Channel Estimation Method and Phase Shift Design for Reconfigurable Intelligent Surface Assisted MIMO Networks

Jawad Mirza, Senior Member, IEEE and Bakhtiar Ali

Abstract

Reconfigurable intelligent surface (RIS) aided multiple-input multiple-output (MIMO) communication is becoming a serious contender for future wireless networks. The reason for this attention is due to reliable communication and low cost deployment offered by the RIS based communication systems compared to the conventional massive MIMO systems. However, the performance of an RIS assisted MIMO system heavily depends on the quality of channel state information available at both ends. As communicating devices and RIS are in close proximity, the desired channels can be modeled as line-of-sight (LOS) channels, where dominant components comes from LOS paths. Thus, leading to ill-conditioned channel matrices. To estimate these channel matrices with high quality, we propose a two stage channel estimation method for RIS aided MIMO time-division duplexing (TDD) communication systems. In particular, we employ the conventional TDD based MIMO channel estimation technique in the first stage to estimate the direct MIMO channel between terminals with no RIS. Whereas, in the second stage of the channel estimation process, we propose to use a recently developed bilinear adaptive vector approximate message passing (BAdVAMP) algorithm to estimate ill-conditioned RIS channels. The BAdVAMP method belongs to the class of approximate message passing (AMP) algorithms and it can be used in wide range of estimation and learning problems in compressed sensing. The BAdVAMP algorithm has been shown to be accurate and robust in ill-conditioned dictionary learning problems. We also propose a phase shift design for the RIS using the estimated channels. Specifically, we formulate an optimization problem that maximizes the total channel gain at the user. For this purpose, a closed-form expression to obtain the phase shift of each passive element in the RIS is also derived. Different from previous studies where single antenna is used at the user, in this study we have assumed multiple antennas at both ends and employ the MIMO eigenmode transmission strategy. Numerical results show that the proposed BAdVAMP based RIS channel estimation performs better than its counterpart i.e., bilinear generalized AMP (BiGAMP) based RIS channel estimation scheme. This performance improvement is not only evident in ill-conditioned scenarios, but also when channel matrices are well-conditioned.
I. Introduction

The number of connected mobile devices and the amount of data traffic through these devices are expected to grow many fold in future communication networks [1]. To support the scale of this huge data traffic, massive multiple-input multiple-output (MIMO) technology is being considered in 5G wireless communication standards. Nevertheless, practically deploying a large number of antennas increases the hardware cost and power consumption of the network [2]. In contrast, the deployment of reconfigurable intelligent surfaces (RISs) to support the wireless communication network is considered to be a cost effective and energy efficient solution that possesses all the known advantages of the massive MIMO technology [3]. A typical RIS consists of a planar array having a large number of reflecting metamaterial elements (e.g., low-cost printed dipoles) which act as phase shifts to assist the wireless communication between transmitter and receiver. In particular, the passive elements or reflect arrays of the RIS interact with the incident signals and reflect them towards the desired user.

Traditionally, reflecting surfaces have been used in radar and satellite communications with fixed phase shifts, therefore, they were not useful in the scenarios where the propagation environment is rapidly changing. However, advancements in RF micro-electromechanical systems (MEMS) and metamaterials have paved the way for reconfigurable reflectors in terrestrial networks, which can be used to enhance received signal power, suppress interference and improve security [4]. An RIS could also assist in controlling the wireless environment, hence giving rise to the concept of a smart environment [5], where a wireless environment is turned into a smart reconfigurable space to improve the performance of the communication system. Note that many future wireless technologies are being integrated with RIS to determine the offered performance gain, such as wireless information and power transfer [6], [7], unmanned aerial vehicle (UAV) communication [8] and non-orthogonal multiple access (NOMA) [9], [10].

Due to its attractive features, the RIS technology has attracted considerable research attention in terrestrial wireless communication systems. However, the quality of channel state information (CSI) available at the terminals plays an important role in determining the performance of the RIS based communication system [11]–[14]. In addition to that, an appropriate RIS phase shift design is also important. First, we briefly discuss some of the seminal studies on phase shift design also known as reflect
or passive beamforming, and later we provide an overview of existing channel estimation techniques in RIS assisted communication systems. For a single-user system, with multiple antennas at the access point (AP) and a single antenna at the user, RIS phase shift designs are proposed in [15], [16] to improve the spectral efficiency of the system. An RIS-aided millimeter-wave (mmWave) MIMO system is considered in [17], [18], where in [18], a joint phase shift and transmit beamforming vector optimization is carried out at the AP to maximize the received signal power at the user. It has been reported in [18] that the performance with discrete phase shifts (low-resolution) is comparable with the performance of the continuous (high resolution) phase shifts. In addition to single-user scenarios, the use of RIS has also been investigated for multiuser (MU) communication scenarios. For a MU MISO system aided with RIS, a hybrid beamforming scheme is proposed in the absence of direct links between the AP and users [19]. According to [19], the minimum number of transmit antennas required at the AP decreases to half of that required in the traditional hybrid beamforming schemes without RIS, thus, decreasing the hardware cost significantly.

As discussed earlier, the quality of CSI available at the AP/user in RIS assisted MIMO communication plays a crucial role in the performance of the system. Due to this reason, the channel estimation of RIS based channels has attracted an enormous amount of research attention recently [11]–[14], [18], [20], [21]. Here, we provide a discussion on the channel estimation work carried out in RIS assisted communication systems. The conventional time-division duplexing (TDD) based MIMO channel estimation method using pilots is employed in [4] to estimate the AP-user and IRS-user channels. It is assumed that the RIS is equipped with receive RF chains that enable it to estimate the channels. For a single antenna each at AP and user, the channel estimation process in [12] minimizes the mean squared error (MSE) of the channel estimation problem using discrete phases in the training phase. For the same antenna settings as in [12], the authors in [13] employ a cost efficient method with no receive RF chains at the RIS, where concatenated user-RIS-AP channels are estimated at the AP. Moreover, the idea of grouping RIS elements to reduce the complexity of the channel estimation process is also introduced in [13].

Prior to [13], the estimation of concatenated user-RIS-AP channels was considered in [11], [14], where ON/OFF switching of RIS elements is implemented for the channel estimation. In [11], the channel estimation procedure for an RIS assisted communication system having multiple antennas at the AP is presented, where each passive element controlled by the AP is powered ON one by one to capture the least square (LS) estimation of RIS channels. The sparse nature of mmWave channel is exploited in [18] to
convert the channel estimation problem into a sparse signal recovery problem, which is then solved by using compressed sensing algorithms such as orthogonal matching pursuit (OMP) and expectation maximization generalized approximate message passing (EMGAMP). According to [18], the training overhead with the EMGAMP algorithm is lower than the classical OMP algorithm. To reduce the training overhead, a three phase channel estimation process is proposed in [20] for the MU system, where correlation among the RIS reflected channels are exploited to reduce the training overhead. In [21], the channel estimation method for the RIS assisted THz massive MIMO system with a hybrid beamforming architecture is proposed which uses a codebook based beam training approach.

The aforementioned seminal studies on the channel estimation problem either assume a single antenna at both AP and user or multiple antennas at the AP. In this study, we investigate RIS channel estimation for the system where both AP and user are equipped with multiple antennas. Since, RIS will be densely deployed in indoor and outdoor locations, the near-field communication channels are likely to experience line-of-sight (LOS) propagation conditions [22]–[24]. This motivates us to investigate the channel estimation of ill-conditioned RIS channels in the MIMO system with multiple transmit and receive antennas. For this purpose, we employ a bilinear adaptive vector approximate message-passing (BAdVAMP) scheme developed in [25], which is originally designed for reconstructing ill-conditioned and mean shifted matrices in dictionary learning problems. This recently developed algorithm has not yet been applied to channel estimation problems in wireless communication systems, therefore, this is the first study to use the BAdVAMP algorithm to estimate RIS based ill-conditioned MIMO channels. Previously, for RIS assisted MIMO systems, the authors in [14] estimate the RIS channels using the approximate message-passing (AMP) based algorithm known as bilinear generalized AMP (BiGAMP). However, the BiGAMP method is designed only for reconstructing well-conditioned matrices, whereas, the BAdVAMP algorithm performs better than the BiGAMP algorithm in both ill-conditioned and well-conditioned matrices [25]. Moreover, the BiGAMP method has higher computational complexity and less accuracy as compared to the recently proposed BAdVAMP method [25]. In this paper, the proposed channel estimation technique has two stages. In the first phase, the direct MIMO channel between user and AP is estimated using the conventional TDD based MIMO channel estimation, whereas, the BAdVAMP based channel estimation for RIS channels is carried out in the second stage. The RIS estimated channels via the BAdVAMP algorithm have inherent

\[1\]In an ill-conditioned channel matrix, some of the singular values of the channel are negligible compared to the largest singular value of the channel. The term condition number is generally used to characterize the ill-conditioned channel matrix, which is the ratio of the largest singular value of that channel matrix to the smallest singular value.
ambiguities out of which the permutation ambiguity has the most destructive effects on the estimated channels. Therefore, to get rid of the permutation ambiguity from the recovered channels, we propose an effective method that exploits the sparsity in the phase shifts during the training phase.

Once the direct and RIS channels are estimated using the proposed two-stage channel estimation scheme, we use these estimated channels to design the phase shifts for RIS elements. In this paper, we design the phase shifts such that the total channel gain at the user is maximized. In particular, we derive a closed form expression for phase shifts that maximizes the total channel gain at the user. The proposed scheme considers continuous or high resolution phase shifts which can be easily transformed to quantized discrete phase shifts. We use channel gain maximization for phase shift design because the fixed eigenmode MIMO [26] based precoding/combining strategy employed in this study is realized using the the estimated channels and designed phase shifts. We assume that a programmable controller is fabricated with RIS which is controlled by the AP to adjust the phases of the RIS based on the proposed channel gain maximization approach. The contributions of the paper are summarized below:

- We propose a two-stage channel estimation method for the RIS assisted MIMO communication system with multiple transmit and receive antennas at the AP and user, respectively. In the first stage of the proposed scheme the direct MIMO channel between the AP and user is estimated, whereas, the RIS channels are estimated in the second stage. The RIS channel matrices are considered to be ill-conditioned, which is a realistic propagation scenario in near-field LOS communications. For the RIS channels, the channel estimation problem is transformed into the well known dictionary learning problem and the recently developed BAdVAMP algorithm is used to estimate the ill-conditioned channels.

- Based on the estimated channels, a phase shift design for the passive elements of RIS is proposed which maximizes the channel gain at the user. Here, we assume a continuous high-resolution phase shift design and derive a closed-form expression to compute the phase shift of each passive element which is assumed independent of phase shifts of other elements in the RIS.

The rest of the paper is organized as follows. Section [II] introduces the RIS assisted MIMO system model used in this study. Section [III] presents the proposed two-stage channel estimation scheme. The ambiguity elimination and the recovery of a unique solution from the estimated channels are discussed in Section [IV]. In Section [V] we derive the phase shift design for the RIS terminal. Numerical results are provided in Section [VI]. Finally, Section [VII] concludes the paper.
**Notations:** We use \((\cdot)^H\), \((\cdot)^*\), \((\cdot)^T\), \((\cdot)^{-1}\) and \((\cdot)^\perp\) to denote the conjugate transpose, the conjugate, the transpose, the inverse and the pseudoinverse operations, respectively. \(\mathbb{E}[.]\) denotes expectation. The complex normal distribution with mean \(\mu\) and variance \(\sigma^2\) is denoted by \(\mathcal{CN}(\mu, \sigma^2)\). The Hadamard product is represented by \(\odot\) and the rank of the matrix is given by \(\text{rank}(\cdot)\). For any given matrix \(A\), the quantity \(A_{i,j}\) denotes the entry of the matrix \(A\) corresponding to the \(i\)th row and \(j\)th column. Similarly, \(a_l\) represents the \(l\)th column of the matrix \(A\).

**II. System Model**

We consider a point-to-point MIMO communication system as shown in Fig. 1, where an Access Point (AP) with \(M\) antennas serves a single user having \(N\) antennas. An RIS is installed in a surrounding area, which consists of \(L\) passive reflecting elements and is connected to the AP via a programmable controller. The downlink channels from the AP to the RIS and from the RIS to the user are denoted by \(H^T \in \mathbb{C}^{L \times M}\) and \(G^T \in \mathbb{C}^{N \times L}\), respectively. In this paper, we consider a TDD transmission mode and assume perfect channel reciprocity. The uplink channels from the user to the RIS and from the RIS to the AP are given by \(G \in \mathbb{C}^{L \times N}\) and \(H \in \mathbb{C}^{M \times L}\), respectively. The direct uplink and downlink channels between the AP and the user are denoted by \(Z \in \mathbb{C}^{M \times N}\) and \(Z^T \in \mathbb{C}^{N \times M}\), respectively.

We assume that all the channels follow a flat quasi-static Rician block fading channel model\(^{2}\), where the entries of the direct channel follow an independent and identically distributed (i.i.d.) \(\mathcal{CN}(\mu_z, \sigma_z^2)\) distribution. Similarly the entries of \(G\) and \(H\) are i.i.d. and follow \(\mathcal{CN}(\mu_g, \sigma_g^2)\) and \(\mathcal{CN}(\mu_h, \sigma_h^2)\) distributions. Furthermore, we assume that the channel matrix \(H\) is an ill-conditioned matrix due to LOS propagation conditions and \(\kappa(H)\) represents the condition number of the channel matrix \(H\). It is assumed that the link between the RIS and user is highly correlated due to a nearly perfect LOS propagation \([14], [29], [30]\), which results in a low-rank channel matrix \(G\).

Each passive element in the RIS acts as a phase shift, which can be dynamically adjusted from the AP via the programmable controller. Let \(s_i \in (0, 1)\) and \(\phi_i \in (0, 2\pi]\) denote the amplitude reflection coefficient (ON/OFF state) and phase shift value at the \(i\)th element then, we can write the phase shift vector for any given time as \(s = [s_1 e^{j\phi_1}, s_2 e^{j\phi_2}, \ldots, s_L e^{j\phi_L}]^T\). Due to the flat quasi static nature of the channel, it is assumed that the channel remains constant for \(T\) samples or channel uses. The total number

\(^{2}\)There are various path loss models developed for RIS-assisted wireless communications \([27], [28]\). However, the path loss model does not directly influences the performance of the proposed channel estimation scheme. Therefore, for simplicity, we do not consider effects of the path loss in this study.
of pilot symbols sent by the user during the UL training stage is denoted by $T_r$. Therefore, the received signal at the AP for time $t$ can be written as:

$$y_a[t] = H (s[t] \odot Gx[t]) + Zx[t] + n_a[t], \quad (1)$$

where, $x[t] \in \mathbb{C}^{N \times 1}$ is the UL pilot signal transmitted by the user at time $t$. The UL pilot symbols satisfy the average power constraint at the user, such that $\mathbb{E}[x[t]x[t]^H] = I_N$. $n_a[t]$ is the complex additive white Gaussian noise (AWGN) vector at the AP having zero mean and covariance matrix $\sigma^2_{n_a} I_M$. For all the $T_r$ samples, we can write the received observation at the AP as

$$Y_a = H (S \odot GX) + ZX + N_a, \quad (2)$$

where, $Y_a \in \mathbb{C}^{M \times T_r}$, $X = [x[1], x[2], \ldots, x[T_r]] \in \mathbb{C}^{N \times T_r}$, $S = [s[1], s[2], \ldots, s[T_r]] \in \mathbb{C}^{L \times T_r}$ and $N_a \in \mathbb{C}^{M \times T_r}$. After the channel estimation process, the downlink received signal at the user is given by

$$r = (G^T \Phi H^T + Z^T) c + n_r, \quad (3)$$

where, $\Phi = \text{diag}(s_1 e^{j\phi_1}, s_2 e^{j\phi_2}, \ldots, s_L e^{j\phi_L})$ is the phase shift matrix. The vector $c = Wu$ of size $M \times 1$ denotes the transmitted signal, where the precoding matrix at the AP is given by $W \in \mathbb{C}^{M \times N_s}$, such that $\mathbb{E}[||w_n||^2] = 1$, where $w_n$ denotes the $n^{th}$ column of the matrix $W$. The total number of data streams is denoted by $N_s$. The data vector for the user is given by $u \in \mathbb{C}^{N_s \times 1}$. The AWGN vector at the user is

---

3In this study, we assume that the ON/OFF operation at the RIS unit is perfect, and therefore, we ignore the imperfections at the RIS unit.
\[ C = \log_2 \det \left( \mathbf{I} + \rho \left( \mathbf{W}^H \left( \mathbf{G}^T \Phi \mathbf{H}^T + \mathbf{Z}^T \right)^H \mathbf{V} \mathbf{V}^H \left( \mathbf{G}^T \Phi \mathbf{H}^T + \mathbf{Z}^T \right) \mathbf{W} \right) \right), \] (5)

given by \( \mathbf{n}_r \) having zero mean and covariance matrix \( \sigma_n^2 \mathbf{I}_N \).

Denoting \( \mathbf{\hat{G}}^T \), \( \mathbf{\hat{H}}^T \) and \( \mathbf{\hat{Z}}^T \) as the estimates of \( \mathbf{G}^T \), \( \mathbf{H}^T \) and \( \mathbf{Z}^T \), respectively, the transmit precoding matrix, \( \mathbf{W} \), is equal to the right singular vectors of the estimated matrix \( \left( \mathbf{\hat{G}}^T \Phi \mathbf{\hat{H}}^T + \mathbf{\hat{Z}}^T \right) \) (obtained via a singular value decomposition), corresponding to the \( N_s \) largest singular values. On the other hand, the user estimates the downlink channel during the downlink training phase and uses the left singular vectors of the estimated channel corresponding to the \( N_s \) largest singular values, and uses them as a receive combining matrix, denoted by a matrix \( \mathbf{V} \in \mathbb{C}^{N \times N_s} \). This transmission is also known as eigenmode transmission as it maximizes the signal-to-noise ratio (SNR) of the system and achieves full diversity order [26]. The resulting signal at the user can be written as

\[ \mathbf{\bar{r}} = \mathbf{V}^H \left( \mathbf{G}^T \Phi \mathbf{H}^T + \mathbf{Z}^T \right) \mathbf{c} + \mathbf{V}^H \mathbf{n}_r. \] (4)

The capacity of the system can be expressed as given in (5), where \( \rho \overset{\Delta}{=} 1/\sigma_n^2 \) denotes the SNR of the link.

In this study, the direct and RIS channels are estimated first, then on the basis of these estimated channels we design the phase shift matrix, \( \Phi \), that maximizes the total channel gain at the user. Once, the phase shift matrix, \( \Phi \), is designed, we use the composite estimated channel \( \left( \mathbf{\hat{G}}^T \Phi \mathbf{\hat{H}}^T + \mathbf{\hat{Z}}^T \right) \) to compute the fixed precoding and combining vectors (\( N_s = 1 \)) or matrices (\( N_s > 1 \)) according to the eigenmode transmission method.

### III. Proposed Channel Estimation Technique

In this section, we present our proposed channel estimation technique for estimating the direct user-AP channel, RIS-AP channel and user-RIS channel. The channel estimation comprises of two stages at the AP. In the first stage, the direct uplink MIMO channel from the user to the AP, given by \( \mathbf{Z} \), is estimated using the conventional TDD based MIMO channel estimation technique [31]. In the second stage, the RIS channels \( \mathbf{H} \) and \( \mathbf{G} \) are estimated using the BAdVAMP scheme [25]. We provide details of the channel estimation at the AP via uplink pilot symbols sent by the user. A similar procedure can be adapted at the
user to estimate the downlink channels. However, we skip the details of the channel estimation process at the user.

Note that the phase shift matrix is optimized after the channel estimation process by using the estimated channels. However, the phase shift design at the RIS unit also plays an important role during the channel estimation process. For the channel estimation phase, we assume that ON/OFF state values \( s_i \) of the RIS passive elements are independent and for the given training duration \( t \), the phase shift vector given by \( s[t] \) is an \( K \)-sparse vector, such that each non-zero value in the vector is equal to one. Note that the sparse nature of the matrix \( S \) in (2) is necessary for channel estimation via the BAdVAMP algorithm. It is assumed that the pilots transmitted by the users over the training duration \( T_r \) results in a full rank matrix, such that \( \text{rank}(X) = \min(N, T_r) \).

\[ \text{A. Conventional TDD based MIMO channel estimation technique} \]

In the first stage of the channel estimation procedure at the AP, the RIS terminal is powered off, i.e., \( \{s_i\} = 0, \forall i \), and the AP estimates the direct uplink channel \( Z \). The estimation process in this stage follows the conventional open-loop TDD based MIMO channel estimation method, where \( T_c \) pilot symbols are transmitted by the user. Using the MMSE criterion at the AP, the estimated direct channel can be expressed as \([31], [32]\)

\[ \hat{Z} = \frac{T_c \rho_u}{T_c \rho_u + 1} Z + \frac{\sqrt{T_c \rho_u}}{T_c \rho_u + 1} N_t, \]  

where \( N_t \) denotes an AWGN matrix whose entries follow a \( \mathcal{CN}(0, 1) \) distribution and \( \rho_u \) is the average SNR of the link.

\[ \text{B. BAdVAMP based channel estimation technique} \]

After the estimation of the direct MIMO channel between the AP and user, the second stage of the channel estimation process begins, where the user sends \( T_r \) pilot symbols to the AP. In this stage, the RIS terminal is switched on and the received signal at the AP is given by (1). The AP removes the direct channel part from (1) by using the estimated channel \( \hat{Z} \) and the known pilot matrix \( X \), such that the resulting signal, after all \( T_r \) pilot symbols are received at the AP, can be expressed as

\[ Y = H (S \odot GX) + N, \]  

(7)
where $N$ is the noise matrix which is equal to the sum of the AWGN part and errors induced while removing the direct path channel, such that $N = N_a + ZX - \tilde{Z}X$. Therefore, the entries of $N$ follow the $\mathcal{CN}(0, \sigma_{n_a}^2 + \sigma_e^2)$ distribution, where $\sigma_{n_a}^2$ and $\sigma_e^2$ denote the noise and error variances, respectively.

In this study, we propose to use the BAdVAMP algorithm to estimate channels $H$ and $G$ from (7). The BAdVAMP algorithm is known to be a computationally efficient method for estimating the matrices $H$ and $D$ from the noisy signal $Y = HD + N$ where, $D = S \odot GX$. This is also known as the dictionary learning problem in the domain of compressive sensing [25], [33]. The BAdVAMP algorithm [25] was shown to be superior to its contenders both in complexity and performance, especially when the matrices $H$ and $G$ are ill-conditioned, which is usually the case when the channel experiences LOS propagation conditions. It is also known that millimeter wave MIMO channels become ill-conditioned when the number of arriving paths are small i.e., less scattering [14].

We aim to estimate the matrix $D \in \mathbb{C}^{L \times T_r}$, from the noisy received matrix $Y \in \mathbb{C}^{M \times T_r}$. As we are dealing with the $L > M$ scenario and $D$ is a sparse matrix, our problem becomes equivalent to the estimation of sparse vectors from underdetermined linear measurements. In addition, as the matrix $H$ is also not known in our problem, the problem becomes a bilinear recovery problem. In this study, we consider an unstructured matrix $H$ which can be expressed as [25]

$$H(\psi_H) = \sum_{m=1}^{M} \sum_{l=1}^{L} \psi_{H,m,l} e_m e_l^T$$

where, $e_m$ denotes the $m^{th}$ basis vector and $\psi_H \in \mathbb{C}^{M \times L}$. We assume that the density of $D$, given by $p_D(.; \theta_d)$, is parameterized by the vector $\theta_d$. In this study, the $n^{th}$ value of the vector $\theta_d$ comprises of mean and variance values of the $n^{th}$ column of the matrix $D$. In order to estimate $H$ and $D$, the BAdVAMP aims to learn the parameters $\Theta \triangleq \{\psi_H, \theta_d\}$ from the matrix $Y$. In this study, it is assumed that the statistical information $\theta_d$ is known at the AP and we focus on recovering $\psi_H$ and $D$ from the noisy measurement matrix $Y$, given by

$$Y = H(\psi_H)D + N,$$

1) Background Estimation Theory: To compute the parameters in $\Theta$ that maximize the received signal distribution, the BAdVAMP algorithm uses maximum likelihood (ML) estimation given by

$$\hat{\Theta}_{\text{ML}} = \text{argmax } p_Y(Y; \Theta).$$
Based on these estimated parameters, the MMSE estimate of $D$ can be expressed as [25]

$$
\hat{D}_{\text{MMSE}} = \mathbb{E}[D | Y; \hat{\Theta}_{\text{ML}}].
$$

(11)

The BAdVAMP algorithm approximates the quantities in (10) and (11), i.e., $\hat{\Theta}_{\text{ML}}$ and $\hat{D}_{\text{MMSE}}$, respectively. To find the expected value in (11), the mean of the posterior density $p_{D|Y}(D|Y; \hat{\Theta}_{\text{ML}})$ is required. Using Bayes’ rule we can write the posterior density as

$$
p_{D|Y}(D|Y; \hat{\Theta}_{\text{ML}}) = \frac{p_{Y|D}(Y|D; \hat{\Theta}_{\text{ML}})p_{D|\Theta}(D; \hat{\Theta}_{\text{ML}})}{p_{Y}(Y; \hat{\Theta}_{\text{ML}})}
$$

(12)

where $p_{Y|D}(Y|D; \hat{\Theta}_{\text{ML}})$ is the likelihood function for (11), $p_{D|\Theta}(D; \hat{\Theta}_{\text{ML}})$ denotes the prior density and the likelihood function of (10) is represented by $p_{Y}(Y; \hat{\Theta}_{\text{ML}})$. Following the decoupling of the posterior density (12) across the columns of $D$, as shown in [25], we can express (12) as

$$
p_{D|Y}(D|Y; \hat{\Theta}_{\text{ML}}) \propto \prod_{t_r=1}^{T_r} p_{d|y}(d_{t_r}; \hat{\Theta}_{\text{ML}}) p_{y|d}(y_{t_r}|d_{t_r}; \hat{\Theta}_{\text{ML}}).
$$

(13)

Using this decoupled posterior density, a vector AMP (VAMP) algorithm [34] can be used to find the matrix $D$ from (9), where each column of $D$ is tracked independently. However, the VAMP algorithm requires complete information of $\Theta$ to estimate $D$. From (10) and (12), the ML estimate of $\Theta$ can be computed as

$$
\hat{\Theta}_{\text{ML}} = \arg \min_{\Theta} \ln \int p_{D}(D; \Theta)p_{Y|D}(Y|D; \Theta) \, dD.
$$

(14)

The BAdVAMP algorithm uses the expectation maximization (EM) approach [35] to solve the problem (14) in an iterative manner. Therefore, the BAdVAMP algorithm interleaves the EM and VAMP algorithms to estimate $H$ and $D$ from the matrix $Y$.

2) The BAdVAMP Procedure: As discussed above, the BAdVAMP algorithm [25] employs the EM approach to estimate $\psi_H$ and the VAMP [36] algorithm is used to estimate the matrix $D$. In particular, the BAdVAMP method performs a joint estimation of parameters using the following optimization problem [25]

$^4$Here, we present the brief overview of the BAdVAMP algorithm. We refer the reader to [25] for a detailed description of it.
where $\mathbf{D}$ is given by

$$\mathbf{D} = \mathbf{D}_2 = \left[ d_{2,1}, d_{2,2}, \ldots, d_{2,T_r} \right].$$

The BA\text{VAMP} algorithm presented in Algorithm 1 recovers $\hat{\mathbf{H}}$ and $\hat{\mathbf{D}}$ up-to certain phase, scalar and permutation ambiguities. For a unique solution, it is important to remove these ambiguities from $\hat{\mathbf{H}}$ and $\hat{\mathbf{D}}$. 

---

1. In Algorithm 1, $g_1(\mathbf{r}_{1,\text{tr}}, \gamma_{1,\text{tr}}; \theta_d)$ represents a denoising function given by

$$g_1(\mathbf{r}_{1,\text{tr}}, \gamma_{1,\text{tr}}; \theta_d) \sim \mathcal{N}(\mathbf{d} + \mathbf{f}, \mathbf{I}/\gamma_{1,\text{tr}}).$$

In Algorithm 1, the quantity $\langle g_1(\mathbf{r}_{1,\text{tr}}, \gamma_{1,\text{tr}}; \theta_d) \rangle$ is $g_1(\mathbf{r}_{1,\text{tr}}, \gamma_{1,\text{tr}}; \theta_d)$ divergence at $\mathbf{r}$. In general, $g_1(\mathbf{r}_{1,\text{tr}}, \gamma_{1,\text{tr}}; \theta_d)$ can be interpreted as denoising of the AWGN-corrupted pseudo-measurement $\mathbf{r} = \mathbf{d} + \mathbf{f}$, with $\mathbf{f} \sim \mathcal{N}(0, \mathbf{I}/\gamma_1)$, which corresponds to the first loop i.e., $\tau_1$ in Algorithm 1 from line 3 till line 7. On the other hand, the second loop $\tau_2$ performs MMSE estimation of $\mathbf{d}$ from the AWGN-corrupted measurements under the pseudo-prior $\mathbf{d} \sim \mathcal{N}(\mathbf{r}_2, \mathbf{I}/\gamma_2)$ and the estimation of $\mathbf{H}(\psi_H)$ as given in line 10 to line 16. The final updated estimate of $\mathbf{H}$ using Algorithm 1 is given by

$$\hat{\mathbf{H}} = \mathbf{H}(\psi_H) = \mathbf{YD}_2^T \left( \mathbf{C}^T + \mathbf{D}_2^T \mathbf{D}_2^T \right)^{-1},$$

where $\mathbf{D}_2 = \left[ d_{2,1}, d_{2,2}, \ldots, d_{2,T_r} \right]$. On the other hand, the estimate of $\hat{\mathbf{D}}$ is given by

$$\hat{\mathbf{D}} = \mathbf{D}_2 = \left[ d_{2,1}, d_{2,2}, \ldots, d_{2,T_r} \right].$$
Algorithm 1 Bilinear Adaptive VAMP (BAdVAMP)

1: Initialization:
   \( \forall \): \( r_{1, tr}^{0}, \gamma_{1, tr}^{0}, \theta_{D}, \psi_{H}^{0} \)
2: for \( t = 0, \ldots, T_{\text{max}} \) do
3:   for \( \tau_1 = 0, \ldots, \tau_{1, \text{max}} \) do
4:     \( \forall tr: d_{1, tr}^{t} \leftarrow g_{1}(r_{1, tr}^{t}, \gamma_{1, tr}^{t}, \theta_{D}) \)
5:     \( \forall tr: 1/\eta_{1, tr}^{t} \leftarrow (g_{1}(r_{1, tr}^{t}, \gamma_{1, tr}^{t}, \theta_{D})) / \gamma_{1, tr}^{t} \)
6:     \( \forall tr: 1/\gamma_{1, tr}^{t} \leftarrow 1/N \|d_{1, tr}^{t} - r_{1, tr}^{t}\|^2 + 1/\eta_{1, tr}^{t} \)
7:   end for
8:   \( \forall tr: \gamma_{2, tr}^{t} = \eta_{1, tr}^{t} - \gamma_{1, tr}^{t} \)
9:   \( \forall tr: r_{2, tr}^{t} = (\eta_{1, tr}^{t} d_{1, tr}^{t} - \gamma_{1, tr}^{t} r_{1, tr}^{t}) / \gamma_{2, tr}^{t} \)
10:  for \( \tau_2 = 0, \ldots, \tau_{2, \text{max}} \) do
11:     \( \forall tr: C_{tr}^{t} \leftarrow (\gamma_{2, tr}^{t} I_N + H(\psi_{H}^{t})^T H(\psi_{H}^{t}))^{-1} \)
12:     \( \forall tr: d_{2, tr}^{t} \leftarrow C_{tr}^{t} (\gamma_{2, tr}^{t} r_{2, tr}^{t} + H(\psi_{H}^{t})^T y_{tr}) \)
13:     \( \forall tr: 1/\eta_{2, tr}^{t} \leftarrow \text{tr}(C_{tr}^{t}) / N \)
14:     \( C^{t} \leftarrow \sum_{tr=1}^{T_{tr}} C_{tr}^{t} \)
15:     \( H(\psi_{H}^{t+1}) = YD_{tr}^{t} (C^{t} + D_{2} D_{2}^{T})^{-1} \)
16:  end for
17: \( H(\psi_{H}^{t+1}) = H(\psi_{H}^{t}) \)
18: \( \forall tr: \gamma_{1, tr}^{t+1} = (\gamma_{2, tr}^{t} - \gamma_{2, tr}^{t}) \)
19: \( \forall tr: r_{1, tr}^{t+1} = (\eta_{2, tr}^{t} d_{2, tr}^{t} - \gamma_{2, tr}^{t} r_{2, tr}^{t}) / \gamma_{1, tr}^{t+1} \)
20: end for

IV. REMOVING PERMUTATION AMBIGUITY AND RECOVERY OF G

As discussed in the previous section the estimated channel matrices obtained via the BAdVAMP algorithm contain ambiguities, which can be classified into three main categories: permutation, scalar and phase ambiguities. It is not possible to avoid inherent scalar and phase ambiguities for the studied system model. However, by leveraging the sparsity in \( D \), the permutation ambiguity can be removed from the estimated matrices \( \hat{H} \) and \( \hat{D} \). In this section, we introduce a framework to remove the permutation ambiguity from the recovered matrices. We also discuss a low-rank matrix completion approach to recover \( G \) after removing the permutation ambiguity from the estimated matrix \( \hat{D} \). Let \( \hat{H} \) and \( \hat{D} \) represent the estimated channel matrices obtained via the BAdVAMP algorithm, then \( \hat{H} \Sigma \Gamma \) and \( \Gamma^T \Sigma^{-1} \hat{D} \) are also valid solutions for any arbitrary permutation matrix \( \Gamma \) and diagonal matrix \( \Sigma \), where \( \Sigma \) comprises of scalar and phase ambiguities.

A. Removing Permutation Ambiguity

The design of phase shifters \( S \) during the training period is crucial for removing the permutation ambiguity that is induced in the estimated channels. Each column of \( S \) is a \( K \)—sparse vector. Therefore,
we define a state matrix for the matrix $\hat{D}$ as

$$S_{l,t} \triangleq \begin{cases} 
1, & \hat{D}_{l,t} \neq 0 \\
0, & \hat{D}_{l,t} = 0 
\end{cases} \quad (24)$$

From (24), the relationship between $\bar{S}$ and $S$, can be expressed as $\bar{S} = \Gamma S$. This relationship shows that the $n^{th}$ row of the state matrix, denoted by $s_n^T$, represents the $n'$ row of $S$ given by $s_{n'}^T$. This suggests that the $(n,n')$ element of the permutation matrix $\Gamma$ must be equal to one, i.e., $\Gamma_{n,n'} = 1$. Using this fact, we can compute the $n^{th}$ row of $\Gamma$ with a non-zero value at the location $\hat{n}$, where

$$\hat{n} = \arg\max_{n'} (s_{n'} \bar{s}_n^T). \quad (25)$$

This approach yields the permutation ambiguity matrix $\Gamma$ which helps in removing the permutation ambiguity from the BAdVAMP estimated channels, such that the recovered channel matrices are given by $\hat{D} = \Gamma^T \hat{D}$ and $\hat{H} = \hat{H} \Gamma$. Note that we rely on maximization of the inner product between these two vectors in problem (25), which is sufficient for obtaining the permutation matrix.

B. Estimation of $G$ via Matrix Completion

To extract the estimate of the channel matrix $G$ from the permutation ambiguity removed matrix $\hat{D}$, similar to [14], we use a low-rank matrix completion approach. Low-rank matrix completion has been widely investigated in the literature due its application in multiple fields such as image processing, machine learning and model reduction. The main idea in the low-rank matrix completion approach is to recover a low-rank matrix from the set of linear measurements. In this study, we use the well-known normalized iterative hard thresholding (NIHT) [37] algorithm to recover $G$ from the matrix $\hat{D}$. The NIHT algorithm works on the principle of projection based gradient descent.

In a matrix completion approach, if a recoverable matrix of size $L \times N$ has a rank of $r$, then $r(L+N-r)$ measurements may be sufficient for recovery [37]. In our problem, we want to recover $G$ from the estimated matrix $\hat{D}$. Therefore, we can write the matrix completion problem as

$$\min_F \frac{1}{2} \| \hat{D} - (S \odot F) \|^2_F \quad (26)$$

$$\text{s. t.} \quad \text{rank}(F) = r_g. \quad (27)$$

where $r_g$ denotes the rank of the matrix $G$. In this study, we assume that the value of $r_g$ is available at
the AP. The NIHT based method to find the estimate of the matrix $G$ is presented in Algorithm 2. In

**Algorithm 2** NIHT based method to recover $G$

1: **Input**: $r_g$, $\tilde{D}$, $S$ and $X$
2: **Set**: $F^0 = 0, j = 0$ and $U_0$ (top $r_g$ left singular vectors of $F^0$)
3: **Repeat**
4:   Set $P_j = U_j U_j^H$
5:   Compute $\alpha_j = \frac{\|P_j S \odot (D - S \odot F_j)\|_2}{\|S \odot (P_j S \odot (D - S \odot F_j))\|_2}$
6:   Set $Q = F_j + \alpha_j S \odot (\tilde{D} - S \odot F_j)$
7:   Set $F_{j+1} = H_{r_g}(Q)$
8:   Set $U_{j+1}$ to top $r_g$ left singular vectors of $F_{j+1}$
9:   $j = j + 1$
10: **Until** $j > \text{max\_iter}$
11: **Output** $\tilde{G} = F^j(XX^H)^{-1}X$

Algorithm 2 the operation $H_{r_g}(\cdot)$ restricts the rank of the input to the $r_g$ by performing SVD on the input and keeping only $r_g$ singular values in the diagonal singular value matrix, while replacing the remaining singular values to zero. Let $\text{SVD}(Q) = U_q \Upsilon V_q^H$, then $\Upsilon_{r_g}$ is the diagonal singular value matrix, $\Upsilon$, with only the dominant $r_g$ singular values remaining and the rest equal to zero. The resulting output of the step 7 in Algorithm 2 is given by $H_{r_g}(Q) = U_q \Upsilon_{r_g} V_q$.

Now we have estimated the RIS channels and the direct channel which are $\{\tilde{H}, \tilde{G}\}$ and $\tilde{Z}$, respectively. Note that scalar and phase ambiguities are still present in the RIS estimated channels. However, the precoding strategy and RIS phase shift design considered in this study depend on the composite channel given by $\tilde{H} \Phi \tilde{G} + \tilde{Z}$. Therefore, there is no need to eliminate these ambiguities. According to [14], we have $\tilde{H} \Phi \tilde{G} + \tilde{Z} = \tilde{H}' \Phi \tilde{G}' + \tilde{Z}$, where $\tilde{H}'$ and $\tilde{G}'$ denotes the estimated channels with perfect elimination of the phase and scalar ambiguities given by a full rank diagonal matrix $\Sigma \in \mathbb{C}^{N \times N}$. Therefore, it is sufficient to have only permutation ambiguity removed from the estimated channels to design the RIS phase shifts and MIMO precoding/combining strategy. In the next section, we present our proposed RIS phase shift design at the RIS during the downlink data transmission stage.

V. OPTIMIZATION OF RIS PHASE SHIFTS

Phase shifts of all the passive elements in the RIS can play an important part in the performance of the RIS assisted MIMO systems. In particular, phase shifts of passive elements in the RIS can be adjusted according to the propagation environment such that the reflected signals add up coherently at the user with the signals arriving from other paths. This adjustment helps improve the channel gain or received
SNR of the signal at the user. For this purpose, in this section, we propose a phase shift design at the RIS that maximizes the gain of the estimated composite channel given by $\|\tilde{H}^T \Phi \tilde{G}^T + \tilde{Z}^T\|_F^2$. Due to the fixed precoding/beamforming strategy at the AP, we have selected the channel gain as an optimization objective. To find the optimal phase shifts, $\Phi^*$, we formulate an optimization problem which maximizes the channel gain at the user, which can be expressed as

$$\max_{\Phi} \|\tilde{H}^T \Phi \tilde{G}^T + \tilde{Z}^T\|_F^2$$

s. t. $\Phi_{l,l} = e^{j\phi_{l,l}}$, $l = 1, \ldots, L$. (28)

Note that the optimization of phase shifts is performed at the AP, which is capable of managing the phase shifts at the RIS. The problem (28) is non-convex optimization problem due to coupled phases. Therefore, in order to solve the optimization problem, we rely on its approximation. Let $A = \tilde{H}^T \Phi \tilde{G}^T + \tilde{Z}^T$, we can write $\|\tilde{H}^T \Phi \tilde{G}^T + \tilde{Z}^T\|_F^2$ in (28) as

$$\|A\|_F^2 = \sum_{i=1}^{M} \sum_{j=1}^{N} |a_{i,j}^{(l)}|^2,$$

where

$$|a_{i,j}^{(l)}|^2 = \left| \sum_{l=1}^{L} \tilde{H}_{i,l}^T \Phi_{l,l} \tilde{G}_{l,j}^T + \tilde{Z}_{i,j}^T \right|^2$$

(31)

$$= \left( \sum_{l=1}^{L} \tilde{H}_{i,l}^T \Phi_{l,l} \tilde{G}_{l,j}^T + \tilde{Z}_{i,j}^T \right)^* \left( \sum_{l=1}^{L} \tilde{H}_{i,l}^T \Phi_{l,l} \tilde{G}_{l,j}^T + \tilde{Z}_{i,j}^T \right)$$

(32)

$$= \left( \sum_{l=1}^{L} \tilde{H}_{i,l}^T \Phi_{l,l} \tilde{G}_{l,j}^T \right)^* \left( \sum_{l=1}^{L} \tilde{H}_{i,l}^T \Phi_{l,l} \tilde{G}_{l,j}^T \right) + (\tilde{Z}_{i,j}^T)^* \tilde{Z}_{i,j}^T +$$

$$2\text{Re} \left\{ (\tilde{Z}_{i,j}^T)^* \left( \sum_{l=1}^{L} \tilde{H}_{i,l}^T \Phi_{l,l} \tilde{G}_{l,j}^T \right) \right\}. $$

(33)

The first term in (33) can be expressed as

$$\left( \sum_{l=1}^{L} \tilde{H}_{i,l}^T \Phi_{l,l} \tilde{G}_{l,j}^T \right)^* \left( \sum_{l=1}^{L} \tilde{H}_{i,l}^T \Phi_{l,l} \tilde{G}_{l,j}^T \right) =$$

$$\sum_{l=1}^{L} |\tilde{H}_{i,l}^T|^{2} |\Phi_{l,l}|^{2} |\tilde{G}_{l,j}^T|^{2} + \sum_{l=1}^{L} \sum_{q \neq l}^{L} (\tilde{H}_{i,l}^T)^* \Phi_{l,l}^* (\tilde{G}_{l,j}^T)^* \tilde{H}_{i,q} \Phi_{q,q} \tilde{G}_{q,j}^T).$$

(34)

We note that the first term in (34) has no phase information because of the fact that $|e^{j\phi_l}|^2 = 1$. On the other hand, the second term which represents the cross terms in (34) has negligible value. Thus, we ignore
the first term in \((33)\) and other terms which are independent of phases and we can approximate \((34)\) as

\[
\left| a_{i,j}^{(l)} \right|^2 \approx 2\text{Re} \left\{ (\tilde{Z}_{i,j}^T)^* \left( \sum_{l=1}^{L} \tilde{H}_{i,l}^T \Phi_{l,l} \tilde{G}_{l,j}^T \right) \right\}.
\]

\((35)\)

We note that the rate of change in \((28)\) and its approximate version \((35)\) with respect to the phases 
\([e^{j\phi_1}, e^{j\phi_1}, \ldots, e^{j\phi_L}]\) is similar. Therefore, to obtain the sub-optimal phase shift matrix in the problem \((28)\), we use the approximation of \((30)\) given by \((35)\), such that

\[
\|A\|^2_F = \sum_{i=1}^{M} \sum_{j=1}^{N} 2\text{Re} \left\{ (\tilde{Z}_{i,j}^T)^* \left( \sum_{l=1}^{L} \tilde{H}_{i,l}^T \Phi_{l,l} \tilde{G}_{l,j}^T \right) \right\}
\]

\((36)\)

\[
= \sum_{i=1}^{L} \sum_{l=1}^{M} \sum_{j=1}^{N} 2\text{Re} \left\{ (\tilde{Z}_{i,j}^T)^* \tilde{H}_{i,l}^T \Phi_{l,l} \tilde{G}_{l,j}^T \right\}.
\]

\((37)\)

From \((37)\), we can obtain the phase of the \(l\)th element in the RIS, given by \(\phi_i^* = \Phi_{l,l}^*\), that maximizes the channel gain in the problem \((28)\) as

\[
\phi_i^* = \tan^{-1}\left( \frac{\text{Im} \left\{ \sum_{i=1}^{M} \sum_{j=1}^{N} (\tilde{Z}_{i,j}^T)^* \tilde{H}_{i,l}^T \Phi_{l,l} \tilde{G}_{l,j}^T \right\}}{\text{Re} \left\{ \sum_{i=1}^{M} \sum_{j=1}^{N} (\tilde{Z}_{i,j}^T)^* \tilde{H}_{i,l}^T \Phi_{l,l} \tilde{G}_{l,j}^T \right\}} \right).
\]

\((38)\)

In this study, we consider the fixed transmit precoding and receive combining scheme as discussed in Section III. Therefore, after obtaining the RIS phase shift for each passive element from \((38)\), we compute the eigenmode based precoding vectors using the composite channel matrix given by \(\tilde{H}^T \Phi^* \tilde{G}^T + \tilde{Z}^T\). It is also possible to perform a joint downlink precoding and RIS phases optimization, if fixed precoding is not used. However, the investigation of the joint optimization problem is out of the scope of this study.

**VI. Numerical Results**

In this section, we evaluate the performance of our proposed channel estimation technique based on the BAdVAMP algorithm for the RIS assisted MIMO system. First, the performance of the proposed BAdVAMP based channel estimation algorithm is compared with the BiGAMP algorithm \([14]\). Later, we present the capacity performance of the RIS assisted MIMO communication system using the proposed channel estimation and optimized RIS phase shift design.

**A. BAdVAMP versus BiGAMP channel estimation**

In this subsection, various experiments are carried out to compare the performance of the proposed BAdVAMP based channel estimation with the BiGAMP channel estimation technique \([14]\). The perfor-
mance of the estimated RIS channels is measured using the normalized minimum squared error (NMSE) metric, for the estimated matrix $\hat{H}$ which is given by

$$\text{NMSE} \left( \hat{H} \right) = \frac{\|H - \hat{H} B\|^2_F}{\|H\|^2_F}, \quad (39)$$

where $B$ denotes the generalized permutation matrix which perfectly removes the permutation, scalar and phase ambiguities. This perfect ambiguity elimination is only considered in this subsection for comparison purposes between BAdVAMP and BiGAMP schemes.

For the first set of simulations, we set the total number of transmit antennas and receive antennas to $M = N = 64$, the number of passive elements in RIS is $L = 70$, the length of the training sequences for RIS channels and direct channel are $T_r = 600$ and $T_c = 30$, respectively. The SNR is defined as $(1/\sigma_n^2)$, where $\sigma_n^2$ is the noise variance. Throughout this section, we have fixed the value of $T_c$ to 30.

For initialization of the BAdVAMP algorithm, we draw $r_{1,tr}^0$ and $\psi_H^0$ from the i.i.d. $\mathcal{CN}(0,10)$ and $\mathcal{CN}(0,1)$ distributions, respectively. The initial value of $\gamma_{1,tr}^0$ is set to $10^{-3}$ and it is assumed that the statistical distribution of matrix $D$, i.e., $\theta_d$, is known at the AP. Inner EM iterations are set to $\tau_{1,\text{max}} = 1$ and $\tau_{2,\text{max}} = 0$. The maximum number of iterations, $T_{\text{max}}$, for both BAdVAMP and BiGAMP schemes is fixed to 3500, where the maximum number of restarts in BAdVAMP is set to 50. In Fig. 2, we plot NMSE
results of the estimated channel matrix for $H$ against various values of SNR for both BAdVAMP and BiGAMP based channel estimation schemes. Here, the channel matrix $H$ is generated with two different condition numbers, which are $\kappa(H) = 30$ and $\kappa(H) = 100$. Similarly, two different cases have been considered for the generation of the channel matrix $G$, where in the first case the the rank of the matrix is fixed to $\text{rank}(G) = 4$, while in the second case the rank is set to $\text{rank}(G) = 7$. Note that both the matrices $H$ and $G$ follow the $CN(0, 1)$ distribution.

It is evident from Fig. 2 that the estimated channel $\hat{H}$ for the proposed BAdVAMP based algorithm is superior than the BiGAMP algorithm [14]. More specifically, when the channel matrix $H$ is highly ill-conditioned, i.e., $\kappa(H) = 100$, it is seen that the performance gap between the two algorithms widens. This is because the VAMP part of the BAdVAMP algorithm is robust to ill-conditioned matrices as compared to AMP based BiGAMP algorithm. Similarly, for the $\kappa(H) = 30$ case, our proposed channel estimation performs better than the BiGAMP algorithm. The channel estimation performance with the proposed BAdVAMP scheme improves in the case of $\text{rank}(G) = 7$ compared to the case when $\text{rank}(G) = 4$. This is due to the fact that BAdVAMP performs well when rank of the matrix is high.

It can be seen from Fig. 2 that at SNR = 10 dB for the $\kappa(H) = 30$, the BAdVAMP algorithm yields NMSE of -38 dB, whereas, NMSE of -25 dB is obtained with the BiGAMP algorithm, therefore, having a performance gap of almost 13 dB. Similarly, at SNR = 5 dB and $\kappa(H) = 100$, the performance gap between the two schemes is about 20 dB. This shows the effectiveness of BAdVAMP algorithm in both ill-conditioned and well conditioned matrices. The performance of BAdVAMP starts to degrade gradually as the SNR increases. This trend is because we have fixed the maximum number of iterations to 3500, whereas the BAdVAMP algorithm requires more iterations to converge at higher SNRs.

A similar trend can be seen from Fig. 3 for the recovery of the channel matrix $G$, where the same simulation parameters as in Fig. 2 are used. The recovery of $G$ is better with the BAdVAMP algorithm as compared to the BiGAMP algorithm. For $\kappa(H) = 100$, the performance gap between the BAdVAMP and the BiGAMP increases, showing a superior performance with the BAdVAMP in the highly ill-conditioned scenario. On the other hand, the performance of both the schemes is similar for the well-conditioned case of $\kappa(H) = 30$. This shows the robustness of the BAdVAMP based channel estimation algorithm over its counterpart BiGAMP in both ill-conditioned and well-conditioned scenarios.

Table I shows the channel estimation performance of BAdVAMP and BiGAMP schemes for various values of training lengths with a fixed SNR of 10 dB, $\kappa(H) = 80$ and $\text{rank}(G) = 7$. All other simulation
parameters are set the same to the ones used for Fig. 2 and Fig. 3. The length of training sequence is varied from $T_r = 200$ to $T_r = 1000$. It is seen that the quality of the channel estimation with the BAdVAMP algorithm improves with increasing $T_r$ for both $H$ and $G$. On the other hand, the performance of the BiGAMP algorithm also improves with increasing $T_r$ but with a slower rate. This trend is due to the fact that $\kappa(H) = 80$ and it is well-known that BiGAMP scheme suffers when the channel matrix is ill-conditioned. From Table I it can be seen that training length plays an important role in determining the performance of channel estimation. Therefore, for a better channel estimation performance the training length should not be kept too low. This requirement of a higher training overhead in approximate message passing based schemes is also noticed in earlier studies [38]. The trade-off between the training overhead and the quality of the channel estimation process is also evident in Table I.
B. MIMO performance with BAdVAMP and optimized RIS phase shifts

For the second set of simulations, we evaluate the capacity performance of the RIS assisted MIMO system using the proposed channel estimation technique and RIS phase shift design. The capacity results of three different configurations are plotted in Fig. 4 where the cases considered are: i) \((M, L, N) = (64, 70, 64)\), ii) \((M, L, N) = (32, 40, 16)\) and iii) \((M, L, N) = (16, 20, 8)\). For the case i) \(T_r = 500\) and \(\text{rank}(G) = 10\), case ii) \(T_r = 300\) and \(\text{rank}(G) = 8\) and case iii) \(T_r = 100\) and \(\text{rank}(G) = 3\). Fig. 4 shows the achievable capacity (given by (5)) results with \(N_s = 1\) for the proposed channel estimation and optimal phase shift scheme. We compare it with the scenario where perfect CSI is available at the AP and RIS phase shift design is based on exhaustive search in problem (28). For a comparison, we also include a scenario where random precoding/combining and random RIS phase shifts are used. Finally, we also plot the capacity results with BiGAMP based channel estimation [14] with our proposed RIS phase shift design. Note that the precoding and combining strategy for all the scenarios are based on the eigenmode precoding scheme as explained in Section II except in the case of random beamforming, where the precoding matrix is drawn randomly from an i.i.d. \(\mathcal{CN}(0, 1)\) distribution and then normalized.

It can be seen from Fig. 4 for the case (i), i.e., 64, 70, 64, that the proposed BAdVAMP scheme has better capacity performance at low SNR values than the scheme where BiGAMP based channel estimation
Fig. 5: Beamforming gain versus SNR results.

is used. For example, at SNR = 0 dB, the capacity gap between the proposed scheme and BiGAMP scheme is nearly equal to 5 bps/Hz. More importantly, it can be seen that the capacity performance with both schemes matches the perfect CSI case at higher SNR values. On the other hand, we observe that the results with the random beamforming case gives the worst capacity performance. Similar performance trend is seen for case (ii) and case (iii). In Fig. 5, we present the beamforming gain versus SNR results for the same cases and network parameters as in Fig. 4. The beamforming gain for a single data stream i.e., $N_s = 1$, is given by

$$BF\ Gain = 10 \log_{10} \left( \| (\hat{H}^T \Phi^* G^T + Z^T) w_1 \|^2 \right),$$

(40)

where for the proposed scheme, $w_1$ denotes the beamforming vector corresponding to the largest singular value of the matrix $\hat{H}^T \Phi^* G^T + Z^T$. For the perfect case $w_1$ denotes the the beamforming vector corresponding to the largest singular value of the matrix $H \Phi_{opt} G + Z$, where $\phi_{opt}$ is the phase shift matrix obtain via an exhaustive search in (28). From Fig. 5, it can be seen that the beamforming gain with the proposed BAAdVAMP scheme is higher than the BiGAMP in the low SNR regime, however, the performance of both the schemes approaches near to the performance of the perfect CSI case at high SNRs. The performance trend is same in all the three cases considered in Fig. 5. However, among the
different cases, we observe that the capacity performance increases by having higher number of transmit antennas $M$, receive antennas $N$ and RIS elements $L$.

In Fig. 6, we plot the capacity versus training length results for the RIS assisted MIMO communication system using the channel estimation technique and optimized phase shift design with $(M, L, N) = (64, 80, 64)$ and SNR = 10 dB. In Fig. 6 we plot two cases with $N_s = 1$ and $N_s = 4$. Similar to the previous figures, we compare the capacity results with perfect CSI and random beamforming. However, here, we also include a case where no RIS is present in the surrounding area. It can be seen in Fig. 6 that the capacity results increase with the increase in the training length, however, after $T_r = 400$ the capacity performance converges and does not increases with $T_r$. This suggests that the training length of 400 is the optimal choice for the given network settings. Moreover, similar to the conventional MIMO case, it is noticed from Fig. 6 that the capacity improves by having higher data streams ($N_s = 4$) compared to the case where only a dominant stream ($N_s = 1$) is used. The proposed scheme performs better than the case where no RIS terminal is available to assist the communication between transmitter and receiver. In fact, the capacity performance with no RIS is close to the random beamforming (with RIS) case, which supports the idea of deploying RIS. Fig. 7 shows the capacity results achieved at SNR=10 dB, by varying the number of passive elements, $L$, in RIS from 40 to 140. Here also, two cases are considered where
the parameters are set to $M = N = 64$ for the first case, and $M = N = 32$ for the second case. It can be observed from Fig. 7 that the capacity improves by increasing the number of passive elements in the RIS. Also, it is evident that the capacity performance with RIS is better than the capacity with no RIS. In fact, for larger value of $L$, the capacity with no RIS falls below the capacity with random beamforming.

**VII. CONCLUSIONS**

In this paper, we have proposed a two stage channel estimation method for the RIS assisted MIMO communication system, where the direct channel between the AP and user is estimated in the first stage and the BAdVAMP algorithm is used to estimate RIS channels in the second stage. In addition to that, we also presented a phase shift design at the RIS which approximately maximizes the channel gain at the user. Through numerical simulations, we show the effectiveness of the proposed channel estimation method which outperforms the BiGAMP based channel estimation method, especially for the case where the channel matrices are ill-conditioned.

There are some key observations in this study which needs to be highlighted. First of all, we can see that the BAdVAMP algorithm outperforms the BiGAMP algorithm in both ill-conditioned and well-conditioned channel matrices. We also observed that for higher SNRs, the BAdVAMP algorithm requires more iterations to converge. Secondly, it is noticed that the rank of $G$ plays a key part in the performance of
the BAdVAMP scheme, i.e., higher the rank of the matrix $G$, better will be the overall channel estimation performance. Thirdly, we noticed that the value of $T_r$ should be large for the better channel estimation performance, but the performance converges at high $T_r$ values. This suggests the need to optimize the training duration $T_r$ which gives the best tradeoff between the estimation quality and training overhead. Lastly, we observed that RIS is beneficial for the MIMO system as it improves the capacity performance compared to the conventional MIMO system with no RIS.

For future work, it will be interesting to investigate the joint MIMO precoding/combining and RIS phase shift optimization. It will also be useful to optimize the training length such that a balance is achieved between the channel estimation quality and training overhead. The channel estimation strategy should also be investigated in the scenario where multiple reflecting surfaces are present in the surrounding areas.

REFERENCES

[1] S. Cisco, “Cisco visual networking index: Global mobile data traffic forecast update, 2017–2022,” *Cisco Public Information: San Jose, CA, USA*, 2017.

[2] S. Zhang, Q. Wu, S. Xu, and G. Y. Li, “Fundamental green tradeoffs: Progresses, challenges, and impacts on 5G networks,” *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 33–56, Jul 2017.

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

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

[5] M. Di Renzo, M. Debbah, D. Phan-Huy, A. Zappone et al., “Smart radio environments empowered by reconfigurable AI meta-surfaces: an idea whose time has come,” *EURASIP J. Wireless commun. and networking*, vol. 129, no. 20, May 2019.

[6] Y. Tang, G. Ma, H. Xie, J. Xu, and X. Han, “Joint transmit and reflective beamforming design for IRS-assisted multiuser MISO SWIPT systems,” [Online] Available https://arxiv.org/abs/1910.07156, 2019.

[7] Q. Wu and R. Zhang, “Joint active and passive beamforming optimization for intelligent reflecting surface assisted SWIPT under QoS constraints,” [Online] Available https://arxiv.org/abs/1910.06220, 2019.

[8] S. Li, B. Duo, X. Yuan, Y.-C. Liang, M. Di Renzo et al., “Reconfigurable intelligent surface assisted UAV communication: Joint trajectory design and passive beamforming,” [Online] Available https://arxiv.org/abs/1908.04082, 2019.

[9] X. Mu, Y. Liu, L. Guo, J. Lin, and N. Al-Dhahir, “Exploiting intelligent reflecting surfaces in multi-antenna aided NOMA systems,” [Online] Available https://arxiv.org/abs/1910.13636, 2019.

[10] M. Fu, Y. Zhou, and Y. Shi, “Reconfigurable intelligent surface empowered downlink non-orthogonal multiple access,” [Online] Available https://arxiv.org/abs/1910.07361, 2019.

[11] D. Mishra and H. Johansson, “Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer,” in *Proc. IEEE ICASSP*, Brighton, UK, May 2019, pp. 4659–4663.

[12] C. You, B. Zheng, and R. Zhang, “Intelligent reflecting surface with discrete phase shifts: Channel estimation and passive beamforming,” [Online] Available https://arxiv.org/abs/1911.03916, 2019.
[13] B. Zheng and R. Zhang, “Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization,” [Online] Available https://arxiv.org/abs/1909.03272, 2019.

[14] Z. He and X. Yuan, “Cascaded channel estimation for large intelligent metasurface assisted massive MIMO,” IEEE Wireless Commun. Lett., to be Published, 2019.

[15] Y. Han, W. Tang, S. Jin, C. Wen, and X. Ma, “Large intelligent surface-assisted wireless communication exploiting statistical CSI,” IEEE Trans. Veh. Technol., vol. 68, no. 8, pp. 8238–8242, Aug 2019.

[16] S. Abeywickrama, R. Zhang, and C. Yuen, “Intelligent reflecting surface: Practical phase shift model and beamforming optimization,” [Online] Available https://arxiv.org/abs/1907.06002, 2019.

[17] J. He, H. Wymeersch, T. Sanguanpuak, O. Silvén, and M. Juntti, “Adaptive beamforming design for mmwave RIS-aided joint localization and communication,” [Online] Available https://arxiv.org/abs/1911.02813, 2019.

[18] P. Wang, J. Fang, and H. Li, “Joint beamforming for intelligent reflecting surface-assisted millimeter wave communications,” [Online] Available https://arxiv.org/abs/1910.08541, 2019.

[19] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, “Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts,” [Online] Available https://arxiv.org/abs/1910.14328, 2019.

[20] Z. Wang, L. Liu, and S. Cui, “Channel estimation for intelligent reflecting surface assisted multiuser communications,” [Online] Available https://arxiv.org/abs/1911.03084, 2019.

[21] B. Ning, Z. Chen, W. Chen, and Y. Du, “Channel estimation and transmission for intelligent reflecting surface assisted THz communications,” [Online] Available https://arxiv.org/abs/1911.04719, 2019.

[22] M. Jung, W. Saad, Y. Jang, G. Kong, and S. Choi, “Reliability analysis of large intelligent surfaces (LISs): Rate distribution and outage probability,” IEEE Wireless Commun. Lett., to be published 2019.

[23] V. Vakilian, H. Mehrpouyan, and Y. Hua, “High rate space-time codes for millimeter-wave systems with reconfigurable antennas,” in Proc. IEEE (WCNC), Istanbul, Turkey, March 2015, pp. 591–596.

[24] E. Torkildson, C. Sheldon, U. Madhow, and M. Rodwell, “Millimeter-wave spatial multiplexing in an indoor environment,” in Proc. IEEE Globecom Workshops, Honolulu, Hawaii, USA, Nov 2009, pp. 1–6.

[25] S. Sarkar, A. K. Fletcher, S. Rangan, and P. Schniter, “Bilinear recovery using adaptive Vector-AMP,” IEEE Trans. Signal Process., vol. 67, no. 13, pp. 3383–3396, July 2019.

[26] A. Paulraj, A. P. Rohit, R. Nabar, and D. Gore, Introduction to Space-Time Wireless Communications. Cambridge University Press, 2003.

[27] W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. Di Renzo, Y. Zeng, S. Jin, Q. Cheng, and T. J. Cui, “Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement,” [Online] Available https://arxiv.org/abs/1911.05326, 2019.

[28] Ö. Özdogan, E. Björnson, and E. G. Larsson, “Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling,” [Online] Available https://arxiv.org/abs/1911.03359, 2019.

[29] S. Hu, F. Rusek, and O. Edfors, “Beyond massive MIMO: The potential of data transmission with large intelligent surfaces,” IEEE Trans. Signal Process., vol. 66, no. 10, pp. 2746–2758, 2018.

[30] M. Jung, W. Saad, Y. Jang, G. Kong, and S. Choi, “Performance analysis of large intelligence surfaces (LISs): Asymptotic data rate and channel hardening effects,” [Online] Available https://arxiv.org/abs/1810.05667, 2018.

[31] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Massive MU-MIMO downlink TDD systems with linear precoding and downlink pilots,” in Proc. Allerton Conf. Commun., Control, and Comput., Monticello, IL, USA, Oct 2013, pp. 293–298.

[32] S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory. Englewood Cliffs, NJ: Prentice Hall, 1993.
[33] J. T. Parker, P. Schniter, and V. Cevher, “Bilinear generalized approximate message passing—Part I: derivation,” IEEE Trans. Signal Process., vol. 62, no. 22, pp. 5839–5853, Sep 2014.

[34] S. Rangan, P. Schniter, and A. K. Fletcher, “Vector approximate message passing,” IEEE Trans. Inf. Theory, May 2019.

[35] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” Journal of the Royal Statistical Society: Series B (Methodological), vol. 39, no. 1, pp. 1–22, Dec 1977.

[36] P. Schniter, S. Rangan, and A. K. Fletcher, “Vector approximate message passing for the generalized linear model,” in Proc. Asilomar Conf. on Sig., Sys. and Comp., Pacific Grove, CA, USA, Nov 2016, pp. 1525–1529.

[37] J. Tanner and K. Wei, “Normalized iterative hard thresholding for matrix completion,” SIAM Journal on Scientific Computing, vol. 35, no. 5, pp. S104–S125, 2013.

[38] J. Mo, P. Schniter, and R. W. Heath, “Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs,” IEEE Trans. Signal Process., vol. 66, no. 5, pp. 1141–1154, March 2018.