Chapter 6 Channel Estimation in RIS-Aided Networks



Fadil Danufane, Placido Mursia, and Jiang Liu

Abstract Reconfigurable intelligent surface (RIS) is a recently emerging transmission technology for application to wireless communications. Regarded to be an emerging solution for the next generation of communications, RIS is a nearly passive device that realizes smart radio environment with low hardware cost and energy consumption. This merit of RIS, on the other hand, imposes a major challenge to the channel estimation of RIS-aided communication systems. Recently, many protocols and algorithms are proposed to handle this challenging problem. In this chapter, we review the problem of channel estimation in RIS-aided systems and survey recent developments on this topic.

6.1 Introduction

The need for high data rates is ever increasing in the future. By 2030, it is forecast that the global data traffic will increase up to the order of thousands of exabytes [1]. In addition, future wireless communication systems such as 6G are expected to deliver these data in a distributed and intelligent way, as well as within some delay and reliability constraints that are more stringent than ever. These requirements cannot be satisfied by the existing technologies and even the newly deployed 5G communication system.

To meet these challenging demands and requirements, a new paradigm on how a communication system is designed is needed. Recently, a vision of smart radio environment (SRE) was proposed to challenge the status quo of communication system design and redefine the performance limit of communication systems. In

F. Danufane · J. Liu (🖂)

Université Paris-Saclay, CentraleSupelec, CNRS, Laboratoire des Signaux et Systèmes, Gif-sur-Yvette, France

e-mail: fadil.danufane@centralesupelec.fr; jiang.liu@centralesupelec.fr

P. Mursia EURECOM, Sophia-Antipolis, France e-mail: placido.mursia@eurecom.fr

[©] The Author(s), under exclusive license to Springer Nature Switzerland AG 2022 J. Rodriguez et al. (eds.), *Enabling 6G Mobile Networks*, https://doi.org/10.1007/978-3-030-74648-3_6

particular, in SRE, the environment is no longer seen as an impairment according to which a system has to be designed, but instead as a component that can be controlled to achieve a specific performance.

RIS is a recently emerging transmission technology for application to wireless communications. Conceptually speaking, RIS is a two-dimensional surface made of metamaterials that is capable of manipulating the incident electromagnetic waves in arbitrary ways. The main selling points of RIS are its near-passive nature, since it does not require a large power source to redirect the waves, and its low cost and low complexity of large-scale deployments. Thanks to these properties, RISs are receiving major attention from the wireless community and are considered to be the key technology to realize the vision of SRE.

A RIS consists of many sub-wavelength unit elements, usually called metaatoms, whose phase shift can be configured independently. By configuring the phase shift of each unit cell, one can manipulate the reflected wave in many ways, e.g., by manipulating the wave by reflecting an incoming beam in any desired direction or by focusing the reflected wave to maximize the electric intensity at a specific location. Therefore, the main property of the RIS is its capability of being reconfigurable even after its deployment in a wireless environment.

Due to the sub-wavelength structure of the RIS, the distance between adjacent unit cells and the size of each unit cell is much smaller than the wavelength. Therefore, the propagation or resonance effects in the direction perpendicular to the surface can be safely ignored in the process of synthesis and analysis of the surface. Thanks to this, a RIS can be modelled through appropriate continuous surface-averaged functions (e.g., susceptibilities), despite being made of discrete elements. This representation of the RIS as a continuous entity allows for convenient performance analysis through some concepts of physics, as demonstrated in [2].

Recently, there have been exciting research activities on the realization of lowcost and practical RIS. Two recent examples of these activities are illustrated in Figs. 6.1 and 6.2. In Fig. 6.1, the RFocus prototype, recently designed by researchers of the Massachusetts Institute of Technology (MIT), USA, is depicted [3]. The prototype is made of 3720 inexpensive antennas arranged on a 6-square meter surface. At scale, each antenna element is expected to have a cost of the order of a few cents or less. In Fig. 6.2, a prototype of smart glass, recently designed by researchers from NTT DOCOMO, Japan, is depicted [4]. The manufactured smart glass is an artificially engineered thin layer (i.e., a metasurface) that comprises a large number of sub-wavelength unit elements placed in a periodic arrangement on a two-dimensional surface covered with a glass substrate. By moving the glass substrate slightly, it is possible to dynamically control the response of the impinging radio waves in three modes: (i) full penetration of the incident radio waves, (ii) partial reflection of the incident radio waves, and (iii) full reflection of all radio waves.

Although RIS is generally capable to modify the impinging electromagnetic wave in any desired way, recently, there are two widely investigated functionalities within the literature of wireless communication technology.



Fig. 6.2 NTT DOCOMO's prototype. (Photo: NTT DOCOMO)

- 1. Anomalous reflection/transmission [5]. Under this setting, the RIS is configued to reflect or refract the impinging radios waves toward specified directions that do not necessarily adhere to the laws of reflection and refraction (i.e., the angle of incident is not necessarily equal to the angle of reflection/transmission). This setting is useful in some setups in which several users are being served simultaneously (e.g., broadcasting application) or when a single user is moving in a constant direction with respect to the RIS (e.g., vehicular application). The limitation of this setting lies on the fact that, in general, the signal-to-noise-ratio is not maximized and thus the system capacity is not achieved.
- 2. *Beamforming/focusing* [3]. Under this setting, the RIS is configured to the reflected/transmitted electromagnetic wave into a specific location such that the intensity is maximized there. Therefore, in this case, the signal-to-noise-ratio is maximized, and thus the system capacity is achieved for a single user in the designated location. The limitation of this setting lies on the potential complexity of the phase-shift reconfiguration of each unit cell of the RIS to accommodate the mobility of the user.

From an application point of view, RIS can be utilized for various use cases. Some examples include but are not limited to the following [6]:

- *Signal engineering.* The RIS provides an additional LOS path between a transmitter and a receiver to mitigate the non-existence of direct link between them.
- *Interference engineering.* The RIS is configured to minimize the signal that comes from an interfering transmitter at the intended receiver.
- Security engineering. The RIS is configured to minimize the signal containing information between a transmitter and a receiver that arrives at a malicious user.
- *Scattering engineering.* The RIS is configured to increase the channel rank between a transmitter and a receiver by means of creating a rich scattering environment (high rank channel) for high data rate transmission.

We end this section by mentioning that despite the study on RIS in the literature, they mostly consider a flat RIS such as internal walls of indoor environments, external facades of buildings, and the glasses of windows; in general, a RIS does not have to be planar. Some applications in wireless communications, for example, include coating several irregularly shaped objects in order to control the reflected/refracted radio waves that impinge it to enhance the overall communication performance. These functions cannot, in general, be realized by using a planar RIS.

6.2 The Channel Estimation Problem in RIS-Aided Networks

Channel estimation in a RIS-assisted wireless system is a much more challenging task than in conventional systems since the passive RIS elements are incapable of sensing and estimating channel information. Such design choice is undoubtedly more appealing due to its extremely low hardware and deployment cost. However, accurate channel state information (CSI) is critical in optimizing the RIS parameters.

Thus, the problem of estimating the channel in RIS-aided networks has gained much attention lately. In particular, the focus is on how to estimate the two cascaded channels between the transmitter and the RIS and between the RIS and the UE with purely passive reflecting elements and an affordable training overhead.

Consider a general multi-user multi-input single-output (MISO) network setup detailed in Fig. 6.3, where a base station (BS) equipped with M antennas communicates with K single-antenna user ends (UEs) with the aid of a RIS made of N reflecting elements. We assume that the transmission takes place over a total of T time slots in which the channel is assumed to be constant, following a quasi-static fading model. The channel between the BS and the RIS is denoted as $G \in \mathbb{C}^{N \times M}$, while $h_k \in \mathbb{C}^{N \times 1}$ denotes the channel between the RIS and UE k. Lastly, $h_{d,k} \in \mathbb{C}^{M \times 1}$ represents the direct channel between the BS and UE k. Hence, the signal received by the k-th UE in the downlink at time t is given by



Fig. 6.3 A model of a RIS-assisted multiuser MISO system

$$y_{k,t} = \left(\boldsymbol{h}_{d,k}^{H} + \boldsymbol{h}_{k}^{H}\boldsymbol{\Phi}_{t}\boldsymbol{G}\right)\mathbf{x}_{t} + n$$
(6.1)

where $\mathbf{\Phi}_t = \text{diag}\left(\beta_{1,t}e^{j\phi_{1,t}}, \dots, \beta_{N,t}e^{j\phi_{N,t}}\right) \in \mathbb{C}^{N \times N}$ is the matrix containing each RIS element absorption coefficient $\beta_{n,t} \in [0, 1]$ and shift $\phi_{n,t} \in [0, 2\pi]$ at time *t*, diag(\mathbf{x}) represents a diagonal matrix with the entries of \mathbf{x} on its main diagonal, $\mathbf{x}_t \in \mathbb{C}^{M \times 1}$ is the signal transmitted by the BS at time *t* with $\mathbb{E}[\|\mathbf{x}_t\|^2] = 1$, and $n \sim \mathcal{CN}\left(0, \sigma_n^2\right)$ is a noise coefficient. Let $\mathbf{y}_k = [y_{k,1}, \dots, y_{k,T}]^T$ be the signal collected at UE *k* after *T* pilot symbols. Similarly, the receive signal at the BS at time *t* is expressed as

$$\mathbf{y}_{t} = \sum_{k=1}^{K} \left(\mathbf{h}_{d,k} + \mathbf{G}^{H} \mathbf{\Phi}_{t} \mathbf{h}_{k} \right) \mathbf{x}_{t,k} + \mathbf{n}$$
(6.2)

where $\mathbf{x}_{t,k}$ is the signal transmitted by UE *k* at time *t* with $\mathbb{E}[|\mathbf{x}_{t,k}|^2] = 1$. Lastly, let $Y = [y_1, \ldots, y_T]$ be the receive signal at the BS after *T* training symbols. Equations (6.1) and (6.2) can be rewritten as

$$y_{k,t} = \left(\boldsymbol{h}_{d,k}^{H} + \boldsymbol{v}_{t}^{H}\overline{\boldsymbol{H}}_{k}\right)\mathbf{x}_{t} + n$$
(6.3)

$$\mathbf{y}_{k,t} = \left(\boldsymbol{h}_{d,k}^{H} + \boldsymbol{v}_{t}^{H}\overline{\boldsymbol{H}}_{k}\right)\mathbf{x}_{t} + n \tag{6.4}$$

where $\boldsymbol{v}_t = [\beta_{1,t}e^{-j\phi_{1,t}}, \dots, \beta_{N,t}e^{-j\phi_{N,t}}]$ contains the RIS configuration at time *t* and $\overline{\boldsymbol{H}}_k = \text{diag}(\boldsymbol{h}_k^H)\boldsymbol{G}$ represents the aggregated effective channel between UE *k* and the BS via the RIS.

6.3 Survey on Channel Estimation

In this section, we review the main existing solutions based on analytical optimization of the channel estimation protocol. In this respect, we identify several main categories depending on the fundamental idea behind each channel estimation procedure. For each category, we point out the main characteristics and drawbacks.

6.3.1 On/Off-Based Channel Estimation

In this section, we review a class of channel estimation protocols based on sequentially activating only one RIS element for each pilot symbol. The full channel is thus estimated in N + 1 training symbols where the first pilot symbol is necessary to estimate the direct channel between the BS and the UEs.

Works such as [7–9] are based on activating only one RIS element for each pilot symbol. In all such works, only the aggregated channels $\{\overline{H}_k\}$ and $\{h_{d,k}\}$ are estimated. Hence, for each UE *k*, the resulting aggregate channel \overline{H}_k is estimated column-wise in a total of *N* training symbols. An extra training symbol is necessary to estimate $h_{d,k}$ with all the RIS elements deactivated. All UEs transmit such pilots concurrently, and interference among them is resolved thanks to the use of orthogonal training sequences. Indeed, we have that $\mathbf{x}_k^H \mathbf{x}_j = 0$ if $k \neq j$ and $\mathbf{x}_k^H \mathbf{x}_k = 1$.

At time t = 1, the received signal at the BS is given by

$$\mathbf{y}_1 = \sum_{k=1}^{K} \mathbf{h}_{d,k} \mathbf{x}_{1,k} + \mathbf{n}$$
(6.5)

while the receive signal at a generic time instant t is given by

$$\mathbf{y}_{t} = \sum_{k=1}^{K} \left(\mathbf{h}_{d,k} + \overline{\mathbf{h}}_{k,t}^{H} v_{t} \right) \mathbf{x}_{t,k} + \mathbf{n}$$
(6.6)

In [9] the channel of each UE *k* is estimated using least squares, i.e., the receive signal *Y* is multiplied by $\overline{\mathbf{x}}_{k}^{*} = [\mathbf{x}_{1,k}^{*}, \mathbf{x}_{2,k}^{*}v_{2}^{*}, \dots, \mathbf{x}_{T,k}^{*}v_{T}^{*}]^{T}$. The first column of *Y* is used to estimate $\mathbf{h}_{d,k}$. The resulting estimate is subtracted from the signal $\mathbf{r} = \mathbf{Y}\overline{\mathbf{x}}_{k}^{*}$ in order to obtain the estimate of $\overline{\mathbf{h}}_{k}$. In [7, 8], such estimate is further refined using the minimum-mean-squared-error (MMSE) principle, i.e., by exploiting the known statistics of the channel and noise.

In practice, to implement the ON/OFF switching of the massive RIS elements is costly. Besides, as only a small portion of its elements is switched ON at each time, the channel estimation accuracy is degraded. To address this issue, [10] proposed an RIS elements-grouping method to reduce the training overhead and estimation

complexity. Instead of controlling the ON/OFF states of a single element each time, the authors applied the ON/OFF method on the grouped RIS elements.

Similarly, in [11], after the superimposed channel is obtained using the least square (LS) estimation, the grouping ON/OFF method is adopted to estimate the direct channel link and the cascaded channel link.

In [12, 13], the idea of grouping ON/OFF method is extended. With the same assumption that the RIS can be divided into multiple sub-surfaces of adjacent strongly correlated reflecting elements that apply the same reflection coefficient, [12] designed the reflection pattern based on discrete Fourier transform (DFT) or Hadamard matrix based on their orthogonality, while the authors in [13] designed the pattern based on the minimum variance unbiased estimation principle, which mimics a series of discrete Fourier transforms.

In [14], the authors propose a three-phase channel estimation protocol based on the observation that each RIS element reflects the signals from all the users to the transmitter via the same channel. The first phase is similar to Eq. (6.5); all the IRS elements are switched off to estimate the direct channel. In the second phase, all the IRS reflection elements are switched on, and merely one typical user transmits nonzero pilot symbols to the BS. In this phase, the BS estimates the cascaded channel of this typical user. The construction of the reflection coefficient matrix can be based on the DFT matrix. In the last phase, the cascaded channels of other users are estimated, where the channel correlations are exploited to reduce complexity. The authors quantified the minimum time to estimate all required channels and show that massive multi-input multi-output (MIMO) may play an important role in reducing the channel estimation overhead in RIS-based communication systems.

6.3.2 Least Squares-Based Channel Estimation

In [15], the authors propose an iterative algorithm for channel estimation that is based on the parallel factor decomposition algorithm. The proposed method is based on an alternating least squares algorithm that iteratively estimates the channel between the transmitter and the RIS **G** as well as the channel between the RIS and the users \mathbf{h}_k . Considering the low resolution of the RIS unit elements, the RIS is assumed to have *P* different phase configuration. Define the $P \times N$ complex-valued matrix $\boldsymbol{\Theta}$ as the configuration matrix; the *p*-th row of $\boldsymbol{\Theta}$ represents the *p*-th RIS phase configuration. Consequently, the end-to-end RIS-based wireless channel can be given by

$$\mathbf{Z}_{p} = \mathbf{H}_{2} \operatorname{diag} \left(\mathbf{\Theta} \left(\boldsymbol{p}, : \right) \right) \mathbf{G}$$
(6.7)

where $\mathbf{H}_2 = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K]^T \in \mathbb{C}^{K \times N}$ is the channel between the RIS and the *K* users. Each (k, m)-th entry of \mathbf{Z}_p with $k = 1, 2, \dots, K$ and $m = 1, 2, \dots, M$ is obtained as

$$\left[\mathbf{Z}_{p}\right]_{k,m} = \sum_{n=1}^{N} \left[\mathbf{H}_{2}\right]_{k,n} \left[\mathbf{G}\right]_{n,m} \left[\mathbf{\Theta}\right]_{p,n}$$
(6.8)

where $[\mathbf{G}]_{n,m}$, $[\mathbf{H}_2]_{k,n}$, and $[\mathbf{\Theta}]_{p,n}$ denote the (n, m)-th entry of \mathbf{G} , (k, n)-th entry of \mathbf{H}_2 , and (p, n)-th entry of $\mathbf{\Theta}$, respectively, with n = 1, 2, ..., N.

The proposed method is based on an alternating least squares algorithm that iteratively estimates the channel between the transmitter and the RIS **G** as well as the channel between the RIS and the users **H**₂. Using the PARAllel FACtor (PARAFAC) decomposition, \mathbf{Z}_p can be represented using three matrix forms. These matrices form the horizontal, lateral, and frontal slices of the tensor composed of Eq. (6.8). The unfolded forms of the mode-1, mode-2, and mode-3 of \mathbf{Z}_p 's are expressed as follows:

$$\mathbf{Mode} - \mathbf{1} : \mathbf{Z}_{\alpha} = (\mathbf{G}^{T^{\circ}} \mathbf{\Theta}) \mathbf{H}_{2}^{T} \in \mathbb{C}^{PM \times K}$$
$$\mathbf{Mode} - \mathbf{2} : \mathbf{Z}_{\beta} = (\mathbf{\Theta}^{\circ} \mathbf{H}_{2}) \mathbf{G} \in \mathbb{C}^{KP \times M}$$
$$\mathbf{Mode} - \mathbf{3} : \mathbf{Z}_{\gamma} = (\mathbf{H}_{2}^{\circ} \mathbf{G}^{T}) \mathbf{\Theta}^{T} \in \mathbb{C}^{MK \times P}$$
(6.9)

where [°] represents the Khatri-Rao (column wise Kronecker) matrix product. Considering AWGN, we define the following three-dimensional matrix:

$$\widetilde{\mathbf{Z}} = \mathbf{Z} + \widetilde{\mathbf{W}}$$
(6.10)

where tensor $\mathbf{\widetilde{W}} \in \mathbb{C}^{K \times M \times P}$ is the AWGN that incorporates all *P* matrices $\mathbf{\widetilde{W}}_{p}$. The proposed iterative channel estimation is expressed as follows:

- 1. First step (Initialization): Initialize with a random feasible phase matrix Θ . $\hat{\mathbf{G}}^{(0)}$ represents the eigenvector matrix corresponding to the *N* nonzero eigenvalues of $\sim_{H}^{H} \sim_{\alpha}^{\sim} \simeq_{\beta} \mathbf{Z}_{\beta}$, where \mathbf{Z}_{β} is the noisy version of Mode-1 form of Eq. (6.10). Similarly, $\hat{\mathbf{H}}_{2}^{(0)}$ is the eigenvector matrix corresponding to the N nonzero eigenvalues of $\sim_{H}^{H} \sim_{\alpha}^{\sim} \mathbf{Z}_{\alpha}$, where \mathbf{Z}_{α} is the noisy version of Mode-2 form of Eq. (6.10). Set the algorithmic iteration i = 1.
- 2. Second and third steps (Iterative Update):

$$\hat{\mathbf{H}}_{2}^{(i)} = \left(\left(\hat{\mathbf{A}}_{1}^{(i-1)} \right)^{+ \widetilde{\mathbf{Z}}'} \right)^{T}$$

$$\hat{\mathbf{A}}_{1}^{(i-1)} = \hat{\mathbf{H}}_{2}^{(i-1)^{\circ}} \boldsymbol{\Theta}.$$

$$\hat{\mathbf{G}}^{(i)} = \left(\hat{\mathbf{A}}_{2}^{(i)}\right)^{+} \widetilde{\mathbf{Z}}^{''}$$

$$\hat{\mathbf{A}}_{2}^{(i)} = \boldsymbol{\Theta}^{\circ} \hat{\mathbf{H}}_{2}^{(i)}$$

where $\widetilde{\mathbf{Z}}' \in \mathbb{C}^{PM \times K}$ is a matrix-stacked form of Eq. (6.10)'s tensor $\widetilde{\mathbf{Z}}, \widetilde{\mathbf{Z}}' \in \mathbb{C}^{KP \times M}$ is another matrix-stacked form of $\widetilde{\mathbf{Z}}$, and (•)⁺ denotes the pseudo-inverse matrix.

3. Fourth step (Iteration Stop Criterion): The proposed iterative algorithm terminates when either the maximum number I_{max} of algorithmic iterations is reached or when between any two algorithmic iterations i - 1 and i hold the following condition for ε being a very small positive real number:

$$\begin{aligned} \left\| \hat{\mathbf{G}}^{(i)} - \hat{\mathbf{G}}^{(i-1)} \right\|_{F}^{2} / \left\| \hat{\mathbf{G}}^{(i)} \right\|_{F}^{2} &\leq \varepsilon \\ \text{or} \\ \frac{\left\| \hat{\mathbf{H}}_{2}^{(i)} - \hat{\mathbf{H}}_{2}^{(i-1)} \right\|_{F}^{2}}{\left\| \left\| \hat{\mathbf{H}}_{2}^{(i)} \right\|_{F}^{2}} &\leq \varepsilon \end{aligned}$$
(6.11)

Thus, the channels G and H_2 are obtained using this alternate LS iteration.

6.3.3 Sparsity-Based Channel Estimation

This section deals with a class of channel estimation methods that rely on the assumption of channel sparsity. Indeed, often, the BS and the RIS are mounted on top of buildings and are in LoS with each other such that the channel **G** can be regarded as being close to rank-one, i.e., dominated by the LoS path. A similar consideration holds for each channel \mathbf{h}_k especially if the latter is a mmWave or TeraHertz channel. However, even in this case, the multipath component typically carries a lower but still noticeable amount of power compared to the LoS path. Leveraging on the sparsity of **G** and the aggregated channels $\overline{\mathbf{H}}_k$, several recent works have proposed to use compressed sensing (CS) [16], beam training (BT) [17, 18], sparse matrix factorization (SMF) [19], matrix calibration [20], or orthogonal matching pursuit (OMP) [21, 22] in order to estimate the channels and reduce the training overhead compared to on/off techniques, as described in the previous section.

6.3.3.1 Compressed Sensing

The work in [16] proposes to exploit the inherent sparsity of the effective channels $\overline{\mathbf{H}}_k$ which is due to the low-scattering link connecting the BS and the RIS via CS. The training period is divided into *BT* symbols. For each one of the *B* blocks, the UEs send mutually orthogonal sequences of length *T* with $T \ge K$. Each UE repeats the same pilot sequence for all the *B* blocks. The RIS is configured following a series of mutually orthogonal sequences which are repeated for the *T* symbols of each block. As a result, the matrix **V** is unitary across the different blocks. Hence, this algorithm is designed to obtain diversity in the received signal across both pilot sequences and RIS configurations.

Assuming that the direct link between the BS and the UEs can be neglected due to low associated power, at each block *b*, the receive signal at the BS is defined as

$$\mathbf{Y}_{b} = \sum_{k=1}^{K} \overline{\mathbf{H}}_{k}^{H} \mathbf{v}_{b} \mathbf{x}_{k}^{H} + \mathbf{n}_{b} \in \mathbb{C}^{M \times T}$$
(6.12)

As a first estimate of the effective channels, the authors propose to use the least squares signal, i.e., $\mathbf{r}_{b,k} = \mathbf{Y}_b \mathbf{x}_k$, for each block *b* and UE *k*. Such initial estimate is then further refined by exploiting its sparsity structure. In particular, the effective channels are modelled using a virtual channel representation as

$$\overline{\mathbf{H}}_k = \mathbf{A}_R \mathbf{X}_k \mathbf{A}_B^H \tag{6.13}$$

where $\mathbf{A}_R \in \mathbb{C}^{N \times N'}$ with N' > N is an over-complete array response at the RIS, $\mathbf{X}_k \in \mathbb{C}^{M' \times N'}$ is the channel coefficient matrix of UE k assumed to be sparse in which each element represents the channel gain along the associated path, and $\mathbf{A}_{R} \in \mathbb{C}^{M \times M'}$ with M' > M is an over-complete array response at the BS. Hence, the problem of channel estimation is reduced to estimating X_k from the least squares signal $\mathbf{r}_{b,k}$ via CS. However, the authors note that the application of the standard OMP algorithm directly to the least squares signal brings substantially two disadvantages: (1) the OMP algorithm requires an accurate sampling of the angular domain to obtain good results, i.e., very large N' and M' which lead to complex matrix operations, and (2) this estimator requires an increasing training overhead in terms of pilot sequences as the channel sparsity increases. Hence, the authors propose to apply OMP on $\mathbf{X}_k \mathbf{A}_R^H$ by exploiting its row block sparsity structure. Note that this significantly reduces computational complexity since typically $M \leq N'$. Moreover, since the link connecting the BS and the RIS is common to all UEs, the aggregated effective channel of all the UEs exhibits both row and column block sparsity which can be leveraged to further enhance the performance of OMP and reduce the training overhead. Note that the column block sparsity is given by the shared **G** channel among all UEs.

6.3.3.2 Beam Training

The authors in [17] study an indoor RIS-assisted network with a massive MIMO BS serving a single receiver equipped with N_u antennas with the aid of a total of N_i RISs operating at THz frequency in the absence of the LoS path. In this case, the sparsity in the channel is given by the large-scale antenna array at the BS and the high pathloss at THz frequencies. The effective channel is thus modelled as

$$\overline{\mathbf{H}} = \sum_{i=1}^{N_i} \overline{\mathbf{H}}_i \tag{6.14}$$

where $\overline{\mathbf{H}}_i$ is the effective channel that is reflected by the RIS *i* via the reflecting coefficients in \mathbf{v}_i . Assuming for simplicity, a uniform linear array (ULA) at both the BS and the receiver, the effective channel relative to RIS *i*, is described as

$$\overline{\mathbf{H}}_{i} = \eta_{i} \mathbf{a}_{N_{u}} \left(\theta_{UR}^{i}\right) \mathbf{a}_{N} \left(\theta_{RU}^{i}\right)^{H} \boldsymbol{\Phi}_{i} \mathbf{a}_{N} \left(\theta_{RB}^{i}\right) \mathbf{a}_{M} \left(\theta_{BR}^{i}\right)^{H}$$
(6.15)

where η_i is the overall path-loss coefficient which depends on the distance from the receiver to the RIS and from the RIS to the BS and $\mathbf{a}_{N_u} \left(\theta_{UR}^i \right)$ is the ULA response vector for the steering angle θ_{UR}^i defined as

$$\mathbf{a}_{N_u}\left(\theta_{UR}^i\right) = \frac{1}{\sqrt{N_u}} \left[1, e^{j2\pi\delta\sin\left(\theta_{UR}^i\right)}, \dots, e^{2\pi\delta\left(N_u-1\right)\sin\left(\theta_{UR}^i\right)}\right]^T$$
(6.16)

with δ being the ratio between the antenna spacing and the signal wavelength. Lastly, note that θ_{UR}^i is the angle of departure (AoD) from the receiver to the *i*-th RIS, θ_{RU}^i is the angle of arrival (AoA) of the same link, θ_{RB}^i is the AoD from the *i*-th RIS to the BS, and θ_{BR}^i is the AoD of the same link. The effective channel is thus estimated via beam training, in which the BS, RIS configuration, and receiver all sweep through a codebook of beam directions, keeping as candidate estimate the direction which gives the strongest received beam power. This is done via a hierarchical search method which greatly reduces the complexity compared to brute-force exhaustive search. In a first stage, only the BS to RIS link is considered. Once the best candidate direction is found, the algorithm considers the RIS to receiver link with the BS to RIS link fixed as the result of the first stage. Note that the direct link between the BS and the UE is estimated in a prior phase via hierarchical beam search with all the RIS elements deactivated.

A similar case assuming a single RIS and both BS and receiver equipped with one antenna only has been studied in [18]. Here, the indoor network is assumed to operate at mmWave frequencies, and both the BS-RIS and RIS-UE links are assumed to be dominated by the LoS only. As in [17], the channel is modelled as depending only on distances AoA and AoD of the two separate links. Hence, in this case, the channel is completely identified by the position of the UE in space.

In order to reduce complexity, the authors propose to divide the RIS into a series of rectangular blocks of reflecting units (RUS). Each RUS is considered as an observation point that is used to estimate the position of the UE via triangulation.

Each RUS is used to sweep through a set of directions in which it is most likely to find the UE. For each RUS, the direction of maximum received power is used as an estimate of the UE position, while the corresponding RUS-UE distance is estimated using classical wideband delay estimation methods. Such estimates are then combined into one refined estimate via triangulation.

6.3.3.3 Sparse Matrix Factorization

The authors in [19] study the joint activity detection and channel estimation problem in a scenario in which a large number of UEs are present in the network, but only a small percentage of them are active in any given time instant. Hence, besides estimating the channels, the goal of this paper is to detect which UEs are actually active during each channel coherence block. In this case, the sparsity is given by the matrix $\mathbf{A} = diag [\alpha_1, \dots, \alpha_K]$ which indicates whether each UE *k* is active or not, i.e., $\alpha_k \in \{0, 1\}$, where $\alpha_k = 1$ indicates that UE *k* is active and $\alpha_k = 0$ indicates that UE *k* is inactive and by a sparse design of the sequences \mathbf{v}_t . Indeed, at each time instant *t*, each RIS element is activated according to a Bernoulli distribution and with uniformly distributed phases shift. Assuming that the direct links between the BS and the UEs are blocked and do not carry a significant amount of power, the receive signal at the BS can be rewritten as

$$\mathbf{Y} = \mathbf{G} \left(\mathbf{V} \odot \left(\mathbf{HAX} \right) \right) + \mathbf{N}$$
(6.17)

where V contains all the T training sequences as columns and H contains the channels from each UE to the RIS as columns. Equation (6.17) can be further simplified as

$$\mathbf{Y} = \mathbf{G}\mathbf{W} + \mathbf{N} \tag{6.18}$$

where we have defined the matrices $\Theta = HA$, $Q = \Theta X$ and $W = P \odot Q$ which are all sparse and can be recovered via the following techniques: SMF is employed to estimate G and W from the observations in Y, and matrix completion is used to complete the missing entries of Q given the estimate of W and the training sequences in V. Lastly, multiple measurement vectors are used to estimate Θ from the estimate of Q and the pilot signals in X. Although simulations show that this method requires three times more pilot sequences than RIS elements to obtain sufficiently accurate channel estimates, it also effectively solves the activity detection problem at the same time.

6.3.3.4 Matrix Calibration

In [20], the authors model the BS-RIS channel according to the Rician fading model, i.e., as being a summation of a deterministic part which represents the LoS link and a random part which represents the fast-fading part. Using a virtual representation of the channel via a grid of sampling angles, the channel G can be modelled as

$$\mathbf{G} = \sqrt{\frac{\gamma}{\gamma+1}} \mathbf{G}_{LoS} + \sqrt{\frac{1}{1+\gamma}} \mathbf{G}_{NLoS}$$
(6.19)

••

where γ is the Rician factor, \mathbf{G}_{LoS} represents the deterministic LoS link, and \mathbf{G}_{NLoS} represents the fast-fading part modelled as

$$\mathbf{G}_{NLoS} = \mathbf{A}_B \mathbf{S} \mathbf{A}_R^H. \tag{6.20}$$

Note that $\mathbf{S} \in \mathbb{C}^{M' \times N'}$ is the channel coefficient matrix assumed to be sparse in which each element represents the channel gain along the associated path, while $\mathbf{A}_B \in \mathbb{C}^{M \times M'}$ and $\mathbf{A}_R \in \mathbb{C}^{N \times N'}$ are as defined above. A similar modelling is used for the channel from each UE *k* to the RIS as

$$\mathbf{h}_k = \mathbf{A}_R \mathbf{h}'_k \tag{6.21}$$

where $\mathbf{h}'_k \in \mathbb{C}^{N' \times 1}$ is a sparse channel coefficients vector. The receive signal at the BS is thus expressed as

$$\mathbf{Y} = \left(\mathbf{H}_d + \mathbf{A}_B \mathbf{S} \mathbf{A}_R^H \mathbf{A}_R\right) \mathbf{H}' \mathbf{X} + \mathbf{N}$$
(6.22)

where the only unknowns are the matrices **S** and **H**['] which are then estimated via posterior mean estimators, i.e., by studying the MMSE, and derived using a sum-product message passing algorithm. Numerical results show that this method requires a number of training symbols that scale linearly with the number of UEs in order to obtain sufficiently good estimation of the channels. Note that since the number of UEs is usually less than the number of RIS elements, this method effectively reduces the training overhead compared to on/off schemes.

6.3.3.5 Orthogonal Matching Pursuit

Lastly, we present a set of works dealing with OMP-based channel estimation. We highlight their characteristics and present a third approach which tries to counteract its limitations.

The authors in [21] study a mmWave cellular system in which the first link G between the BS and the RIS is assumed to be dominated by the LoS part and thus

known a priori. The channels $\{\mathbf{h}_k\}$ from each UE to the RIS are assumed to be sparse and are recovered using CS. In particular, using a virtual representation of the channel as in [20], an OMP algorithm is designed to recover the sparse coefficient vector.

The authors also study the design of the RIS configurations for each one of the T pilot sequences. Such sequences comprise both the phase shifts introduced by the RIS and the baseband part which is implemented at the BS. The design choice in this case is to match the BS-to-RIS major channel directions and to uniformly spread the signal along the angular dimension for the RIS-to-UEs channels. In such a way, the authors intend to exploit the known strong channel directions from the BS to the RIS which are dictated by the LoS path and to accurately sound the channel from the RIS to the UEs.

However, as it is well-known in the literature, the OMP algorithm may fail in case the sampling grid taken to sound the signal is not precise enough, i.e., if there are not enough degrees of freedom (e.g., in the form of antennas) at both the RIS and the BS.

The authors in [22] find a sparse representation for both channels **G** and \mathbf{h}_k by exploiting the properties of the Kronecker and Khatri-Rao products. In particular, the BS-RIS channel is modelled as

$$\mathbf{G} = \mathbf{A}_B \boldsymbol{\Sigma} \mathbf{A}_R^H \tag{6.23}$$

where A_B and A_R are the over-complete pre-discretized grids of directions at the BS and RIS, respectively, while Σ is the sparse channel coefficient matrix. Similarly, the channel between the RIS and the UE is expressed as

$$\mathbf{h}_k = \mathbf{A}_R \boldsymbol{\alpha} \tag{6.24}$$

where α is the sparse channel coefficient vector. The effective channel $\overline{\mathbf{H}}_k$ can be thus expressed as

$$\overline{\mathbf{H}}_{k} = \mathbf{D}_{U} \mathbf{\Lambda} \mathbf{A}_{B}^{H} \tag{6.25}$$

where \mathbf{D}_U is a matrix constructed by taking the first N' columns of the matrix $\mathbf{A}_R^* \mathbf{A}_R$ with \circ representing the Khatri-Rao product and $\mathbf{\Lambda} = (\boldsymbol{\alpha}^* \otimes \boldsymbol{\Sigma})$ with \otimes representing the Kronecker product. Hence, all the relevant channel information of both the BS-RIS and RIS-UE channels is contained in $\mathbf{\Lambda}$ which can be estimated via a conventional OMP algorithm.

Again, the authors assume that the true AoAs and AoDs are contained within the pre-discretized grids A_B and A_R thus neglecting possible mismatches which may cause the OMP algorithm to fail.

To overcome the aforementioned limitations, the authors in [23] propose an iterative reweighted method where the channel estimation is performed by sending in the downlink a series of T random training matrices each of which are reflected by

the RIS with a random phase-shift matrix and combined (with a random combining matrix) by a single multi-antenna user. The composite channel between the BS and the RIS and the RIS and the user is assumed to be entirely LoS. Hence, the only parameters to be estimated are the instantaneous propagation path gains and AoA and AoD of the LoS links.

The channel estimation problem is formulated as the minimization of the sum over all training symbols of the matrix norm difference between the received signal and its parametric model which depends on the product of the instantaneous path gains between the BS and the RIS and between the RIS and the user and the corresponding AoAs and AoDs plus a regularization term which ensures sparsity of the estimated channel vector. In this first step, the output of the algorithm is an estimated product of instantaneous pathloss gains and an estimated difference of directional sine between the AoA and AoD, i.e., the difference between the sine of the AoD and the sine of the AoA. In a second step, both such parameters are further refined using gradient descent.

The authors compare their method over conventional OMP-based approaches demonstrating that it guarantees a higher sum rate performance. However, the gradient descent-based second step of their proposed method may result in a slow convergence of the overall algorithm.

6.3.3.6 Machine Learning

In [24], a fully connected artificial neural network is adopted in a RIS-aided wireless system to estimate the channels and phase angles from a reflected signal received through an RIS. The proposed deep network consists of four hidden layers, each of which is a fully connected layer followed by a hyperbolic tangent (tanh) activation function. The numbers of neurons in the fully connected layers are given following a test and trial method. To avoid overfitting of the network, the channel and additive white Gaussian noise (AWGN) intensities are shuffled at each iteration. The network maps the effects of the channel and phase angles on the transmitted signal using the nonlinear function approximation in its hidden layers. The proposed deep network yields an improved performance compared with the conventional LS and MMSE estimators.

In [25], a supervised deep learning framework is used for channel estimation in a RIS-assisted massive MIMO system. The authors designed a twin convolutional neural network (CNN) for the estimation of direct (BS-user) and cascaded (BS-RISuser) channels. The CNN is fed with the received pilot signals, and it constructs a nonlinear relationship between the received signals and the channel data. First, all of the RIS elements are turned off using the BS backhaul link, and the deep network to estimate the direct channel is trained. Then, each of the RIS elements are turned on one by one to finally estimate the cascaded channel. In the deep network, real, imaginary, and the absolute value of each entry of the received signal is fed as input, because the use of "three-channel" data ameliorates the performance by enriching the features inherited in the input data. The approach is compared against state-ofthe-art deep learning-based techniques, and performance gains are shown.

In [26–29], the authors adopted a design of a small portion of active elements on the RIS. In [28], to improve the channel estimation performance, the authors proposed to utilize deep learning to reduce the angle offset rate. While in [29], a complex-valued de-noising convolution neural network is further proposed to enhance performance.

6.4 The Road Ahead

In this section, we provide a non-exhaustive list of major open research problems that we consider to be of great importance for unveiling the potential benefits of RISs.

- EM-based circuit models. Current studies on RIS mostly rely on simplified models of RIS. To obtain accurate characteristic of RIS functionalities, it is therefore imperative to develop basic understanding of the working principles of RIS by taking a physics-based approach on the analysis. In particular, the effect of the spatial coupling among the meta-atoms needs to be taken into account.
- 2. *Path-loss and channel modeling.* In order to obtain accurate performance limits of RIS in wireless networks, realistic models for the propagation of the signals scattered by the RIS are required. Additionally, one needs to consider not only the far-field regime, which is commonly assumed in a large portion of RIS analysis, but also in the near-field regime in which the benefits of RISs deployment may arise. Along this line of research, some fundamental works such as [30] have been proposed.
- 3. *Fundamental performance limits.* Depending on how a RIS is utilized, difference performance limits may be obtained. Therefore, it is important to develop theoretical frameworks that can capture these performance limits which are still largely unknown to date.
- 4. Large-scale networks: deployment, analysis, and optimization. Thanks to its low cost, low energy, and low complexity of deployments, RIS has an advantage over its competing technologies to be implemented in a large-scale environment. However, unfortunately, most studies in the literature are limited to "small-size" system models where usually one or only a few RISs are considered. To investigate the potential of large-scale RIS deployments, more studies need to be conducted that take into account large-scale networks with hundreds or possibly thousands of RIS elements.
- 5. Low-complexity channel estimation. Due to its passive nature, RIS lacks the ability to "sense" the wireless environment, and thus channel estimation is an integral part in designing a reliable system based on RIS. In this chapter, we introduced the state of the art of channel estimation in RIS-based systems such as on/off-based algorithm and machine-learning-based methods. The complexity

of these methods increases with the number of the RIS elements. Since RIS is normally made up hundreds or thousands of elements, an improvement on lowcomplexity channel estimation method is essential in order to bring RIS into realization.

Acknowledgments The work of F. Danufane and J. Liu was supported in part by the European Commission through the H2020 5GstepFWD project under grant agreement number 722429.

References

- ITU-R (2015) IMT traffic estimates for the years 2020 to 2030. http://www.itu.int/pub/R-REP-M.2370
- Danufane, F. H., Di Renzo, M., De Rosny, J., & Tretyakov, S. (2020). On the path-loss of reconfigurable intelligent surfaces: An approach based on green's theorem applied to vector fields. https://arxiv.org/abs/2007.13158
- Arun, V., & Balakrishnan, H. (2020). RFocus: Beamforming using thousands of passive antennas. In 17th USENIX symposium on networked systems design and implementation (NSDI 20), pp. 104 7–1061.
- NTT DOCOMO. (2000). DOCOMO conducts world's first successful trial of transparent dynamic metasurface. https://www.nttdocomo.co.jp/english/info/media_center/pr/2020/ 0117_00.html. Accessed 26 Aug 2000.
- Yu, N., Genevet, P., Kats, M. A., et al. (2011). Light propagation with phase discontinuities: Generalized laws of reflection and refraction. *Science*, 334(6054), 333–337. https://doi.org/ 10.1126/science.1210713
- Di Renzo, M., Ntontin, K., Song, J., Danufane, F. H., Qian, X., et al. (2020). Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison. *IEEE Open Journal of the Communications Society*, 1, 798–807. https://doi.org/10.1109/ OJCOMS.2020.3002955
- Nadeem, Q. U. A., Kammoun, A., Chaaban, A., et al. (2019). Intelligent reflecting surface assisted wireless communication: Modeling and channel estimation. https://arxiv.org/abs/ 1906.02360
- Lin, J., Wang, G., & Fan, R., et al. (2019). Channel estimation for wireless communication systems assisted by large intelligent surfaces. https://arxiv.org/abs/1911.02158
- Mishra, D., & Johansson, H. (2019). Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp 4659– 4663. https://doi.org/10.1109/ICASSP.2019.8683663
- Yang, Y., Zheng, B., Zhang, S., et al. (2020). Intelligent reflecting surface meets OFDM: Protocol design and rate maximization. *IEEE Transactions on Communications*, 68(7), 4522– 4535. https://doi.org/10.1109/TCOMM.2020.2981458
- Zheng, B., & Zhang, R. (2019). Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization. *IEEE Wireless Communications Letters*, 9(4), 518–522. https://doi.org/10.1109/LWC.2019.2961357
- You, C., Zheng, B., & Zhang, R. (2019). Intelligent reflecting surface with discrete phase shifts: Channel estimation and passive beamforming. In *IEEE International Conference on Communications (ICC)*, Dublin, Ireland, pp. 1–6. https://doi.org/10.1109/ICC40277.2020.9149292
- Jensen, T. L., & De Carvalho, E. (2020). An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5000–5004. https:// doi.org/10.1109/ICASSP40776.2020.9053695

- Wang, Z., Liu, L., & Cui, S. (2020). Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis. *IEEE Transactions* on Wireless Communications. https://doi.org/10.1109/TWC.2020.3004330
- Wei, L., Huang, C., Alexandropoulos, G. C., et al. (2020). Parallel factor decomposition channel estimation in RIS-assisted multi-user MISO communication. In *IEEE 11th sensor* array and multichannel signal processing workshop (SAM), pp. 1–5. https://doi.org/10.1109/ SAM48682.2020.9104305
- Chen, J., Liang, Y. C., Cheng, H. V., et al. (2019). Channel estimation for reconfigurable intelligent surface aided multi-user MIMO systems. https://arxiv.org/abs/1912.03619
- Ning, B., Chen, Z., Chen, W., et al. (2019). Channel estimation and transmission for intelligent reflecting surface assisted THz communications. In *IEEE international conference* on communications (ICC), pp. 1–7. https://doi.org/10.1109/ICC40277.2020.9149153
- Cui, Y., & Yin, H. (2019). An efficient CSI acquisition method for intelligent reflecting surfaceassisted mmwave networks. https://arxiv.org/abs/1912.12076
- Xia, S., & Shi, Y. (2020). Intelligent reflecting surface for massive device connectivity: Joint activity detection and channel estimation. In *IEEE international conference on* acoustics, speech and signal processing (ICASSP), pp. 5175–5179. https://doi.org/10.1109/ ICASSP40776.2020.9054415
- Liu, H., Yuan, X., & Jun, Y. (2020). Matrix-calibration-based cascaded channel estimation for reconfigurable intelligent surface assisted multiuser MIMO. *IEEE Journal on Selected Areas* in Communications.https://doi.org/10.1109/JSAC.2020.3007057
- Wan Z, Gao Z, Alouini M S (2020) Broadband channel estimation for intelligent reflecting surface aided mmWave massive MIMO systems. https://arxiv.org/abs/2002.01629
- Wang, P., Fang, J., Duan, H., et al. (2020). Compressed channel estimation and joint beamforming for intelligent reflecting surface-assisted millimeter wave systems. *IEEE Signal Processing Letters*, 27, 905–909. https://doi.org/10.1109/LSP.2020.2998357
- He, J., Leinonen, M., & Wymeersch, H., et al. (2020). Channel estimation for RIS-aided mmWave MIMO channels. https://arxiv.org/abs/2002.06453
- 24. Khan, S., & Shin, S. Y. (2019). Deep-learning-aided detection for reconfigurable intelligent surfaces. https://arxiv.org/abs/1910.09136
- Elbir, A. M., Papazafeiropoulos, A., Kourtessis, P., et al. (2020). Deep channel learning for large intelligent surfaces aided mm-wave massive MIMO systems. *IEEE Wireless Communications Letters*.https://doi.org/10.1109/LWC.2020.2993699
- 26. Taha, A., Alrabeiah, M., & Alkhateeb, A. (2019). Enabling large intelligent surfaces with compressive sensing and deep learning. https://arxiv.org/abs/1904.10136
- Taha, A., Zhang, Y., Mismar, F. B., et al. (2020). Deep reinforcement learning for intelligent reflecting surfaces: Towards standalone operation. In 2020 IEEE 21st international workshop on signal processing advances in wireless communications (SPAWC), Atlanta, GA, USA, pp. 1–5. https://doi.org/10.1109/SPAWC48557.2020.9154301
- Jiang, F., Yang, L., da Costa, D. B., et al. (2020). Channel estimation via direct calculation and deep learning for RIS-aided mmWave systems. https://arxiv.org/abs/2008.04704
- Liu, S., Gao, Z., Zhang, J., et al. (2020). Deep denoising neural network assisted compressive channel estimation for mmWave intelligent reflecting surfaces. *IEEE Transactions on Vehicular Technology*, 69(8), 9223–9228. https://doi.org/10.1109/TVT.2020.3005402
- 30. Danufane, F. H., Di Renzo, M., de Rosny, J., et al. (2020). On the path-loss of reconfigurable intelligent surfaces: An approach based on green's theorem applied to vector fields. https:// arxiv.org/abs/2007.13158