On Channel Estimation Overhead in Beyond Diagonal RIS Aided Communication

Liang LIU

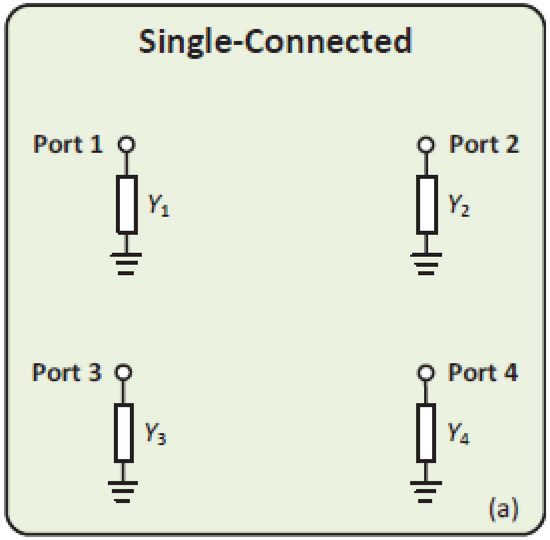
Department of Electrical and Electronic Engineering
The Hong Kong Polytechnic University

Reconfigurable Intelligent Surface

- Reconfigurable intelligent surface (RIS) is powerful technique to wireless systems
- Via properly designing RIS scattering strategy, user rates can be significantly improved
- Limitation: each RIS atom can merely affect its own signals

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \phi_M \end{bmatrix}, \quad |\phi_m|^2 = 1, m = 1, \dots, M$$

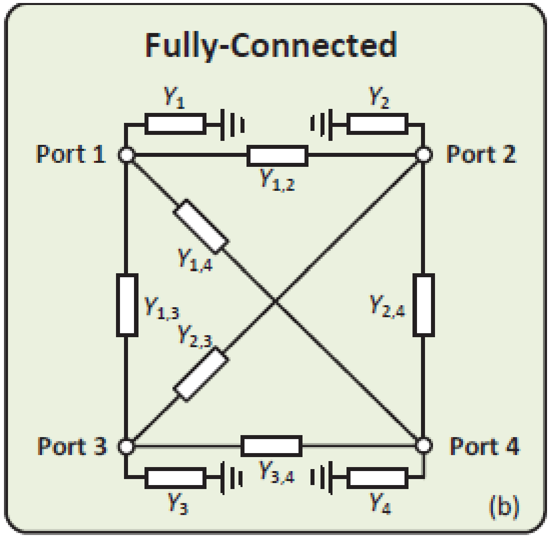
- Can we have more design freedom



Beyond Diagonal RIS

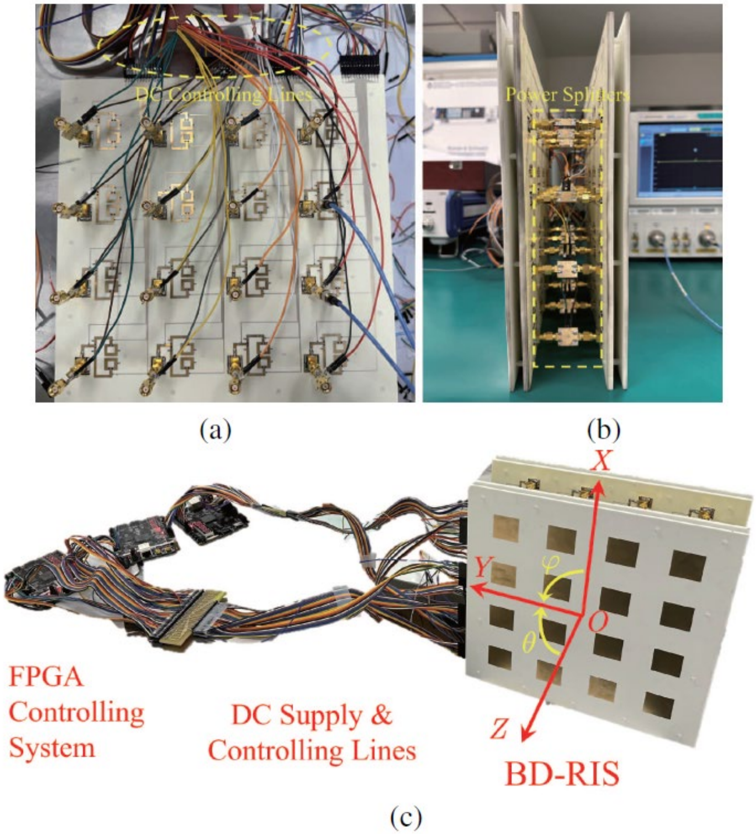
- ❑ RIS 2.0: Beyond diagonal RIS (BD-RIS)
- ❑ RIS atoms are connected via circuits
  - ➢ Signals received by one atom will flow to other atoms and affect their scattered signals
  - ➢ Non-diagonal scattering matrix

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi_{1,1} & \cdots & \phi_{1,M} \\ \vdots & \ddots & \vdots \\ \phi_{M,1} & \cdots & \phi_{M,M} \end{bmatrix}, \quad \boldsymbol{\Phi}^H \boldsymbol{\Phi} = \boldsymbol{I}$$

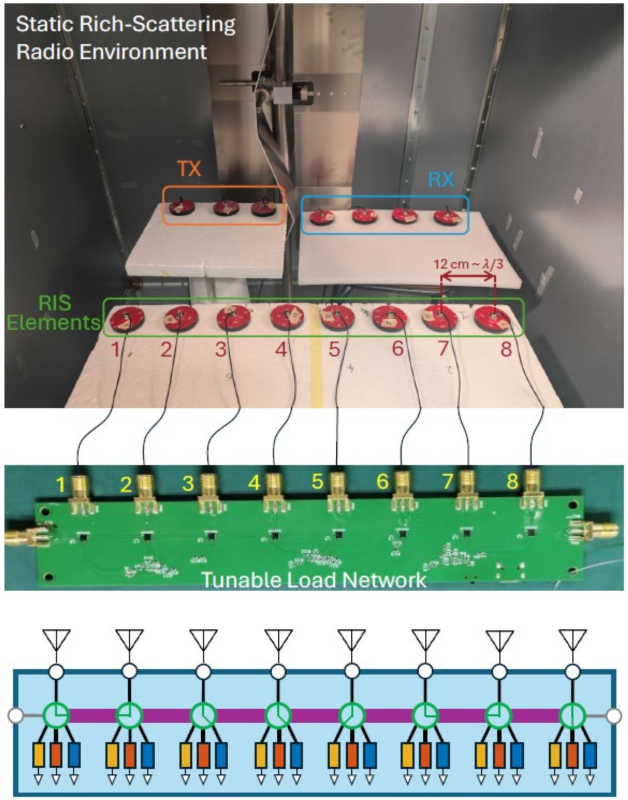


[1] H. Li, M. Nerini, S. Shen, and B. Clerckx, "A tutorial on beyond-diagonal reconfigurable intelligent surfaces: Modeling, architectures, system design and optimization, and applications". [Online] https://arxiv.org/abs/2505.16504

Beyond Diagonal RIS Prototypes



[2] Z. Ming, S. Shen, J. Rao, Z. Li, J. Zhang, C. Y. Chiu, and R. Murch, "A hybrid transmitting and reflecting beyond diagonal reconfigurable intelligent surface with independent beam control and power splitting," 2025. [Online] Available: https://arxiv.org/abs/2504.09618



[3] J. Tapie, M. Nerini, B. Clerckx, and P. del Hougne, "Beyond-diagonal RIS prototype and performance evaluation," 2025. [Online] Available: https://arxiv.org/abs/2505.13392

Beyond Diagonal RIS

- Scattering matrix of conventional RIS

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \phi_M \end{bmatrix}, \quad |\phi_m|^2 = 1, m = 1, \dots, M$$

- Scattering matrix of fully connected BD-RIS

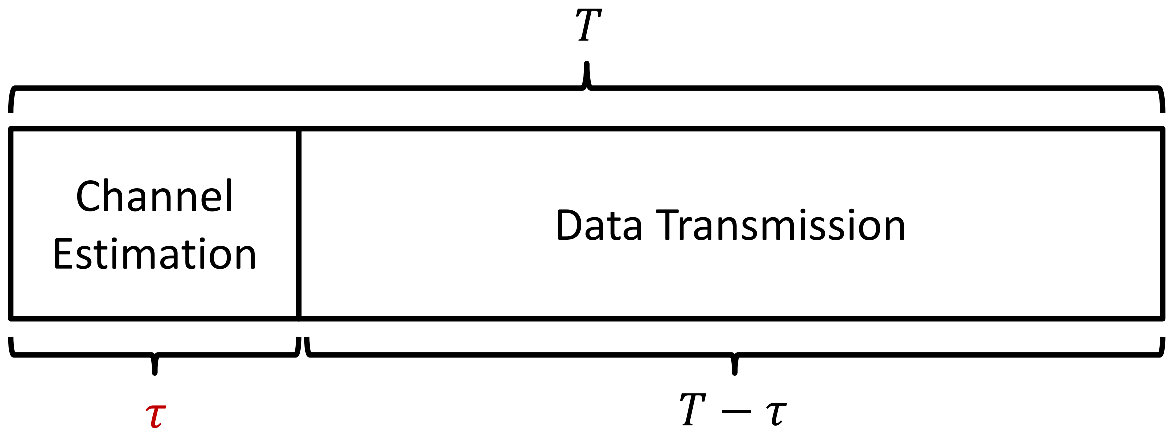
$$\boldsymbol{\Phi} = \begin{bmatrix} \phi_{1,1} & \cdots & \phi_{1,M} \\ \vdots & \ddots & \vdots \\ \phi_{M,1} & \cdots & \phi_{M,M} \end{bmatrix}, \quad \boldsymbol{\Phi}^H \boldsymbol{\Phi} = \boldsymbol{I}$$

- More design freedom for BD-RIS

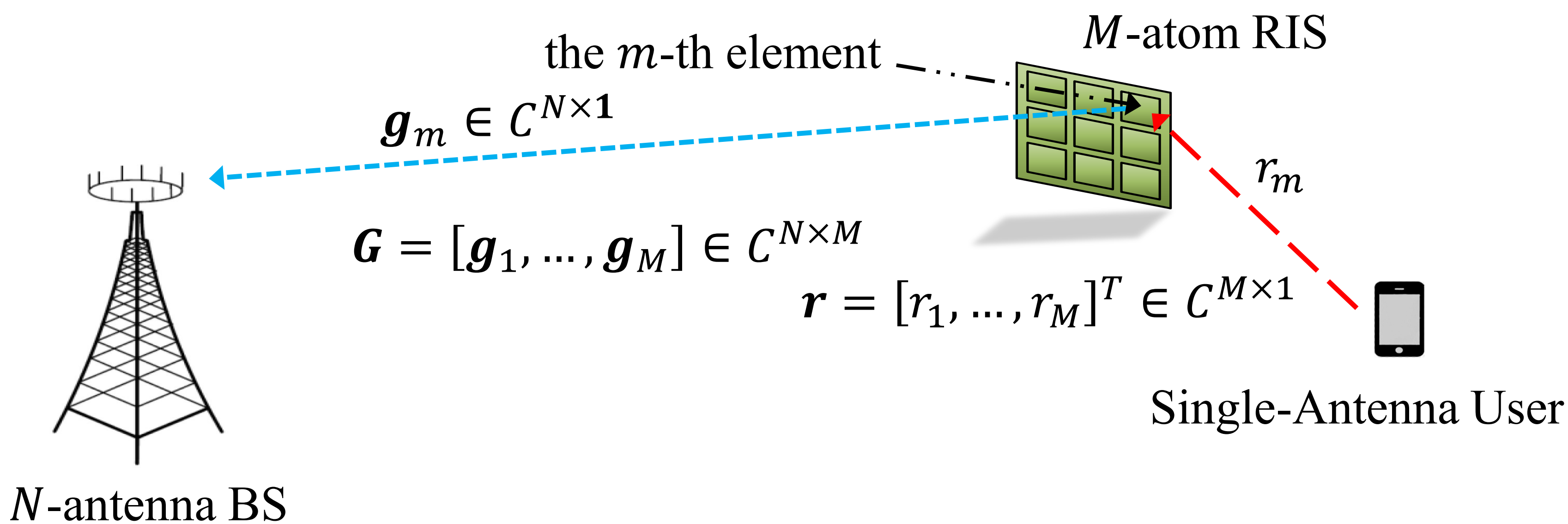
BD-RIS > conventional RIS? Not necessarily

Challenge of BD-RIS: Channel Estimation

- ❑ How much data can be transmitted depends on
  - ➢ Data transmission rate (bits/second)
  - ➢ Data transmission duration (second)
- ❑ There are many lessions about channel estimation overhead
  - ➢ Downlink massive MIMO in FDD systems
- ❑ Pros and cons of BD-RIS v.s. conventional RIS
  - ➢ Higher beamforming gain
  - ➢ Higher channel estimation overhead



Example on Channel Estimation Overhead



❑ Consider RIS-aided single-user uplink communication

- One BS with N antennas
- One single-antenna user
- One RIS with M atoms

Channel Estimation with Conventional RIS

- ❑ Received signals

$$\boldsymbol{y}_t = \boldsymbol{G}\boldsymbol{\Phi}_t\boldsymbol{r}\sqrt{p}a_t + \boldsymbol{z}_t = \sum_{m=1}^{M}\boldsymbol{h}_m\phi_{m,t}\sqrt{p}a_t + \boldsymbol{z}_t, t = 1, \dots, T$$

  - ➢ $\boldsymbol{h}_m = r_m\boldsymbol{g}_m \in C^{N\times 1}, m = 1, \dots, M$

- ❑ Channels that are needed for beamforming design

$$\boldsymbol{h}_1 = r_1\boldsymbol{g}_1, \dots, \boldsymbol{h}_M = r_M\boldsymbol{g}_M$$

- ❑ Number of unknowns in these channels: MN
- ❑ Number of linear equations: NT
- ❑ Minimum overhead to estimate channels without noise

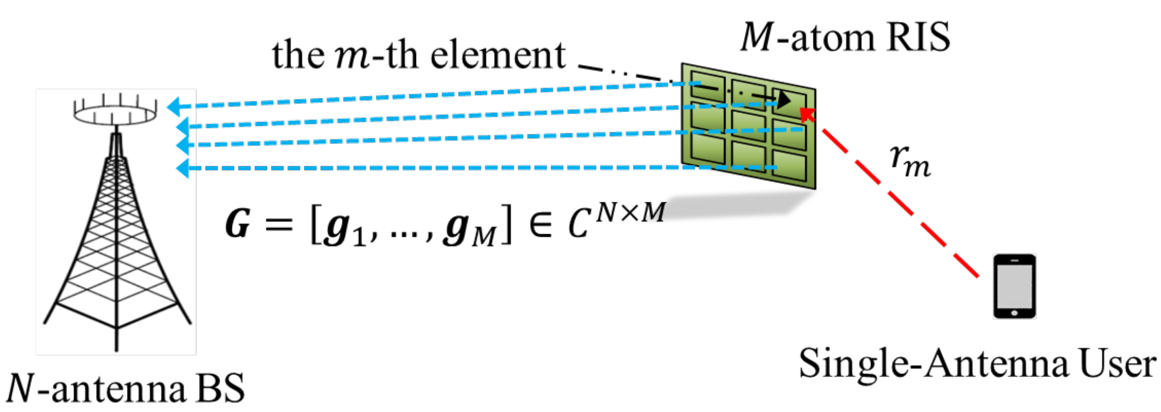
$$T = M$$

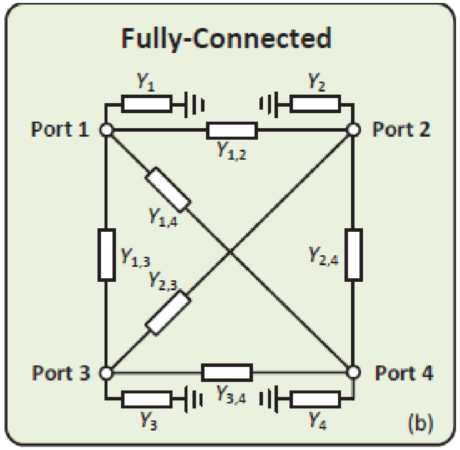
Channel Estimation with Fully Connected BD-RIS

❑ Received signals

$$\boldsymbol{y}_t = \boldsymbol{G}\boldsymbol{\Phi}_t\boldsymbol{r}\sqrt{p}a_t + \boldsymbol{z}_t = \sum_{m=1}^{M} \boldsymbol{Q}_m\boldsymbol{\phi}_{t,m}\sqrt{p}a_t + \boldsymbol{z}_t, t = 1, \dots, T$$

➢ $\boldsymbol{Q}_m = r_m\boldsymbol{G} \in C^{N\times M}, m = 1, \dots, M$, and $\boldsymbol{\Phi}_t = [\boldsymbol{\phi}_{t,1}, \dots, \boldsymbol{\phi}_{t,M}]$





Need to estimate $r_m\boldsymbol{g}_{m-1}$ ...

Channel Estimation with Fully Connected BD-RIS

- Received signals

$$\boldsymbol{y}_t = \boldsymbol{G}\boldsymbol{\Phi}_t\boldsymbol{r}\sqrt{p}a_t + \boldsymbol{z}_t = \sum_{m=1}^{M} \boldsymbol{Q}_m \boldsymbol{\phi}_{t,m}\sqrt{p}a_t + \boldsymbol{z}_t, t = 1, \dots, T$$

  - $\boldsymbol{Q}_m = r_m\boldsymbol{G} \in C^{N\times M}, m = 1, \dots, M$, and $\boldsymbol{\Phi}_t = [\boldsymbol{\phi}_{t,1}, \dots, \boldsymbol{\phi}_{t,M}]$
- Channels that are needed for beamforming design

$$\boldsymbol{Q}_1 = r_1\boldsymbol{G}, \dots, \boldsymbol{Q}_M = r_M\boldsymbol{G}$$

- Number of unknowns in these channels: NM^2
- Number of linear equations: NT
- Overhead to estimate channels without noise [4]

$$T = M^2 \gg M$$

Is this the best we can do?

[4] H. Li, S. Shen, Y. Zhang, and B. Clerckx, "Channel estimation and beamforming for beyond diagonal reconfigurable intelligent surfaces," *IEEE Trans. Signal Process.*, vol. 72, pp. 3318–3332, Jul. 2024.

Objective of Our Talk

- Numbers of RIS atoms and BS antennas: M and N
- Numbers of users and antennas at each user: K and U
- Rank of BS-RIS channel: $q = \mathrm{rank}(\boldsymbol{G})$ (to be estimated)

$$\boldsymbol{G} = [\boldsymbol{g}_1, \dots, \boldsymbol{g}_M]$$

Types of RISs	Channel Estimation Overhead
BD-RIS	$2M + \lceil M(KU-1)/q \rceil$ [5]
BD-RIS	KUM^2 [4]
Conventional RIS	$M + \lceil M(KU-1)/q \rceil$ [6]

BD-RIS > conventional RIS? Yes

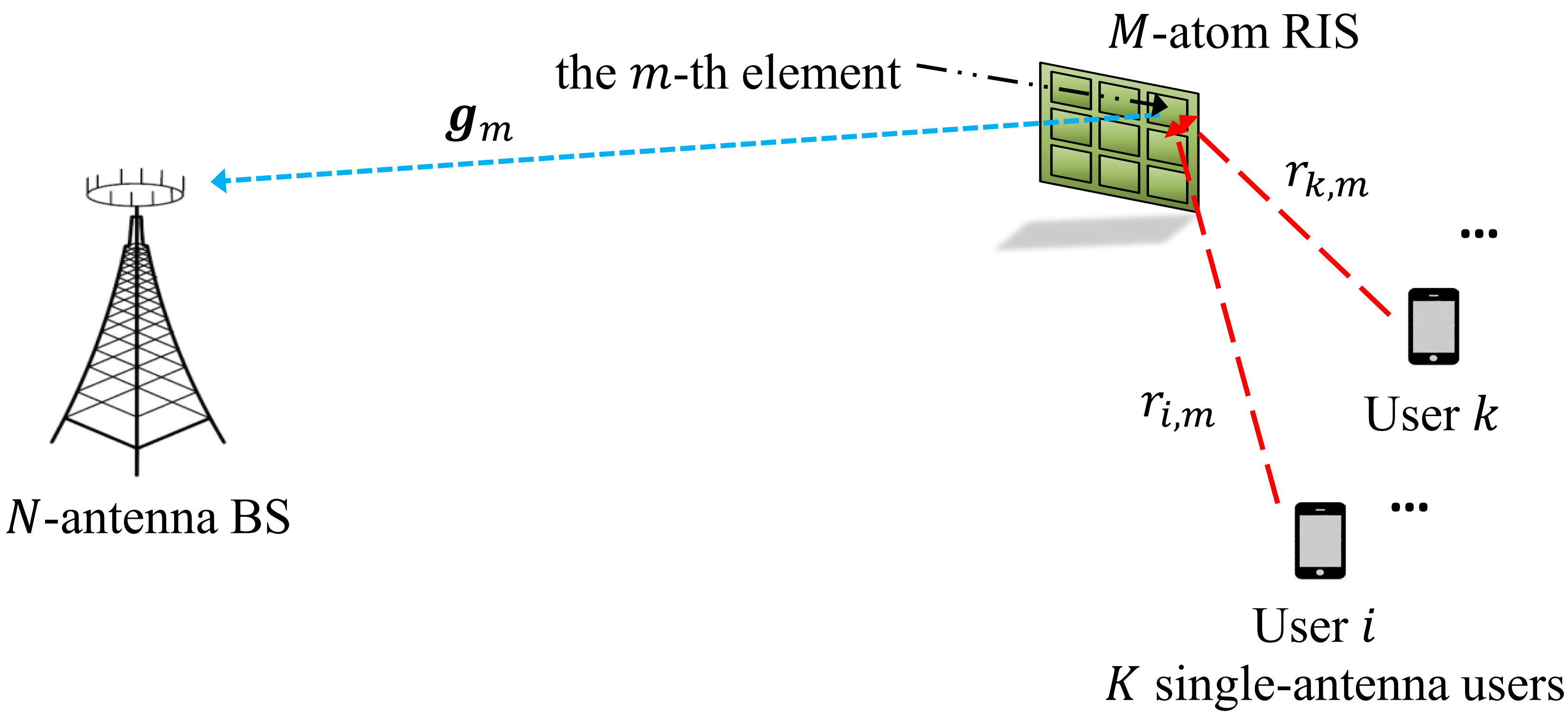
[5] R. Wang, S. Zhang, B. Clerckx, and **L. Liu**, "Low-overhead channel estimation framework for beyond diagonal reconfigurable intelligent surface assisted multi-user MIMO communication," under major revision, *IEEE Trans. Signal Process.*, 2025. [Online] Available: https://arxiv.org/abs/2504.10911
[6] Z. Wang, **L. Liu**, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6607-6620, Oct. 2020.

Outline

- ❑ **Part I: Conventional RIS [6] (Brief Overview)**
- ❑ Part II: Fully Connected BD-RIS [5]

[5] R. Wang, S. Zhang, B. Clerckx, and **L. Liu**, "Low-overhead channel estimation framework for beyond diagonal reconfigurable intelligent surface assisted multi-user MIMO communication," under major revision, *IEEE Trans. Signal Process.*, 2025. [Online] Available: https://arxiv.org/abs/2504.10911
[6] Z. Wang, **L. Liu**, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6607-6620, Oct. 2020.

Signal Model in Channel Estimation Phase



- Pilot signal received by BS at time instant t (ignore noise)

$$\boldsymbol{y}_t = \sum_{k=1}^{K} \sum_{m=1}^{M} \boldsymbol{h}_{k,m} \phi_{m,t} \sqrt{p} a_{k,t}, t = 1, \ldots, T.$$

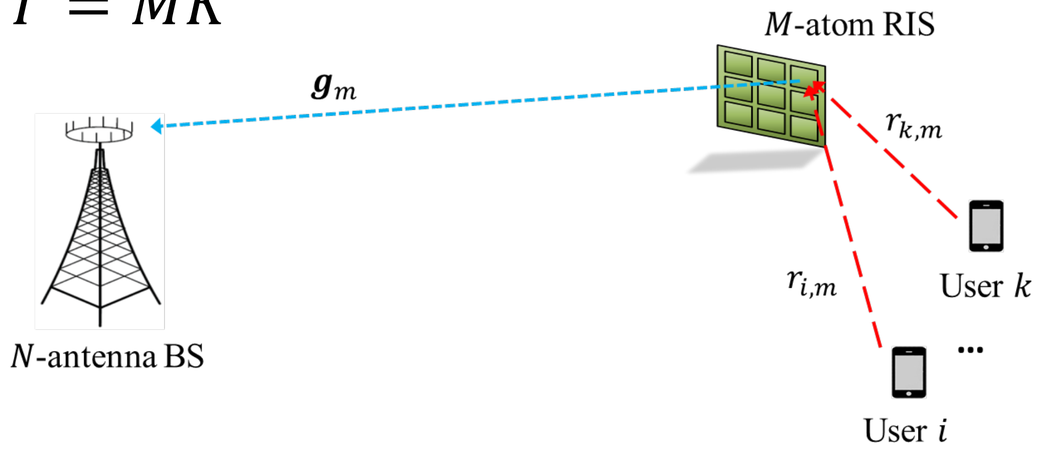
$$\boldsymbol{h}_{k,m} = r_{k,m} \boldsymbol{g}_m \in C^{N \times 1}, \ \forall k, m$$

Discussion on Channel Estimation Overhead

- Channel estimation overhead under conventional method

$$\boldsymbol{y}_t = \sum_{k=1}^{K}\sum_{m=1}^{M} \boldsymbol{h}_{k,m}\phi_{m,t}\sqrt{p}a_{k,t} \in C^{N\times 1}, t = 1, \dots, T.$$

  - Number of linear equations: NT
  - Number of unknowns in $\boldsymbol{h}_{k,m} \in C^{N\times 1}, \forall k, m$: KMN
  - Minimum overhead: $T = MK$



What is missing here?

$\boldsymbol{h}_{1,m} = r_{1,m}\boldsymbol{g}_m, \dots, \boldsymbol{h}_{K,m} = r_{K,m}\boldsymbol{g}_m$ are correlated over k

Signal Model in Channel Estimation Phase

- Relation among cascaded channels

$$\boldsymbol{h}_{1,m} = r_{1,m}\boldsymbol{g}_m, \forall m$$
$$\boldsymbol{h}_{k,m} = r_{k,m}\boldsymbol{g}_m, \forall k \geq 2, \forall m$$

⇒

$$\boldsymbol{h}_{k,m} = \beta_{k,m}\boldsymbol{h}_{1,m}, \forall k \geq 2, \forall m$$
where $\beta_{k,m} = \frac{r_{k,m}}{r_{1,m}}, \forall k \geq 2, \forall m$

- **Independent** channel coefficients to be estimated

$$\boldsymbol{h}_{1,m}, m = 1, \ldots, M,$$
$$\boldsymbol{h}_{k,m}, k = 2, \ldots, K, \forall m$$

⇒

$$\boldsymbol{h}_{1,m}, m = 1, \ldots, M$$
$$\beta_{k,m}, k = 2, \ldots, K, m = 1, \ldots, M$$

- Number of unknowns

$$KMN \Rightarrow MN + (K-1)M$$

A New Look at Received Pilot Signals

- Received pilot signals

$$\boldsymbol{y}_t = \sum_{k=1}^{K} \sum_{m=1}^{M} \boldsymbol{h}_{k,m} \phi_{m,t} \sqrt{p} a_{k,t}$$

$$= \sum_{m=1}^{M} \boldsymbol{h}_{1,m} \phi_{m,t} \sqrt{p} a_{1,t} + \sum_{k=2}^{K} \sum_{m=1}^{M} \beta_{k,m} \boldsymbol{h}_{1,m} \phi_{m,t} \sqrt{p} a_{k,t}\,, t = 1, \ldots, T$$

- Goal: estimate channels with minimum number of equations
- Pathway: design RIS scattering and user pilot
- Challenge: non-linear functions

A New Look at Received Pilot Signals

- Received pilot signals

$$\boldsymbol{y}_t = \sum_{k=1}^{K} \sum_{m=1}^{M} \boldsymbol{h}_{k,m} \phi_{m,t} \sqrt{p} a_{k,t}$$

$$= \sum_{m=1}^{M} \boldsymbol{h}_{1,m} \phi_{m,t} \sqrt{p} a_{1,t} + \sum_{k=2}^{K} \sum_{m=1}^{M} \beta_{k,m} \boldsymbol{h}_{1,m} \phi_{m,t} \sqrt{p} a_{k,t}\,, t = 1, \ldots, T$$

- Our approach: sequential estimation

A New Look at Received Pilot Signals

- Received pilot signals

$$\boldsymbol{y}_t = \sum_{k=1}^{K} \sum_{m=1}^{M} \boldsymbol{h}_{k,m} \phi_{m,t} \sqrt{p} a_{k,t}$$

$$= \sum_{m=1}^{M} \boldsymbol{h}_{1,m} \phi_{m,t} \sqrt{p} a_{1,t} + \sum_{k=2}^{K} \sum_{m=1}^{M} \beta_{k,m} \boldsymbol{h}_{1,m} \phi_{m,t} \sqrt{p} a_{k,t}, t = 1, \dots, T$$

- Our approach: sequential estimation
  - Phase 1: only user one transmits (linear in $\boldsymbol{h}_{1,m}$)

A New Look at Received Pilot Signals

- Received pilot signals

$$\boldsymbol{y}_t = \sum_{k=1}^{K}\sum_{m=1}^{M} \boldsymbol{h}_{k,m}\phi_{m,t}\sqrt{p}a_{k,t}$$

$$= \sum_{m=1}^{M} \boldsymbol{h}_{1,m}\phi_{m,t}\sqrt{p}a_{1,t} + \sum_{k=2}^{K}\sum_{m=1}^{M} \beta_{k,m}\boldsymbol{h}_{1,m}\phi_{m,t}\sqrt{p}a_{k,t}, t = 1, \dots, T$$

- Our approach: sequential estimation
  - Phase 1: only user one transmits (linear in $\boldsymbol{h}_{1,m}$)
  - Phase 2: other users transmit (linear in $\beta_{k,m}$ given $\boldsymbol{h}_{1,m}$)

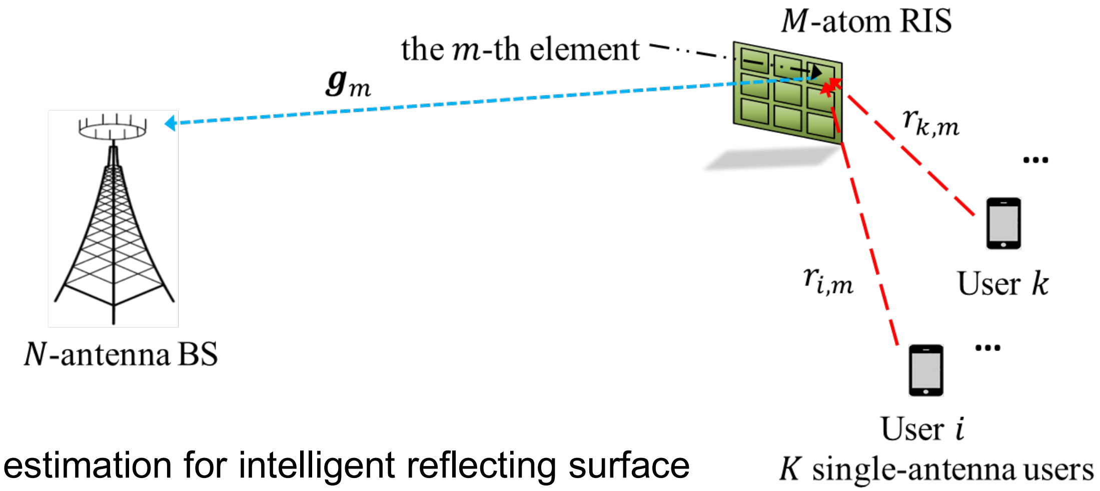
Main Result

- Under our approach, channel estimation overhead is

$$M + \left\lceil \frac{M(K-1)}{q} \right\rceil \ll KM, q = \mathrm{rank}([\boldsymbol{g}_1, \dots, \boldsymbol{g}_M])$$

- Generalization: U-antenna users

$$M + \left\lceil \frac{M(KU-1)}{q} \right\rceil$$



[6] Z. Wang, **L. Liu**, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6607-6620, Oct. 2020.

Outline

- ❑ Part I: Conventional RIS [6] (Brief Overview)
- ❑ **Part II: Fully Connected BD-RIS [5]**

[5] R. Wang, S. Zhang, B. Clerckx, and **L. Liu**, "Low-overhead channel estimation framework for beyond diagonal reconfigurable intelligent surface assisted multi-user MIMO communication," under major revision, *IEEE Trans. Signal Process.*, 2025. [Online] Available: https://arxiv.org/abs/2504.10911

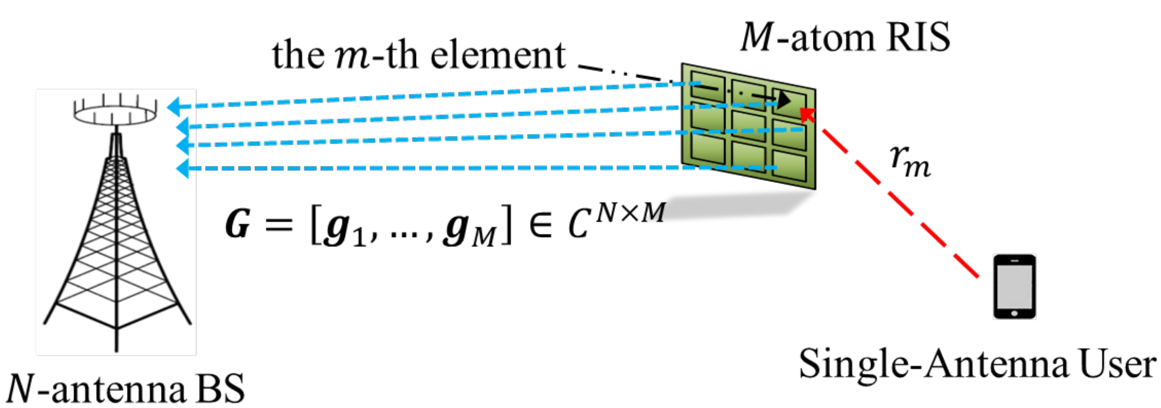
[6] Z. Wang, **L. Liu**, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6607-6620, Oct. 2020.

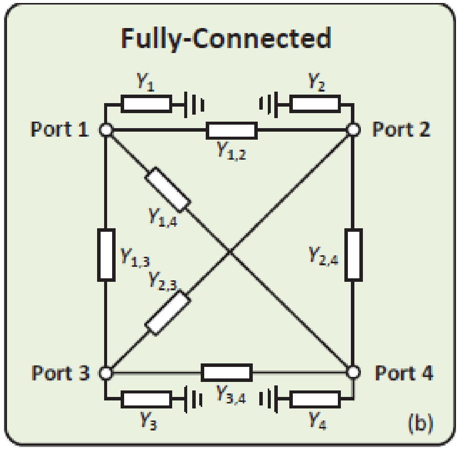
One Single-Antenna User Case

- Received signals with a single-antenna user

$$\boldsymbol{y}_t = \boldsymbol{G}\boldsymbol{\Phi}_t\boldsymbol{r}\sqrt{p}a_t = \sum_{m=1}^{M} \boldsymbol{Q}_m\boldsymbol{\phi}_{t,m}\sqrt{p}a_t, t = 1, \dots, T$$

  - $\boldsymbol{Q}_m = r_m\boldsymbol{G} \in C^{N\times M}, m = 1, \dots, M$, and $\boldsymbol{\Phi}_t = [\boldsymbol{\phi}_{t,1}, \dots, \boldsymbol{\phi}_{t,M}]$





One Single-Antenna User Case

- Received signals with a single-antenna user

$$\boldsymbol{y}_t = \boldsymbol{G}\boldsymbol{\Phi}_t\boldsymbol{r}\sqrt{p}a_t = \sum_{m=1}^{M} \boldsymbol{Q}_m\boldsymbol{\phi}_{t,m}\sqrt{p}a_t, t = 1, \ldots, T$$

  - $\boldsymbol{Q}_m = r_m\boldsymbol{G} \in C^{N\times M}, m = 1, \ldots, M$, and $\boldsymbol{\Phi}_t = [\boldsymbol{\phi}_{t,1}, \ldots, \boldsymbol{\phi}_{t,M}]$

- Channels that are needed for beamforming design

$$\boxed{\boldsymbol{Q}_1 = r_1\boldsymbol{G}, \ldots, \boldsymbol{Q}_M = r_M\boldsymbol{G}}$$

- Number of unknowns in these channels: NM^2
- Number of linear equations: NT
- Minimum overhead to directly estimate $\boldsymbol{Q}_m, m = 1, \ldots, M$

$$T = M^2 \gg M$$

Is this the best we can do?

Signal Model in Channel Estimation Phase

- Relation among cascaded channels

$$\boldsymbol{Q}_1 = r_1 \boldsymbol{G}$$
$$\boldsymbol{Q}_m = r_m \boldsymbol{G}, \forall m \geq 2$$

$$\boldsymbol{Q}_m = \beta_m \boldsymbol{Q}_1, \forall m \geq 2$$
where $\beta_m = \frac{r_m}{r_1}, \forall m \geq 2$

- **Independent** channel coefficients to be estimated

$\boldsymbol{Q}_1$
$\boldsymbol{Q}_m, m = 2, \dots, M$

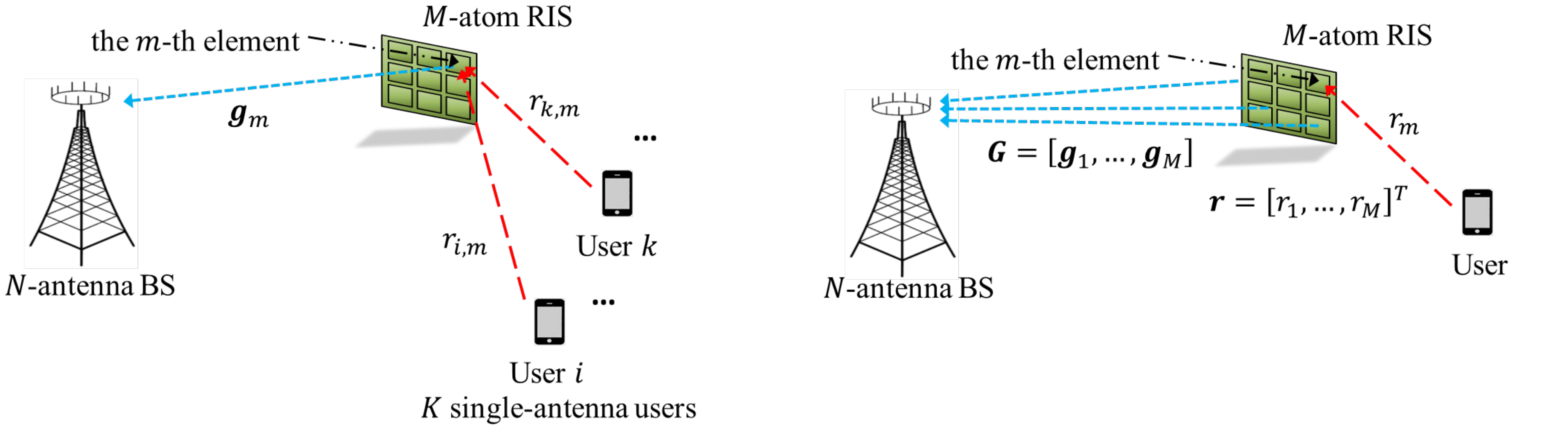
$\boldsymbol{Q}_1$
$\beta_m, m = 2, \dots, M$

- Number of unknowns

NM^2 → $MN + M - 1$

Similarity Between Conventional and BD RIS

- ❑ Multi-user communication with conventional RIS
  - ➢ $\boldsymbol{g}_m$ common in estimating $\boldsymbol{h}_{1,m} = r_{1,m}\boldsymbol{g}_m, \ldots, \boldsymbol{h}_{K,m} = r_{K,m}\boldsymbol{g}_m$
- ❑ Single-user communication with BD-RIS
  - ➢ $\boldsymbol{G}$ common in estimating $\boldsymbol{Q}_1 = r_1\boldsymbol{G}, \ldots, \boldsymbol{Q}_M = r_M\boldsymbol{G}$



(a) Conventional RIS

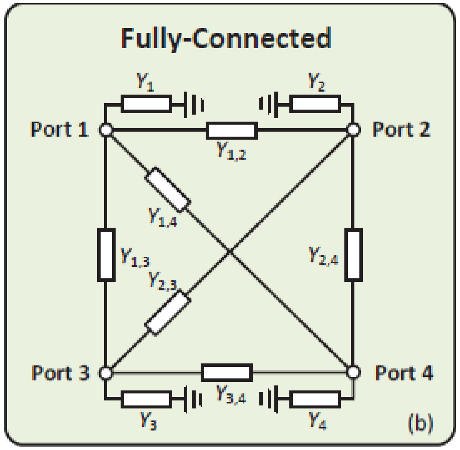
(b) BD-RIS

A New Look at Received Pilot Signals

- Received pilot signals

$$\boldsymbol{y}_t = \sum_{m=1}^{M} \boldsymbol{Q}_m \boldsymbol{\phi}_{t,m} \sqrt{p} a_t = \boldsymbol{Q}_1 \boldsymbol{\phi}_{t,1} \sqrt{p} a_t + \sum_{m=2}^{M} \beta_m \boldsymbol{Q}_1 \boldsymbol{\phi}_{t,m} \sqrt{p} a_t$$

- Goal: estimate channels with minimum number of equations
- Pathway: design RIS scattering and user pilot
- Challenge 1: non-linear functions
- Challenge 2: $\boldsymbol{\Phi}^H \boldsymbol{\Phi} = \boldsymbol{I}$ (new challenge)



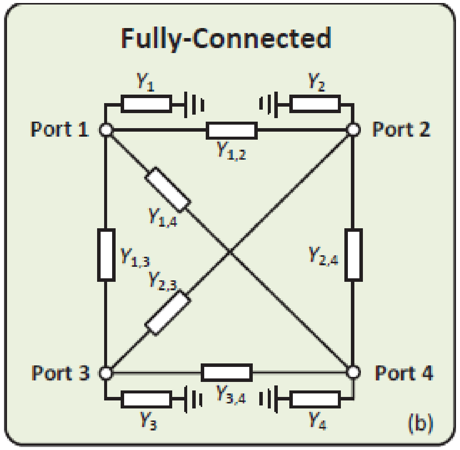
A New Look at Received Pilot Signals

- ❑ Received pilot signals

$$\boldsymbol{y}_t = \sum_{m=1}^{M} \boldsymbol{Q}_m \boldsymbol{\phi}_{t,m} \sqrt{p} a_t = \boldsymbol{Q}_1 \boldsymbol{\phi}_{t,1} \sqrt{p} a_t + \sum_{m=2}^{M} \beta_m \boldsymbol{Q}_1 \boldsymbol{\phi}_{t,m} \sqrt{p} a_t$$

- ❑ Dose the following sequential estimation approach work?
  - ➢ Phase 1: $\boldsymbol{\phi}_{t,1} \neq \boldsymbol{0}, \boldsymbol{\phi}_{t,2} = \cdots = \boldsymbol{\phi}_{t,M} = \boldsymbol{0}$ (linear in $\boldsymbol{Q}_1$)
    - ▪ Signals of atoms $2-M$ are not reflected
  - ➢ Phase 2: signals of all atoms are reflected (linear in β_m given $\boldsymbol{Q}_1$)

- ▪ $[\boldsymbol{\phi}_{t,1}, \boldsymbol{0}, \dots, \boldsymbol{0}]$ is not unitary matrix
- ▪ No BD-RIS circuit can achieve this



Our Approach

- Main result: $T = 2M$ time slots are sufficient
- Divide into two blocks, each with M time slots
  - Block 1: at each time slot $t = 1, \ldots, M$
    - User pilot: arbitrary a_t with $|a_t| = 1$
    - RIS scattering: $\boldsymbol{\phi}_{t,m} = \boldsymbol{p}_{((m+t-1) \bmod M)+1}, m = 1, \ldots, M$

  where $\boldsymbol{P} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_M] \in C^{M \times M}$ is an arbitrary unitary matrix

⇓

$$\boldsymbol{\Phi}_1 = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_M], \boldsymbol{\Phi}_2 = [\boldsymbol{p}_2, \ldots, \boldsymbol{p}_M, \boldsymbol{p}_1], \ldots, \boldsymbol{\Phi}_M = [\boldsymbol{p}_M, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_{M-1}]$$

$t = 1$ $\qquad$ $t = 2$ $\qquad$ $t = M$

⇓

$$\boldsymbol{\Phi}_1^H \boldsymbol{\Phi}_1 = \boldsymbol{I}, \ldots, \boldsymbol{\Phi}_M^H \boldsymbol{\Phi}_M = \boldsymbol{I}$$

Our Approach

- Main result: $T = 2M$ time slots are sufficient
- Divide into two blocks, each with M time slots
  - Block 2: at each time slot $t = M + 1, \dots, 2M$
    - User pilot: $a_t = a_{t-M}$
    - RIS scattering:

$$\boldsymbol{\phi}_{t,m} = \begin{cases} e^{j\theta} \boldsymbol{\phi}_{t-M,1} & m = 1 \\ \boldsymbol{\phi}_{t-M,m} & m = 2, \dots, M \end{cases}$$ where $\theta \in (0, 2\pi)$

⇩

$$\underset{t = M+1}{\boldsymbol{\Phi}_{M+1} = \left[e^{j\theta}\boldsymbol{p}_1, \dots, \boldsymbol{p}_M\right]}, \dots, \underset{t = 2M}{\boldsymbol{\Phi}_{2M} = \left[e^{j\theta}\boldsymbol{p}_M, \boldsymbol{p}_1, \dots, \boldsymbol{p}_{M-1}\right]}$$

⇩

$$\boldsymbol{\Phi}_{M+1}^H \boldsymbol{\Phi}_{M+1} = \boldsymbol{I}, \dots, \boldsymbol{\Phi}_{2M}^H \boldsymbol{\Phi}_{2M} = \boldsymbol{I}$$

Our Approach

- Main result: $T = 2M$ time slots are sufficient
- Divide into two blocks, each with M time slots
  - Received signals at Block 1

$$\boldsymbol{y}_t = \boldsymbol{Q}_1 \boldsymbol{\phi}_{t,1} \sqrt{p} a_t + \boxed{\sum_{m=2}^{M} \beta_m \boldsymbol{Q}_1 \boldsymbol{\phi}_{t,m} \sqrt{p} a_t}, t = 1, \ldots, M$$

  - Received signals at Block 2

$$\boldsymbol{y}_{M+t} = \boldsymbol{Q}_1 \boldsymbol{\phi}_{M+t,1} \sqrt{p} a_{M+t} + \sum_{m=2}^{M} \beta_m \boldsymbol{Q}_1 \boldsymbol{\phi}_{M+t,m} \sqrt{p} a_{M+t},$$

$$= e^{j\theta} \boldsymbol{Q}_1 \boldsymbol{\phi}_{t,1} \sqrt{p} a_t + \boxed{\sum_{m=2}^{M} \beta_m \boldsymbol{Q}_1 \boldsymbol{\phi}_{t,m} \sqrt{p} a_t}, t = 1, \ldots, M$$

Estimation of Q_1

- Main result: $T = 2M$ time slots are sufficient
- Difference between $\boldsymbol{y}_t$ and $\boldsymbol{y}_{M+t}$

$$\Delta \boldsymbol{y}_t = \boldsymbol{y}_{M+t} - \boldsymbol{y}_t = (e^{j\theta} - 1)\boldsymbol{Q}_1 \boldsymbol{\phi}_{t,1} \sqrt{p} a_t, t = 1, \dots, M$$

⇓

$$[\Delta \boldsymbol{y}_1, \dots, \Delta \boldsymbol{y}_M] = (e^{j\theta} - 1)\boldsymbol{Q}_1 [a_1 \boldsymbol{\phi}_{1,1}, \dots a_M \boldsymbol{\phi}_{M,1}] \sqrt{p}$$

$$= (e^{j\theta} - 1)\boldsymbol{Q}_1 [a_1 \boldsymbol{p}_1, \dots a_M \boldsymbol{p}_M] \sqrt{p}$$

$$\boldsymbol{\Phi}_1 = [\boldsymbol{p}_1, \dots, \boldsymbol{p}_M], \dots, \boldsymbol{\Phi}_M = [\boldsymbol{p}_M, \boldsymbol{p}_1, \dots, \boldsymbol{p}_{M-1}]$$

⇓

$\mathrm{rank}([a_1 \boldsymbol{p}_1, \dots a_M \boldsymbol{p}_M]) = M$ and we can estimate $\boldsymbol{Q}_1$
and $q = \mathrm{rank}(\boldsymbol{Q}_1) = \mathrm{rank}(\boldsymbol{G})$

$$\boldsymbol{Q}_1 = r_1 \boldsymbol{G} = r_1 [\boldsymbol{g}_1, \dots, \boldsymbol{g}_M]$$

Estimation of Q_1

- If we know $q = \mathrm{rank}(\boldsymbol{G})$ ahead of time, it is possible to use $T < 2M$ time slots to estimate $\boldsymbol{Q}_1$ when $q < N$

$$[\Delta \boldsymbol{y}_1, \dots, \Delta \boldsymbol{y}_t] = \left(e^{j\theta} - 1\right)\boldsymbol{Q}_1[a_1\boldsymbol{p}_1, \dots a_t\boldsymbol{p}_t]\sqrt{p}$$

- But before we estimate $\boldsymbol{Q}_1$, we do not know its rank
- We need to estimate $\boldsymbol{Q}_1$ no matter what its rank is

$$[\Delta \boldsymbol{y}_1, \dots, \Delta \boldsymbol{y}_M] = \left(e^{j\theta} - 1\right)\boldsymbol{Q}_1[a_1\boldsymbol{p}_1, \dots a_M\boldsymbol{p}_M]\sqrt{p}$$

- After we perfectly estimate $\boldsymbol{Q}_1$, we estimate its rank

Estimation of $\beta_2, \dots, \beta_M$

- Main result: $T = 2M$ time slots are sufficient
- Remove effect of RIS atom 1 on $\boldsymbol{y}_t$

$$\bar{\boldsymbol{y}}_t = \boldsymbol{y}_t - \left(e^{j\theta} - 1\right)\boldsymbol{Q}_1 \boldsymbol{\phi}_{t,1}\sqrt{p}a_t = \sum_{m=2}^{M} \beta_m \boldsymbol{Q}_1 \boldsymbol{\phi}_{t,m}\sqrt{p}a_t$$

$$= \boldsymbol{F}_t \boldsymbol{\beta}, t = 1, \dots, M$$

$$\boldsymbol{F}_t = \boldsymbol{Q}_1\left[\boldsymbol{\phi}_{t,2}, \dots, \boldsymbol{\phi}_{t,M}\right]\sqrt{p}a_t \in C^{N\times(M-1)}, \boldsymbol{\beta} = [\beta_2, \dots, \beta_M]^T$$

⇩

$[\bar{\boldsymbol{y}}_1^T, \dots, \bar{\boldsymbol{y}}_M^T]^T = [\boldsymbol{F}_1^T, \dots, \boldsymbol{F}_M^T]^T \boldsymbol{\beta}$, where $[\boldsymbol{F}_1^T, \dots, \boldsymbol{F}_M^T]^T \in C^{MN\times(M-1)}$

⇩

We can show $\text{rank}([\boldsymbol{F}_1^T, \dots, \boldsymbol{F}_M^T]^T) = M - 1$ and we can estimate $\boldsymbol{\beta}$

Impact of Unitary Matrix Constraint

- Received pilot signals

N equations, but q independent ones

$$\boldsymbol{y}_t = \sum_{m=1}^{M} \boldsymbol{Q}_m \boldsymbol{\phi}_{t,m} \sqrt{p} a_t = \boldsymbol{Q}_1 \boldsymbol{\phi}_{t,1} \sqrt{p} a_t + \sum_{m=2}^{M} \boxed{\beta_m \boldsymbol{Q}_1 \boldsymbol{\phi}_{t,m} \sqrt{p} a_t}$$

- Suppose there is no unitary matrix constraint on $\boldsymbol{\Phi}$
  - Phase 1: $\boldsymbol{\phi}_{t,1} \neq \boldsymbol{0}, \boldsymbol{\phi}_{t,2} = \cdots = \boldsymbol{\phi}_{t,M} = \boldsymbol{0}$ (linear in $\boldsymbol{Q}_1$)
    - Overhead to estimate $\boldsymbol{Q}_1$ (also q): M $\quad q = \text{rank}(G) \leq \min(M, N)$
  - Phase 2: $\boldsymbol{\phi}_{t,1} \neq \boldsymbol{0}, m = 2, \ldots, M$ (linear in β_m given $\boldsymbol{Q}_1$)
    - Overhead to estimate $\beta_2, \ldots, \beta_M$ given $\boldsymbol{Q}_1$: $\lceil (M-1)/q \rceil$
  - Overall overhead: $M + \lceil (M-1)/q \rceil$
- Overhead with unitary matrix constraint: $2M$

Multiple Multi-Antenna Users Case

- Received pilot signals

$$\boldsymbol{y}_t = \sum_{k,u,m} \beta_{k,u,m} \boldsymbol{Q}_{k,u,m} \boldsymbol{\phi}_{t,m} \sqrt{p} a_{k,u,t}$$

$$= \boldsymbol{Q}_{1,1,1} \boldsymbol{\phi}_{t,1} \sqrt{p} a_{1,1,t} + \sum_{m=2}^{M} \beta_{1,1,m} \boldsymbol{Q}_{1,1,1} \boldsymbol{\phi}_{t,1} \sqrt{p} a_{1,1,t}$$

$$+ \sum_{(k,u)\neq(1,1)} \sum_{m=1}^{M} \beta_{k,u,m} \boldsymbol{Q}_{1,1,1} \boldsymbol{\phi}_{t,m} \sqrt{p} a_{k,u,t}$$

$$\boldsymbol{Q}_{k,u,m} = r_{k,u,m} \boldsymbol{G}$$

$$\boldsymbol{Q}_{k,u,m} = \beta_{1,1,1} \boldsymbol{Q}_{1,1,1}$$

$$\beta_{k,u,m} = r_{k,u,m} / r_{1,1,1}$$

- Two-phase approach
  - Phase 1: only antenna 1 of user 1 transmits
    - $2M$ time slots to estimate $\boldsymbol{Q}_{1,1,1}$ and $\beta_{1,1,m}, m = 2, \ldots, M$
    - $q = \mathrm{rank}(G)$ is also estimated

Multiple Multi-Antenna Users Case

- Received pilot signals

$$\boldsymbol{y}_t = \sum_{k,u,m} \beta_{k,u,m} \boldsymbol{Q}_{k,u,m} \boldsymbol{\phi}_{t,m} \sqrt{p} a_{k,u,t}$$

$$= \boldsymbol{Q}_{1,1,1} \boldsymbol{\phi}_{t,1} \sqrt{p} a_{1,1,t} + \sum_{m=2}^{M} \beta_{1,1,m} \boldsymbol{Q}_{1,1,1} \boldsymbol{\phi}_{t,1} \sqrt{p} a_{1,1,t}$$

$$+ \sum_{(k,u)\neq(1,1)} \sum_{m=1}^{M} \boxed{\beta_{k,u,m} \boldsymbol{Q}_{1,1,1} \boldsymbol{\phi}_{t,m} \sqrt{p} a_{k,u,t}}$$

$$\boldsymbol{Q}_{k,u,m} = r_{k,u,m} \boldsymbol{G}$$
$$\boldsymbol{Q}_{k,u,m} = \beta_{1,1,1} \boldsymbol{Q}_{1,1,1}$$
$$\beta_{k,u,m} = r_{k,u,m} / r_{1,1,1}$$

- Two-phase approach
  - Phase 2: all other antennas transmit
    - $\lceil M(KU-1)/q \rceil$ time slots to estimate $\beta_{k,u,m}, \forall (k,u) \neq (1,1)$

N equations, but q independent ones

Main Result

- One single-antenna case: channel estimation overhead is

$$2M$$

- Generalization: K U-antenna users [5]

$$2M + \left\lceil \frac{M(KU-1)}{q} \right\rceil$$

- Overhead with conventional RIS [6]

$$M + \left\lceil \frac{M(KU-1)}{q} \right\rceil$$

- Surprisingly, same order
- The cost is from M to $2M$

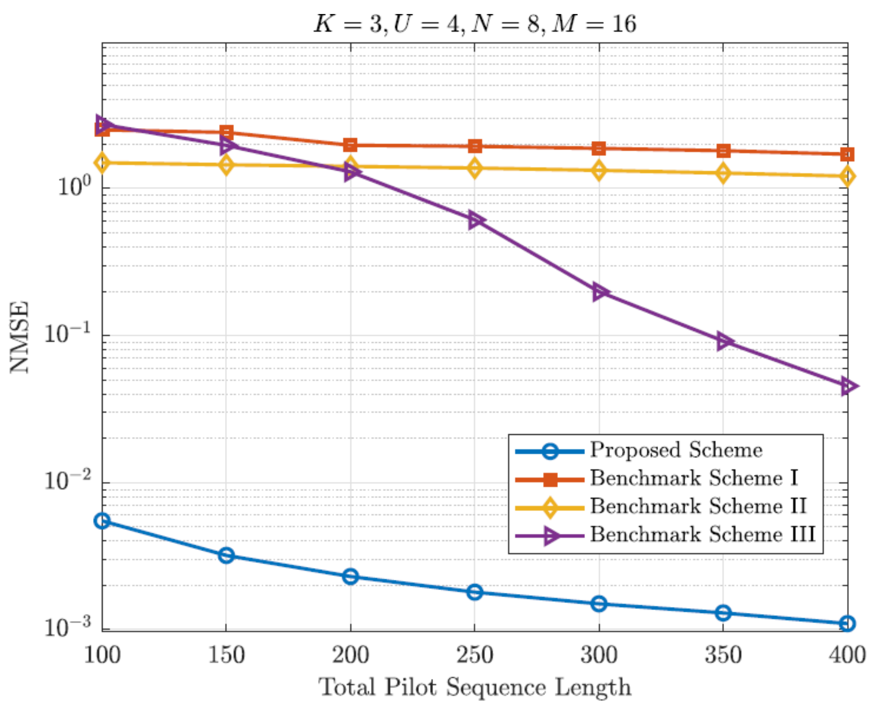
- Remark: we also have efficient algorithms to estimate channels in noisy systems

[5] R. Wang, S. Zhang, B. Clerckx, and **L. Liu**, "Low-overhead channel estimation framework for beyond diagonal reconfigurable intelligent surface assisted multi-user MIMO communication," under major revision, *IEEE Trans. Signal Process.*, 2025. [Online] Available: https://arxiv.org/abs/2504.10911

[6] Z. Wang, **L. Liu**, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6607-6620, Oct. 2020.

Numerical Results

- ❑ Our scheme: LMMSE to $\boldsymbol{Q}_{1,1,1}$ and $\beta_{k,u,m}, \forall(k,u) \neq (1,1), \forall m$
- ❑ Benchmark Scheme I: LS algorithm [4] to $\boldsymbol{Q}_{k,u,m}, \forall k, m$
- ❑ Benchmark Scheme II: BTKF algorithm [7] to $\boldsymbol{Q}_{k,u,m}, \forall k, m$
- ❑ Benchmark Scheme III: BTALS algorithm [7] to $\boldsymbol{Q}_{k,u,m}, \forall k, m$



[4] H. Li, S. Shen, Y. Zhang, and B. Clerckx, "Channel estimation and beamforming for beyond diagonal reconfigurable intelligent surfaces," *IEEE Trans. Signal Process.*, vol. 72, pp. 3318–3332, Jul. 2024.
[5] R. Wang, S. Zhang, B. Clerckx, and L. Liu, "Low-overhead channel estimation framework for beyond diagonal reconfigurable intelligent surface assisted multi-user MIMO communication," under major revision, *IEEE Trans. Signal Process.*, 2025.
[7] A. L. F. de Almeida, B. Sokal, H. Li, and B. Clerckx, "Channel estimation for beyond diagonal RIS via tensor decomposition," *IEEE Trans. Signal Process.*, early access, 2025.

Conclusions

- We characterize channel estimation overhead with BD-RIS
  - # of independent unknowns is far fewer than we expect
  - Challenges: non-linear equations and unitary matrix
- Main message:

overhead with BD-RIS ≈ overhead with conventional RIS

throughput with BD-RIS > throughput with conventional RIS

Types of RISs	Channel Estimation Overhead
BD-RIS	$2M + \lceil M(KU-1)/q \rceil$ [5]
BD-RIS	KUM^2 [4]
Conventional RIS	$M + \lceil M(KU-1)/q \rceil$ [6]

Thank You!