

Improved Differential Evolution for Enhancing the Aggregated Channel Estimation of RIS-Aided Cell-Free Massive MIMO

Trinh Van Chien, Nguyen Hoang Viet, Symeon Chatzinotas, *IEEE Fellow*, and Lajos Hanzo, *IEEE Life Fellow*

Abstract—Cell-Free Massive multiple-input multiple-output (MIMO) systems are investigated with the support of a reconfigurable intelligent surface (RIS). The RIS phase shifts are designed for improved channel estimation in the presence of spatial correlation. Specifically, we formulate the channel estimate and estimation error expressions using linear minimum mean square error (LMMSE) estimation for the aggregated channels. An optimization problem is then formulated to minimize the average normalized mean square error (NMSE) subject to practical phase shift constraints. To circumvent the problem of inherent nonconvexity, we then conceive an enhanced version of the differential evolution algorithm that is capable of avoiding local minima by introducing an augmentation operator applied to some high-performing Differential Evolution (DE) individuals. Numerical results indicate that our proposed algorithm can significantly improve the channel estimation quality of the state-of-the-art benchmarks.

Index Terms—Cell-free massive MIMO, reconfigurable intelligent surface, differential evolution

I. INTRODUCTION

Cell-Free Massive MIMO systems constitute a promising technology for next-generation (NG) wireless networks. In these system architectures, a large number of access points (APs) equipped with multiple antennas are spread across the coverage area to serve the users. The spatial diversity gain of Cell-Free Massive MIMO technology results in significant performance enhancements [1]. Despite this appeal, Cell-Free Massive MIMO systems still encounter challenges such as pilot contamination and the support of remote users. The detector of Cell-Free Massive MIMO systems relies on instantaneous channel state information (CSI), which is obtained by the access points (APs) from the uplink pilot training. With many users admitted to the network, having mutually orthogonal pilots for each user is impractical, partly because the bandwidth available limits the number of orthogonal pilot sequences. However even if all pilots were orthogonal, the orthogonality may be eroded by the limited coherence interval. Therefore, several users may share the same pilot signals that introduce interference [2].

Integrating RISs into Cell-Free Massive MIMO systems can tackle these limitations by dynamically controlling the radio

signals through reflection, refraction or absorption. Thanks to the reconfigurable phase shift elements, RIS beneficially ameliorates the incoming waves to obtain a constructive signal combination at the receiver [3]. RIS-aided Cell-Free Massive MIMO systems offer various advantages for NG networks by enhancing signal strength, reducing mutual interference, and extending system coverage, even in challenging propagation environments [4]. Acquiring the perfect CSI in RIS-aided networks is nontrivial due to the passive nature of elements [5]. The pilot overhead required for channel estimation may become excessive for the support of the RIS if the cascaded two-hop channels are estimated separately [6]. Thus, the aggregated channels involved in both links are jointly estimated without extra pilot overhead [7]–[9]. Accurate phase shift designs are paramount and require sophisticated techniques capable of operating in the face of uncertainty. Previous research in [10], [11] and the reference therein have demonstrated the benefits of deterministic optimization based on the first- and second-order derivatives of the objective function and constraints. We emphasize that it is nontrivial to compute these derivatives and update the solution along with imperfect channel state information and in the presence of spatial correlation.

As a sub-field of artificial intelligence and optimization, evolutionary algorithms (EAs) have successfully been harnessed for solving challenging nonconvex problems, regardless of the nature of constraints and variables. For example, EAs have been successfully harnessed for maximizing the sum rate [12] or for minimizing the number of RISs [13]. In the family of EAs, Differential Evolution (DE) excels in terms of providing an effective mechanism of tackling continuous-valued optimization problems. For example, if in our system there are more users than the number of unique orthogonal pilot sequences, because their length is limited in a given bandwidth, then pilot contamination is encountered. Thanks to the raw power of DE, this phenomenon can be mitigated by appropriately adjusting its evolutionary operators. However, the conventional DE still has some issues, since its performance hinges on the specific choice of its hyperparameters. Since these parameters do not readily lend themselves to nimble adaptation, the fixed values may not be effective in the face of variations in the input factors. The standard DE often faces the challenge of getting stuck at a local optimum owing to its premature convergence, resulting in suboptimality.

Against this backdrop, we study the aggregated channel estimation of RIS-aided Cell-Free Massive MIMO systems in the presence of spatial correlation. The main contributions are concisely summarized as follows:

- We derive the closed-form expression of the channel estimate and estimation error by using the LMMSE estimator in the presence of spatial correlation at both the transmitter and receiver. Furthermore, we introduce

This research is funded by the Vietnam Ministry of Education and Training under project number B2025-BKA04 for Trinh Van Chien. L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council (EPSRC) projects under grant EP/Y026721/1, EP/W032635/1 and EP/X04047X/1 as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028).

T. V. Chien and N. H. Viet are with the School of Information and Communications Technology, Hanoi University of Science and Technology, Hanoi 100000, Vietnam (e-mail: chientv@soict.hust.edu.vn and nguyen-hoangviet.hsgs@gmail.com). S. Chatzinotas is with the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, L-1855 Luxembourg, Luxembourg (email: symeon.chatzinotas@uni.lu). L. Hanzo is with the Department of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (email: lh@ecs.soton.ac.uk)

the LMMSE of the channel estimate for each user as a criterion to assess the channel estimation quality;

- We formulate the optimization problem of minimizing the average NMSE of the overall network subject to specific phase shift constraints. Because the problem is NP-hard, we propose an improved DE-based algorithm for obtaining a high-quality solution in polynomial time. Our proposed algorithm is capable of escaping local minima, hence preventing premature convergence;
- Numerical results demonstrate that the proposed approach enhances the channel estimation quality by up to 33% compared to the canonical DE.

Notation: Matrices and vectors are denoted by bold letters. The notation $\text{diag}(\mathbf{x})$ is the diagonal matrix with the elements of vector $\mathbf{x}$ on the main diagonal. The symbols $\text{tr}(\cdot)$ and $\|\cdot\|$ are trace and norm operators, respectively. Meanwhile, the superscript $(\cdot)^H$ denote the Hermitian transpose. $\mathcal{CN}(\cdot, \cdot)$ and $\mathcal{U}([a, b])$ denote the circularly symmetric Gaussian distribution and the uniform distribution in $[a, b]$. The expectation of a random variable is $\mathbb{E}\{\cdot\}$. Finally, $\mathbf{I}_N$ is the n -dimensional identity matrix.

II. SYSTEM MODEL, PILOT TRAINING, AND AGGREGATED CHANNEL ESTIMATION

This section presents the system considered associated with an arbitrary pilot reuse factor. We then formulate an optimization problem for improving the quality of channel estimation.

A. Channel Model

Let us consider a system consisting of L APs each equipped with M antennas serving K single-antenna users. We assume that all the APs and users are randomly distributed in the coverage area riddled with obstacles. Thus, the network is assisted by an RIS that has N controllable reflecting elements. The phase shift matrix of the RIS is $\Phi = \text{diag}\{e^{i\theta_1}, \dots, e^{i\theta_N}\} \in \mathbb{C}^{N \times N}$, where $\theta_k \in [-\pi, \pi], \forall k = 1, N$ is the phase shift applied by the k -th element of the RIS. The direct channel between AP m and user k , denoted by $\mathbf{g}_{mk} \in \mathbb{C}^M$, follows the Rayleigh fading distribution, i.e., we have $\mathbf{g}_{mk} \sim \mathcal{CN}(0, \mathbf{G}_{mk})$, where $\mathbf{G}_{mk} \in \mathbb{C}^{M \times M}$ is the covariance matrix that describes the corresponding spatial correlation and the large-scale fading.¹ Thanks to the support of the RIS, the channel between APs and the RIS, as well as the RIS and users, is modelled by the Rician distribution. We denote the channel between AP m and the RIS by $\mathbf{H}_m \in \mathbb{C}^{M \times N}$, which can be formulated as [14] $\mathbf{H}_m = \bar{\mathbf{H}}_m + \mathbf{R}_{m,\text{AP}}^{1/2} \mathbf{X}_m \mathbf{R}_{m,\text{RIS}}^{1/2}$, where $\bar{\mathbf{H}}_m$ denotes the LoS channel. The elements of $\mathbf{X}_m$ are identically and independently distributed (iid) as $\mathcal{CN}(0, 1)$ with $\mathbf{R}_{m,\text{AP}} \in \mathbb{C}^{M \times M}$ and $\mathbf{R}_{m,\text{RIS}} \in \mathbb{C}^{N \times N}$ including the spatial correlation and large-scale fading effects. $\mathbf{z}_k \in \mathbb{C}^N$ stands for the channel between the RIS and user k , given as $\mathbf{z}_k \sim \mathcal{CN}(\bar{\mathbf{z}}_k, \mathbf{R}_k)$, where $\bar{\mathbf{z}}_k$ involves the LoS components and $\mathbf{R}_k$ denotes the spatial correlation and large-scale fading.

¹Cell-free networks are often deployed in, for example, urban areas associated with many scatterers. Hence, the Rayleigh fading distribution is well suited for such rich scattering propagation environments.

B. Pilot Training with an Arbitrary Pilot Reuse Factor

We assume that the pilot signals received in the uplink from the users are formed as a unit vector in the τ_p -dimensions complex field. Let $\phi_k \in \mathbb{C}^{\tau_p}$ associated with $\|\phi_k\| = 1$ denote the pilot signal allocated to user k . We denote the set of indices of the users that share the same pilot as user k by $\mathcal{P}_k$. The pilot signals are mutually orthogonal, in which $\phi_{k'}^H \phi_k = 1$ if $k' \in \mathcal{P}_k$. Otherwise, $\phi_{k'}^H \phi_k = 0$. If $\tau_p < K$, the pilot contamination appears as coherent interference. The aggregated channel between user k and AP m is defined as $\mathbf{u}_{mk} = \mathbf{g}_{mk} + \mathbf{H}_m \Phi \mathbf{z}_k$, which is encapsulated in the received training signal as

$$\mathbf{Y}_{pm} = \sqrt{p\tau_p} \sum_{k=1}^K \mathbf{u}_{mk} \phi_k^H + \mathbf{W}_{pm}, \quad (1)$$

where p is the normalized signal-to-noise ratio (SNR) of each pilot symbol, and $\mathbf{W}_{pm}$ is the additive white Gaussian noise. Upon projecting $\mathbf{Y}_{pm}$ onto ϕ_k^H , one can obtain

$$\mathbf{y}_{pmk} = \sqrt{p\tau_p} \sum_{k' \in \mathcal{P}_k} \mathbf{u}_{mk'} + \mathbf{w}_{pmk}, \quad (2)$$

where $\mathbf{w}_{pmk} \sim \mathcal{CN}(0, \mathbf{I}_M)$. The aggregated channel between AP m and user k impinging through the RIS is $\mathbf{u}_{mk} = \mathbf{g}_{mk} + \mathbf{H}_m \Phi \mathbf{z}_k$. Based on the signal observation in (2) and applying the linear MMSE estimator of [15], the estimated version of the aggregated channel $\hat{\mathbf{u}}_{mk}$ can be formulated as

$$\hat{\mathbf{u}}_{mk} = \bar{\mathbf{u}}_{mk} + \mathbb{E}\{\tilde{\mathbf{u}}_{mk} \tilde{\mathbf{y}}_{pmk}^H\} (\mathbb{E}\{\tilde{\mathbf{y}}_{pmk} \tilde{\mathbf{y}}_{pmk}^H\})^{-1} \tilde{\mathbf{y}}_{pmk}, \quad (3)$$

where $\bar{\mathbf{u}}_{mk} = \mathbb{E}\{\mathbf{u}_{mk}\}$, $\tilde{\mathbf{u}}_{mk} = \mathbf{u}_{mk} - \bar{\mathbf{u}}_{mk}$, and $\tilde{\mathbf{y}}_{pmk} = \mathbf{y}_{pmk} - \mathbb{E}\{\mathbf{y}_{pmk}\}$. The closed-form expression of the channel estimate $\hat{\mathbf{u}}_{mk}$ is obtained by computing the expectations in (3), as shown in Lemma 1.

Lemma 1. *If the LoS components are not blocked completely and the LMMSE estimator is used, then the aggregated channel estimate $\hat{\mathbf{u}}_{mk}, \forall m, k$, can be expressed in closed form as*

$$\hat{\mathbf{u}}_{mk} = \bar{\mathbf{H}}_m \Phi \bar{\mathbf{z}}_k + \Gamma_{mk} \Psi_{mk}^{-1} \tilde{\mathbf{y}}_{pmk}, \quad (4)$$

where $\Delta_{mk} = \mathbf{G}_{mk} + \bar{\mathbf{H}}_m \Phi \mathbf{R}_k \Phi^H \bar{\mathbf{H}}_m^H + \text{tr}(\mathbf{R}_{m,\text{RIS}} \Phi \mathbf{R}_k \Phi^H) \mathbf{R}_{m,\text{AP}} \Gamma_{mk} = \sqrt{p\tau_p} \left(\Delta_{mk} + \text{tr} \left(\mathbf{R}_{m,\text{RIS}} \Phi \bar{\mathbf{z}}_k \left(\sum_{k' \in \mathcal{P}_k} \bar{\mathbf{z}}_{k'}^H \right) \Phi^H \right) \mathbf{R}_{m,\text{AP}} \right)$, $\Psi_{mk} = p\tau_p \text{tr}(\mathbf{R}_{m,\text{RIS}} \Phi (\sum_{k' \in \mathcal{P}_k} \bar{\mathbf{z}}_{k'}) (\sum_{k' \in \mathcal{P}_k} \bar{\mathbf{z}}_{k'}^H) \Phi^H) \mathbf{R}_{m,\text{AP}} + p\tau_p \sum_{k' \in \mathcal{P}_k} \Delta_{mk'} + \mathbf{I}_M$, $\tilde{\mathbf{y}}_{pmk} = \sqrt{p\tau_p} \sum_{k' \in \mathcal{P}_k} \tilde{\mathbf{u}}_{mk'} + \mathbf{w}_{pmk}$. In addition, the channel estimation error defined as $\mathbf{e}_{mk} = \mathbf{u}_{mk} - \hat{\mathbf{u}}_{mk}$ has zero mean and the covariance matrix equals to

$$\mathbb{E}\{\mathbf{e}_{mk} \mathbf{e}_{mk}^H\} = \Delta_{mk} + \alpha_{mk} \mathbf{R}_{m,\text{AP}} - \Gamma_{mk} \Psi_{mk}^{-1} \Gamma_{mk}^H, \quad (5)$$

where $\alpha_{mk} = \text{tr}(\mathbf{R}_{m,\text{RIS}} \Phi \bar{\mathbf{z}}_k \bar{\mathbf{z}}_k^H \Phi^H)$.

Proof. See Appendix B. $\square$

In contrast to the channel estimate in (3), the closed-form expressions in (4) and (5) explicitly unveil the influences of spatial correlation, pilot reuse pattern, and phase shift design on the channel estimation quality. We also stress that the aggregated channel estimate in (4) does not require any additional pilot training overhead compared to that of the conventional Cell-Free Massive MIMO systems. Having highly accurate

channel estimation is of paramount importance in boosting the achievable rate of the uplink and downlink data transmission, especially in the time division duplex mode.

C. Problem formulation

We first denote the normalized mean square error (NMSE) of the estimated channel between user k and AP m as

$$\text{NMSE}_{mk} = \frac{\mathbb{E} \{ \|\mathbf{e}_{mk}\|^2 \}}{\mathbb{E} \{ \|\mathbf{u}_{mk}\|^2 \}} \stackrel{(a)}{=} \frac{\mathbb{E} \{ \|\mathbf{e}_{mk}\|^2 \}}{\mathbb{E} \{ \|\hat{\mathbf{u}}_{mk}\|^2 \} + \mathbb{E} \{ \|\mathbf{e}_{mk}\|^2 \}}, \quad (6)$$

where (a) is obtained by exploiting the independence of the channel estimate and the resultant estimation error. The second equality of (6) indicates that the NMSE is constrained to the interval $[0, 1]$. By exploiting Lemma 1, one can formulate the closed-form expression of the NMSE as

$$\text{NMSE}_{mk} = \frac{\text{tr} \left(\Delta_{mk} + \alpha_{mk} \mathbf{R}_{m,\text{AP}} - \mathbf{\Gamma}_{mk} \mathbf{\Psi}_{mk}^{-1} \mathbf{\Gamma}_{mk}^H \right)}{\text{tr} \left(\bar{\mathbf{H}}_m \mathbf{\Phi} \bar{\mathbf{z}}_k \bar{\mathbf{z}}_k^H \mathbf{\Phi}^H \bar{\mathbf{H}}_m^H + \Delta_{mk} + \alpha_{mk} \mathbf{R}_{m,\text{AP}} \right)} = 1 - \frac{\text{tr} \left(\bar{\mathbf{H}}_m \mathbf{\Phi} \bar{\mathbf{z}}_k \bar{\mathbf{z}}_k^H \mathbf{\Phi}^H \bar{\mathbf{H}}_m^H + \mathbf{\Gamma}_{mk} \mathbf{\Psi}_{mk}^{-1} \mathbf{\Gamma}_{mk}^H \right)}{\text{tr} \left(\bar{\mathbf{H}}_m \mathbf{\Phi} \bar{\mathbf{z}}_k \bar{\mathbf{z}}_k^H \mathbf{\Phi}^H \bar{\mathbf{H}}_m^H + \Delta_{mk} + \text{tr} \left(\mathbf{R}_{m,\text{RIS}} \mathbf{\Phi} \bar{\mathbf{z}}_k \bar{\mathbf{z}}_k^H \mathbf{\Phi}^H \right) \mathbf{R}_{m,\text{AP}} \right)}. \quad (7)$$

Observe from (7) that the NMSE can be arbitrarily small if orthogonal pilot signals are harnessed and the pilot power p is sufficiently high. In the case of pilot reuse, the NMSE will no longer tend to zero. We stress that the NMSE obtained in (7) is a generic version of the previous work where was only applied for correlated Rayleigh fading [7] or without the presence of pilot contamination [16]. Consequently, we propose to design the phase shifts by minimizing the average NMSE of the overall network as follows:

$$\begin{aligned} & \underset{\{\theta_1, \theta_2, \dots, \theta_N\}}{\text{minimize}} \quad \frac{1}{LK} \sum_{m=1}^L \sum_{k=1}^K \text{NMSE}_{mk} \\ & \text{subject to} \quad -\pi \leq \theta_n \leq \pi, \forall 1 \leq n \leq N. \end{aligned} \quad (8)$$

Problem (8) is non-convex and it is nontrivial to obtain the first and second derivatives of the objective function under spatial correlation. Hence, solving (8) is challenging.

III. IMPROVED DE-BASED SOLUTION

A. DE Relying on Augmentation Operator

To tackle the complex optimization of the NMSE in (7) and the non-convexity of problem (8), we conceive an improved version of DE by integrating it with an augmentation strategy, termed as Augmentation DE (ADE). To elaborate, initially, a population of individuals is generated, which is then evolved through consecutive generations. The mutation and crossover operators are harnessed for generating new individuals in each generation. Then the fitness of each new offspring is evaluated and compared directly to that of its parent. The specific individuals having a lower average NMSE are then passed on to the next generation. When this happens, the parent individual will be stored in an archive set $\mathcal{A}$ for later use in order to maintain search diversity. The current solution of problem (8) is the particular individual having the

best fitness value, i.e. the lowest average NMSE. Again, to avoid getting stuck at a local optimum as the conventional DE, we harness an augmentation operator to improve the current solution. Additionally, unlike in the conventional DE, the hyperparameters in the mutation and crossover operators are dynamically adjusted. This method enables ADE to update the solution based on the search behavior after each iteration, resulting in improved solutions in subsequent generations. A detailed description of Algorithm 1 is provided below.

Algorithm 1 Augmentation Differential Evolution (ADE)

Input: The channel $\mathbf{g}_{mk}, \mathbf{H}_m, \mathbf{z}_k$ with $k = \overline{1, K}, m = \overline{1, L}$; the maximum number of generations G_{MAX} ; the augmentation search condition ε and augmentation parameter β .

Output: The phase shift matrix $\mathbf{\Phi}$.

- 1: Set the generation index $G \leftarrow 1$. Randomly initialize a population $\mathcal{P}^{(1)}$ with I individuals and set archive set $\mathcal{A} \leftarrow \mathcal{P}^{(1)}$.
 - 2: **while** $G < G_{\text{MAX}}$ **do**
 - 3: Initialize the next generation $Q^{(G)} \leftarrow \emptyset$.
 - 4: **for** $i = 1 : I$ **do**
 - 5: Generate $\mathbf{F}^{(iG)}$ and $\mathbf{CR}^{(iG)}$ as the SHADE method [17]. Create mutant vector $\mathbf{u}^{(iG)}$ using DE/pbest/1 strategy in (9). Create trial vector $\mathbf{\omega}^{(iG)}$ using (11).
 - 6: **if** $f(\mathbf{\omega}^{(iG)}) \leq f(\mathbf{x}^{(iG)})$ **then**
 - 7: $Q^{(G)} \leftarrow Q^{(G)} \cup \mathbf{\omega}^{(iG)}, \mathcal{A}_j \leftarrow \mathbf{x}^{(iG)}$
 - 8: **else**
 - 9: $Q^{(G)} \leftarrow Q^{(G)} \cup \mathbf{x}^{(iG)}$
 - 10: **end if**
 - 11: **if** $f^{(\text{best}G)} - f^{(\text{best}Q)} < \varepsilon$ **then**
 - 12: **for** $j = 1 : \lambda$ **do**
 - 13: Create augmented vector $\tilde{\mathbf{\omega}}^{(jG)}$ using (12).
 - 14: **if** $f(\tilde{\mathbf{\omega}}^{(jG)}) < f(\mathbf{\omega}^{(jG)})$ **then**
 - 15: $Q_j^{(G)} \leftarrow \tilde{\mathbf{\omega}}^{(jG)}$
 - 16: **else if** $f(\tilde{\mathbf{\omega}}^{(jG)}) < f(\mathcal{A}_j)$ **then**
 - 17: $\mathcal{A}_j \leftarrow \tilde{\mathbf{\omega}}^{(jG)}$
 - 18: **end if**
 - 19: **end for**
 - 20: **end if**
 - 21: **end for**
 - 22: $\mathcal{P}^{(G)} \leftarrow Q^{(G)}, G \leftarrow G + 1$.
 - 23: **end while**
-

1) *Solution representation and population initialization:* The population $\mathcal{P}^{(1)}$ contains I individuals, which are randomly generated. The i -th individual of the G -th generation is denoted by $\mathbf{x}^{(iG)} \in \mathbb{R}^N$, whose elements range from $-\pi$ to π .

2) *Mutation:* In the G -th generation, a mutant vector $\mathbf{u}^{(iG)} \in \mathbb{R}^N$ is created corresponding to its parent $\mathbf{x}^{(iG)}$. In this paper, we employ the popular DE/pbest/1 strategy to create a mutant vector, which is formulated as

$$\mathbf{u}^{(iG)} = \mathbf{x}^{(p\text{best}G)} + \mathbf{F}^{(iG)} (\mathbf{x}^{(r_1G)} - \mathbf{x}^{(r_2G)}), \quad (9)$$

where $\mathbf{x}^{(p\text{best}G)}$ is chosen from the $[I \times p]$ individuals with the best fitness values in the current population $\mathcal{P}^{(G)}$, whereas $\mathbf{x}^{(r_1G)}$ and $\mathbf{x}^{(r_2G)}$ is selected randomly from $\mathcal{P}^{(G)}$ and $\mathcal{P}^{(G)} \cup \mathcal{A}$, respectively. The term $\mathbf{x}^{(p\text{best}G)}$ is employed

for ensuring that the search behaviour is not be dominated by any particular individual. Meanwhile, the scale factor $F^{(iG)}$ of the i -th individual will be adapted according to the specific nature of the search behaviour instead of using a fixed value throughout all generations, as proposed in [17]. Nonetheless, the update in (9) may result in having some elements of $\mathbf{u}^{(iG)}$ to fall outside the feasible domain. To tackle this risk, we adjust the mutant vector formulated mathematically as

$$[\mathbf{u}^{(iG)}]_t = \begin{cases} (1 + [\mathbf{x}^{(iG)}]_t)/2, & \text{if } [\mathbf{u}^{(iG)}]_t > 1, \\ (-1 + [\mathbf{x}^{(iG)}]_t)/2, & \text{if } [\mathbf{u}^{(iG)}]_t < -1, \end{cases} \quad (10)$$

where $[\mathbf{a}]_t$ is the t -th element of vector $\mathbf{a}$.

3) *Crossover*: After the mutation, $\mathbf{u}^{(iG)}$ is combined with its parents to generate a trial vector $\boldsymbol{\omega}^{(iG)} \in \mathbb{R}^N$, which is

$$[\boldsymbol{\omega}^{(iG)}]_t = \begin{cases} [\mathbf{u}^{(iG)}]_t, & \text{if } \mathcal{U}[0, 1] \leq \text{CR}^{(iG)} \text{ or } t = t_{\text{rand}}, \\ [\mathbf{x}^{(iG)}]_t, & \text{otherwise,} \end{cases} \quad (11)$$

where $\mathcal{U}[0, 1]$ represents a uniform random number range between 0 and 1, $\text{CR}^{(iG)}$ is the crossover rate, and t_{rand} is chosen randomly between 1 and N to ensure that $\boldsymbol{\omega}^{(iG)}$ has at least one element from the mutant vector $\mathbf{u}^{(iG)}$.

4) *Selection*: The trial vector $\boldsymbol{\omega}^{(iG)}$ will replace $\mathbf{x}^{(iG)}$ in the next generation, if it has better fitness value. Meanwhile, $\mathbf{x}^{(iG)}$ will replace the i -th member in the archive set $\mathcal{A}$ for ensuring that $\mathcal{P}^{(G)}$ and $\mathcal{A}$ have the same cardinality.

5) *Dispensing with Augmentation*: The conventional DE algorithm has a major drawback. Explicitly, it tends to promptly evolve all individuals in the population towards a local minimum without searching for a better solution. To circumvent the problem of premature convergence, we introduce a local search operator, where all phase shift elements of each selected individual will be augmented by a random parameter to ensure that some individuals escape from the local minimum. In particular, we modify the first λ ($\lambda \in [1, I]$) members of the population if $|f^{(\text{bestQ})} - f^{(\text{bestG})}| < \varepsilon$, where $f^{(\text{bestG})}$ is the best fitness value obtained in the current generation, $f^{(\text{bestQ})}$ is the best fitness value after selection and ε is a fixed tolerance threshold. The augmented vector $\tilde{\boldsymbol{\omega}}^{(jG)}$ ($j \in [1, \lambda]$), which has a minimum value of $\boldsymbol{\omega}^{(jG)}$ is formulated as

$$\tilde{\boldsymbol{\omega}}^{(jG)} = \boldsymbol{\omega}^{(jG)} + \beta \mathbf{1}_N, \quad (12)$$

where $\mathbf{1}_N$ is a vector having N elements equal to 1 and $\beta \sim \mathcal{N}(0, \sigma^2)$ has a low variance σ^2 . Note that $\tilde{\boldsymbol{\omega}}^{(jG)}$ will replace $\boldsymbol{\omega}^{(jG)}$ in the next generation if $f(\tilde{\boldsymbol{\omega}}^{(jG)}) < f(\boldsymbol{\omega}^{(jG)})$. Otherwise, the augmented vector $\tilde{\boldsymbol{\omega}}^{(jG)}$ will replace the j -th individual in the archive set $\mathcal{A}$.

6) *Parameter adaption*: The performance of Algorithm 1 significantly depends on the choice of the control parameter $F^{(iG)}$ in (9) and $\text{CR}^{(iG)}$ in (11) because they directly affect both the mutation and crossover mechanisms. It is adjusted throughout the generations relying on the so-called success-history-based parameter adaptation (SHADE) of [17].²

²Thanks to intelligent computation, Algorithm 1 can potentially be extended to scenarios of multiple-RIS-aided systems by including the RIS indices as optimization variables and extending the evolutionary dimensions. Due to the limited space in this compactness, this promising research direction is left for future research.

TABLE I: The complexity of ADE, DE, and GA.

Benchmark	Complexity
ADE	$\mathcal{O}[GI \log(I) + GIN + \lambda GI]$
DE	$\mathcal{O}[GI \log(I) + GIN]$
GA	$\mathcal{O}[GI \log(I) + GIN]$

B. Computational Complexity

As a function of the phase shift matrix Φ , the channel state information $\mathbf{u}_{mk}$, $\hat{\mathbf{u}}_{mk}$, $\mathbf{e}_{mk}$, and the average NMSE in (7) are updated accordingly. Specifically, the update of the covariance matrix of aggregated channel $\mathbf{u}_{mk}$ has a computational complexity order of $\mathcal{O}(MN^2 + M^2N + N^3)$. By contrast, the matrices Γ_{mk} and Ψ_{mk} impose a complexity order of $\mathcal{O}(MN^2 + M^2N + N^3 + K)$. Hence, the complexity of evaluating the objective function (8) is $\mathcal{O}[LK(MN^2 + M^2N + N^3 + K)]$. Regarding the phase shift optimization, the initialization imposes the computational complexity of $\mathcal{O}(IN)$. In each generation, sorting the population for use in (9) has the cost of $\mathcal{O}[I \log(I)]$, while both the mutation and crossover have the complexity of $\mathcal{O}(IN)$. The selection, augmentation search, and parameter adaptation have the cost of $\mathcal{O}(I)$, $\mathcal{O}(\lambda I)$, and $\mathcal{O}(I)$, respectively. Thus, the computational complexity of Algorithm 1 is $\mathcal{O}[GI \log(I) + GIN + \lambda GI]$ with G being the number of generations. Basically, at each iteration, ADE shares the same core stages as the Genetic Algorithm (GA) and DE, which includes mutation, crossover, selection, and population sorting. The significant distinction between the proposed algorithm and the two canonical benchmarks is that ADE activates the Augmentation Operator after the selection and updates the hyper-parameters according to the search behavior. These enhancements lead to the difference in the complexity of three algorithms considered, as summarized in Table I.

IV. NUMERICAL RESULTS

A system is considered that includes 40 APs, each equipped with 16 antennas and a single RIS, serving ten users in the square area of 1 km². These APs and users are uniformly distributed in 0.25×0.25 km² squares at the southwest and southeast corners, respectively, while the RIS is at the centre of the region considered. The height of the APs, RIS, and users is 15 [m], 30 [m], and 1.65 [m], respectively, while the spatial correlation matrices are defined in [18]. To simulate a harsh communication environment, it is assumed that the direct links between APs and users, i.e. $\mathbf{g}_{mk}, \forall m, k$, are unblocked with a probability of $\tilde{p}$ [7]. Specifically, the covariance matrix $\mathbf{G}_{mk}$ will be multiplied by a scalar α_{mk} , where $\alpha_{mk} = 1$ with probability $\tilde{p}$, and $\alpha_{mk} = 0$ otherwise. In our setup, $\tilde{p} = 0.2$, meaning that, on average, each AP can directly transmit signal to two user. Our proposed phase shift design is compared both to the classic GA [19] and DE [20].

In Fig. 1, we plot the convergence of all the evolutionary benchmarks considered. While both GA and DE could yield a good solution, SDE outperforms these baselines. Indeed, the results demonstrate the superior performance of our proposed algorithm, which incorporates an augmentation strategy, compared to the standard DE and GA. In the first evolutions, SDE continuously explores the search space for finding a superior

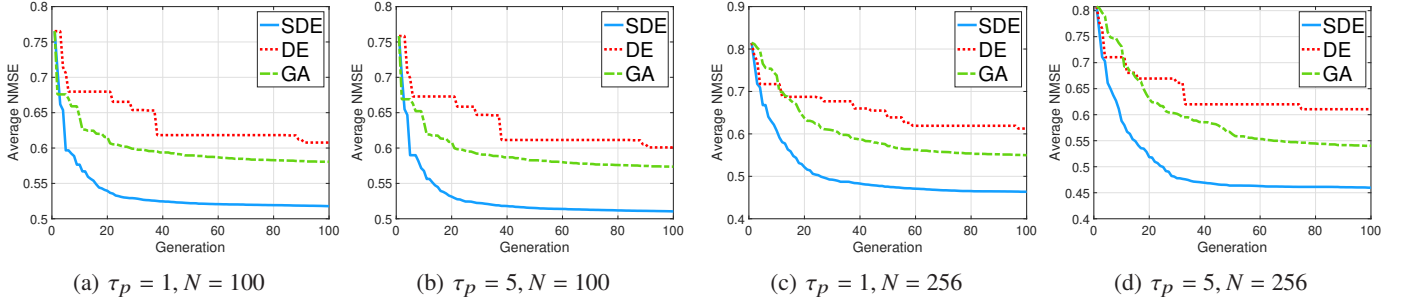


Fig. 1: The convergence trend of SDE and other evolutionary algorithms, including canonical DE and GA.

TABLE II: Average NMSE with different benchmarks.

Scenarios	Average NMSE				
	SDE	GA [19]	DE [20]	RPS [19]	EPS [7]
$\tau_p = 1$ $N = 100$	0.5178	0.5805	0.6077	0.9096	0.9925
$\tau_p = 5$ $N = 100$	0.5104	0.5737	0.6009	0.8918	0.9856
$\tau_p = 1$ $N = 256$	0.4636	0.5498	0.6126	0.9071	0.9940
$\tau_p = 5$ $N = 256$	0.4597	0.5401	0.6105	0.9003	0.9872

TABLE III: Ergodic SE per user with different benchmarks.

Scenarios	$\tau_p = 1$	$\tau_p = 2$	$\tau_p = 5$
RPS	0.93	1.51	2.38
SDE	1.25	1.93	2.88
Gain	34%	28%	21%

solution that DE and GA cannot achieve. Nonetheless, after many generations, all the selected individuals may correspond to a local minimum. Thus, the augmentation mechanism promotes efficiency by preventing premature convergence and continuing exploration to avoid these local minima. For completeness, two additional benchmarks, namely the equal phase shift design (EPS) [7] and random phase shift design (RPS) [19] are also harnessed for comparison. In the RPS design, each element is drawn from the uniform distribution in the interval $[-\pi, \pi]$. We evaluate the average NMSE for 1000 different realizations of the random phase shift elements, as reported in TABLE II. We now evaluate the ergodic spectral efficiency (SE) obtained with the phase shifts designed by SDE and RPS with $N = 100$ and the pilot signal duration $\tau_p \in \{1, 2, 5\}$. Specifically, the uplink SE of user k is formulated as $R_k = B(1 - \tau_p/\tau_c) \log_2(1 + \gamma_k)$ [Mbps], where B [MHz] is the system bandwidth and the effective signal-to-interference-plus-noise ratio γ_k is

$$\gamma_k = \frac{|\text{DS}_{uk}|^2}{\mathbb{E}\{|\text{BU}_{uk}|^2\} + \sum_{k'=1, k' \neq k}^K \mathbb{E}\{|\text{UI}_{uk'k}|^2\} + \mathbb{E}\{|\text{NO}_{uk}|^2\}}, \quad (13)$$

where $\text{DS}_{uk} = \sqrt{p_u \eta_k} \mathbb{E}\{\sum_{m=1}^L \hat{\mathbf{u}}_{mk}^H \mathbf{u}_{mk}\}$, $\text{BU}_{uk} = \sqrt{p_u \eta_k} (\sum_{m=1}^L \hat{\mathbf{u}}_{mk}^H \mathbf{u}_{mk} - \mathbb{E}\{\sum_{m=1}^L \hat{\mathbf{u}}_{mk}^H \mathbf{u}_{mk}\})$, $\text{UI}_{uk'k} = \sqrt{p_u \eta_k} \sum_{m=1}^L \hat{\mathbf{u}}_{mk}^H \mathbf{u}_{mk'}$, and $\text{NO}_{uk} = \sum_{m=1}^L \hat{\mathbf{u}}_{mk}^H \mathbf{w}_{um}$. The ergodic SE per user measured in Mbps are shown in TABLE III for a system bandwidth $B = 10$ [MHz].

V. CONCLUSIONS

Improved channel estimation was conceived for RIS-assisted MIMO systems by optimizing the phase shift elements. The proposed algorithm has shown significant channel estimation error reduction. The solutions were found in polynomial time without relying on the gradient of the objective function and outperformed the state-of-the-art benchmarks found in the literature.

APPENDIX

A. Useful Corollary and Lemma

Corollary 1. For a positive semi-definite matrix $\mathbf{D} \in \mathbb{C}^{N \times N}$ and $\tilde{\mathbf{H}}_m = \mathbf{R}_{m, \text{AP}}^{1/2} \mathbf{X}_m \mathbf{R}_{m, \text{RIS}}^{1/2}$, it holds that $\mathbb{E}\{\tilde{\mathbf{H}}_m \mathbf{D} \tilde{\mathbf{H}}_m^H\} = \text{tr}(\mathbf{R}_{m, \text{RIS}} \mathbf{D}) \mathbf{R}_{m, \text{AP}}$.

Proof. The left-hand side of (1) can be written as follows

$$\begin{aligned} \mathbb{E}\{\tilde{\mathbf{H}}_m \mathbf{D} \tilde{\mathbf{H}}_m^H\} &= \mathbb{E}\{\mathbf{R}_{m, \text{AP}}^{1/2} \mathbf{X}_m \mathbf{R}_{m, \text{RIS}}^{1/2} \mathbf{D} \mathbf{R}_{m, \text{RIS}}^{1/2} \mathbf{X}_m \mathbf{R}_{m, \text{AP}}^{1/2}\} \stackrel{(a)}{=} \\ &\mathbf{R}_{m, \text{AP}}^{1/2} \text{tr}(\mathbf{R}_{m, \text{RIS}}^{1/2} \mathbf{D} \mathbf{R}_{m, \text{RIS}}^{1/2}) \mathbf{R}_{m, \text{AP}}^{1/2} \stackrel{(b)}{=} \text{tr}(\mathbf{R}_{m, \text{RIS}} \mathbf{D}) \mathbf{R}_{m, \text{AP}}, \end{aligned} \quad (14)$$

where (a) is obtained by the independence of propagation channels between the AP and the RIS together with lemma [21, Lemma 2]; (b) is obtained by the property $\text{tr}(\mathbf{X}\mathbf{Y}) = \text{tr}(\mathbf{Y}\mathbf{X})$ for size-matched matrices $\mathbf{X}$ and $\mathbf{Y}$. $\square$

Lemma 2. The first moment of the aggregated channel $\mathbf{u}_{mk}$ is $\mathbb{E}\{\mathbf{u}_{mk}\} = \bar{\mathbf{H}}_m \Phi \bar{\mathbf{z}}_k$. Additionally, the covariance matrix of the NLoS component $\tilde{\mathbf{u}}_{mk}$ is

$$\mathbb{E}\{\tilde{\mathbf{u}}_{mk} \tilde{\mathbf{u}}_{mk}^H\} = \Delta_{mk} + \alpha_{mk} \mathbf{R}_{m, \text{AP}}. \quad (15)$$

Furthermore, the second moment of $\mathbf{u}_{mk}$ is computed as

$$\mathbb{E}\{\mathbf{u}_{mk} \mathbf{u}_{mk}^H\} = \bar{\mathbf{H}}_m \Phi \bar{\mathbf{z}}_k \bar{\mathbf{z}}_k^H \Phi^H \bar{\mathbf{H}}_m^H + \Delta_{mk} + \alpha_{mk} \mathbf{R}_{m, \text{AP}}. \quad (16)$$

The correlation between the NLoS components of the two aggregated channels $\mathbf{u}_{mk}$ and $\mathbf{u}_{mk'}$, $\forall k' \neq k$, is

$$\mathbb{E}\{\tilde{\mathbf{u}}_{mk} \tilde{\mathbf{u}}_{mk'}^H\} = \text{tr}(\mathbf{R}_{m, \text{RIS}} \Phi \bar{\mathbf{z}}_k \bar{\mathbf{z}}_{k'}^H \Phi^H) \mathbf{R}_{m, \text{AP}}. \quad (17)$$

Proof. The first moment of the aggregated channel $\mathbf{u}_{mk}$ is $\mathbb{E}\{\mathbf{u}_{mk}\} = \mathbb{E}\{\mathbf{g}_{mk} + (\bar{\mathbf{H}}_m + \tilde{\mathbf{H}}_m) \Phi (\bar{\mathbf{z}}_k + \tilde{\mathbf{z}}_k)\} = \mathbb{E}\{\mathbf{g}_{mk} + \bar{\mathbf{H}}_m \Phi \bar{\mathbf{z}}_k + \tilde{\mathbf{H}}_m \Phi \bar{\mathbf{z}}_k + \bar{\mathbf{H}}_m \Phi \tilde{\mathbf{z}}_k + \tilde{\mathbf{H}}_m \Phi \tilde{\mathbf{z}}_k\} = \bar{\mathbf{H}}_m \Phi \bar{\mathbf{z}}_k$. Moreover, the second moment of the NLoS component of $\mathbf{u}_{mk}$ is

$$\begin{aligned} \mathbb{E}\{\tilde{\mathbf{u}}_{mk} \tilde{\mathbf{u}}_{mk}^H\} &= \mathbf{G}_{mk} + \text{tr}(\mathbf{R}_{m, \text{RIS}} \Phi \bar{\mathbf{z}}_k \bar{\mathbf{z}}_k^H \Phi^H) \mathbf{R}_{m, \text{AP}} + \bar{\mathbf{H}}_m \times \\ &\Phi \mathbf{R}_k \Phi^H \bar{\mathbf{H}}_m^H + \text{tr}(\mathbf{R}_{m, \text{RIS}} \Phi \mathbf{R}_k \Phi^H) \mathbf{R}_{m, \text{AP}} = \Delta_{mk} + \alpha_{mk} \mathbf{R}_{m, \text{AP}}, \end{aligned} \quad (18)$$

by virtue of Corollary 1. The second moment of the channel $\mathbf{u}_{mk}$ is $\mathbb{E}\{\mathbf{u}_{mk}\mathbf{u}_{mk}^H\} = \mathbb{E}\{(\bar{\mathbf{u}}_{mk} + \tilde{\mathbf{u}}_{mk})(\bar{\mathbf{u}}_{mk} + \tilde{\mathbf{u}}_{mk})^H\} \stackrel{(a)}{=} \bar{\mathbf{u}}_{mk}\bar{\mathbf{u}}_{mk}^H + \mathbb{E}\{\tilde{\mathbf{u}}_{mk}\tilde{\mathbf{u}}_{mk}^H\} = \bar{\mathbf{H}}_m\Phi\bar{\mathbf{z}}_k\bar{\mathbf{z}}_k^H\Phi^H\bar{\mathbf{H}}_m + \Delta_{mk} + \alpha_{mk}\mathbf{R}_{m,AP}$, where (a) is obtained because of the independence between the LoS and NLoS components of the channel $\mathbf{u}_{mk}$. Moreover, the correlation between the two NLoS parts of the aggregated channels $\mathbf{u}_{mk}$ and $\mathbf{u}_{mk'}$, $k' \neq k$, is as $\mathbb{E}\{\tilde{\mathbf{u}}_{mk}\tilde{\mathbf{u}}_{mk'}^H\} = \text{tr}(\mathbf{R}_{m,RIS}\Phi\bar{\mathbf{z}}_k\bar{\mathbf{z}}_{k'}^H\Phi^H)\mathbf{R}_{m,AP}$, by virtue of Corollary 1. $\square$

B. Proof of Lemma 1

We first define the term $\tilde{\mathbf{y}}_{pmk}$, $\forall m, k$, in (3) as

$$\begin{aligned} \tilde{\mathbf{y}}_{pmk} &= \mathbf{y}_{pmk} - \mathbb{E}\{\mathbf{y}_{pmk}\} = \sqrt{p\tau_p} \sum_{k' \in \mathcal{P}_k} \mathbf{u}_{mk'} + \mathbf{w}_{pmk} \\ &\quad - \left(\sqrt{p\tau_p} \sum_{k' \in \mathcal{P}_k} \bar{\mathbf{u}}_{mk'} \right) = \sqrt{p\tau_p} \sum_{k' \in \mathcal{P}_k} \tilde{\mathbf{u}}_{mk'} + \mathbf{w}_{pmk}. \end{aligned} \quad (19)$$

By exploiting the above expression of $\tilde{\mathbf{y}}_{pmk}$, the expectation $\Gamma_{mk} = \mathbb{E}\{\tilde{\mathbf{u}}_{mk}\tilde{\mathbf{y}}_{pmk}^H\}$ can be formulated as

$$\begin{aligned} \Gamma_{mk} &= \mathbb{E}\left\{ \tilde{\mathbf{u}}_{mk} \left(\sqrt{p\tau_p} \sum_{k' \in \mathcal{P}_k} \tilde{\mathbf{u}}_{mk'}^H + \mathbf{w}_{pmk} \right)^H \right\} = \sqrt{p\tau_p} \times \\ &\quad \sum_{k' \in \mathcal{P}_k \setminus \{k\}} \mathbb{E}\{\tilde{\mathbf{u}}_{mk}\tilde{\mathbf{u}}_{mk'}^H\} + \sqrt{p\tau_p} \mathbb{E}\{\tilde{\mathbf{u}}_{mk}\mathbf{w}_{pmk}^H\} \stackrel{(a)}{=} \sqrt{p\tau_p} \times \\ &\quad \sum_{k' \in \mathcal{P}_k} \text{tr}(\mathbf{R}_{m,RIS}\Phi\bar{\mathbf{z}}_k\bar{\mathbf{z}}_{k'}^H\Phi^H)\mathbf{R}_{m,AP} + \sqrt{p\tau_p}\Delta_k \stackrel{(b)}{=} \sqrt{p\tau_p} \times \\ &\quad \times \text{tr}(\mathbf{R}_{m,RIS}\Phi\bar{\mathbf{z}}_k \left(\sum_{k' \in \mathcal{P}_k} \bar{\mathbf{z}}_{k'}^H \right) \Phi^H)\mathbf{R}_{m,AP} + \sqrt{p\tau_p}\Delta_{mk}, \end{aligned} \quad (20)$$

where (a) follows Lemma 2; and (b) is obtained by exploiting the linearity of the trace of matrices. Furthermore, the covariance matrix $\Psi_{mk} = \mathbb{E}\{\tilde{\mathbf{y}}_{pmk}\tilde{\mathbf{y}}_{pmk}^H\}$ can be formulated as

$$\begin{aligned} \Psi_{mk} &= \mathbb{E}\left\{ \left(\sqrt{p\tau_p} \sum_{k' \in \mathcal{P}_k} \tilde{\mathbf{u}}_{mk'} + \mathbf{w}_{pmk} \right) \left(\sqrt{p\tau_p} \sum_{k' \in \mathcal{P}_k} \tilde{\mathbf{u}}_{mk'}^H + \mathbf{w}_{pmk}^H \right)^H \right\} \\ &\stackrel{(a)}{=} \sum_{\substack{k', k'' \in \mathcal{P}_k \\ k' \neq k''}} \text{tr}(\mathbf{R}_{m,RIS}\Phi\bar{\mathbf{z}}_{k'}\bar{\mathbf{z}}_{k''}^H\Phi^H)\mathbf{R}_{m,AP} + \sum_{k' \in \mathcal{P}_k} p\tau_p\Delta_{mk} + p\tau_p \\ &\quad \times \sum_{k' \in \mathcal{P}_k} \text{tr}(\mathbf{R}_{m,RIS}\Phi\bar{\mathbf{z}}_{k'}\bar{\mathbf{z}}_{k'}^H\Phi^H)\mathbf{R}_{m,AP} + \mathbf{I}_M = p\tau_p \sum_{k' \in \mathcal{P}_k} \Delta_{mk'} \\ &\quad + \mathbf{I}_M + p\tau_p \text{tr} \left(\mathbf{R}_{m,RIS}\Phi \left(\sum_{k' \in \mathcal{P}_k} \bar{\mathbf{z}}_{k'} \right) \left(\sum_{k' \in \mathcal{P}_k} \bar{\mathbf{z}}_{k'}^H \right) \Phi^H \right) \mathbf{R}_{m,AP}, \end{aligned} \quad (21)$$

where (a) follows by Lemma 2 and the channel estimate obtained as in the lemma.

REFERENCES

- [1] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.
- [2] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, 2011.
- [3] J. An, L. Wang, C. Xu, L. Gan, and L. Hanzo, "Optimal pilot power based channel estimation improves the throughput of intelligent reflective surface assisted systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 16 202–16 206, 2020.
- [4] E. Shi, J. Zhang, H. Du, B. Ai, C. Yuen, D. Niyato, K. B. Letaief, and X. Shen, "RIS-aided cell-free massive MIMO systems for 6G: Fundamentals, system design, and applications," *Proceedings of the IEEE*, vol. 112, no. 4, pp. 331–364, 2024.
- [5] J. Xu, C. Yuen, C. Huang, N. Ul Hassan, G. C. Alexandropoulos, M. Di Renzo, and M. Debbah, "Reconfiguring wireless environments via intelligent surfaces for 6G: Reflection, modulation, and security," *Science China Information Sciences*, vol. 66, no. 3, p. 130304, 2023.
- [6] C. Xu, J. An, T. Bai, S. Sugiura, R. G. Maunder, Z. Wang, L.-L. Yang, and L. Hanzo, "Channel estimation for reconfigurable intelligent surface assisted high-mobility wireless systems," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 1, pp. 718–734, 2022.
- [7] T. Van Chien, H. Q. Ngo, S. Chatzinotas, M. Di Renzo, and B. Ottersten, "Reconfigurable intelligent surface-assisted cell-free massive MIMO systems over spatially-correlated channels," *IEEE Transactions on Wireless Communications*, vol. 21, no. 7, pp. 5106–5128, 2021.
- [8] Q. Li, M. El-Hajjar, I. Hemadeh, A. Shojaeifard, and L. Hanzo, "Low-Overhead Channel Estimation for RIS-Aided Multi-Cell Networks in the Presence of Phase Quantization Errors," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 5, pp. 6626 – 6641, 2023.
- [9] J. An, C. Xu, L. Wang, Y. Liu, L. Gan, and L. Hanzo, "Joint training of the superimposed direct and reflected links in reconfigurable intelligent surface assisted multiuser communications," *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 2, pp. 739–754, 2022.
- [10] G. Yan, L. Zhu, and R. Zhang, "Passive reflection optimization for IRS-aided multicast beamforming with discrete phase shifts," *IEEE Wireless Communications Letters*, vol. 12, no. 8, pp. 1424–1428, 2023.
- [11] F. Jiang, A. Abrardo, K. Keykhosravi, H. Wymeersch, D. Dardari, and M. Di Renzo, "Two-timescale transmission design and RIS optimization for integrated localization and communications," *IEEE Transactions on Wireless Communications*, vol. 22, no. 12, pp. 8587–8602, 2023.
- [12] J. Dai, J. Ge, K. Zhi, C. Pan, Z. Zhang, J. Wang, and X. You, "Two-timescale transmission design for RIS-aided cell-free massive MIMO systems," *IEEE Transactions on Wireless Communications*, 2023.
- [13] P.-Q. Huang, Y. Zhou, K. Wang, and B.-C. Wang, "Placement optimization for multi-IRS-aided wireless communications: An adaptive differential evolution algorithm," *IEEE Wireless Communications Letters*, vol. 11, no. 5, pp. 942–946, 2022.
- [14] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on communications*, vol. 48, no. 3, pp. 502–513, 2000.
- [15] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993.
- [16] K. Zhi, C. Pan, H. Ren, K. Wang, M. El-kashlan, M. Di Renzo, R. Schober, H. V. Poor, J. Wang, and L. Hanzo, "Two-timescale design for reconfigurable intelligent surface-aided massive MIMO systems with imperfect CSI," *IEEE Transactions on Information Theory*, vol. 69, no. 5, pp. 3001–3033, 2022.
- [17] R. Tanabe and A. Fukunaga, "Success-history based parameter adaptation for differential evolution," in *2013 IEEE congress on evolutionary computation*. IEEE, 2013, pp. 71–78.
- [18] Ö. T. Demir and E. Björnson, "Is channel estimation necessary to select phase-shifts for RIS-assisted massive MIMO?" *IEEE Transactions on Wireless Communications*, vol. 21, no. 11, pp. 9537–9552, 2022.
- [19] K. Zhi, C. Pan, H. Ren, and K. Wang, "Power scaling law analysis and phase shift optimization of RIS-aided massive MIMO systems with statistical CSI," *IEEE Transactions on Communications*, vol. 70, no. 5, pp. 3558–3574, 2022.
- [20] M. B. Taha, C. Talhi, H. Ould-Slimane, S. Alrabaei, and K.-K. R. Choo, "A multi-objective approach based on differential evolution and deep learning algorithms for VANETs," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 3, pp. 3035–3050, 2022.
- [21] H. A. Le, T. Van Chien, W. Choi *et al.*, "Double RIS-assisted MIMO systems over spatially correlated Rician fading channels and finite scatterers," *IEEE Transactions on Communications*, vol. 71, no. 8, pp. 4941–4956, 2023.