# An optimal scheme of two-timescale channel estimation framework for RIS-assisted wireless communications

Kai She[1*] and Yichuang Sun[2]

[1*]School of Physics and Electronic Science and Key Laboratory of Intelligent Sensors and Advanced Sensing Materials of Hunan Province, Hunan University of Science and Technology, Taoyuan Road, Xiangtan, 411201, Hunan Province, China.
[2]School of Engineering and Computer Science, University of Hertfordshire, Hatfield AL10 9AB, Hertfordshire, U.K.

*Corresponding author(s). E-mail(s): shekai@126.com;
Contributing authors: Y.Sun@herts.ac.uk;

**Abstract**

The high pilot overhead associated with channel estimation (CE) is a major challenge in reconfigurable intelligent surfaces (RIS)-assisted wireless systems. A two-timescale CE framework has been proven to be able to significantly reduce the pilot overhead. This framework employs a coordinate descent-based iterative refinement (CDIR) optimization algorithm based on full-duplex base station (BS) assumptions, yet it suffers from high computational complexity. In this paper, we reconsider the optimization algorithm and two working modes along with full-duplex BS. Firstly, the CE equations are reformulated as a second-order nonlinear rank-one optimization problem after data matrix completion. Then, the received complex-valued pilot matrix is partitioned and represented in real form, and an optimization method based on principal eigenvector approximation is proposed. Lastly, a semi-closed-form expression is used to construct the relation between the received pilots and the channel parameters. Discussions and simulation results demonstrate that the proposed method exhibits lower computational complexity, lower pilot overhead, and acceptable estimation accuracy compared to various benchmark schemes. Especially, when the signal-to-noise ratio of the received pilots is less than 20 dB for the case of two pilot transmitting antenna, the estimation accuracy surpasses that of the referenced CDIR optimization algorithm.

**Keywords:** Reconfigurable intelligent surface, Channel estimation, Two-timescale framework, Principle eigenvector approximation.

# 1 Introduction

Reconfigurable intelligent surface (RIS) is a novel technology for fifth generation (5G) beyond. It is expected to improve the receiver signal-to-noise ratio (SNR), the band and energy efficiency, and the network capacity by way of intelligently modifying the wireless propagation environment [1, 2]. RIS, typically composed of a passive surface array, intelligently adjusts the reflection coefficient of each reflection unit and the overall beam direction to combine several in-phase multipath signals at the receiver. To realize the expectations, it is undoubtedly important to develop an efficient and accurate channel estimation (CE) method for RIS-assisted wireless systems, especially for the massive multi-input-multi-output (MIMO) orthogonal frequency division multiplexing (OFDM) systems to be compatible with 5G beyond [3, 4]. Different from current CE methods, the CE framework for RIS-assisted wireless systems has to deal with the newly added cascaded channel among base station (BS)-RIS-Users with a low pilot overhead [5, 6]. The newly added cascaded channels, originate from the BS, then are reflected by the RIS, and finally reach different users. This is challenging because the RIS is passive and does not have the ability of signal processing. Hence, there is much literature on it.

Initially, classical CE methods were applied. The authors in [5, 6] considered the available fundamental algorithms and provided a general CE framework based on the channel reciprocity between the BS and users. Reference [7] proposed an minimum variance unbiased (MVU) CE method by optimizing the reflection coefficients of RIS elements to a discrete Fourier transform (DFT) matrix, thereby improving estimation accuracy and reducing computational complexity. Then, reference [8] addressed the non-ideal relationship between the phase and amplitude of the RIS reflection coefficients and proposed a CE method based on the discrete phase model. Recently, machine learning methods were also utilized to further enhance the CE accuracy. For instances, the underestimated channel was treated as an image on the time-frequency plane, and convolutional deep neural networks (DNNs) with the ability of super-resolution image reconstruction were used to achieve higher CE performance [9, 10]. Additionally, reference [11] proposed an untrained DNN to improve estimation accuracy, while simultaneously accounting for the nonlinear effects of RIS hardware.

However, another more critical aspect of CE research of RIS-assisted wireless systems is reducing pilot overhead, particularly when RIS is applied to massive MIMO systems, with huge underestimated channel parameters. Reference [12] demonstrated the significance of reducing pilot overhead as the number of BS antennas and RIS reflective elements increases, especially for multi-user RIS assisted systems [13]. Therefore, semi-blind CE methods were explored to mitigate this pilot overhead [14]. Reference [15, 16] used compressed CE methods based on sparse channel assumptions and orthogonal matching pursuit (OMP) algorithms. Reference [17] proposed a semi-active RIS model to overcome the problem. Additionally, reference [18] leveraged user clustering to further reduce pilot overhead.

Most papers estimated the channel among BS-RIS-User as a whole cascaded channel. Different from these papers, the authors in [19] proposed a novel method, called the *two-timescale dual link CE framework*, which estimates the BS-RIS and RIS-Users channels separately. It estimates the mobile RIS-User channel with a high frequency and the quasi-static BS-RIS channel with a low frequency. Then the pilot overhead can be significantly reduced. This framework has received wide attention. For instance, reference [17] studied a novel double-RIS-aided system and proposed a Bayesian channel estimation (CE) method. Reference [20] investigated a beam training method to achieve the maximum beamforming gain. These works are all based on the two-timescale CE framework. References [17, 21–24] exploited the two-timescale property of the cascaded channel and proposed various CE methods based on different system models and assumptions. Reference [17] proposed a separate CE method, similar to the two-timescale framework, but it considered a hybrid active and passive RIS system and sparse channels. Reference [21] proposed a novel two-stage CE strategy to reduce pilot overhead for RIS-assisted millimeter wave systems. The main difference is that in Stage One, it estimates the BS-RIS channels by utilizing correlation factors between different paths of the RIS-BS channel when all users send pilots to the BS instead of relying on the BS full-duplex assumption. Reference [22] decomposed the RIS-reflected channels into three major components: static BS-RIS angles, quasi-static RIS-UE angles, and time-varying BS-RIS-UE path gains, and then estimated them on different timescales. This method imposes stricter limitations on the time-varying characteristics of the propagation channels and is therefore only applicable to more specific scenarios. Reference [23] developed a two-stage adaptive CE scheme without making assumptions about specific time variance properties. It incorporated a check process for the initial channel status and used it in the second CE stage. However, the CE accuracy of the check process requires further improvement. Recently, reference [24] also analyzed the characteristics of the cascaded channel and proposed a method that combines amplitude correlation with existing structured sparsity.

The key to the two-timescale CE framework is to estimate the quasi-static backscattered cascaded channel between BS-RIS accurately and efficiently. Reference [19] adopts an iterative optimization method based on gradient. The computation cost at the BS is high, especially for the massive MIMO-OFDM system. Moreover, the description of the algorithm is complicated, and it is hard to manifest the intrinsic relationship between the channel parameters and the received pilots. In this paper, we propose an optimal CE scheme to reduce the computation cost of this two-timescale CE framework. The contributions and novelties are summarized as follows.

- We mainly focus on the CE framework of the backscattered cascaded channel between BS and RIS for the reason that the CE methods of BS-Users and RIS-Users channels have no difference with the CE methods in conventional wireless systems.
- We reconsider the assumption of full-duplex BS, and classify the working modes to monostatic and mixed modes. Different from the equation (14) in [19], we represent the optimization problem as a new form after simple matrix completion. It facilitates obtaining a semi-closed-form expression.
- We propose a CE method based on principle eigenvector approximation (PEA). Compared with the iterative optimization method, named coordinate descent-based iterative refinement (CDIR) in [19], our method has lower computational complexity, a semi-closed-form expression, and a comparable estimation accuracy while reducing the pilot overhead simultaneously.

Notations: Lower-case and upper-case boldface letters $\mathbf{a}$ and $\mathbf{A}$ denote a vector and a matrix, respectively; $\mathbf{A}^T$, $\mathbf{A}^*$, $\mathbf{A}^H$ and $\mathrm{Tr}(\mathbf{A})$ denote the transpose, conjugate, conjugate transpose, and trace of matrix $\mathbf{A}$, respectively. $\mathbf{0}_{n\times n}$ and $\mathbf{I}_n$ respectively represent $n\times n$ zero and identity matrices, and $j=\sqrt{-1}$. $\mathbf{diag}\left(\mathbf{a}\right)$ denotes the diagonal matrix with the vector $\mathbf{a}$ on its diagonal. Lastly, $\mathbb{R}$ and $\mathbb{C}$ respectively denote the real and complex number sets, $\mathcal{CN}\left(\boldsymbol{\mu},\mathbf{C}\right)$ denotes complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$.

# 2 The two-timescale CE framework

## 2.1 System model and assumptions

The typical system model of a RIS-assisted wireless MIMO system is shown in Fig.1. The BS is a full-duplex BS, and has $M$ antennas. RIS has $N$ reflection elements. With the assistance of RIS, the BS communicates with $K$ single-antenna mobile users. For each subcarrier frequency within every OFDM symbol, which is one slot in the time-frequency plane, the received signal of BS $\mathbf{y} \in \mathbb{C}^{M\times 1}$ will be the sum of the direct $K$ user signals and the RIS reflected user signals, and can be expressed as the system model equation [5]

$$\mathbf{y} = \sum_{k=1}^{K} (\mathbf{h}_{\mathrm{D},k} + \mathbf{G}\mathrm{diag}\left(\boldsymbol{\phi}\right)\mathbf{h}_{\mathrm{A},k})s_k + \mathbf{w}, \tag{1}$$

where $\mathbf{h}_{\mathrm{D},k} \in \mathbb{C}^{M\times 1}$ is the direct propagation channel between the $k$-th user and BS, $\mathbf{G} \in \mathbb{C}^{M\times N}$ is the channel between RIS and BS, $\mathbf{h}_{\mathrm{A},k} \in \mathbb{C}^{N\times 1}$ is the channel between the $k$-th user and RIS, $s_k \in \mathbb{C}$ is the pilot symbol sent by the $k$-th user, $\boldsymbol{\phi} = [\phi_1, \phi_2, \cdots, \phi_N]^T$ is a vector composed of the reflection coefficients of RIS, and $\phi_n \in \mathbb{C}$ is the reflection coefficient of the $n$-th element of RIS, $\mathbf{w} \in \mathbb{C}^{M\times 1}$ is the additive white Gaussian noise (AWGN) at BS. Reflection coefficient $\phi_n$ can be expressed as $\phi_n = \rho_n e^{j\theta_n}$, and $\rho_n \in [0,1]$ is the amplitude, $\theta_n \in [0, 2\pi]$ is the phase.

Note that RIS-assisted wireless systems, equation (1), have two new channels, $\mathbf{G}$ and $\mathbf{h}_{\mathrm{A},k}$. The RIS does not have the capability of signal processing due to the necessity of low cost and low complexity of system implementation. So most papers define the cascaded channel among BS-RIS-Users as $\mathbf{C}_k \triangleq \mathbf{G}\mathbf{diag}\left(\mathbf{h}_{\mathrm{A},k}\right)$, and estimate it as one channel.

As most references do, this paper assumes the RIS-assisted wireless system is a time division multiplexing system, and the uplink and downlink channels are reciprocal.

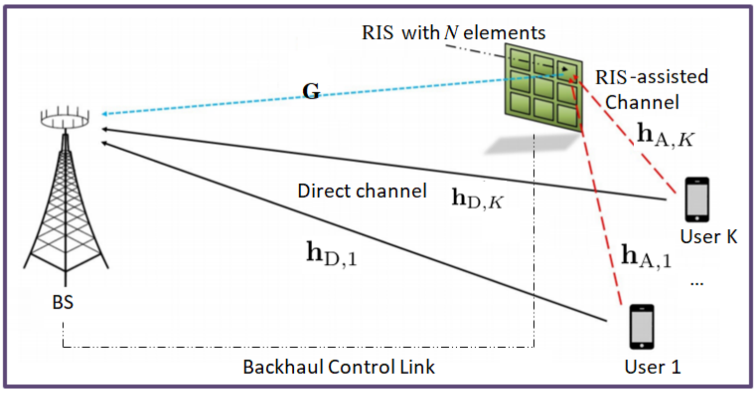


**Fig. 1** System model.

The $k$-th user will transmit the pilots, and the BS will estimate the channels $\mathbf{C}_k$ and $\mathbf{h}_{\mathrm{A},k}$ using the classical algorithms in MIMO OFDM system [5, 25, 26]. This paper also assumes other necessary parts, the synchronization and cyclic prefix, work well. The MIMO channel between BS and RIS is modeled as a backscattered cascaded Rician fading channel, which is used in the deduction of the optimal algorithm and simulations.

## 2.2 The protocol of CE

In the two-timescale CE framework, the BS and the RIS are mostly placed in fixed positions, and the BS-RIS channel $\mathbf{G}$ is quasi-static. We only need to estimate $\mathbf{G}$ in a large-timescale. The RIS-User channel and the BS-User channel are time-varying due to the mobility of the users, so we need to estimate them in a small-timescale. We will obtain $\mathbf{C}_k$ using an unchanged high-dimensional $\mathbf{G}$ and a changed low-dimensional $\mathbf{h}_{\mathrm{A},k}$ during the two-timescale pilot CE frame. Therefore, the pilot overhead can be significantly reduced.

Fig. 2 shows the pilot transmission frame structure. In the dual-link large-timescale CE stage, the BS transmits pilots to the RIS via the downlink channel using a selected antenna. Simultaneously, the RIS reflects pilots to the BS via the uplink channel using a set of pre-designed reflection coefficients. Concurrently, the BS receives pilots using the remaining antennas (or includes the transmitting antenna itself via a radio frequency (RF) circulator). After self-interference mitigation, the BS-RIS channel is estimated at the BS. Regarding the mobile RIS-User and BS-User channels, since they are low-dimensional, they can be estimated using a classical CE algorithm in the uplink small-timescale pilot transmission stage. The pilot frame structure for this stage will be explained in Section 4.

In the dual-link large-timescale CE stage, the BS needs to work in full-duplex mode. And it is divided into two working modes in this paper, depending on whether the pilot transmitting antenna simultaneously receives RIS-reflected signals. They are the monostatic antenna mode, which is consistent with the paper [19], and the bistatic plus monostatic antenna mixed mode. Both modes require full-duplex BS communication, but the methods for achieving self-interference cancellation differ [27]. In the monostatic mode, self-interference originates from the air interface signals propagated from the nearby transmitting antenna. Besides RF domain and digital domain techniques, self-interference elimination can also be achieved through signal isolation and zero traps in the spatial domain. However, in the mixed mode, there are newly added internal RF self-interferences in the transmitting and receiving channel, which are affected by various factors such as circulator leakage. Since the full-duplex BS hardware implementation structures differ, we consider the above two modes separately [28, 29]. Handling self-interference can be main challenges in practical systems, but we also assume that the noises caused by self-interference are 3 dB higher after applying various elimination techniques, as referenced in paper [28].

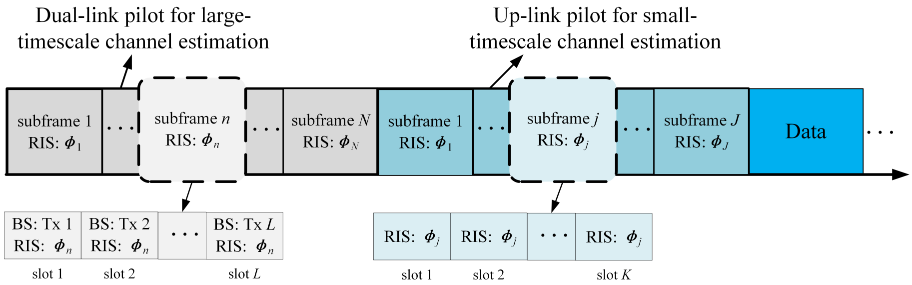


**Fig. 2** The pilot frame structure.

# 3 The proposed optimal CE method for the BS-RIS dual-link

## 3.1 Problem formulation

In Fig.2, the pilot frame consists of $N$ subframes, and each subframe lasts for $L$ time slots. In the $n$-th ($n = 1, 2, \cdots, N$) subframe, the reflection coefficient vector at the RIS is $\boldsymbol{\phi}_n \in \mathbb{C}^{N\times 1}$. In the $m_1$-th ($m_1 = 1, 2, \cdots, L$) time slot of the $n$-th subframe, the $m_1$-th BS antenna transmits a pilot $s_{m_1,n}$, and the rest $M-1$ BS antennas (or $M$ antennas for mixed mode) receive the RIS reflected pilots. The received pilots at the rest BS antennas can be written as

$$y_{m_1,m_2,n} = \left[(\mathbf{g}_{m_1} \odot \mathbf{g}_{m_2})^T \boldsymbol{\phi}_n\right] s_{m_1,n} + i_{m_1,m_2,n} + z_{m_1,m_2,n}, \tag{2}$$

where $m_2 = 1, 2, \cdots, M$, and only for monostatic antenna mode $m_2 \neq m_1$, which means the receiver antenna is different from the transmitter antenna. $\mathbf{g}_{m_1} \triangleq \mathbf{G}(m_1, :)^T \in \mathbb{C}^{N\times 1}$, $\mathbf{g}_{m_2} \triangleq \mathbf{G}(m_2, :)^T \in \mathbb{C}^{N\times 1}$, are the vectors that denote the uplink channels from the RIS to the $m_1$-th and $m_2$-th BS antenna, respectively. $z_{m_1,m_2,n} \sim \mathcal{CN}\left(0, \sigma_z^2\right)$ is the received AWGN at the $m_2$-th BS antenna. $i_{m_1,m_2,n}$ is the self-interference after mitigation, and it can be mitigated to as low as about only

3 dB higher than the receiver noise. Therefore, we assume $i_{m_1,m_2,n} \sim \mathcal{CN}\left(0,\sigma_i^2\right)$ and there is no considerable differences in orders of magnitude between $\sigma_i^2$ and $\sigma_z^2$.

Denote the index tuple set $\mathcal{S}_1 \triangleq \{(m_1,m_2)\,|1 \leq m_1 \leq L,\ \ 1 \leq m_2 \leq M, m_1 \neq m_2\}$ for monostatic mode and $\mathcal{S}_2 \triangleq \{(m_1,m_2)\,|1 \leq m_1 \leq L,\ \ 1 \leq m_2 \leq M\}$ for mixed mode. After $N$ subframes, we can get all the received pilots $\{y_{m_1,m_2,n}|\,(m_1,m_2) \in \mathcal{S}_{1,2},\quad 1 \leq n \leq N\}$ in the dual-link pilot transmission frame. The received pilots corresponding to the $m_1$-th transmitting antenna and the $m_2$-th receiving antenna throughout $N$ subframes are defined as $\mathbf{y}_{m_1,m_2}^T \triangleq [y_{m_1,m_2,1}, y_{m_1,m_2,2}, \cdots, y_{m_1,m_2,N}]$.

Assuming the transmitted pilot $s_{m_1,n} = \sqrt{P_{\text{BS}}}$ without loss of generality, where $P_{\text{BS}}$ is the transmitted power of the single BS transmitting antenna. We can write the model in the vector form

$$\mathbf{y}_{m_1,m_2}^T = \sqrt{NP_{\text{BS}}}\left(\mathbf{g}_{m_1} \odot \mathbf{g}_{m_2}\right)^T \Phi + \mathbf{i}_{m_1,m_2}^T + \mathbf{z}_{m_1,m_2}^T, \tag{3}$$

where $\Phi = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \cdots, \boldsymbol{\phi}_N] \in \mathbb{C}^{N\times N}$, $\mathbf{i}_{m_1,m_2} = [i_{m_1,m_2,1}, \cdots, i_{m_1,m_2,N}]^T$, $\mathbf{z}_{m_1,m_2} = [z_{m_1,m_2,1}, \cdots,\ \ z_{m_1,m_2,N}]^T$. And $\odot$ is the Hadmard product. For the convenience of signal processing, the reflection coefficient vectors are designed as a unitary DFT matrix $\Phi$ [7].

For all $(m_1,m_2) \in \mathcal{S}_{1,2}$, we can easily obtain that

$$\begin{aligned}\left(\mathbf{g}_{m_1} \odot \mathbf{g}_{m_2}\right)^T &= [g_{m_1,1}g_{m_2,1}, \cdots, g_{m_1,N}g_{m_2,N}] \\ &= \frac{1}{\sqrt{NP_{\text{BS}}}}\mathbf{y}_{m_1,m_2}^T \Phi^{-1} \\ &- \frac{1}{\sqrt{NP_{\text{BS}}}}\left(\mathbf{i}_{m_1,m_2}^T + \mathbf{z}_{m_1,m_2}^T\right)\Phi^{-1}.\end{aligned} \tag{4}$$

where $\frac{1}{\sqrt{NP_{\text{BS}}}}\mathbf{y}_{m_1,m_2}^T\Phi^{-1}$ is the estimation of $\left(\mathbf{g}_{m_1} \odot \mathbf{g}_{m_2}\right)^T$ with an error $\frac{1}{\sqrt{NP_{\text{BS}}}}\left(\mathbf{i}_{m_1,m_2}^T + \mathbf{z}_{m_1,m_2}^T\right)\Phi^{-1}$.

Now that for a specific $n$, $1 \leq n \leq N$, the products $g_{m_1,n}g_{m_2,n}$ are dependent on the $n$-th RIS element but independent of other RIS elements. That is to say, the problem of estimating the quasi-static BS-RIS channel can thus be divided into $N$ independent subproblems. Since $|\mathcal{S}_{1,2}| \geq M$ is required to guarantee the solvability of the subproblems, we need to have $L \geq 2$ for monostatic mode, but $L \geq 1$ for mixed mode because of $|\mathcal{S}_2| = M$ when $L = 1$, that is the transmitting antenna also receives pilots simultaneously.

For simplifying the deduction and obtaining a semi-closed form of the proposed CE method, we use a simple matrix completion equation to construct some new redundant equations for monostatic mode only. That is for the case of $m_1 = m_2$,

$$y_{m_1,m_1,n} = \frac{y_{m_1,m_2,n}y_{m_1,m_3,n}}{y_{m_2,m_3,n}}, \tag{5}$$

where $n = 1, 2, \cdots, N$, $(m_1, m_2) \in S_1$, $(m_1, m_3) \in S_1$, $(m_2, m_3) \in S_1$, and $m_2 \neq m_3$. Note that there is phase ambiguity of $180^0$ in the above equation, but it will be eliminated automatically. It will be explained in Section 5.1.

Hence we define the $n$-th RIS element related channel coefficients as $\mathbf{g}_n \triangleq (g_{1,n}, g_{2,n}, \cdots, g_{M,n})^T$, which is the $n$-th column of BS-RIS channel matrix $\mathbf{G}$. We also define $[a_{m_1,m_2,1}, a_{m_1,m_2,2}, \cdots, a_{m_1,m_2,N}] \triangleq \frac{1}{\sqrt{NP_{\text{BS}}}}\mathbf{y}_{m_1,m_2}^T\Phi^{-1}$, and reshape all the $a_{m_1,m_2,n}$ to $N$ matrices, with $n$-th matrix $\mathbf{A}_n^T \in \mathbb{C}^{L\times M}$. Its element of $m_1$-th row and $m_2$-th column is $a_{m_1,m_2,n}$. Note that the diagonal elements of $\mathbf{A}_n$ are constructed using matrix completion in equation (5) for monostatic mode.

So we can formulate the $n$-th subproblem for both modes as

$$\begin{aligned} \mathbf{g}_n &= \arg\min_{\mathbf{g}_n} \Theta(\mathbf{g}_n) \\ \Theta(\mathbf{g}_n) &= \left|\left|\mathbf{A}_n - \mathbf{g}_n\mathbf{g}_n^T\mathbf{E}_L\right|\right|_F^2, \end{aligned} \tag{6}$$

where $\mathbf{E}_L = [\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_L] \in \mathbb{C}^{M \times L}$, and $\mathbf{e}_l \in \mathbb{C}^{M \times 1}$ is a vector, with its $l$-th element being one, and others being zero. And $\| \cdot \|_F$ is the Frobenius norm.

For the description simplification in the following sections, we change $\mathbf{A}_n$ to $\mathbf{A}$ and $\mathbf{g}_n$ to $\mathbf{g}$ for independent optimization problems with a specification $n$. To enhance readability, the symbols for the above system model are listed in Table 1. The symbols and functions needed to describe the subsequent CE optimal algorithms are provided in Table 2.

**Table 1** Summary of the system model related symbols.

| Symbols | Meaning |
|---|---|
| $M$ | the number of BS antennas |
| $N$ | the number of reflection elements of RIS, or the number of pilot subframes |
| $K$ | the number of mobile users with single antenna |
| $L$ | the number of BS transmitted antennas, or the number of slots in one subframe |
| $J$ | the number of subframes in small-scale CE |
| $j$ | the sequence number of subframe in small-scale CE |
| $\mathbf{y}$ | the received signal of BS, and $\mathbf{y} \in \mathbb{C}^{M \times 1}$ |
| $\mathbf{h}_{\mathrm{D},k}$ | the direct propagation channel between the $k$-th user and BS, and $\mathbf{h}_{\mathrm{D},k} \in \mathbb{C}^{M \times 1}$ |
| $\mathbf{h}_{\mathrm{A},k}$ | the channel between the $k$-th user and RIS, and $\mathbf{h}_{\mathrm{A},k} \in \mathbb{C}^{N \times 1}$ |
| $\mathbf{G}$ | the channel between RIS and BS, and $\mathbf{G} \in \mathbb{C}^{M \times N}$ |
| $s_k$ | the pilot symbol sent by the $k$-th user, and $s_k \in \mathbb{C}$ |
| $s_{m_1,n}$ | the BS transmitted pilot using $m_1$-th antenna in $n$-th subframe |
| $y_{m_1,m_2,n}$ | the BS received pilot with $m_1$-th transmitting antenna and $m_2$-th receiving antenna in $n$-th subframe |
| $\boldsymbol{\phi}$ | the vector composed of the reflection coefficients of RIS, and $\boldsymbol{\phi} = [\phi_1, \phi_2, \cdots, \phi_N]^T$ |
| $\phi_n$ | the reflection coefficient of the $n$-th element of RIS, and $\phi_n \in \mathbb{C}$ |
| $\rho_n$ | the amplitude of $\phi_n$, and $\rho_n \in [0, 1]$ |
| $\theta_n$ | the phase of $\phi_n$, and $\theta_n \in [0, 2\pi]$ |
| $\mathbf{w}$ | the AWGN at BS, and $\mathbf{w} \in \mathbb{C}^{M \times 1}$ |
| $\mathbf{C}_k$ | the cascaded channel among $k$-th user, RIS and BS, and $\mathbf{C}_k \triangleq \mathbf{G}\mathrm{diag}\left(\mathbf{h}_{\mathrm{A},k}\right)$ |
| $n$ | the $n$-th subframe, and $n = 1, 2, \cdots, N$ |
| $m_1$ | the sequence number of BS transmitting antennas |
| $m_2$ | the sequence number of BS receiving antennas |
| $\mathbf{g}_{m_i}$ | the vector of the uplink channel from the RIS to the $m_i$-th BS antenna $(i = 1, 2)$ |
| $\mathbf{g}_n$, $\mathbf{g}$ | the $n$-th column of the BS-RIS channel matrix $\mathbf{G}$ |
| $z_{m_1,m_2,n}$ | the received AWGN at the $m_2$-th BS antenna, and $z_{m_1,m_2,n} \sim \mathcal{CN}\left(0, \sigma_z^2\right)$ |
| $i_{m_1,m_2,n}$ | the self-interference after mitigation, and $i_{m_1,m_2,n} \sim \mathcal{CN}\left(0, \sigma_i^2\right)$ |
| $\mathcal{S}_1$ | the index tuple set for monostatic mode, and $\mathcal{S}_1 \triangleq \{(m_1, m_2) \,\vert 1 \le m_1 \le L,\ \ 1 \le m_2 \le M, m_1 \ne m_2\}$ |
| $\mathcal{S}_2$ | the index tuple set for mixed mode, and $\mathcal{S}_2 \triangleq \{(m_1, m_2) \,\vert 1 \le m_1 \le L,\ \ 1 \le m_2 \le M\}$ |
| $\mathbf{y}_{m_1,m_2}$ | the received pilots corresponding to the $m_1$-th transmitting antenna and the $m_2$-th receiving antenna |
| $P_{\mathrm{BS}}$ | the transmitted power of a single BS transmitting antenna |
| $\Phi$ | the reflection DFT matrix of RIS |
| $\mathbf{A}_n$, $\mathbf{A}$ | the reshaped and constructed matrix from $[a_{m_1,m_2,1}, a_{m_1,m_2,2}, \cdots, a_{m_1,m_2,N}] \triangleq \frac{1}{\sqrt{NP_{\mathrm{BS}}}} \mathbf{y}_{m_1,m_2}^T \Phi^{-1}$ |
| $\Theta(\cdot)$ | the objective function, defined in equation (6) |
| $\mathbf{E}_L$ | the constant matrix, and $\mathbf{E}_L = [\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_L] \in \mathbb{C}^{M \times L}$ |
| $\mathbf{e}_l$ | the vector with its $l$-th element being one and others being zero, and $\mathbf{e}_l \in \mathbb{C}^{M \times 1}$ |

**Table 2** Summary of the symbols and functions in CE optimal algorithms.

| Symbols | Meaning |
|---|---|
| $\mathbf{A}_{\mathrm{E}}$ | the symmetric form of the data matrix $\mathbf{A}$, defined in equation (10) |
| $\mathbf{A}_{\mathrm{E}_L}$ | the first $L$ rows and first $L$ columns of $\mathbf{A}_{\mathrm{E}}$, and $\mathbf{A}_{\mathrm{E}_L} \in \mathbb{C}^{L\times L}$ |
| $\mathbf{A}_{\mathrm{E}_{\bar{L}}}$ | the last $M-L$ rows and $L$ columns of $\mathbf{A}_{\mathrm{E}}$, and $\mathbf{A}_{\mathrm{E}_{\bar{L}}} \in \mathbb{C}^{(M-L)\times L}$ |
| $\mathbf{g}_L$ | the first $L$ elements of $\mathbf{g}$ |
| $\mathbf{g}_{\bar{L}}$ | the remaining $M-L$ elements of $\mathbf{g}$ |
| $\mathbf{g}_{RI_L}$ | the real-valued representation of $\mathbf{g}_L$ |
| $\mathbf{g}_{RI_{\bar{L}}}$ | the real-valued representation of $\mathbf{g}_{\bar{L}}$ |
| $\hat{\mathbf{g}}$ | the estimated value of $\mathbf{g}$ |
| $\mathbf{v}_{\mathrm{Z}_{1_L}}$ | the dominant eigenvector of $\mathbf{Z}_{1_L}$, and $\mathbf{v}_{\mathrm{Z}_{1_L}} \in \mathbb{R}^{2L\times 1}$ |
| $\lambda_{\mathrm{Z}_{1_L}}$ | the maximum eigenvalue of $\mathbf{Z}_{1_L}$ |
| $\mathrm{sym}(\cdot)$ | the symmetric operation, defined in equation (10) |
| $\mathrm{lu}(\cdot)$ | the matrix partition operation, defined in equation (12) |
| $\mathrm{ld}(\cdot)$ | the matrix partition operation, defined in equation (12) |
| $\Omega\{\cdot\}$ | the complex-to-real transformation operation, defined in equation (15) |
| $\mathrm{PEA}(\cdot)$ | the primary eigenvalue approximation function for the transformed real matrix, defined in equation (21) |
| $\mathbf{Z}_{1_L}$ | the real-valued representation of $\mathbf{A}_{\mathrm{E}_L}$, and $\mathbf{Z}_{1_L} \triangleq \Omega\{\mathbf{A}_{\mathrm{E}_L}\} \in \mathbb{R}^{2L\times 2L}$ |
| $\mathbf{Z}_{2_L}$ | the real-valued representation of $\mathbf{A}_{\mathrm{E}_{\bar{L}}}$, and $\mathbf{Z}_{2_L} \triangleq \Omega\{\mathbf{A}_{\mathrm{E}_{\bar{L}}}\} \in \mathbb{R}^{2(M-L)\times 2L}$ |
| $\Psi(\cdot)$ | the primary eigenvalue approximation function, defined in equation (23) |
| $\mathcal{A}$ | the constructed tensor from all the $\mathbf{A}_n$, and $\mathcal{A} \in \mathbb{C}^{L\times M\times N}$ |
| $\tilde{\mathbf{y}}_{k,j}$ | the equivalent received pilot for the $k$-th User in $j$-th subframe, and $\tilde{\mathbf{y}}_{k,j} \in \mathbb{C}^{M\times 1}$ |
| $P_{\mathrm{UE}}$ | the transmitted power of each User |
| $\mathbf{B}$ | the data matrix defined in equation (28) |

It is worth noting that equation (4) represents the least squares solution for $\mathbf{g}_{m_1} \odot \mathbf{g}_{m_2}$, and pilot overhead can be further reduced if we adopt a sparse representation for the channel $\mathbf{G}$, which involves formulating and solving an underdetermined system of equations. This will be left for future work.

## 3.2 Optimal solution based on the PEA

In this section, we present a novel approach to obtain the global minimization of the optimization problem, as defined in equation (6). $\Theta\{\mathbf{g}\}$ can be firstly expanded as

$$\Theta\{\mathbf{g}\} = \mathrm{Tr}\left\{\mathbf{A}\mathbf{A}^{\mathrm{H}} - \mathbf{A}\,\mathbf{E}_L^{\mathrm{T}}\mathbf{g}^*\mathbf{g}^{\mathrm{H}} - \mathbf{g}\mathbf{g}^{\mathrm{T}}\mathbf{E}_L\mathbf{A}^{\mathrm{H}} + \mathbf{g}\mathbf{g}^{\mathrm{T}}\mathbf{E}_L\mathbf{E}_L^{\mathrm{T}}\mathbf{g}^*\mathbf{g}^{\mathrm{H}}\right\}. \tag{7}$$

Then we derivate it with respect to $\mathbf{g}$ and set it to zero to characterize the critical points

$$\frac{\partial\Theta\{\mathbf{g}\}}{\partial\mathbf{g}} = -\mathbf{g}^{\mathrm{T}}\left(\mathbf{E}_L\mathbf{A}^{\mathrm{H}} + \left(\mathbf{E}_L\mathbf{A}^{\mathrm{H}}\right)^{\mathrm{T}}\right)$$

$$+\,\mathbf{g}^{\mathrm{T}}\left(\mathbf{g}^{*}\,\mathbf{g}^{\mathrm{H}}\,\mathbf{E}_{L}\,\mathbf{E}_{L}^{\mathrm{T}}+\mathbf{E}_{L}\,\mathbf{E}_{L}^{\mathrm{T}}\,\mathbf{g}^{*}\,\mathbf{g}^{\mathrm{H}}\right)=\mathbf{0}_{1\times M}. \tag{8}$$

After simplifications, we get

$$\mathbf{A}_{\mathrm{E}}\,\mathbf{g}=\left(\mathbf{g}^{*}\,\mathbf{g}^{\mathrm{H}}\,\mathbf{E}_{L}\,\mathbf{E}_{L}^{\mathrm{T}}+\mathbf{E}_{L}\,\mathbf{E}_{L}^{\mathrm{T}}\,\mathbf{g}^{*}\,\mathbf{g}^{\mathrm{H}}\right)\mathbf{g}, \tag{9}$$

where $\mathbf{A}_{\mathrm{E}}\in\mathbb{C}^{M\times M}$ is a symmetric matrix and is defined as

$$\mathbf{A}_{\mathrm{E}}\triangleq\mathrm{sym}\left\{\mathbf{A}^{*}\mathbf{E}_{L}^{\mathrm{T}}\right\}=\left(\mathbf{A}^{*}\mathbf{E}_{L}^{\mathrm{T}}\right)^{\mathrm{T}}+\mathbf{A}^{*}\mathbf{E}_{L}^{\mathrm{T}}. \tag{10}$$

Equation (9) is essentially a system of equations, consisting of $M$ complex nonlinear equations. The unknown variables are the $M$ complex entries of $\mathbf{g}$. Solving (9) directly is very expensive in computation when the number of BS antennas is large ($M\gg 1$) [12]. Next, we will solve it efficiently after using a partitioned matrix representation of (10) and an alternative real domain representation.

Now we express $\mathbf{A}_{\mathrm{E}}$ as a new form of the partitioned matrix,

$$\mathbf{A}_{\mathrm{E}}=\begin{bmatrix}\mathbf{A}_{\mathrm{E}_{L}} & \mathbf{A}_{\mathrm{E}_{\bar{L}}}^{T}\\ \mathbf{A}_{\mathrm{E}_{\bar{L}}} & \mathbf{0}_{(M-L)\times(M-L)}\end{bmatrix}, \tag{11}$$

where $\mathbf{A}_{\mathrm{E}_{L}}\in\mathbb{C}^{L\times L}$ is the first $L$ rows and first $L$ columns of $\mathbf{A}_{\mathrm{E}}$. And similarly, the last $M-L$ rows and $L$ columns of $\mathbf{A}_{E}$ are represented as a $(M-L)\times L$ matrix $\mathbf{A}_{\mathrm{E}_{\bar{L}}}$. They can be represented as left and right multiplications of specific constant matrices applied to $\mathbf{A}_{E}$. For subsequent description clarity, we define this matrix partitioning operation as functions $\mathrm{lu}(\cdot)$ and $\mathrm{ld}(\cdot)$.

$$\begin{aligned}\mathbf{A}_{\mathrm{E}_L} &= \mathrm{lu}\left(\mathbf{A}_{\mathrm{E}}\right) \triangleq \begin{bmatrix}\mathbf{I}_L & \mathbf{0}_{L\times(M-L)}\end{bmatrix} \mathbf{A}_{\mathrm{E}} \begin{bmatrix}\mathbf{I}_L & \mathbf{0}_{L\times(M-L)}\end{bmatrix}^T \\ \mathbf{A}_{\mathrm{E}_{\bar{L}}} &= \mathrm{ld}\left(\mathbf{A}_{\mathrm{E}}\right) \triangleq \begin{bmatrix}\mathbf{0}_{(M-L)\times L} & \mathbf{I}_{M-L}\end{bmatrix} \mathbf{A}_{\mathrm{E}} \begin{bmatrix}\mathbf{I}_{M-L} & \mathbf{0}_{(M-L)\times L}\end{bmatrix}^T\end{aligned} \tag{12}$$

We denote the first $L$ entries of $\mathbf{g} \in \mathbb{C}^{M\times 1}$ by a $L\times 1$ column vector $\mathbf{g}_L \triangleq [[\mathbf{g}]_1, [\mathbf{g}]_2, \cdots, [\mathbf{g}]_L]^{\mathrm{T}}$ and the remaining $M-L$ entries by a $(M-L)\times 1$ column vector $\mathbf{g}_{\bar{L}} \triangleq [[\mathbf{g}]_{L+1}, [\mathbf{g}]_{L+2}, \cdots, [\mathbf{g}]_M]^{\mathrm{T}}$. So equation (9) can be represented as

$$\begin{bmatrix}\mathbf{A}_{\mathrm{E}_L} & \mathbf{A}_{\mathrm{E}_{\bar{L}}}^T \\ \mathbf{A}_{\mathrm{E}_{\bar{L}}} & \mathbf{0}_{(M-L)\times(M-L)}\end{bmatrix} \begin{bmatrix}\mathbf{g}_L \\ \mathbf{g}_{\bar{L}}\end{bmatrix} = \|\mathbf{g}_L\|^2 \begin{bmatrix}\mathbf{g}_L^* \\ \mathbf{g}_{\bar{L}}^*\end{bmatrix} + \|\mathbf{g}\|^2 \begin{bmatrix}\mathbf{g}_L^* \\ \mathbf{0}_{(M-L)\times 1}\end{bmatrix}. \tag{13}$$

Then the real representation of the complex matrix is used on both sides, an alternate system of $2M$ real equations can be defined as

$$\begin{aligned}&\begin{bmatrix}\Omega\left\{\mathbf{A}_{\mathrm{E}_L}\right\} & \Omega\left\{\mathbf{A}_{\mathrm{E}_{\bar{L}}}^T\right\} \\ \Omega\left\{\mathbf{A}_{\mathrm{E}_{\bar{L}}}\right\} & \mathbf{0}_{2(M-L)\times 2(M-L)}\end{bmatrix} \begin{bmatrix}\mathbf{g}_{RI_L} \\ \mathbf{g}_{RI_{\bar{L}}}\end{bmatrix} = \\ &\begin{bmatrix}\left(\|\mathbf{g}_L\|^2 + \|\mathbf{g}\|^2\right)\mathbf{I}_{2L} & \mathbf{0}_{2L\times 2(M-L)} \\ \mathbf{0}_{2(M-L)\times 2L} & \|\mathbf{g}_L\|^2\mathbf{I}_{2(M-L)}\end{bmatrix} \begin{bmatrix}\mathbf{g}_{RI_L} \\ \mathbf{g}_{RI_{\bar{L}}}\end{bmatrix},\end{aligned} \tag{14}$$

where $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ represent the real and imaginary parts of a matrix or vector,

$$\Omega\left\{\cdot\right\} \triangleq \begin{bmatrix}\mathrm{Re}\{\cdot\} & -\mathrm{Im}\{\cdot\} \\ -\mathrm{Im}\{\cdot\} & -\mathrm{Re}\{\cdot\}\end{bmatrix} \tag{15}$$

represents the transformation of complex matrix to real matrix, and $\mathbf{g}_{RI_L} = [\mathrm{Re}\{\mathbf{g}_L\} \quad \mathrm{Im}\{\mathbf{g}_L\}]^T \in \mathbb{R}^{2L\times 1}$, $\mathbf{g}_{RI_{\bar{L}}} = [\mathrm{Re}\left\{\mathbf{g}_{\bar{L}}\right\} \quad \mathrm{Im}\left\{\mathbf{g}_{\bar{L}}\right\}]^T \in \mathbb{R}^{2(M-L)\times 1}$.

After further simplifications, the initial system of $M$ complex equations (9) is lastly translated to the following two systems of real nonlinear equations

$$\big(\|\mathbf{g}_L\|^2 + \|\mathbf{g}\|^2\big)\mathbf{g}_{RI_L} = \mathbf{Z}_{1_L}\, \mathbf{g}_{RI_L} + \mathbf{Z}_{2_L}^{\mathrm{T}}\, \mathbf{g}_{RI_{\bar{L}}}, \tag{16}$$

$$\mathbf{Z}_{2_L}\, \mathbf{g}_{RI_L} = \|\mathbf{g}_L\|^2 \mathbf{g}_{RI_{\bar{L}}}, \tag{17}$$

where $\mathbf{Z}_{1_L} \triangleq \Omega\left\{\mathbf{A}_{\mathrm{E}_L}\right\} \in \mathbb{R}^{2L\times 2L}$ and $\mathbf{Z}_{2_L} \triangleq \Omega\left\{\mathbf{A}_{\mathrm{E}_{\bar{L}}}\right\} \in \mathbb{R}^{2(M-L)\times 2L}$. Here $\Omega\left\{\cdot\right\}$ means the complex to real transformation which is defined in (15). From (17), we get

$$\mathbf{g}_{RI_{\bar{L}}} = \begin{bmatrix} \mathrm{Re}\{\mathbf{g}_{\bar{L}}\} \\ \mathrm{Im}\{\mathbf{g}_{\bar{L}}\} \end{bmatrix} = \frac{1}{\|\mathbf{g}_L\|^2}\, \mathbf{Z}_{2_L}\, \mathbf{g}_{RI_L}. \tag{18}$$

Finally using another deduction from (14),

$$\mathbf{Z}_{2_L}^{\mathrm{T}}\, \mathbf{g}_{RI_{\bar{L}}} = \big(\|\mathbf{g}\|^2 - \|\mathbf{g}_L\|^2\big)\mathbf{g}_{RI_L}, \tag{19}$$

we obtain the following key result after simplifications

$$\mathbf{Z}_{1_L}\mathbf{g}_{RI_L} = 2\|\mathbf{g}_L\|^2\, \mathbf{g}_{RI_L}. \tag{20}$$

Equation (20) is in the form of a conventional eigenvalue problem, the solution to (20) for $\mathbf{g}_{RI_L}$ is the eigenvector corresponding to the positive eigenvalue $2\|\mathbf{g}_L\|^2$ of the matrix $\mathbf{Z}_{1_L}$. This is also the estimate for the global minimum problem (6). Furthermore, since $\mathbf{g} = \mathrm{Re}\{\mathbf{g}\} + j\,\mathrm{Im}\{\mathbf{g}\} \in \mathbb{C}^{M\times 1}$, the real and imaginary components for the first $L$ entries are obtained from the eigenvector $\mathbf{v}_{\mathrm{Z}_{1_L}} \in \mathbb{R}^{2L\times 1}$ corresponding to the maximum eigenvalue $\lambda_{\mathrm{Z}_{1_L}}$ of $\mathbf{Z}_{1_L}$. We define this operation as the function

PEA($\cdot$),

$$\widehat{\mathbf{g}}_{RI_L} = \text{PEA}(\text{Z}_{1_L}) \triangleq \pm\sqrt{\frac{\lambda_{\text{Z}_{1_L}}}{2}}\frac{\mathbf{v}_{\text{Z}_{1_L}}}{\|\mathbf{v}_{\text{Z}_{1_L}}\|} \tag{21}$$

Lastly the remaining $M - L$ entries of vector $\mathbf{g}$ can be calculated based on equation (18).

### 3.3 Semi-closed form of $\widehat{\mathbf{g}}$

Henforth, we can represent $\widehat{\mathbf{g}}$ as

$$\begin{aligned}
\widehat{\mathbf{g}} &= \text{Re}\left\{\begin{bmatrix}\widehat{\mathbf{g}}_L \\ \widehat{\mathbf{g}}_{\bar{L}}\end{bmatrix}\right\} + j\text{Im}\left\{\begin{bmatrix}\widehat{\mathbf{g}}_L \\ \widehat{\mathbf{g}}_{\bar{L}}\end{bmatrix}\right\} \\
&= \begin{bmatrix}\mathbf{I}_L \ \mathbf{0}_L & \\ & \mathbf{I}_{M-L} \ \mathbf{0}_{M-L}\end{bmatrix}\begin{bmatrix}\widehat{\mathbf{g}}_{RI_L} \\ \widehat{\mathbf{g}}_{RI_{\bar{L}}}\end{bmatrix} + j\begin{bmatrix}\mathbf{0}_L \ \mathbf{I}_L & \\ & \mathbf{0}_{M-L} \ \mathbf{I}_{M-L}\end{bmatrix}\begin{bmatrix}\widehat{\mathbf{g}}_{RI_L} \\ \widehat{\mathbf{g}}_{RI_{\bar{L}}}\end{bmatrix} \\
&= \begin{bmatrix}\mathbf{I}_L \ j\mathbf{I}_L & \\ & \mathbf{I}_{M-L} \ j\mathbf{I}_{M-L}\end{bmatrix}\begin{bmatrix}\widehat{\mathbf{g}}_{RI_L} \\ \widehat{\mathbf{g}}_{RI_{\bar{L}}}\end{bmatrix}
\end{aligned} \tag{22}$$

Moreover, as noted from the definitions for sym $\{\cdot\}$, matrix partitioning and $\Omega\{\cdot\}$ in (10), (12) and (15), respectively, the estimate $\widehat{\mathbf{g}}$ is actually a function of $\mathbf{A}_{\text{E}} = \mathbf{A}^*\text{E}_L^{\text{T}}$. Based on the equations (18) and (21), we can summarize that the proposed optimal CE of the BS-RIS channel vector $\mathbf{g}$ for a specific $n$ can be expressed in semi-closed form.

$$\widehat{\mathbf{g}} = \Psi\left(\mathbf{A}^*\mathbf{E}_L^{\text{T}}\right) \tag{23}$$

where

$$\Psi(\cdot) \triangleq \begin{bmatrix}\mathbf{I}_L \ j\mathbf{I}_L & \\ & \mathbf{I}_{M-L} \ j\mathbf{I}_{M-L}\end{bmatrix}\begin{bmatrix}\text{PEA}\left(\Omega\left(\text{lu}\left(\text{sym}\left(\cdot\right)\right)\right)\right) \\ \frac{\Omega(\text{ld}(\text{sym}(\cdot))}{\|\text{PEA}(\Omega(\text{lu}(\text{sym}(\cdot))))\|^2}\text{PEA}\left(\Omega\left(\text{lu}\left(\text{sym}\left(\cdot\right)\right)\right)\right)\end{bmatrix} \tag{24}$$

## 3.4 Special cases

Now we consider two special cases of $L = 1$ and $L = M$. It means only one BS antenna or all the BS antennas are used as pilot transmitting antenna respectively. Obviously it is necessary that the full-duplex BS is working on mixed mode.

### 3.4.1 $L = 1$

For this case, solving (20) reduces to solving the following two equations

$$\begin{aligned} &\mathrm{Re}\left\{[\mathbf{A}_{\mathrm{E}}]_{11}\right\}\mathrm{Re}\left\{[\widehat{\mathbf{g}}]_1\right\} - \mathrm{Im}\left\{[\mathbf{A}_{\mathrm{E}}]_{11}\right\}\mathrm{Im}\left\{[\widehat{\mathbf{g}}]_1\right\} = 2\|\,[\widehat{\mathbf{g}}]_1\,\|^2\,\mathrm{Re}\left\{[\widehat{\mathbf{g}}]_1\right\}, \\ &-\mathrm{Im}\left\{[\mathbf{A}_{\mathrm{E}}]_{11}\right\}\mathrm{Re}\left\{[\widehat{\mathbf{g}}]_1\right\} - \mathrm{Re}\left\{[\mathbf{A}_{\mathrm{E}}]_{11}\right\}\mathrm{Im}\left\{[\widehat{\mathbf{g}}]_1\right\} = 2\|\,[\widehat{\mathbf{g}}]_1\,\|^2\,\mathrm{Im}\left\{[\widehat{\mathbf{g}}]_1\right\}, \end{aligned} \tag{25}$$

where $[\mathbf{A}_{\mathrm{E}}]_{11}$ denotes the first element of matrix $\mathbf{A}_{\mathrm{E}}$, $[\widehat{\mathbf{g}}]_i$ is the $i$-th element of $\widehat{\mathbf{g}}$, and $\|\,[\widehat{\mathbf{g}}]_1\,\|^2 = \left(\mathrm{Re}\left\{[\widehat{\mathbf{g}}]_1\right\}\right)^2 + \left(\mathrm{Im}\left\{[\widehat{\mathbf{g}}]_1\right\}\right)^2$. Solving the above equation (25) and substituting the resultant into (18), yields the desired estimate for $L = 1$ as

$$\begin{aligned} &\mathrm{Re}\left\{[\widehat{\mathbf{g}}]_1\right\} \triangleq \pm\sqrt{\tfrac{\left|[\mathbf{A}_{\mathrm{E}}]_{11}\right| + \mathrm{Re}\left\{[\mathbf{A}_{\mathrm{E}}]_{11}\right\}}{2}}, \\ &\mathrm{Im}\left\{[\widehat{\mathbf{g}}]_1\right\} \triangleq \pm\tfrac{\mathrm{Im}\left\{[\mathbf{A}_{\mathrm{E}}]_{11}\right\}}{\sqrt{2\left(\left|[\mathbf{A}_{\mathrm{E}}]_{11}\right| + \mathrm{Re}\left\{[\mathbf{A}_{\mathrm{E}}]_{11}\right\}\right)}}, \\ &\begin{bmatrix} \mathrm{Re}\left\{[\widehat{\mathbf{g}}]_i\right\} \\ \mathrm{Im}\left\{[\widehat{\mathbf{g}}]_i\right\} \end{bmatrix} \triangleq \tfrac{\left[\mathbf{Z}_{2_L}\right]_{i1}\mathrm{Re}\left\{[\widehat{\mathbf{g}}]_1\right\} + \left[\mathbf{Z}_{2_L}\right]_{i2}\mathrm{Im}\left\{[\widehat{\mathbf{g}}]_1\right\}}{\left|[\widehat{\mathbf{g}}]_1\right|^2}, \end{aligned} \tag{26}$$

for $i = 2, 3, \cdots, M$.

### 3.4.2 $L = M$

Here using the fact that $\mathbf{E}_L = \mathbf{I}_M$ in equation (6), the optimal estimation of $\mathbf{g}$ is given by (21). Further, with $L = M$, the real and imaginary component of $\mathbf{g}$ can be directly

**Algorithm 1** Proposed PEA-based optimal solution to the $N$ CE subproblems (6) for obtaining the estimation of BS-RIS channel, $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_N]$.

**Require:** The parameters of system model: $L$, $M$, $N$, $\Phi$, $P_{\mathrm{BS}}$ and the received pilots: $\mathbf{y}_{m_1,m_2}$ for all the $(m_1, m_2) \in \mathcal{S}_{1,2}$.

**Ensure:** The estimated BS-RIS channel, $\hat{\mathbf{G}} = [\hat{\mathbf{g}}_1, \hat{\mathbf{g}}_2, \cdots, \hat{\mathbf{g}}_N]$.

1: **if** the full-duplex BS is working on monostatic mode ($L \geq 2$) **then**
2: using (5) to complete received pilots $\mathbf{y}_{m_1,m_2}$ for the cases of $m_1 = m_2$.
3: **end if**
4: For all $(m_1, m_2) \in \mathcal{S}_1$, construct tensor $\mathcal{A} \in \mathbb{C}^{L\times M\times N}$ using $\mathcal{A}(m_1, m_2, :) = \frac{1}{\sqrt{NP_{\mathrm{BS}}}} \mathbf{y}^T_{m_1,m_2} \Phi^{-1}$.
5: **for** $n = 1, 2, \cdots, N$ **do**
6: Reshape $\mathbf{A}^T = \mathcal{A}(:, :, n)$.
7: Compute $\mathbf{A}_{\mathrm{E}}$ based on equation (10).
8: Compute $\mathbf{Z}_{1_L}$ and $\mathbf{Z}_{2_L}$ based on equation (15) and (11).
9: Compute the first $L$ elements of $\hat{\mathbf{g}}$ based on equation (21).
10: Compute the left $M - L$ elements of $\hat{\mathbf{g}}$ based on equation (18).
11: **end for**
**return** $\hat{\mathbf{G}} = [\hat{\mathbf{g}}_1, \hat{\mathbf{g}}_2, \cdots, \hat{\mathbf{g}}_N]$.

obtained using the maximum eigenvalue of $\mathbf{Z}_{1_L}$ as

$$\begin{bmatrix} \mathrm{Re}\{[\hat{\mathbf{g}}]_i\} \\ \mathrm{Im}\{[\hat{\mathbf{g}}]_i\} \end{bmatrix} = \pm \sqrt{\frac{\lambda_{\mathrm{Z}_{1_L}}}{2}} \frac{\left[\mathbf{v}_{\mathrm{Z}_{1_L}}\right]_i}{\|\mathbf{v}_{\mathrm{Z}_{1_L}}\|}, \tag{27}$$

where $\forall i = 1, 2, \cdots, M$, and $\mathbf{v}_{\mathrm{Z}_{1_L}}$ represents the eigenvector corresponding to the maximum eigenvalue of $\mathbf{Z}_{1_L}$.

### 3.5 Algorithm and computation complexity

After the above derivatives and nontrivial complex to real transformations, we have finally reduced the whole CE process to a simple semi-closed-form expression. We just mainly need to compute the principle eigenvalue and eigenvector of a $2L \times 2L$ square matrix $\mathbf{Z}_{1_L}$ and a matrix multiplication between $\mathbf{Z}_{2_L}$ and $\mathbf{g}_{RI_L}$. The proposed optimal CE algorithm is described in Algorithm (1).

In steps 1-3, the parameters initialization and data completion of received pilots are finished. The computation complexity is $2LN$ complex multiplication or $8LN$ real multiplication. In step 4, the tensor $\mathcal{A}$ is constructed, and $\Phi$ is assigned as a DFT matrix as referenced in [7]. The calculation is indeed an inverse DFT of $\mathbf{y}_{m_1,m_2}^T$ and it needs $\frac{N}{2}\log_2 N$ complex multiplications for the fast Fourier transform will be utilized. Thus this step needs $2LMN\log_2 N$ real multiplications. In steps 6-8, the sym $\{\cdot\}$ and $\Omega\{\cdot\}$ are simple matrix operations. They can be operated through special addressing of the data memory. In step 9, the equation (21) is used to calculate the first $L$ values of $\hat{\mathbf{g}}$. The $\Psi\{\cdot\}$ operation is actually to calculate the principle eigenvalue and its corresponding eigenvector of $2L \times 2L$ square matrix $\mathbf{Z}_{1_L}$. Using the symmetric QR iterative method, the computation complexity is $\mathcal{O}\left(4L^2N\right)$ real multiplications. In step 10, based on the equation (18), we calculate the last $M-L$ values of $\hat{\mathbf{g}}$, which is a matrix multiplication and needs $4L(M-L)N$ real multiplications. Therefore, the total algorithm complexity is $\mathcal{O}(4LMN + 2LMN\log_2 N + 8LN)$. For typical system model parameters, this is better than the optimization algorithm in [19] where the computation complexity is $\mathcal{O}\left(N^3 + LMN^2 + LMNI_{max}\right)$.

# 4 The CE method for the uplinks of RIS-UE and BS-UE

Given the estimate of the quasi-static BS-RIS channel $\mathbf{G}$, we can estimate the mobile RIS-User channel $\{\mathbf{h}_{\mathrm{A},k}|1 \le k \le K\}$ and the BS-User channel $\{\mathbf{h}_{\mathrm{D},k}|1 \le k \le K\}$ with the conventional uplink pilot transmission scheme and CE method.

As depicted in Fig.2, the uplink pilot transmission frame consists of $J$ subframes, and each subframe lasts for $K$ time slots. In the $j$-th subframe $(j = 1, 2, \cdots, J)$, the reflection coefficient vector at the RIS is set to be $\phi_j \in \mathbb{C}^{N\times 1}$. The complex reflection coefficient vectors are randomly selected from a DFT matrix of dimension $N \times N$. During the $K$ time slots in a subframe, the Users transmit uplink orthogonal pilot sequences, i.e., $\mathbf{s}_k \in \mathbb{C}^{K\times 1}$, $k = 1, 2, \cdots, K$.

We use $\tilde{\mathbf{y}}_{k,j} \in \mathbb{C}^{M\times 1}$ denotes the equivalent received pilot for the $k$-th User in $j$-th subframe. Then based on the system model (1), we can distinguish the channels of different User by right multiplying the conjugate of the pilot sequences:

$$\begin{aligned}
\tilde{\mathbf{y}}_{k,j} &= \frac{1}{KP_{\mathrm{UE}}}\mathbf{Y}_j\mathbf{s}_k^* \\
&= \frac{1}{KP_{\mathrm{UE}}}\sum_{k'=1}^{K}\left[\mathbf{G}\mathrm{diag}\left(\mathbf{h}_{\mathrm{A},k'}\right)\phi_j + \mathbf{h}_{\mathrm{D},k'}\right]\mathbf{s}_{k'}^T\mathbf{s}_k^* + \frac{\mathbf{W}_j\mathbf{s}_k^*}{KP_{\mathrm{UE}}} \\
&= \left[\mathbf{G}\mathrm{diag}\left(\phi_j\right)\mathbf{h}_{\mathrm{A},k} + \mathbf{h}_{\mathrm{D},k}\right] + \tilde{\mathbf{w}}_{k,j} \\
&= \mathbf{B}_j\begin{bmatrix}\mathbf{h}_{\mathrm{A},k}\\ \mathbf{h}_{\mathrm{D},k}\end{bmatrix} + \tilde{\mathbf{w}}_{k,j}, \quad k = 1, 2, \cdots, K,
\end{aligned} \tag{28}$$

where $P_{\mathrm{UE}}$ denotes the transmitted power of each User, $\mathbf{Y}_j \in \mathbb{C}^{M\times K}$ is the matrix of received pilots at the BS in $j$-th subframe, and each column of $\mathbf{Y}_j$ is the received pilots in a single time slot, $\mathbf{W}_j \in \mathbb{C}^{M\times K}$ is the noise, and $\tilde{\mathbf{w}}_{k,j} = \frac{\mathbf{W}_j\mathbf{s}_k^*}{KP_{\mathrm{UE}}}$, $\mathbf{B}_j \triangleq \left[\mathbf{G}\mathrm{diag}\left(\tilde{\phi}_j\right)\ \mathbf{I}_M\right]$.

Then, by collecting the equivalent received pilots in $J$ subframes, we have

$$\tilde{\mathbf{y}}_k = \mathbf{B}\begin{bmatrix}\mathbf{h}_{\mathrm{A},k}\\ \mathbf{h}_{\mathrm{D},k}\end{bmatrix} + \tilde{\mathbf{w}}_k, \quad k = 1, 2, \cdots, K, \tag{29}$$

where $\tilde{\mathbf{y}}_k \triangleq \left[\tilde{\mathbf{y}}_{k,1}^T, \tilde{\mathbf{y}}_{k,2}^T, \cdots, \tilde{\mathbf{y}}_{k,J}^T\right]^T$, $\mathbf{B} \triangleq \left[\mathbf{B}_1^T, \mathbf{B}_2^T, \quad \cdots, \mathbf{B}_J^T\right]^T$, and $\tilde{\mathbf{w}}_k \triangleq \left[\tilde{\mathbf{w}}_{k,1}^T, \tilde{\mathbf{w}}_{k,2}^T, \cdots, \tilde{\mathbf{w}}_{k,J}^T\right]^T$.

In the above equation, $\mathbf{B}$ is determined by the exact quasi-static channel $\mathbf{G}$ and the predesigned $\left\{\boldsymbol{\phi}_j \middle| 1 \le j \le J\right\}$. Based on the estimated quasi-static channel matrix

$\hat{\mathbf{G}}$, we can define $\hat{\mathbf{B}}$ as

$$\hat{\mathbf{B}} \triangleq \begin{bmatrix} \hat{\mathbf{G}}\text{diag}\left(\boldsymbol{\phi}_1\right) & \mathbf{I}_M \\ \vdots & \vdots \\ \hat{\mathbf{G}}\text{diag}\left(\boldsymbol{\phi}_J\right) & \mathbf{I}_M \end{bmatrix}. \tag{30}$$

Finally, the mobile channels can be estimated by the conventional least square (LS) estimation as

$$\begin{bmatrix} \hat{\mathbf{h}}_{\mathrm{A},k} \\ \hat{\mathbf{h}}_{\mathrm{D},k} \end{bmatrix} = \hat{\mathbf{B}}^{\dagger}\tilde{\mathbf{y}}_k = \left(\hat{\mathbf{B}}^H\hat{\mathbf{B}}\right)^{-1}\hat{\mathbf{B}}^H\tilde{\mathbf{y}}_k, \quad k = 1, 2, \cdots, K, \tag{31}$$

where $\hat{\mathbf{B}}^{\dagger}$ is the pseudo-inverse of $\hat{\mathbf{B}}$. $\hat{\mathbf{h}}_{\mathrm{A,k}}$ and $\hat{\mathbf{h}}_{\mathrm{D},k}$ are the estimates of the RIS-UE channel and the BS-UE channel of the $k$-th User, respectively.

# 5 Discussion

## 5.1 Phase ambiguity

Notice that the estimator $\widehat{\mathbf{g}}$ yielding the global minimizer involves an unresolvable 180° phase ambiguity and hence is not unique. Note that $\mathbf{C}_k = \mathbf{Gdiag}\left(\mathbf{h}_{\mathrm{A},k}\right)$, and the aim is to save pilot overhead by estimating $\mathbf{G}$ and $\mathbf{h}_{\mathrm{A},k}$ separately. Once $\mathbf{G}$ is estimated, unchanged during the large-timescale though with phase ambiguity, the BS will estimate the new small-timescale $\mathbf{h}_{\mathrm{A},k}$ and calculate $\mathbf{C}_k$ using the pilots transmitted from the $k$-th user. This will also result in a 180° phase ambiguity to $\mathbf{h}_{\mathrm{A},k}$. However the BS only uses the cascaded channel $\mathbf{C}_k$ to realize the maximum ratio SNR or RF beamforming. In other words, this phase ambiguity will disappear automatically when we use the product $\mathbf{Gdiag}\left(\mathbf{h}_{\mathrm{A},k}\right)$.

### 5.2 Pilot overhead

As shown in Fig.2, for the large-timescale CE stage, the pilots are contained in $N$ subframes, and each subframe contains $L$ time slots, thus the pilot overhead is $\tau_1 = NL$. For the mobile small-timescale CE stage of the channels between $K$ users and the BS or RIS, there are $NK$ and $MK$ parameters for $\mathbf{h}_{\mathrm{A},k}$ and $\mathbf{h}_{\mathrm{D},k}$ need to be estimated, respectively. Since the BS of $M$ antennas receives $MK$ pilot symbols in each timeslot, so $J \geq \lceil (M+N)\, K/MK \rceil = \lceil N/M \rceil + 1$ subframes are required, where $\lceil \cdot \rceil$ indicates that it is rounded up. Therefore, the pilot overhead for the small-scale CE is $\tau_2 = K\lceil N/M \rceil + K$. On the other hand, the pilot overhead of the two-timescale CE framework depends on the ratio of coherence time between the quasi-static channel and the mobile dynamic channel $\alpha = T_L/T_S$. Therefore, the total pilot overhead is $\tau = NL/\alpha + K\lceil N/M \rceil + K$.

## 6 Experiments

### 6.1 Simulation setup

To compare effectively with the existing methods, we use the same channel models and system parameters as in [19] to simulate the proposed method. And the Monte Carlo method is used to evaluate its performance under different system model parameters. Hence the normalized mean square error (NMSE) of the estimated channel matrix $\mathbf{G}$ and cascaded channel for $k$-th User $\mathbf{C}_k$ are approximately calculated through the arithmetic mean as

$$\mathrm{NMSE}(\mathbf{G}) \triangleq \frac{\|\hat{\mathbf{G}} - \mathbf{G}\|_F^2}{\|\mathbf{G}\|_F^2} \tag{32}$$

and

$$\mathrm{NMSE}(\mathbf{C}_k) \triangleq \frac{\|\hat{\mathbf{C}}_k - \mathbf{C}_k\|_F^2}{\|\mathbf{C}_k\|_F^2}. \tag{33}$$

In our simulations, the number of BS antennas $M = 16$, the number of RIS reflection units $N = 64$, the number of pilot transmitted BS antenna $L \in \{1, 2, 4, M\}$, and the number of users $K = 8$. The large-scale fading between BS and RIS is modeled as $\rho_g = \rho_0 \left(d_g/d_0\right)^{\alpha_g}$, where $\rho_0 = -20$ dB is the large-scale fading factor at the reference distance of $d_0 = 1$ m, and we set path loss exponent $\alpha_g = 2.1$. The BS-RIS distance is $d_g = 20$ m. Rician channel model is used as the small-scale fading, and the noise power is computed for different SNRs by SNR $= P_{BS}\rho_g^2/(\sigma_i^2 + \sigma_z^2)$, where the self-interference is set to 3 dB higher than the receiver noise, that is $\sigma_i^2 = 2\sigma_z^2$. In the simulation of the CE algorithm in [19], the outer iterative times $I_{max}$ is set to 5, for a quite low NMSE from the results of Fig.7 in [19].

## 6.2 Simulation results

The purpose of two-timescale CE framework is to save pilot overhead. Figures 3 to 5 present the curves of pilot overhead as $M$, $N$ and $K$ increase, respectively. Among these, the pilot overhead of the MVU estimation method is $NK + K$, while the pilot overhead of multi-user channel estimation is $K + N + \max\left\{K - 1, \left\lceil \frac{(K-1)N}{M} \right\rceil\right\}$[12, 13]. The pilot overhead of sparse CE methods based on compressed sensing is $\mathcal{O}\left\{KS \log N\right\}$, where $S$ is sparsity level [13, 15]. And in reference [13], the pilot overhead is also determined by the number of common paths $L_c$ between the RIS and users for the modified OMP algorithm. Notably, the pilot overhead of the proposed PEA-based CE method is $\tau = NL/\alpha + KN/M + K$, whereas that of [19] is $\tau = (N+1)L/\alpha + K\lceil N/M \rceil + K$. The minor difference arises because the proposed PEA-based method eliminates pilots for estimating environmental reflections.

The results show that the proposed PEA-based method and the two-timescale strategy can significantly reduce pilot overhead. It is worth noting that $\alpha$ significantly affects this reduction. For instance, the pilot overhead of the multi-user CE method in [13] is lower than that of the proposed method when $\alpha \leq 2$. In other words, the proposed method saves pilot overhead through utilization of the quasi-static property of the channel between the BS and RIS and exploitation of large coherence time conditions.

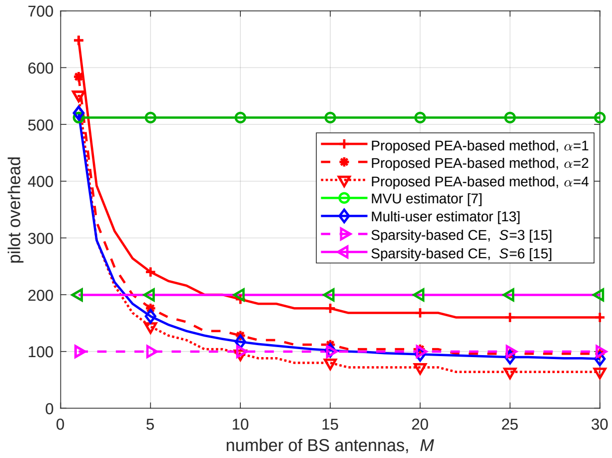


**Fig. 3** Pilot overhead with different $M$.

The pilot overhead of sparsity-based CE is closely related to the sparsity level $S$. In RIS-assisted systems, $S$ typically represents the number of effective paths in cascaded channels, as used in millimeter-wave systems. In order to meet sparsity channel conditions, the number of paths of the actual propagation channel is limited to single digits in the simulations. Fig. 3 shows that the pilot overhead of sparse estimation methods does not depend on $M$. Although channel sparsity may relate to $M$ in practice, simulations here consider only $S = 3$ and $S = 6$, yielding $M$-independent curves. Figs. 4 and 5 simulate how this pilot overhead varies with $N$ and $K$. Notably, pilot overhead for $S = 3$ is lower than that of other methods or approximately equal under some values of $K$ and $N$. This is expected since the proposed method, MVU, and Multi-user method target general channels, not dedicated sparse channels in millimeter-wave angular domains. Thus, each method has distinct application scenarios. It is worth noting that only the general compressive sensing method in reference [15] has been simulated and evaluated here, while the more complicated modified OMP algorithm in reference [16] has not been simulated.

Fig.6 illustrates the estimation accuracy for the special cases of $L = 1$ and $L = M$ when the full-duplex BS is working on mixed mode. Simulation results illustrate the impact of received SNR and $L$ on the estimation accuracy of $\mathbf{G}$. And we observe that

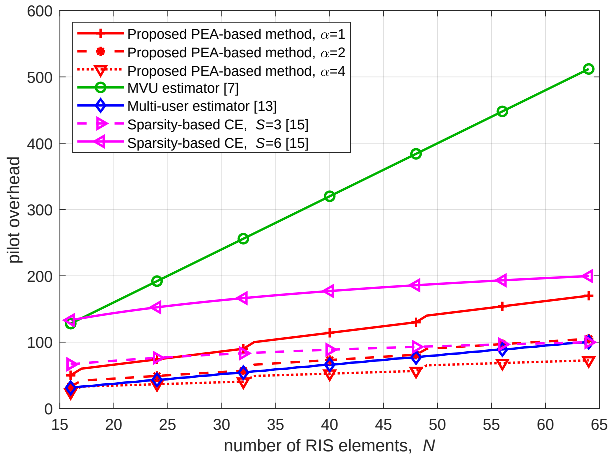


**Fig. 4** Pilot overhead with different $N$.

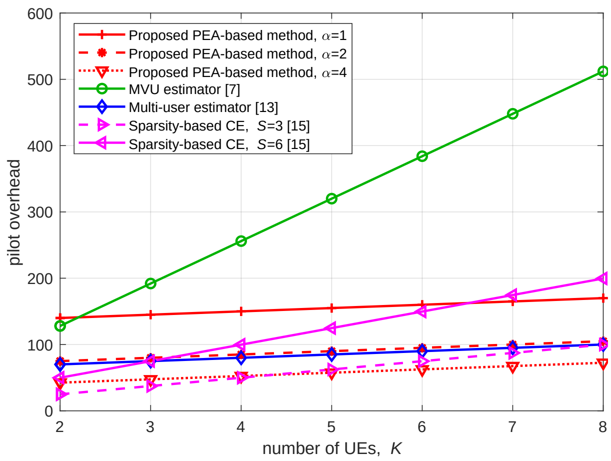


**Fig. 5** Pilot overhead with different $K$.

as the SNR increases, the difference of NMSE between the two special cases becomes larger. For instances, as the SNR increases from -10 dB to 20 dB, the difference in NMSE increases from approximately 5 dB to 10 dB. Although the NMSE is better for $L = M$ compared to $L = 1$, the pilot overhead is much higher. Therefore, careful consideration should be given to selecting the number of pilot transmitting antennas $L$.

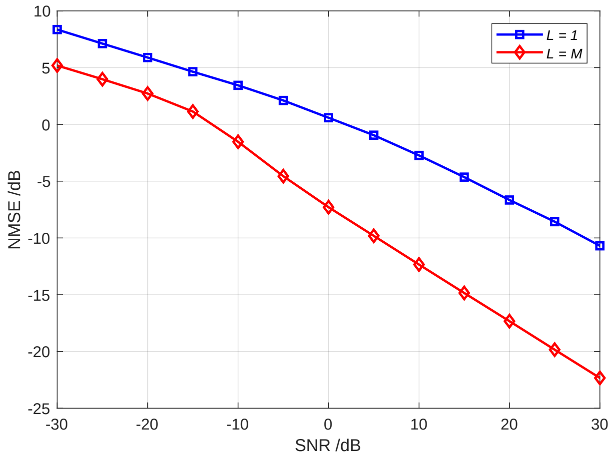


**Fig. 6** Estimation error of $\mathbf{G}$ for special cases $L = 1, M$ (only for BS mixed mode).

For the monostatic mode and mixed mode of the full-duplex BS, Figures 7 and 8 compare the estimation errors of the proposed PEA-based method with the reference method presented in [19], across different values of $M$ and SNR when $L = 2$. Simulation results show that the proposed PEA-based method with BS working on mixed mode has the best performance. The reason is that there are more data equations available than in the other two cases. For the monostatic mode, the proposed PEA-based method has a better performance when the SNR is below about 20 dB.

When the SNR exceeds 20 dB, the proposed PEA-based method in monostatic mode yields higher but acceptable NMSE than the CDIR method. This is because we deliberately introduced $L$ redundant data equations to simplify computational complexity and unify the CE algorithm with a semi-closed-form expression for both working modes. These equations are constructed from actual pilot data via matrix completion, and will cause minor estimation errors by assigning slightly larger weights to these redundant equations. The CDIR method adopts an iterative optimization approach designed specifically for monostatic mode. Its drawbacks include empirically determined iteration counts and high complexity. This method achieves marginally better accuracy at high SNR, but convergence slows at low SNR. Considering

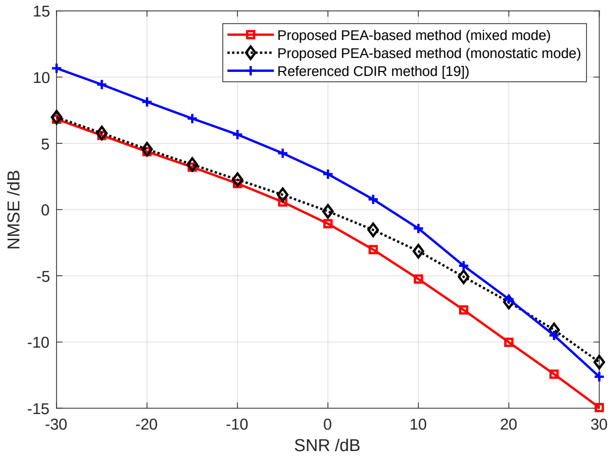


**Fig. 7** Estimation error of $\mathbf{G}$ with different SNR, where $L = 2$

RIS deployment scenarios—typically user-proximal and BS-distanced [1, 4]—received pilots frequently operate in low-SNR regimes. Therefore, the proposed PEA-based method demonstrates stronger suitability.

Using the proposed PEA-based method and the reference method, Fig. 8 demonstrates how the estimation error (NMSE) varies with the number of transmitting antennas $L$ when the SNR is -10, 0, and 10 dB. The experimental results show that the number of pilot transmitting antenna $L$ has a certain degree of influence on the estimation error. Specifically, as $L$ increases, the estimation error decreases. This is because a larger $L$ results in more overdetermined equations for CE. Furthermore, the impact of SNR on the estimation error is also significant. For instance, the curve representing the estimation error drops more rapidly at 10 dB compared to 0 dB, and it drops more slowly at -10 dB. Therefore, in practical applications, a reasonable selection of the number of pilot transmitting antennas is crucial. This involves making appropriate compromises between SNR, pilot overhead, and CE errors.

Optimizing this problem involves numerous parameters, such as antenna configurations, pilot overhead, channel conditions, beamforming algorithms, and so on. Some papers study this issue by examining the curve of estimation accuracy versus

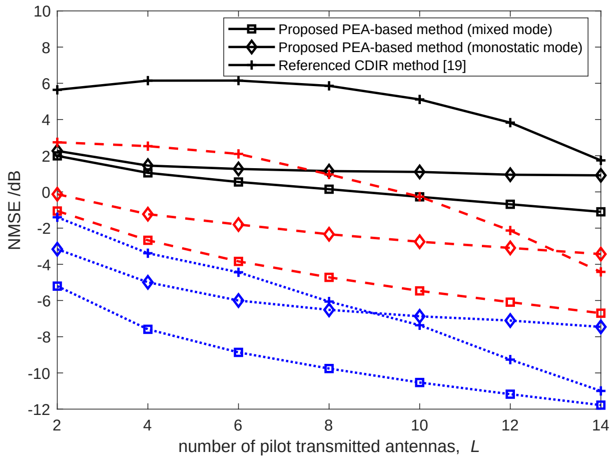


**Fig. 8** Estimation error of $\mathbf{G}$ with different number of transmitting antenna $L$.

pilot overhead or by evaluating the ultimate channel capacity [5, 15]. The former approach presents overly simplistic results and fails to reflect the weighting of various parameters, while the latter requires consideration of beamforming algorithms.

Given that this problem falls into multi-objective optimization, where CE error and pilot overhead form a pair of mutually conflicting objective functions, this paper employs a Pareto analysis based optimization method and presents numerical results for optimizing the selection of the number of BS and RIS antennas [30]. In the Pareto analysis, the channel conditions are fixed at typical values: the SNR of the received pilots is 10 dB, the ratio of coherent time is $\alpha = 2$, and the number of users is $K = 8$. Objective function 1 is the pilot overhead, which is a function of $N$, $M$, and $L$. Objective function 2 is the estimation error, represented by the numerical results of the proposed PEA-based optimization method. After applying the Pareto search algorithm, the optimization results are presented in Fig. 9 and Fig. 10. The Pareto frontier is illustrated with the horizontal axis representing the pilot overhead and the vertical axis representing the estimation error, along with the corresponding antenna configurations. In practical system design, considering factors such as overall system requirements and hardware implementation costs, an optimized antenna configuration is selected to ensure that both objective functions lie on the Pareto frontier.

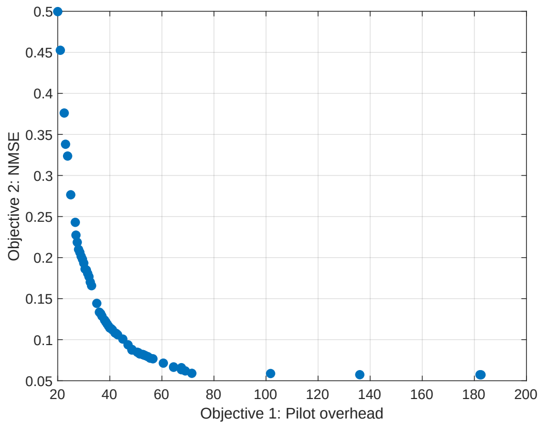


**Fig. 9** Pareto frontier (SNR=10 dB, $\alpha = 2$, $K = 8$).

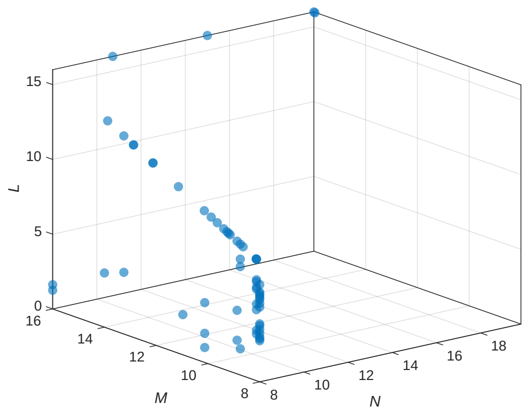


**Fig. 10** Parameter space (SNR=10 dB, $\alpha = 2$, $K = 8$).

In order to compare the estimation accuracy with other CE methods, Fig.11 shows the NMSE of the cascaded channel $\mathbf{C}_k$ with different SNR. We can see that the NMSE decreases as SNR becomes large except for the sparsity-based CE method [15, 16]. This is because the modified sparsity-based method and its modified OMP algorithm cannot achieve a meaningful CE when the sparsity is not available in Rician channels. The results show that the proposed PEA-based method obtains lower NMSE than the multi-user CE method and the referenced CDIR method. The MVU CE method in [7]

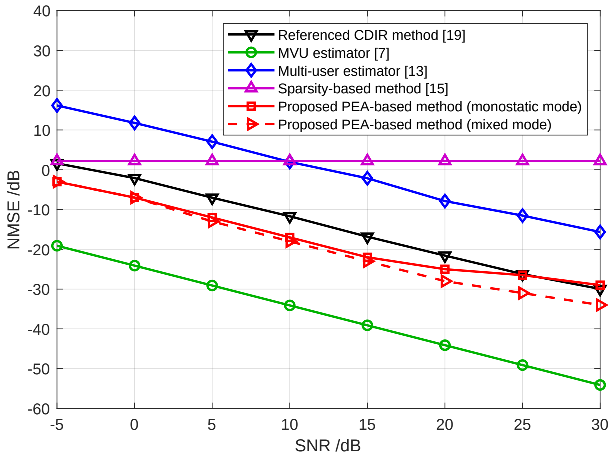

**Fig. 11** Estimation error of the cascaded channel $\mathbf{C}_k$ with different uplink SNR.

is more accurate. This is mainly because the pilot overhead is an order of magnitude higher than those of other methods.

# 7 Conclusion

In this paper, we propose an optimal scheme of two-timescale CE framework for RIS-assisted wireless system. The primary contribution includes reformulating the CE problem into a set of optimum equations for two full-duplex working modes, and introducing a novel PEA-based optimal algorithm. Simulations demonstrate the proposed algorithm has lower computation complexity and acceptable estimation error compared to the referenced CE method. In future work, we plan to explore the statistical properties of the BS-RIS channel and leverage machine learning methods to improve CE accuracy. Specifically, to explore the channel's statistical properties, we believe that integrating the sparse channel assumption into the system model and applying a sparse representation to the BS-RIS backscattered channels, followed by specific mathematical derivation and experimental validation of the PEA optimization method, represents a promising direction for future research.

## Declarations


- The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.
- The authors have no relevant financial or non-financial interests to disclose.
- The datasets generated during the current study are not publicly available but are available from the corresponding author on reasonable request.
- Kai She conceived the two-timescale CE framework optimization concept, formulated the rank-one nonlinear problem, and proposed the principal eigenvector approximation method. Kai She additionally implemented the semi-closed-form estimator algorithm and designed comparative simulations. All authors collaborated on the derivation, conducted critical analysis of optimization methods, and jointly wrote the manuscript through iterative revisions.


## References


[1] Wu, Q., Zhang, R.: Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network. IEEE Communications Magazine **58**(1), 106–112 (2020) https://doi.org/10.1109/MCOM.001.1900107

[2] ElMossallamy, M.A., Zhang, H., Song, L., Seddik, K.G., Han, Z., Li, G.Y.: Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities. IEEE Transactions on Cognitive Communications and Networking **6**(3), 990–1002 (2020) https://doi.org/10.1109/TCCN.2020.2992604

[3] Basar, E., Poor, H.V.: Present and future of reconfigurable intelligent surface-empowered communications [perspectives]. IEEE Signal Processing Magazine **38**(6), 146–152 (2021) https://doi.org/10.1109/MSP.2021.3106230

[4] Basar, E., Di Renzo, M., De Rosny, J., Debbah, M., Alouini, M.-S., Zhang, R.:

Wireless communications through reconfigurable intelligent surfaces. IEEE Access **7**, 116753–116773 (2019) https://doi.org/10.1109/ACCESS.2019.2935192

[5] Wei, X., Shen, D., Dai, L.: Channel estimation for RIS assisted wireless communications - Part I: Fundamentals, solutions, and future opportunities. IEEE Communications Letters **25**(5), 1398–1402 (2021) https://doi.org/10.1109/LCOMM.2021.3052822

[6] Swindlehurst, A.L., Zhou, G., Liu, R., Pan, C., Li, M.: Channel estimation with reconfigurable intelligent surfaces - A general framework. Proceedings of the IEEE **110**(9), 1312–1338 (2022) https://doi.org/10.1109/JPROC.2022.3170358

[7] Jensen, T.L., De Carvalho, E.: An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator. In: ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5000–5004 (2020). https://doi.org/10.1109/ICASSP40776.2020.9053695

[8] You, C., Zheng, B., Zhang, R.: Intelligent reflecting surface with discrete phase shifts: Channel estimation and passive beamforming. In: ICC 2020 - 2020 IEEE International Conference on Communications (ICC), pp. 1–6 (2020). https://doi.org/10.1109/ICC40277.2020.9149292

[9] Jin, Y., Zhang, J., Huang, C., Yang, L., Xiao, H., Ai, B., Wang, Z.: Multiple residual dense networks for reconfigurable intelligent surfaces cascaded channel estimation. IEEE Transactions on Vehicular Technology **71**(2), 2134–2139 (2022) https://doi.org/10.1109/TVT.2021.3132305

[10] Shen, W., Qin, Z., Nallanathan, A.: Deep learning for super-resolution channel estimation in reconfigurable intelligent surface aided systems. IEEE Transactions

on Communications **71**(3), 1491–1503 (2023) https://doi.org/10.1109/TCOMM.2023.3239621

[11] Ginige, N., Shashika Manosha, K.B., Rajatheva, N., Latva-aho, M.: Untrained DNN for channel estimation of RIS-assisted multi-user OFDM system with hardware impairments. In: 2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), pp. 561–566 (2021). https://doi.org/10.1109/PIMRC50174.2021.9569694

[12] He, Z.-Q., Yuan, X.: Cascaded channel estimation for large intelligent metasurface assisted massive MIMO. IEEE Wireless Communications Letters **9**(2), 210–214 (2020) https://doi.org/10.1109/LWC.2019.2948632

[13] Wang, Z., Liu, L., Cui, S.: Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis. IEEE Transactions on Wireless Communications **19**(10), 6607–6620 (2020) https://doi.org/10.1109/TWC.2020.3004330

[14] He, Z.-Q., Liu, H., Yuan, X., Zhang, Y.-J.A., Liang, Y.-C.: Semi-blind channel estimation for RIS-aided massive MIMO: A trilinear amp approach. In: 2021 IEEE International Symposium on Information Theory (ISIT), pp. 2822–2827 (2021). https://doi.org/10.1109/ISIT45174.2021.9518185

[15] Wang, P., Fang, J., Duan, H., Li, H.: Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems. IEEE Signal Processing Letters **27**, 905–909 (2020) https://doi.org/10.1109/LSP.2020.2998357

[16] Wei, X., Shen, D., Dai, L.: Channel estimation for RIS assisted wireless communications - Part II: An improved solution based on double-structured sparsity. IEEE Communications Letters **25**(5), 1403–1407 (2021) https://doi.org/10.1109/LCOMM.2021.3052787

[17] Yang, S., Lyu, W., Xiu, Y., Zhang, Z., Yuen, C.: Active 3D double-RIS-aided multi-user communications: Two-timescale-based separate channel estimation via Bayesian learning. IEEE Transactions on Communications **71**(6), 3605–3620 (2023) https://doi.org/10.1109/TCOMM.2023.3265115

[18] Zhou, G., Pan, C., Ren, H., Popovski, P., Swindlehurst, A.L.: Channel estimation for RIS-aided multiuser millimeter-wave systems. IEEE Transactions on Signal Processing **70**, 1478–1492 (2022) https://doi.org/10.1109/TSP.2022.3158024

[19] Hu, C., Dai, L., Han, S., Wang, X.: Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications. IEEE Transactions on Communications **69**(11), 7736–7747 (2021) https://doi.org/10.1109/TCOMM.2021.3072729

[20] Huang, H., Zhang, C., Zhang, Y., Ning, B., Gao, H., Fu, S., Qiu, K., Han, Z.: Two-timescale-based beam training for RIS-aided millimeter-wave multi-user MISO systems. IEEE Transactions on Vehicular Technology **72**(9), 11884–11897 (2023) https://doi.org/10.1109/TVT.2023.3269153

[21] Peng, Z., Pan, C., Zhou, G., Ren, H., Jin, S., Popovski, P., Schober, R., You, X.: Two-stage channel estimation for RIS-aided multiuser mmwave systems with reduced error propagation and pilot overhead. IEEE Transactions on Signal Processing **71**, 3607–3622 (2023) https://doi.org/10.1109/TSP.2023.3317729

[22] Kim, S., Wu, J., Shim, B.: Efficient channel probing and phase shift control for mmwave reconfigurable intelligent surface-aided communications. IEEE Transactions on Wireless Communications **23**(1), 231–246 (2024) https://doi.org/10.1109/TWC.2023.3276975

[23] You, Y., Chen, F., Zhang, L., Huang, Y., Zhang, C.: Adaptive channel estimation

for RIS-assisted systems in time-varying mmwave channels. IEEE Transactions on Communications, 1–1 (2024) https://doi.org/10.1109/TCOMM.2024.3425592

[24] Jin, W., Zhang, J., Wen, C.-K., Jin, S.: Amplitude correlation and structured sparsity inspired compressed sensing for channel estimation in RIS-aided MU-MISO systems. IEEE Transactions on Wireless Communications, 1–1 (2025) https://doi.org/10.1109/TWC.2025.3592787

[25] Mishra, D., Larsson, E.G.: Optimal channel estimation for reciprocity-based backscattering with a full-duplex MIMO reader. IEEE Transactions on Signal Processing **67**(6), 1662–1677 (2019) https://doi.org/10.1109/TSP.2019.2893859

[26] Delestre, F., Owojaiye, G., Sun, Y.: Efficient space-frequency block coded pilot-aided channel estimation method for multiple-input-multiple-output orthogonal frequency division multiplexing systems over mobile frequency-selective fading channels. IET Communications **8**(6), 841–851 (2014) https://doi.org/10.1049/iet-com.2013.0068

[27] Shen, H., Ding, T., Xu, W., Zhao, C.: Beamformig design with fast convergence for IRS-aided full-duplex communication. IEEE Communications Letters **24**(12), 2849–2853 (2020) https://doi.org/10.1109/LCOMM.2020.3014201

[28] Ahmed, E., Eltawil, A.M.: All-digital self-interference cancellation technique for full-duplex systems. IEEE Transactions on Wireless Communications **14**(7), 3519–3532 (2015) https://doi.org/10.1109/TWC.2015.2407876

[29] Masmoudi, A., Le-Ngoc, T.: Channel estimation and self-interference cancelation in full-duplex communication systems. IEEE Transactions on Vehicular Technology **66**(1), 321–334 (2017) https://doi.org/10.1109/TVT.2016.2540538

[30] Le, T.A., Van Chien, T., Nakhai, M.R., Le-Ngoc, T.: Pareto-optimal pilot design

for cellular massive mimo systems. IEEE Transactions on Vehicular Technology **69**(11), 13206–13215 (2020) https://doi.org/10.1109/TVT.2020.3021766