Intelligent and Converged Networks ISSN 2708-6240
2022, 3(1): 64−73 DOI: 10.23919/ICN.2022.0004

Maximizing dirty-paper coding rate of RIS-assisted multi-user MIMO broadcast channels

(*invited paper*)

Mohamed A. Elmossallamy, Radwa Sultan*, Karim G. Seddik, Geoffery Ye Li, and Zhu Han

Abstract: We consider a downlink multi-user scenario and investigate the use of reconfigurable intelligent surfaces (RISs) to maximize the dirty-paper-coding (DPC) sum rate of the RIS-assisted broadcast channel. Different from prior works, which maximize the rate achievable by linear precoders, we assume a capacity-achieving DPC scheme is employed at the transmitter and optimize the transmit covariances and RIS reflection coefficients to directly maximize the sum capacity of the broadcast channel. We propose an optimization algorithm that iteratively alternates between optimizing the transmit covariances using convex optimization and the RIS reflection coefficients using Riemannian manifold optimization. Our results show that the proposed technique can be used to effectively improve the sum capacity in a variety of scenarios compared to benchmark schemes.

Key words: broadcast channels; dirty-paper coding; multiple-input-multiple-output (MIMO); reconfigurable intelligent surfaces

1 Introduction

Reconfigurable intelligent surfaces (RISs) challenge the conventional wisdom that wireless propagation environment is uncontrollable. Some surfaces in the wireless environment can be coated by electromagnetic materials that can intelligently interact with electromagnetic waves incident upon it, ultimately controlling the propagation characteristics and giving rise to intelligent and reconfigurable radio environments[1, 2].

The ability to control propagation can be leveraged to engineer the channel realizations observed by the communicating nodes to optimize different performance metrics in numerous scenarios leading to impressive gains[3−8]. Of most practical relevance are multi-user multiple-access (MAC) and broadcast channels (BCs) corresponding to the uplink and downlink of some commercially prevalent wireless systems. Downlink scenarios, in particular, have received a lot of attention[3−6]. It has been demonstrated that the use of RISs results in substantial gains both in terms of energy efficiency[3, 4] and achievable rates[5, 6], and could achieve gains similar to massive multiple-input-multiple-output (MIMO) systems but with drastically more efficient hardware.

In this paper, we investigate the use of RISs to improve the dirty-paper coding rate (DPC), i.e., the capacity of a multi-user MIMO Gaussian BC. Our contributions can be summarized as follows.

- Unlike prior works in Refs. [5, 6], which considered a multi-user multiple-input-single-output (MISO) system with linear precoding at the transmitter, we assume a multi-user MIMO Gaussian broadcast channel with a non-linear capacity-achieving DPC

• Mohamed A. Elmossallamy is with Qualcomm Inc., Boxborough, MA 01719, USA. E-mail: m.ali@ieee.org.
• Zhu Han is with the Electrical and Computer Engineering Department and the Computer Science Department, University of Houston, Houston, TX 77004, USA. E-mail: zhan2@uh.edu.
• Radwa Sultan is with the Electrical and Computer Engineering Department, Manhattan College, Riverdale, NY 10471, USA. E-mail: rsultan02@manhattan.edu.
• Karim G. Seddik is with the Electronics and Communications Engineering Department, American University in Cairo, New Cairo 11835, Egypt. E-mail: kseddik@aucegypt.edu.
• Geoffery Ye Li is with the Department of Electrical and Electronic Engineering, Imperial College London, London, SW7 2AZ, UK. E-mail: Geoffrey.Li@imperial.ac.uk.
* To whom correspondence should be addressed.
Manuscript received: 2021-11-24; revised: 2022-01-19; accepted: 2022-03-09

scheme.

- Unlike prior works in Refs. [4, 9], which rely on the signal-to-noise-plus-interference ratios (SINRs) achieved by linear precoders, we assume a non-linear capacity-achieving DPC scheme and aim to maximize the sum rate capacity.
- We exploit universal channel knowledge with BC-MAC duality to map the broadcast problem into a computationally friendlier dual-MAC problem that achieves the same sum rate capacity[10]. Then, we develop an alternating optimization algorithm to solve the dual problem by iteratively optimizing the dual covariances and the RIS reflection coefficients independently. After solving the dual problem, the dual reflection coefficients solution will be equal to the original problem's reflection coefficients solution, i.e., the duality gap is zero. On the other hand, the dual covariances solutions will be transformed to the original problem covariances in closed-form[10].
- We provide numerical simulation results demonstrating that the proposed optimization scheme can be used to effectively improve the sum rate capacity in a variety of scenarios compared to benchmark schemes.

2 System model

We consider a downlink multi-user broadcast scenario with a base station with M antennas and K multi-antenna users with N antennas each, where $M \geqslant KN$. As shown in Fig. 1, the link is assisted by a single L-element reflectarray-based RIS[2]. Let $\boldsymbol{x} = \sum_{k=1}^{K} \boldsymbol{x}_k$ denote the transmitted vector from the base station,

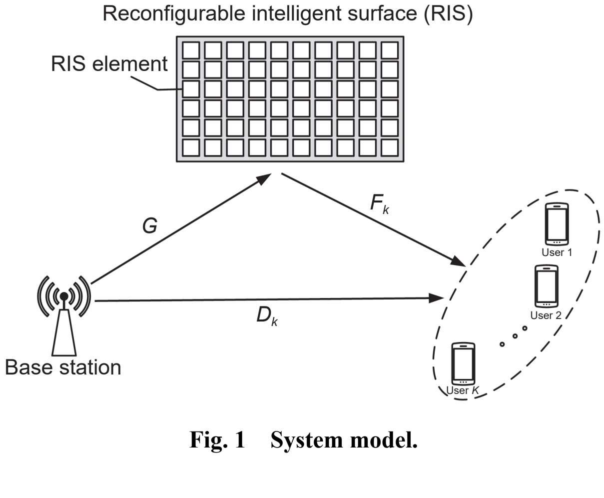

Fig. 1 System model.

where $\boldsymbol{x}_k$ is the DPC vector transmitted to the k-th user whose covariance $\boldsymbol{R}_k$ is to be optimized. Hence, the baseband received signal at the k-th user can be written as

$$\boldsymbol{y}_k = \boldsymbol{H}_k \boldsymbol{x} + \boldsymbol{n}_k \tag{1}$$

where $\boldsymbol{H}_k$ is the $N \times M$ effective channel matrix from the base station to the k-th user including the effects of the RIS, and $\boldsymbol{n}_k$ is the $N \times 1$ noise vector at the k-th user whose elements are distributed as independent and identically distributed (i.i.d.) zero mean complex Gaussian random variables with σ^2 variance, i.e., $\mathcal{CN}\left(0, \sigma_n^2\right)$.

Unlike a traditional system, the effective channel, $\boldsymbol{H}_k$, between the base station and the users can be controlled by changing the RIS configuration. We assume a reflectarray-based RIS, which is the simplest way to implement an RIS. In reflectarray-based RIS, a passive reflectarray whose elements are traditional antennas is connected to phase shifters that can be controlled electronically to phase-shift the incident signal[2]. Therefore, we can write the RIS-assisted channel for the k-th user as

$$\boldsymbol{H}_k = \boldsymbol{D}_k + \boldsymbol{F}_k \boldsymbol{Q} \boldsymbol{G} \tag{2}$$

where $\boldsymbol{D}_k$ denotes the $N \times M$ direct, i.e., not controllable by the RIS, channel between the base station and the k-th user, $\boldsymbol{F}_k$ denotes the $N \times L$ channel between the RIS and the k-th user, $\boldsymbol{G}$ denotes the $L \times M$ channel between the base station and the RIS, and $\boldsymbol{Q}$ denotes the RIS controllable interaction matrix. We assume the elements of $\boldsymbol{D}_k$, $\boldsymbol{F}_k$, and $\boldsymbol{G}_k$ are i.i.d. with distributions $\mathcal{CN}(0, \beta_d)$, $\mathcal{CN}\left(0, \beta_f\right)$, and $\mathcal{CN}\left(0, \beta_g\right)$, respectively, where parameters β_d, β_f, and β_g are determined by the path loss model to be discussed in Section 4.

Assuming no coupling between the RIS antenna elements, the interaction matrix can be written as

$$\boldsymbol{Q} = \text{diag}\ (\psi_1, \psi_2, \ldots, \psi_L) \tag{3}$$

where ψ_i represents the complex reflection coefficient of the i-th element. Since the RIS does not possess any amplification capabilities and can only reflect and shift the phase of the incident signals, we assume $|\psi_i| = 1$, $\forall i = 1, 2, \ldots, L$. Finally, similar to Refs. [4−6, 9], we assume both the base station and all users have prior knowledge of all relevant channels, which can be obtained using the techniques in the literature, e.g., Refs. [11, 12].

3 Capacity maximization for RIS-assisted broadcast MIMO transmissions

In this paper, we assume a capacity-achieving DPC scheme is used at the base station and aim to jointly optimize the transmit covariances, $\{\boldsymbol{R}_k\}_{k=1}^K$, and the RIS interaction matrix, $\boldsymbol{Q}$, to maximize the DPC sum rate, i.e., the sum capacity, of the RIS-assisted MIMO-BC channel.

Assuming the users are encoded in the order of their indices, the achievable MIMO-BC sum rate using a DPC scheme, for arbitrary transmit covariances and RIS configuration, can be written as

$$R_{\text{sum}}^{\text{BC}} = \sum_{k=1}^{K} \log_2 \det\left(\boldsymbol{I} + \frac{1}{\sigma_n^2}\boldsymbol{H}_k\boldsymbol{R}_k\boldsymbol{H}_k^\dagger\left(\boldsymbol{I} + \boldsymbol{H}_k \sum_{k=1}^{K}\frac{1}{\sigma_n^2}\boldsymbol{R}_k\,\boldsymbol{H}_k^\dagger\right)^{-1}\right) \tag{4}$$

where $\boldsymbol{I}$ is an identity matrix, $\boldsymbol{H}_k = \boldsymbol{D}_k + \boldsymbol{F}_k\boldsymbol{Q}\boldsymbol{G}$, and $(\cdot)^\dagger$ denotes the matrix Hermitian transpose. Our goal is to find a set of transmit covariances, $\{\boldsymbol{R}_k\}_{k=1}^K$, and an RIS configuration $\boldsymbol{Q}$ that approximately maximizes the achievable sum rate in Eq. (4) given a transmit power constraint P. This problem can be formulated as

$$\begin{aligned} \underset{\{\boldsymbol{R}_k\}_{k=1}^K,\boldsymbol{Q}}{\text{maximize}} \quad & R_{\text{sum}}^{\text{BC}}, \\ \text{subject to} \quad & \boldsymbol{Q} = \text{diag}\ (\psi_1,\psi_2,\ldots,\psi_L), \\ & |\psi_\ell| = 1,\ \forall \ell = 1,2,\ldots,L; \\ & \boldsymbol{R}_k \succcurlyeq \boldsymbol{0}, \\ & \sum_{k=1}^{K} \text{tr}\ (\boldsymbol{R}_k) \leqslant P \end{aligned} \tag{5}$$

Unfortunately, the objective in Formula (5) is not convex in $\{\boldsymbol{R}_k\}_{k=1}^K$ and $\boldsymbol{Q}$. Additionally, the first and the second constraints constitute non-convex sets which render the problem very hard to solve. In the sequel, we propose an alternating optimization scheme to obtain a sub-optimal solution for Formula (5)※. In particular, as shown in Fig. 2, we break Formula (5) into two sub-problems over the RIS reflection coefficients, $\boldsymbol{Q}$, and the transmit covariances, $\{\boldsymbol{R}_k\}_{k=1}^K$, individually. The two problems are then alternately solved until convergence to a locally optimal solution.

※Alternating optimization algorithms have been widely utilized in Refs. [4, 6, 9]. It was shown that, in general, alternating optimization algorithms converge to a locally optimal point and can achieve quite good performance.

3.1 RIS interaction matrix optimization

In this section, we present the RIS reflection coefficients optimization problem given a set of fixed transmit covariances. Again, it is advantageous to leverage duality to work with the more tractable dual MAC sum rate expression. Given a fixed set of covariances, the problem of optimizing the RIS reflection coefficients can be written as

$$\begin{aligned} \underset{\boldsymbol{Q}}{\text{maximize}} \quad & \log_2 \det\left(\boldsymbol{I} + \frac{1}{\sigma_n^2}\sum_{k=1}^{K}\boldsymbol{H}_k^\dagger\boldsymbol{R}_k^{\text{dMAC}}\boldsymbol{H}_k\right), \\ \text{subject to} \quad & \boldsymbol{Q} = \text{diag}\ (\psi_1,\psi_2,\ldots,\psi_L), \\ & |\psi_\ell| = 1,\ \forall \ell = 1,2,\ldots,L \end{aligned} \tag{6}$$

We transform the dual MAC into an equivalent single-user (SU)-MIMO channel with a block diagonal constraint on the transmit covariance. In particular, let $\boldsymbol{H} = [\boldsymbol{H}_1, \boldsymbol{H}_2, \ldots, \boldsymbol{H}_K]$ denote the composite channel of all users and $\boldsymbol{R} = \text{blkdiag}\,(\boldsymbol{R}_1,\boldsymbol{R}_2,\ldots,\boldsymbol{R}_K)$ denote the block diagonal matrix whose k-th diagonal block is the k-th user transmit covariance, $\boldsymbol{R}_k$. Hence, we can rewrite Formula (6) as

$$\begin{aligned} \underset{\boldsymbol{Q}}{\text{minimize}} \quad & -\log_2 \det\left(\boldsymbol{I} + \frac{1}{\sigma_n^2}\boldsymbol{H}^\dagger\boldsymbol{R}\boldsymbol{H}\right), \\ \text{subject to} \quad & \boldsymbol{Q} = \text{diag}\ (\psi_1,\psi_2,\ldots,\psi_L), \\ & |\psi_\ell| = 1,\ \forall \ell = 1,2,\ldots,L \end{aligned} \tag{7}$$

Note that the above objective function resembles the mutual information of a linear vector Gaussian channel whose gradient with respect to the channel matrix, $\boldsymbol{H}$, has been obtained in closed-form in Ref. [13], and the constraint set can be geometrically interpreted as restricting the solution to lie on the complex circle manifold, which is a smooth Riemannian submanifold of $\mathbb{C}^L$[14] formally defined as

$$\mathcal{M} = \left\{\boldsymbol{\psi} \in \mathbb{C}^L : |\psi_1| = |\psi_2| = \cdots = |\psi_L| = 1\right\} \tag{8}$$

Hence, the above problem can be cast as an unconstrained manifold optimization problem. Riemannian manifold optimization has been shown to perform well in solving hybrid/constant envelope precoding problems[15] and SU multiple-input-single-output (MISO) RIS-assisted channels[16]. Many of the classical gradient based unconstrained optimization

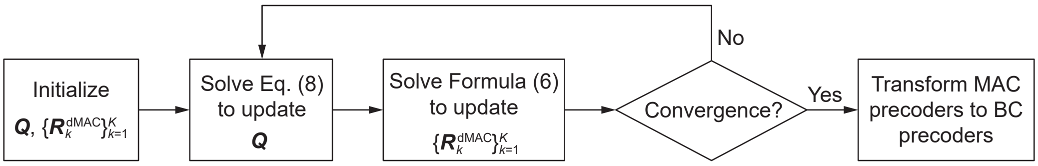

Fig. 2 Proposed alternating optimization scheme.

algorithms in Euclidean space have generalized counterparts in Riemannian manifolds. In this article, in Algorithm 1, we use the Riemannian Polak-Reibére conjugate gradient algorithm[14], which we briefly summarize for the complex circle manifold case next.

Starting from a point $\boldsymbol{\psi}_m$ on the surface of the manifold, we find a descent direction. The Riemannian gradient, $\nabla_{\mathcal{M}} f(\boldsymbol{\psi}_m)$, is defined as the direction of the greatest increase of the objective function at a given point $\boldsymbol{\psi}_m$ on the manifold but restricted to its tangent space, which is defined for any rector $\boldsymbol{v} \in \mathbb{C}^L$, for the complex circle manifold as

$$T_{\boldsymbol{\psi}_m}\mathcal{M} = \left\{\boldsymbol{v} \in \mathbb{C}^L : \Re\left(\boldsymbol{v} \odot \boldsymbol{\psi}_m^{\dagger} = \mathbf{0}_L\right)\right\} \tag{9}$$

where $\Re\{\cdot\}$ and $\odot$ denote the element-wise real-part of a complex vector and the Hadamard element-wise multiplication, respectively. Numerically, the Riemannian gradient is found by projecting the Euclidean gradient into the tangent space, which for the complex circle manifold is expressed as

$$\nabla_{\mathcal{M}} f(\boldsymbol{\psi}_m) = \nabla f(\boldsymbol{\psi}_m) - \Re\left\{\nabla f(\boldsymbol{\psi}_m) \odot \boldsymbol{\psi}_m^{\dagger}\right\} \odot \boldsymbol{\psi}_m \tag{10}$$

Hence, we first find the Euclidean gradient of the objective in Formula (7) with respect to the RIS reflection coefficients. Let f denote the objective of Formula (7), which is a real-valued function that depends on reflection coefficient ψ_ℓ through $\boldsymbol{H}_k$, which are both complex. Thus, the chain rule from Ref. [13] can be used to find the ℓ-th element of the gradient, $\nabla f(\boldsymbol{\psi}_m)$, i.e., $\dfrac{\partial f}{\partial \psi_\ell}$ to be

$$\frac{\partial f}{\partial \psi_\ell} = \mathrm{tr}\left(\nabla_{\boldsymbol{H}} f \cdot \frac{\partial \boldsymbol{H}^{\dagger}}{\partial \psi_\ell}\right) + \mathrm{tr}\left(\nabla_{\boldsymbol{H}}^{\dagger} f \cdot \frac{\partial \boldsymbol{H}}{\partial \psi_\ell}\right) \tag{11}$$

where $\mathrm{tr}(\cdot)$ denotes the matrix trace operator.

Now, we find $\nabla_{\boldsymbol{\psi}} \boldsymbol{H}$ and $\nabla_{\boldsymbol{H}} f$. Note that the k-th user channel can be rewritten as

$$\boldsymbol{H}_k = \boldsymbol{D}_k + \sum_{\ell=1}^{L} \psi_\ell [\boldsymbol{F}_k]_{:,\ell} \otimes [\boldsymbol{G}]_{\ell,:} \tag{12}$$

where $[\boldsymbol{F}_k]_{:,\ell}$ denotes the ℓ-th column of $\boldsymbol{F}_k$, $[\boldsymbol{G}]_{\ell,:}$ denotes the ℓ-th row of $\boldsymbol{G}$, and $\otimes$ denotes the Kronecker product. Hence, for the k-th user, the matrix of partial derivatives with respect to the ℓ-th phase shift is given by

$$\frac{\partial \boldsymbol{H}_k}{\partial \psi_\ell} = [\boldsymbol{F}_k]_{:,\ell} \otimes [\boldsymbol{G}]_{\ell,:} \tag{13}$$

and the composite matrix of partial derivatives can be written as

$$\frac{\partial \boldsymbol{H}}{\partial \psi_\ell} = \left[\frac{\partial \boldsymbol{H}_1}{\partial \psi_\ell}, \frac{\partial \boldsymbol{H}_2}{\partial \psi_\ell}, \ldots, \frac{\partial \boldsymbol{H}_K}{\partial \psi_\ell}\right] \tag{14}$$

From Ref. [13], $\nabla_{\boldsymbol{H}} f$ can be expressed as

$$\nabla_{\boldsymbol{H}} f = \frac{1}{\sigma_n^2} \boldsymbol{H}\boldsymbol{R}\left(\boldsymbol{I} + \frac{1}{\sigma_n^2} \boldsymbol{H}^{\dagger}\boldsymbol{H}\boldsymbol{R}\right)^{-1} \tag{15}$$

Starting at any point on the manifold, $\boldsymbol{\psi}_m$, the Riemannian gradient can be used to find a descent direction $\boldsymbol{\eta}_m$ in the tangent space $T_{\boldsymbol{\psi}_m}\mathcal{M}$ such that $\langle \boldsymbol{\eta}_m, \nabla f(\boldsymbol{\psi}_m)\rangle < 0$. As alluded to above, we use the Polak-Reibére conjugate gradient algorithm¶. In Euclidean space, the Polak-Reibére descent direction at the $(m+1)$-th iteration is given by

$$\boldsymbol{\eta}_{m+1} = -\nabla_{\mathcal{M}} f(\boldsymbol{\psi}_{m+1}) + \beta_{m+1}^{\mathrm{PR}} \boldsymbol{\eta}_m \tag{16}$$

where $\beta_{m+1}^{\mathrm{PR}}$ is the conjugate parameter for the Polak-

Algorithm 1 Riemannian Polak-Reibére conjugate gradient algorithm

Data: all CSI information, α_1

(1) **Initially:** Iterations $m = 1$, step size = $\alpha_m = \alpha_1$, initial point = $\boldsymbol{\psi}_m = \boldsymbol{\psi}_1$

(2) Compute the Riemannian gradient in Eq. (10);

(3) Calculate the Polak-Reibére conjugate parameter for the manifold generalization $\beta_{m+1}^{\mathrm{PR}-\mathcal{M}}$ in Eq. (19);

(4) Calculate the Polak-Reibére descent direction at the $(m+1)$-th iteration $\boldsymbol{\eta}_{m+1}^{\mathcal{M}}$ in Eq. (22);

(5) **while:** No Convergence **do**

$\quad\quad \boldsymbol{\psi}_{m+1} = \mathcal{R}\left(\boldsymbol{\psi}_m + \alpha_m \boldsymbol{\eta}_{m+1}^{\mathcal{M}}\right)$;

$\quad\quad m = m+1$;

$\quad\quad$ Return to Step (2);

end

¶The Polak-Reibére conjugate gradient algorithm is known to converge superlinearly with little added complexity compared to steepest-descent[14]. In general, the convergence of the conjugate gradient algorithms has been addressed in Refs. [17−19].

Reibére algorithm, which is given by[14]

$$\beta_{m+1}^{\mathrm{PR}}=\frac{\nabla f(\boldsymbol{\psi}_{m+1})^{\dagger}\left(\nabla f(\boldsymbol{\psi}_{m+1})-\nabla f(\boldsymbol{\psi}_m)\right)}{\|\nabla f(\boldsymbol{\psi}_m)\|^2} \tag{17}$$

where $\|\nabla f(\boldsymbol{\psi}_m)\|^2$ denotes the square of the Euclidean norm of $\nabla f(\boldsymbol{\psi}_m)$ and is given by $\|\nabla f(\boldsymbol{\psi}_m)\|^2=\nabla f(\boldsymbol{\psi}_m)^{\dagger}\nabla f(\boldsymbol{\psi}_m)$.

The update equation above cannot be directly applied to manifold optimization. In particular, $\nabla_{\mathcal{M}} f(\boldsymbol{\psi}_{m+1})$ lies in the tangent space at point $\boldsymbol{\psi}_{m+1}$, i.e., $T_{\boldsymbol{\psi}_{m+1}}\mathcal{M}$, while $\boldsymbol{\eta}_m$ and $\nabla_{\mathcal{M}} f(\boldsymbol{\psi}_m)$ lie in the tangent space at point $\boldsymbol{\psi}_m$, i.e., $T_{\boldsymbol{\psi}_m}\mathcal{M}$; hence, they cannot be directly added/subtracted. This problem is circumvented by means of a vector transport operation, $\mathcal{T}_{\boldsymbol{\psi}_m\mapsto\boldsymbol{\psi}_{m+1}}: T_{\boldsymbol{\psi}_m}\mathcal{M}\mapsto T_{\boldsymbol{\psi}_{m+1}}\mathcal{M}$, which takes a vector from the tangent space at one point on the manifold to the tangent space at another point on the manifold. In the case of the complex circle manifold, the vector transport operation can be expressed as[14]

$$\mathcal{T}_{\boldsymbol{\psi}_m\mapsto\boldsymbol{\psi}_{m+1}}(\boldsymbol{v})=\boldsymbol{v}-\Re\left\{\boldsymbol{v}\odot\boldsymbol{\psi}_{m+1}^{\dagger}\right\}\odot\boldsymbol{\psi}_{m+1} \tag{18}$$

which is equivalent to an additional projection from the old tangent space to the new tangent space. Leveraging vector transport operator, we define the conjugate parameter for the manifold generalization of Polak-Reibére as

$$\beta_{m+1}^{\mathrm{PR}-\mathcal{M}}=\nabla f(\boldsymbol{\psi}_{m+1})^{\dagger}\frac{\left(\nabla f(\boldsymbol{\psi}_{m+1})-\mathcal{T}_{\boldsymbol{\psi}_m\mapsto\boldsymbol{\psi}_{m+1}}(\nabla f(\boldsymbol{\psi}_m))\right)}{\|\nabla f(\boldsymbol{\psi}_m)\|^2} \tag{19}$$

Finally, note that an updated point $\boldsymbol{\psi}_{m+1}$ lies in the tangent space $T_{\boldsymbol{\psi}_m}\mathcal{M}$ and not necessarily on the surface of the manifold $\mathcal{M}$. Hence, an additional mapping is needed to ensure the updated point remains on the manifold. This mapping is referred to as a retraction and for the case of the complex circle manifold can be expressed as[14]

$$\mathcal{R}(\boldsymbol{v})=\boldsymbol{v}\oslash|\boldsymbol{v}| \tag{20}$$

where $|\cdot|$ and $\oslash$ denote element-wise absolute value and element-wise Hadamard division, respectively. Combining everything we described earlier, the solution is iteratively updated using the formula

$$\boldsymbol{\psi}_{m+1}=\mathcal{R}\left(\boldsymbol{\psi}_m+\alpha_m\boldsymbol{\eta}_{m+1}^{\mathcal{M}}\right) \tag{21}$$

where

$$\boldsymbol{\eta}_{m+1}^{\mathcal{M}}=-\nabla_{\mathcal{M}} f(\boldsymbol{\psi}_{m+1})+\beta_{m+1}^{\mathrm{PR}-\mathcal{M}}\mathcal{T}_{\boldsymbol{\psi}_m\mapsto\boldsymbol{\psi}_{m+1}}(\boldsymbol{\eta}_m) \tag{22}$$

and α_m is the step size, which can be chosen using any one-dimensional line search algorithm. In particular, we use the well-known Armijo backtracking line search algorithm[14], which ensures that the objective function is non-increasing, i.e., $f(\boldsymbol{\psi}_{m+1})\leqslant f(\boldsymbol{\psi}_m)$. The proposed alternating solution algorithm is summarized in Algorithm 2.

3.2 Transmit covariances optimization

In this section, we present the transmit covariances optimization problem with fixed RIS reflection coefficients matrix $\boldsymbol{Q}$. Given a fixed $\boldsymbol{Q}$, the effective channel matrix will be fixed, and the optimization problem in Foumula (5) reduces to the traditional MIMO-BC covariance optimization problem. Even if the problem remains non-convex even for a fixed $\boldsymbol{Q}$, it can be cast into a convex dual problem using the BC-MAC duality relationship. In particular, the MIMO-BC is transformed into a dual MIMO MAC with a sum power constraint over all users. This problem can be readily solved using convex optimization techniques to obtain the set of dual MAC covariances. The sum capacity of the dual MIMO MAC has been shown to be equal to the original MIMO-BC, i.e., the duality gap is zero, and the BC covariances can be readily obtained using simple closed-form transformations[10].

Algorithm 2 Proposed alternating optimization scheme

Data: all CSI information.

(1) **Initially:** Iterations $I=1$.

(2) Initialize $\boldsymbol{Q}$, $\{\boldsymbol{R}^{\mathrm{dMAC}}\}_{k=1}^{K}$;

(3) **RIS interaction matrix optimization**

  (a) For fixed $\{\boldsymbol{R}^{\mathrm{dMAC}}\}_{k=1}^{K}$, formulate the RIS reflection coefficients optimization problem in Formula (6);

  (b) Transform the dual MAC problem into an equivalent single-user (SU)-MIMO channel problem in Formula (7);

  (c) Apply the Riemannian Polak-Reibére conjugate gradient algorithm to solve Formula (7);

  (d) Update $\boldsymbol{Q}$;

(4) **Transmit covariances optimization**

  (a) For the value of $\boldsymbol{Q}$ obtained in Steps (3) and (4), formulate and solve the transmit covariances optimization problem in Formula (23);

  (b) Update $\{\boldsymbol{R}^{\mathrm{dMAC}}\}_{k=1}^{K}$;

(5) **If** Solution Converges **then**

Transform MAC precoders to BC precoders;

else

$I=I+1$; Return to Step (3);

end

Let $\left\{\boldsymbol{R}_k^{\mathrm{dMAC}}\right\}_{k=1}^{K}$ denote the dual MAC transmit covariances. The problem of optimizing the transmit covariances of the dual MIMO MAC with sum power constraint can be written as

$$\begin{aligned} \underset{\left\{\boldsymbol{R}_k^{\mathrm{dMAC}}\right\}_{k=1}^{K}}{\text{maximize}} \quad & \log_2 \det\left(\boldsymbol{I} + \frac{1}{\sigma_n^2} \sum_{k=1}^{K} \boldsymbol{H}_k^{\dagger} \boldsymbol{R}_k^{\mathrm{dMAC}} \boldsymbol{H}_k\right), \\ \text{subject to} \quad & \boldsymbol{R}_k^{\mathrm{dMAC}} \geqslant 0, \\ & \sum_{k=1}^{K} \mathrm{tr}\left(\boldsymbol{R}_k^{\mathrm{dMAC}}\right) \leqslant P \end{aligned} \tag{23}$$

which can be easily solved to obtain $\left\{\boldsymbol{R}_k^{\mathrm{dMAC}}\right\}_{k=1}^{K}$. The BC transmit covariances, $\{\boldsymbol{R}_k\}_{k=1}^{K}$, can then be readily obtained using the transformations in Ref. [10].

To prove the convergence of Algorithm 2, we need to show that first, the optimal solution of Formula (5) is bounded, second, that the objective function in Formula (5) is non-decreasing over the iterations of Algorithm 2. It is clear that the optimal solution of Formula (5) is bounded from above by the transmit power constraint. Therefore, the convergence of Algorithm 2 is guaranteed if the objective function in Formula (5) is non-decreasing over the iterations of Algorithm 2, which is proved by the following proposition.

Proposition 1 The objective function in Formula (5) is non-decreasing over the iterations of Algorithm 2.

Proof: Denote $R_{\mathrm{sum}}^{\mathrm{BC}}$ as $R(\{\boldsymbol{R}_k\}_{k=1}^{K}, \boldsymbol{Q})$ for a feasible solution $(\{\boldsymbol{R}_k\}_{k=1}^{K}, \boldsymbol{Q})$, and the solution obtained at the I-th iteration is $([\{\boldsymbol{R}_k\}_{k=1}^{K}]^I, [\boldsymbol{Q}]^I)$.

(1) From step (3), for a fixed $[\{\boldsymbol{R}_k\}_{k=1}^{K}]^I$, there exists a feasible solution of the problem in Formula (6) using the Riemannian Polak-Reibére conjugate gradient algorithm, i.e., $R([\{\boldsymbol{R}_k\}_{k=1}^{K}]^I, [\boldsymbol{Q}]^{I+1}) \geqslant R([\{\boldsymbol{R}_k\}_{k=1}^{K}]^I, [\boldsymbol{Q}]^I)$

(2) From step (4), for a fixed $[\boldsymbol{Q}]^{I+1}$, the optimal solution of the problem in Formula (23), and accordingly, the optimal solution for transmit covariances optimization problem can be found, i.e., $R([\{\boldsymbol{R}_k\}_{k=1}^{K}]^{I+1}, [\boldsymbol{Q}]^{I+1}) \geqslant R([\{\boldsymbol{R}_k\}_{k=1}^{K}]^I, [\boldsymbol{Q}]^I)$.

4 Numerical results

In this section, we present numerical simulation results to show the efficacy of the proposed optimization technique. As shown in Fig. 3, we assume the base station and the RIS are 200 m apart on the same horizontal line, and the users are randomly and uniformly distributed on a circle whose center is 30 m away from the RIS vertically and whose radius is 10 m. 3GPP path-loss models[20] at 2.5 GHz are used to calculate the path loss parameters, i.e., β_d, β_f, and β_g. We assume the direct channels, $\{\boldsymbol{D}_k\}_{k=1}^{K}$, to be non line-of-sight (NLOS) while the channels through the RIS, $\boldsymbol{G}$ and $\{\boldsymbol{F}_k\}_{k=1}^{K}$, are assumed to be line-of-sight (LOS). Simulation parameters are summarized in Table 1. Results are obtained over 10^4 independent channel realizations covering 100 random locations with 100 small scale fading realization in each location.

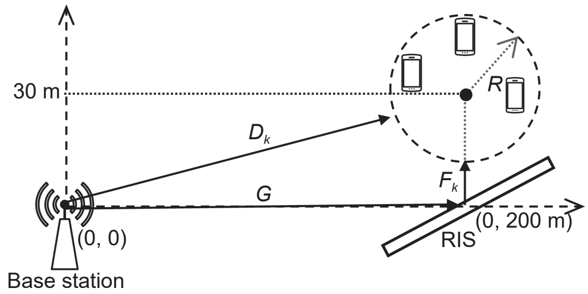

Fig. 3 Illustration of the geometric scenario used for calculating the distance dependent path loss parameters.

First, we start by verifying the proposed algorithm's convergence. Figure 4 shows the achievable MIMO-BC sum rate variation with the algorithm's iterations. As shown from the results in Fig. 4, the proposed algorithm can converge to a locally optimal solution after a limited number of iterations. Additionally, when increasing the number of the RIS reflective elements, L, the algorithm still converges but with more iterations‡. Additionally, when the transmission power is changed, the algorithm still converges.

Figure 5 shows the empirical cumulative distribution function (CDF) of the sum rate capacities achieved by the proposed scheme compared to two benchmarks. The first benchmark, Random (no RIS), refers to the

Table 1 Simulation parameters.

Parameter	Value
LOS path loss (dB)	$35.6 + 22.0\log_{10} d$
NLOS path loss (dB)	$32.6 + 36.7\log_{10} d$
Transmission power (P) (dBm)	23
Number of RIS elements (L)	100
Channel bandwidth (kHz)	180
Thermal noise PSD (dBm/Hz)	−170

‡ Increasing the number of the RIS reflective elements increases the number of optimization parameters in Formula (5), and accordingly, more iterations will be needed to reach a locally-optimal solution.

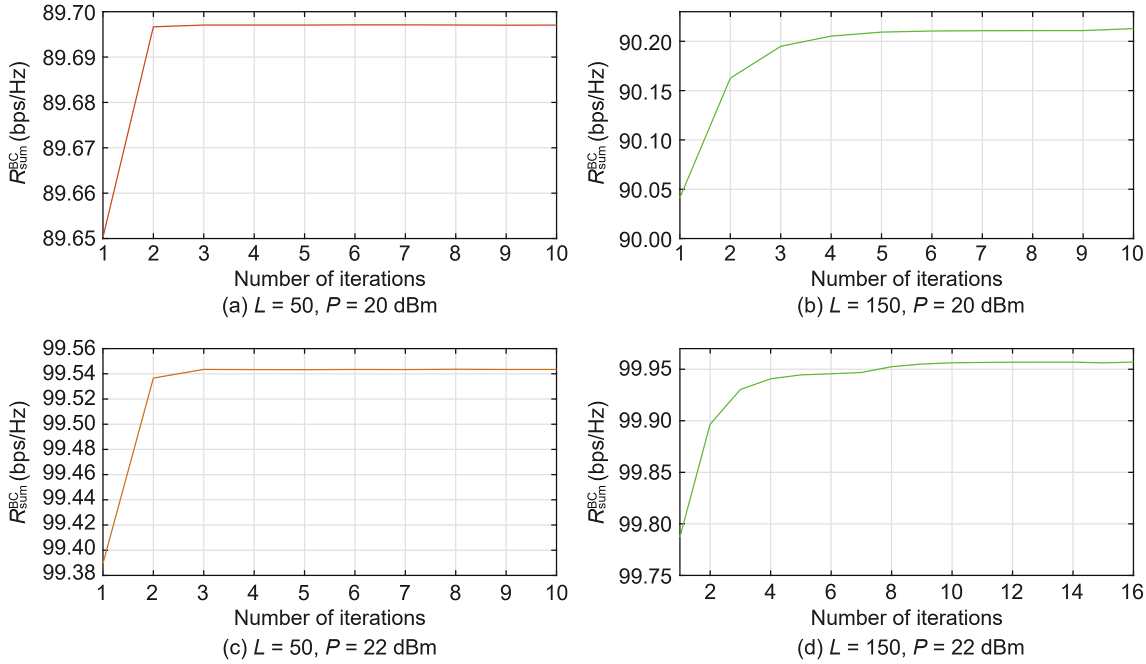

Fig. 4 Convergence of the sum rate with the proposed algorithm's iterations.

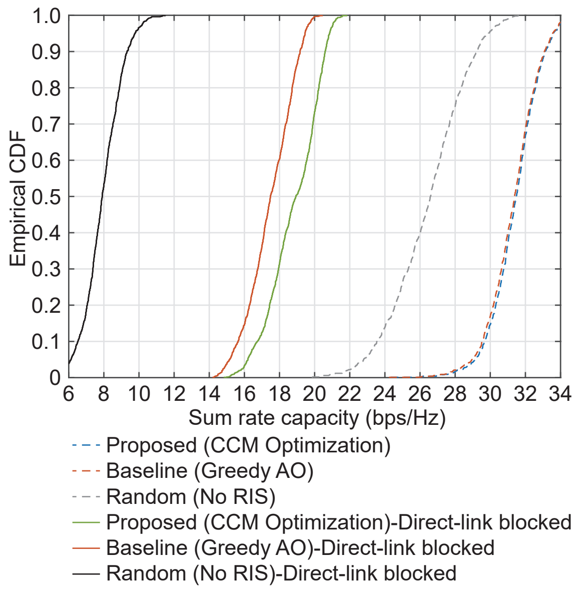

Fig. 5 Empirical cumulative distribution of the sum rate capacities achieved by the proposed scheme compared to benchmarks. Baseline (Greedy AO) refers to using the scheme from Ref. [21] to optimize the reflection coefficients.

case where a reflecting/scattering surface exists but is not intelligent, i.e., we cannot control the reflection coefficients and they are set randomly[§]. The second benchmark, Greedy AO, refers to using the scheme from Ref. [21]. It should be noted that the Greedy AO iteratively adjusts each phase-shift individually to optimize the RIS configuration. The authors in Ref. [21] have proved that the Greedy AO has a polynomial complexity/iteration in terms of the number of the RIS reflection coefficients. On the other hand, the proposed Riemannian Polak-Reibére conjugate gradient algorithm has a linear complexity/iteration in terms of the number of the RIS reflection coefficients[22, 23]. We also consider the case when the direct link is completely blocked, which is a rather common assumption in the literature, e.g., Ref. [7]. From Fig. 5, when the direct link exists, the proposed technique improves sum rate capacity by about 25% compared to the random configuration case; however, it is only very slightly better than the algorithm from Ref. [21]. When the direct link is blocked, the channel is completely controlled by the RIS and its configuration quality plays an important role in determining performance. In this case, the proposed technique nearly doubles the sum rate capacity of the channel compared to the random configuration case and significantly improves upon the performance of the Greedy AO technique from Ref. [21] by about 10%.

Figure 6 shows the average sum rate capacities achieved by the proposed scheme versus the number of RIS elements compared to the same two benchmarks discussed above. From Fig. 6, when the direct link

[§]Since the reflecting surfaces are not intelligent, i.e., they are not reconfigurable then, we can assume this case without RIS or with a random configuration of the reflecting coefficients.

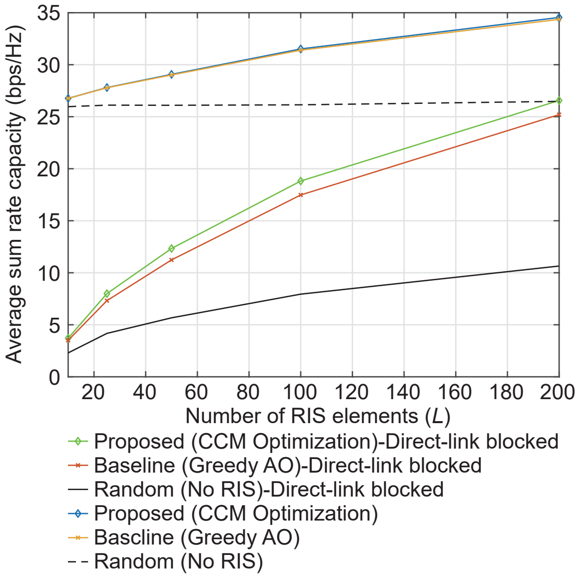

Fig. 6 Average sum rate capacities achieved by the proposed scheme compared to benchmarks for different number of RIS elements. Baseline (Greedy AO) refers to using the scheme from Ref. [21] to optimize the reflection coefficients.

exists, the performance gain of the proposed technique over the random configuration widens as the number of the RIS elements increases; however, its performance advantage over the baseline from Ref. [21] appears negligible for the considered numbers of RIS elements. When the direct link is blocked, the proposed technique performance gain over the random configuration benchmark grows faster with the increasing number of RIS elements. Moreover, a significant improvement over the performance of the Greedy AO technique from Ref. [21] is observed and it widens as the number of RIS elements increases.

5 Conclusion

In this article, we have investigated the use of reconfigurable intelligent surfaces to maximize the sum capacity of the RIS-assisted BC. We assumed a capacity-achieving dirty-paper precoding scheme is employed at the transmitter and proposed an alternating optimization scheme that iteratively optimizes the transmit covariances and RIS reflection coefficients using convex optimization and Riemannian manifold optimization, respectively. Numerical simulation results have shown that the proposed technique can be used to effectively improve the sum rate capacity compared to benchmark schemes.

Acknowledgment

This work was partially supported by the National Science Foundation (Nos. CNS-2107216 and CNS-2128368).

References

[1] E. Basar, M. D. Renzo, J. D. Rosny, M. Debbah, M. S. Alouini, and R. Zhang, Wireless communications through reconfigurable intelligent surfaces, *IEEE Access*, vol. 7, pp. 116753–116773, 2019.

[2] M. A. Elmossallamy, H. Zhang, L. Song, K. G. Seddik, Z. Han, and G. Y. Li, Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities, *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 3, pp. 990–1002, 2020.

[3] M. A. Elmossallamy, Passive reflection techniques for wireless communications, Doctoral dissertation, Electrical and Computer Engineering Department, University of Houston, Houston, TX, USA, 2020.

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

[5] H. Guo, Y. C. Liang, J. Chen, and E. G. Larsson, Weighted sum rate maximization for reconfigurable intelligent surface aided wireless networks, *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3064–3076, 2020.

[6] C. Huang, R. Mo, and C. Yuen, Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning, *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1839–1850, 2020.

[7] A. Taha, M. Alrabeiah, and A. Alkhateeb, Deep learning for large intelligent surfaces in millimeter wave and massive MIMO systems, presented at 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 2019.

[8] Q. U. A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M. S. Alouini, Asymptotic max-min SINR analysis of reconfigurable intelligent surface assisted MISO systems, *IEEE Transactions on Wireless Communications*, vol. 19, no. 12, pp. 7748–7764, 2020.

[9] Q. Wu and R. Zhang, Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts, *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1838–1851, 2020.

[10] S. Vishwanath, N. Jindal, and A. Goldsmith, Duality, achievable rates, and sum rate capacity of Gaussian MIMO broadcast channels, *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, 2003.

[11] Z. He and X. Yuan, Cascaded channel estimation for large intelligent metasurface assisted massive mimo, *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp.

210–214, 2020.

[12] X. Wei, D. Shen, and L. Dai, Channel estimation for RIS assisted wireless communications—Part I: Fundamentals, solutions, and future opportunities, *IEEE Communications Letters*, vol. 25, no. 5, pp. 1398–1402, 2021.

[13] D. P. Palomar and S. Verdu, Gradient of mutual information in linear vector Gaussian channels, *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 141–154, 2006.

[14] P. A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton, NJ, USA: Princeton University Press, 2008.

[15] X. Yu, J. Shen, J. Zhang, and K. B. Letaief, Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems, *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 3, pp. 485–500, 2016.

[16] X. Yu, D. Xu, and R. Schober, MISO wireless communication systems via intelligent reflecting surfaces, in *Proc. 2019 IEEE/CIC International Conference on Communications in China (ICCC)*, Changchun, China, 2019, pp. 735–740.

[17] Z. J. Shi and J. Shen, Convergence of the Polak-Ribiére-Polyak conjugate gradient method, *Nonlinear Analysis*:*Theory*, *Methods* & *Applications*, vol. 66, no. 6, pp. 1428–1441, 2007.

[18] Y. F. Hu and C. Storey, Global convergence result for conjugate gradient methods, *Journal of Optimization Theory and Applications*, vol. 71, no. 2, pp. 399–405, 1991.

[19] A. I. Cohen, Rate of convergence of several conjugate gradient algorithms, *SIAM Journal on Numerical Analysis*, vol. 9, no. 2, pp. 248–259, 1972.

[20] Further advancements for e-utra physical layer aspects (release 9), Tech. Rep. TR 36.814, 3GPP, 2010.

[21] S. Zhang and R. Zhang, Capacity characterization for intelligent reflecting surface aided MIMO communication, *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1823–1838, 2020.

[22] J. R. Shewchuk, An introduction to the conjugate gradient method without the agonizing pain, Tech. Rep., Carnegie Mellon University, Pittsburgh, PA, USA, 1994.

[23] R. Hauser, The Conjugate Gradient Method, Lecture 5, Continuous Optimisation, Oxford University Computing Laboratory, https://www.numerical.rl.ac.uk/people/nimg/course/lectures/raphael/lectures/lec5slides.pdf, 2006.

Mohamed A. Elmossallamy received the PhD degree in electrical engineering from University of Houston, Houston, TX, USA, in 2020, and the MSc degree in electronics and communications engineering from American University in Cairo, Cairo, Egypt, in 2017. He was the recipient of the 2019 IEEE ComSoc TCGCC Best Conference Paper Award and an NSF INTERN award in 2018. He has held engineering internships in Qualcomm Inc., Boxborough, MA, USA, MediaTek USA Inc., San Jose, CA, USA, and Skylark Wireless, Houston, TX, USA. He is currently a senior engineer at Qualcomm Inc., Boxborough, MA, USA. His research interests include transceiver design and performance analysis for the Internet of Things, massive MIMO-OFDM systems, and machine learning for wireless communications.

Radwa Sultan received the BSc (with highest honors) and MS degrees in electrical engineering from Alexandria University, Alexandria, Egypt, in 2009 and 2013, respectively, and the PhD degree from University of Houston, Houston, TX, USA, in 2017. Currently, she is an assistant professor in the Electrical and Computer Engineering Department, Manhattan College, Riverdale, NY, USA. She has served on the technical program committees of several IEEE journals and conferences in the areas of wireless communication. Her primary research interests are intelligent reflecting surface, Internet-of-Things, software defined radio systems, mmWave networks, small cell networks, full-duplex communication, massive MIMO, signal processing, wireless resource allocation, and big data analysis in wireless communication.

Karim G. Seddik received the BSc with an honor and MSc degrees in electrical engineering from Alexandria University, Alexandria, Egypt, in 2001 and 2004, respectively, and the PhD degree from University of Maryland, College Park, MD, USA, in 2008. He is currently a professor in the Electronics and Communications Engineering Department at the American University in Cairo (AUC) and the associate dean for graduate studies and research in School of Sciences and Engineering (SSE) at the AUC. Before joining AUC, he was an assistant professor at Alexandria University. His research interests include applications of machine learning in communication networks, intelligent reflecting surfaces, age of information, cognitive radio communications, and layered channel coding. He has served on the technical program committees of numerous IEEE conferences in the areas of wireless networks and mobile computing. He is the recipient of the American University in Cairo Faculty Merit Award for Excellence in Research and Creative Endeavors in 2021. He is a recipient of the State Encouragement Award in 2016, the State Medal of Excellence in 2017, and the certificate of honor from the Egyptian President for being ranked the first among all departments in the College of Engineering, Alexandria University in 2002. He received the Graduate School Fellowship from the University of Maryland in 2004 and 2005, and the Future Faculty Program Fellowship from the University of Maryland in 2007. He also co-authored a conference paper that received the best conference paper award from the IEEE Communication Society Technical Committee on Green Communications and Computing in 2019.

Geoffrey Ye Li is currently a chair professor at Imperial College London, UK. Before moving to Imperial in 2020, he was a professor with Georgia Institute of Technology, USA, for 20 years and a principal technical staff member with ATT Labs - Research in New Jersey, USA, for five years.

His general research interests include statistical signal processing and machine learning for wireless communications. In the related areas, he has published over 600 journal and conference papers in addition to over 40 granted patents and several books. His publications have been cited over 50 000 times with an H-index over 100 and he has been recognized as a highly cited researcher, by Thomson Reuters, almost every year. He was awarded the IEEE fellow and IET fellow for his contributions to signal processing for wireless communications. He won several prestigious awards from IEEE Signal Processing Society (Donald G. Fink Overview Paper Award in 2017), IEEE Vehicular Technology Society (James Evans Avant Garde Award in 2013 and Jack Neubauer Memorial Award in 2014), and IEEE Communications Society (Stephen O. Rice Prize Paper Award in 2013, Award for Advances in Communication in 2017, and Edwin Howard Armstrong Achievement Award in 2019). He also received the 2015 Distinguished ECE Faculty Achievement Award from Georgia Tech. He has been involved in editorial activities for over 20 technical journals, including the founding editor-in-chief of *IEEE JSAC* special series on ML in communications and networking. He has organized and chaired many international conferences, including technical program vice-chair of the IEEE ICC'03, general co-chair of the IEEE GlobalSIP'14, the IEEE VTC'19 (Fall), and the IEEE SPAWC'20.

Zhu Han received the BS degree in electronic engineering from Tsinghua University, in 1997, and the MS and PhD degrees in electrical and computer engineering from University of Maryland, College Park, in 1999 and 2003, respectively. From 2000 to 2002, he was an R&D engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a research associate at the University of Maryland. From 2006 to 2008, he was an assistant professor at Boise State University, Idaho. Currently, he is a John and Rebecca Moores Professor in the Electrical and Computer Engineering Department as well as in the Computer Science Department at the University of Houston, Texas. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. He received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the *Journal on Advances in Signal Processing* in 2015, IEEE Leonard G. Abraham Prize in the field of communications systems (best paper award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. He was an IEEE communications society distinguished lecturer during 2015−2018, AAAS fellow since 2019, and ACM distinguished member since 2019. He is a 1% highly cited researcher since 2017 according to Web of Science. He is also the winner of the 2021 IEEE Kiyo Tomiyasu Award, for outstanding early to mid-career contributions to technologies holding the promise of innovative applications, with the following citation: "for contributions to game theory and distributed management of autonomous communication networks".