only, their operating ranges are usually much shorter than those of active BSs/relays, which makes it easier to practically deploy IRSs without interfering each other. In the following, we provide more detailed discussion on this issue.

First, from the viewpoint of optimizing the performance in a single-cell setup, the IRS should be intuitively deployed at a location with clear LoS from the BS in order to maximize its received signal power for passive beamforming. However, when the IRS needs to support simultaneous transmissions between the BS and multiple users in its coverage region, such a straightforward deployment strategy may not work well. This is because one single LoS path between the IRS and the BS results in a low-rank MIMO channel that cannot support spatial multiplexing for the transmission to multiple users via the IRS [4]. Therefore, the deployed location for the IRS is practically preferable to possess a strong LoS path with the BS as well as a sufficiently large number of non-LoS paths for enabling a high-rank MIMO channel, so as to resolve the above trade-off. Besides, the deployment of IRS should also take into consider-
ation the spatial user density, i.e., with a high priority to be deployed in hot-spot areas with a large number of users, as well as the inter-cell interference issue, e.g., when there is an urgent need to deploy an IRS near the boundary of two adjacent cells to help cancel the co-channel interference between them, as shown in Fig. 1.

In practice, the propagation environment may be complicated and each IRS can be associated with multiple BSs. In such scenarios, using good heuristics alone for deploying IRS may be ineffective, while an exhaustive search for the optimal location requires the global CSI at all locations, which is practically difficult to obtain. Ray-tracing based methods can be used to estimate such CSI, but they are computationally costly and also require site-specific information (such as building/floor layout for indoor communication). As such, how to achieve autonomous deployment of IRSs by identifying the most suitable locations for them is a new problem of high practical interest. One promising approach to solve this problem is by leveraging machine learning techniques, such as deep learning (DL). For example, in the training phase, we can deploy IRSs at some properly selected reference locations and collect key performance indicators such as received signal strength measured at different user locations. Such IRS locations and corresponding performance indicators are then used to train a DL-based neural network as the input and output, respectively. Next, in the deployment phase, with the desired performance indicators as the input, the trained DL network is used to predict a set of locations for deploying IRSs. After deploying IRSs at these locations, a new set of performance indicators can be collected and used to further train the DL network to improve its prediction accuracy in the future.

**Numerical Results**

We consider a BS with M antennas, an IRS with N elements, and one single-antenna user, with their locations shown in Fig. 3. Denote the horizontal distance between the BS and user by \( d \) meter (m). It is assumed that the BS-IRS channel is dominated by the LoS link with the path loss exponent of 2.2, whereas both the BS-user and IRS-user channels are assumed to follow Rayleigh fading with path loss exponent of 3.2. The receiver noise power is \(-80\) dBm.

**Signal Power Enhancement and Scaling Law**

To demonstrate the signal power enhancement capability of IRS, we assume that the user in Fig. 3 needs to be served by the BS with the IRS’s help, similar to the first scenario in Fig. 1. We compare the following four schemes under the setup of \( M = 5 \) and \( N = 40 \):

- Joint optimization as in [4]
- BS-user maximum-ratio transmission (MRT) where the BS beams toward the BS-user channel
- BS-IRS MRT where the BS beams toward the BS-IRS rank-one channel
- Benchmark scheme without the IRS where the BS beams toward the BS-user channel.

As shown in Fig. 4, by varying the value of \( d \), we examine the minimum transmit power required at the BS for achieving a target user signal-to-noise ratio (SNR) of 20 dB. First, it is observed that for the scheme without IRS, moving the user farther away from the BS leads to higher transmit power due to the increased signal attenuation. However, this issue is alleviated by deploying the IRS, which helps significantly improve the SNR when the user is near to it. As a result, the user near either the BS (e.g., \( d = 25 \) m) or IRS (e.g., \( d = 50 \) m) requires lower transmit power than a user far away from both of them (e.g., \( d = 40 \) m). This demonstrates the practical usefulness of IRS in creating a “signal hotspot” in its vicinity. Furthermore, compared to other heuristic BS transmit beamforming schemes, the joint active and passive beamforming design achieves substantial power saving at the BS.

In Fig. 5, we show the performance of the IRS where each element reflects with unit amplitude, but using a practical \( b \)-bit uniformly quantized phase shifter. The BS transmit power is plotted versus \( N \) when \( d = 50 \) m. First, it is interesting to observe that for the ideal continuous phase, the BS transmit power scales down with \( N \) approximately in the order of \( N^3 \). For example, for the
foresee that the integration of IRSs into future wireless networks will fundamentally change their architecture from the traditional one with active components solely to a new hybrid one with both active and passive components co-working in an intelligent way, thus opening fertile directions for future research.

INTERFERENCE SUPPRESSION

Next, we demonstrate the interference suppression capability of IRS, by considering now the BS in Fig. 3 is a neighboring transmitter that causes co-channel interference to the user when $d = 50$ m, and the IRS is deployed to help suppress its received interference from this BS, like the third scenario in Fig. 1. This setup also resembles the physical layer security scenario in Fig. 6, where the user is an eavesdropper and its received signal from the legitimate transmitter (BS) needs to be canceled with the help of the IRS. For simplicity, we assume $M = 1$ and the transmit power of the BS is $30$ dBm. For comparison, we plot the interference power at the user versus $N$ for three schemes as shown in Fig. 6. It is first observed that as compared to the scheme without IRS, the interference power is substantially reduced even by adjusting the phase shifts of the IRS’s elements solely. Moreover, with jointly optimized amplitude and phase shifts, it is observed that the co-channel interference can be more effectively canceled as compared to the case with fixed amplitude, especially when $N$ is sufficiently large. This is because with the additional amplitude control, the IRS is able to impose an opposite interference signal at the user to perfectly cancel that from the BS-user link, thus creating a virtually “interference-free zone”.

CONCLUSIONS

In this article, we provide an overview of the promising IRS technology for achieving a smart and reconfigurable environment in future wireless networks. Notably, the IRS can sense the wireless environment and accordingly adjust its reflection coefficients dynamically to achieve different functions by leveraging advanced signal processing and machine learning techniques. As IRS-aided wireless networks are new and remain largely unexplored, it is hoped that this article would provide a useful and effective guide for the future research on them. In particular, we foresee that the integration of IRSs into future wireless networks will fundamentally change their architecture from the traditional one with active components solely to a new hybrid one with both active and passive components co-working in an intelligent way, thus opening fertile directions for future research.

REFERENCES


BIOGRAPHIES

QINGQING WU is a research fellow in the ECE Department of the National University of Singapore and an editor for IEEE Communications Letters.

RUI ZHANG [F’17] received the Ph.D. degree from Stanford University and is now a Dean’s Chair Associate Professor in the ECE Department of National University of Singapore. He has been listed as a Highly Cited Researcher by Thomson Reuters since 2013.