A Unified Framework for STAR-RIS Coefficients Optimization

Hancheng Zhu, *Student Member, IEEE,* Yuanwei Liu, *Fellow, IEEE,* Yik-Chung Wu, *Senior Member, IEEE,* Vincent K. N. Lau, *Fellow, IEEE*

***Abstract*—Simultaneously transmitting and reflecting (STAR) reconfigurable intelligent surface (RIS) has recently emerged as a promising enhancement to the traditional reflective only RIS. In view of the difficulty of comparing wireless systems equipped with different modes of STAR-RIS and the performance degradation caused by the constraints involving discrete selection, this paper proposes a unified optimization framework for handling the constraints arising from various STAR-RIS operating modes and discrete phase coefficients. With a judiciously introduced penalty term, this framework transforms the original problem into two iterative subproblems, with one containing the selection-type constraints, and the other subproblem handling other wireless resource. Convergent point of the whole algorithm is found to be at least a stationary point under mild conditions. As an illustrative example, the proposed framework is applied to a sum-rate maximization problem in the downlink transmission. Simulation results show that the algorithms from the proposed framework not only outperform other existing algorithms tailored for different STAR-RIS scenarios, but also facilitate a fair and unified comparison among different operating modes of STAR-RIS. Furthermore, it is found that 4 or even 2 discrete phases STAR-RIS could achieve almost the same sum-rate performance as the continuous phase setting, showing for the first time that discrete phase is not necessarily a cause of significant performance degradation.**

***Index Terms*—simultaneously transmitting and reflecting reconfigurable intelligent surface (STAR-RIS), operating mode constraint, discrete phase constraint, unified framework.**

I. INTRODUCTION

The development of meta-surface/tunnel diode technology provides a feasible avenue for the realization of reflective intelligent surfaces (RIS) in the forthcoming communication systems [1], [2]. RIS, characterized by its cost-effectiveness and remarkable scalability, leverages an array of reflective meta-surfaces affixed to walls or physical infrastructures. These meta-surfaces are endowed with the capability to manipulate its phase shifts and amplitudes, thereby changing the propagation channels [3]–[6]. By redirecting incoming signals towards the receiver, the RIS establishes a virtual direct pathway connecting the transmitter and receiver, effectively circumventing physical obstructions [7]. This innovation therefore facilitates wireless transmissions across a multitude of challenging scenarios [8], [9].

Traditionally, RIS solely functions to reflect signals, limiting its coverage to receivers positioned on the same side as the transmitters. A recent ground-breaking approach to achieve full 360° coverage is the concept of Simultaneously Transmitting and Reflecting (STAR) RIS [10]–[12]. This innovative paradigm enables the concurrent reflection and transmission (refraction) of incident signals, effectively catering to users situated on both sides of the surface.

To fully unlock the potential in STAR-RIS technology, the transmitting and reflecting coefficients need to be properly optimized. Compared to conventional RISs, the optimization of the STAR-RIS coefficients is subjected to a set of intricate constraints. For example, there are three distinct operational modes - energy splitting (ES), mode switching (MS) and time switching (TS). In ES mode, the principles of energy conservation and lossless power considerations dictate that the sum power of transmitting and reflecting coefficients confines to unity. On the other hand, operation under MS mode allocates each constituent element of the STAR-RIS to either transmission or reflection, resulting in a mixed-integer programming (MIP) problem, which is NP-hard even under the simple quadratic objective function [13]. To address this intricacy, a prevalent strategy involves the conversion of integer constraints into a penalty term and adds it to the objective function [14]–[16]. As the penalty weight increases, each STAR-RIS element is forced to either transmission or reflection. In TS mode, all elements of the STAR-RIS are dedicated to fully transmission or reflection at any given time. Therefore, the amplitude of the STAR-RIS transmission and reflection coefficients are constrained to be one. In this mode, the optimization problem is to determine the phase of the STAR-RIS coefficients and the optimal time allocation for reflection and transmission [15].

Furthermore, recent investigations have unveiled a distinctive phase coupling phenomenon in ES STAR-RIS. This intriguing property dictates that the difference between the phases of reflection and transmission is consistently maintained at $\pi/2$ [17]–[19]. This constraint does not exists in reflective only RIS so its exploration is still at its infancy. The earliest attempt to incorporate this constraint within STAR-RIS optimization employs the element-wise alternating optimization (AO) method [19]. In this approach, optimization is carried out for individual RIS elements sequentially so

Manuscript received 01 September 2023; revised 10 March 2024 and 11 May 2024; accepted 28 May 2024.This work was supported in part by Hong Kong Research Grants Council under Grant 16214122. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Foad Sohrabi. *(Corresponding author: Yik-Chung Wu.)*

Hancheng Zhu and Yik-Chung Wu are with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong (email: u3006551@hku.hk, ycwu@eee.hku.hk).

Yuanwei Liu are with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K (email: yuanwei.liu@qmul.ac.uk).

Vincent K. N. Lau is with Hong Kong University of Science and Technology, Hong Kong, China (email:eeknlau@ee.ust.hk).

that each subproblem has only two potential choices (either transmission phase surpasses the reflection phase by $\pi/2$ or vice versa), and the solution with better objective value is then chosen as the subsequent iteration point. Although this leads to a duplication in the number of optimization subproblems, it circumvents the combinatorial challenge of jointly determining the coupled-phase options of all STAR-RIS elements. More recently, [20] proposes to transform the coupled-phase constraint and the previously mentioned sum power constraint into penalty terms. Then, AO was employed to alternatively optimize the amplitude and phase of the STAR-RIS coefficients. Compared with elementwise-AO method, the solution quality of this penalty-based algorithm is guaranteed when the problem satisfies the Mangasarian-Fromovitz constraint qualification (MFCQ) conditions.

While the prevailing model in existing STAR-RIS research adopts continuous phase shifts, it is essential to acknowledge that practical limitations from finite control signal resolution or hardware constraints leads to a finite number of permissible phases [21]. Such discrete phase phenomenon has emerged as one of the basic models in reflective only RIS. It is reasonable to anticipate that STAR-RIS encounters the same discrete phase constraints. However, the combinatorial nature of discrete phase optimization has an exponential compleixty order with the number of STAR-RIS elements. This inherent complexity renders exhaustive searching approach computationally infeasible for STAR-RIS with massive elements. Therefore, numerous extant studies would simply ignore this discrete phase constraint during optimization, and then apply quantization to the resultant continuous phase design [22]. Regrettably, such quantization-based strategy lacks a guarantee in solution quality, and leads to a significant performance loss if the number of allowable phases is small (e.g., 2 or 4).

Through the preceding deliberations, it is evident that different constraints in STAR-RIS were handled with a multitude of optimization techniques, which makes the comparison of different types of STAR-RIS difficult. Furthermore, when multiple mentioned constraints appear concurrently, it is unclear which of the existing methods can be generalized to such scenario. To fill this gap, this paper for the first time establishes an innovative unified framework for handling the operating mode and discrete phase constraints. Through the introduction of auxiliary variables for the STAR-RIS coefficients, we strategically transform the original problem into two distinct subproblems. One subproblem is dedicated to constraints involving discrete selection, while the other addresses the sum power constraint along with the additional wireless resource constraints. This strategic decoupling enables the derivation of a closed-form global optimal solution for the subproblem associated with selection constraints, which facilitates the proof of solution quality of the proposed framework. Specifically, the converged solution is guaranteed to be at least a stationary point under mild conditions. To the best of our knowledge, this is an inaugural work that provides solution quality guarantee in various STAR-RIS configurations, even under discrete phases.

To illustrate the efficiency of the proposed framework, we apply it to a downlink STAR-RIS assisted sum-rate maximization system. Through this framework, the solutions for various types of STAR-RIS can be obtained simultaneously. Simulation results show that the proposed framework outperforms other existing methods. Furthermore, the proposed framework enables performance of discrete phase setting akin to that of continuous phase even as sparse as four or two discrete phases, which overthrows the conventional notion that discrete phase is to be blamed for performance degradation. Since the proposed framework separates the handling of selection-type constraints arising from STAR-RIS and the optimization of other wireless resource, it can be easily extended to communication scenarios involving other objective functions and radio resources. Such versatility substantially mitigates the anticipated challenges associated with future resource allocation problem under STAR-RIS.

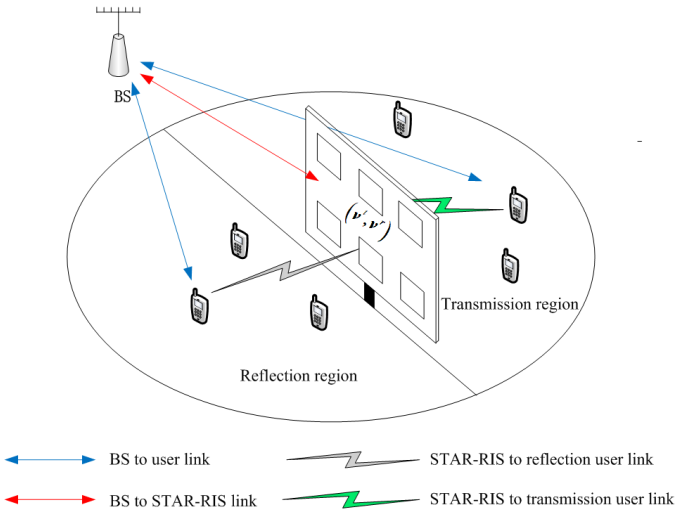


Fig. 1: A typical STAR-RIS assisted communication system

The rest of the paper is organized as follows. The general STAR-RIS aided communication model, its penalty formulation, and the conditions for solution quality guarantee are presented in Section II. Then the closed-form solution of the selection related subproblem is derived in Section III. The downlink STAR-RIS assisted sum-rate maximization is formulated and solved according to the proposed penalty framework in Section IV. Simulation results are provided in Section V and conclusions are drawn in Section VI.

II. A General STAR-RIS Optimization Model

We consider a communication system assisted by a STAR-RIS as shown in Fig. 1. Users are distributed on both sides of STAR-RIS with M elements. We denote the transmission coefficient and the reflection coefficient at the m^{th} STAR-RIS element as $v_m^t \in \mathbb{C}$ and $v_m^r \in \mathbb{C}$ respectively. Depending on the types of STAR-RIS, the STAR-RIS coefficients are subjected to different constraints, which are discussed below.

A. Modeling of STAR-RIS Constraints

There are three types of constraints for STAR-RISs, namely operating mode constraint, lossless power constraint and phase constraint. These constraints are detailed as follows.

TABLE I: Typical STAR-RIS models

STAR-RIS Case Index	Operating mode	Amplitude constraint	Coupled-phase constraint $\angle v_m^t - \angle v_m^r = \{\pi/2\,(\mathrm{mod}2\pi), -\pi/2\,(\mathrm{mod}2\pi)\}$	Discrete phase constraint $\{0, 2\pi/L, \cdots 2\pi\,(L-1)/L\}$	Time allocation constraint $\lambda^t + \lambda^r = 1$	Existing works employing this model
1	TS	1	✘	✘	✔	[15], [23], [24]
2			✘	✔	✔	[25], [26]
3	MS	$\{0,1\}$	✘	✘	✘	[14], [24], [27]
4			✘	✔	✘	No existing work considered this model yet
5	ES	$[0,1]$	✘	✘	✘	[28]–[30]
6			✘	✔	✘	[31], [32]
7			✔	✘	✘	[20], [23], [33] [19], [34]
8			✔	✔	✘	[34]

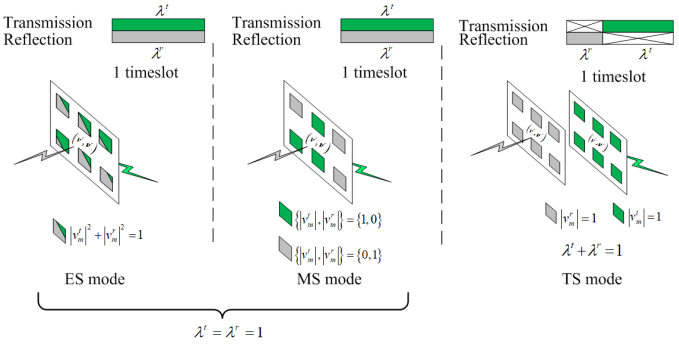


Fig. 2: Three operating modes of the STAR-RIS

1) **Operating mode constraint**. There are commonly three operating modes of STAR-RIS [15]. The first one is ES mode, where the incident signal is split between reflection and transmission, giving rise to the constraint $\left|v_m^t\right|, |v_m^r| \in [0,1]$. The second one is MS mode where each STAR-RIS element is dedicated to reflection or transmission, giving rise to the constraint $\left|v_m^t\right|, |v_m^r| \in \{0,1\}$. The third one is TS mode which divides the transmission interval into two sub-intervals, with one sub-interval dedicated to reflection while the other to transmission. Let $\lambda^t \geq 0$ and $\lambda^r \geq 0$ denoting the percentage of the time allocated to transmission period and reflection period respectively, we have $\lambda^t + \lambda^r = 1$. On the other hand, for ES and MS modes, as both transmission and reflection occupy the whole interval, we can set $\lambda^t = \lambda^r = 1$.
2) **Lossless power constraint**. It is usually assumed that the metasurface is lossless. Hence, in ES and MS mode, the reflected energy plus transmitted energy must be equal to the incident signal energy. This gives rise to the constraint $\left|v_m^t\right|^2 + |v_m^r|^2 = 1$. On the other hand, in TS mode, since transmission or reflection is the only operation in a certain time interval, lossless constraint means $\left|v_m^t\right| = |v_m^r| = 1$. Fig. 2 illustrates the STAR-RIS operating in different modes and the corresponding constraints.
3) **Phase constraint**. It is known that the STAR-RIS phase may not take infinite resolution in practice. In this case, $\angle v_m^t,\ \angle v_m^r \in \{0, 2\pi/L, \cdots 2\pi\,(L-1)/L\}$, where L is the number of allowable phases. More recently, a new coupled-phase model is proposed in [17]–[19], which states that a necessary condition for a physically realizable STAR-RIS should satisfy $\left|v_m^t\right| |v_m^r| \cos\left(\angle v_m^t - \angle v_m^r\right) = 0$. In ES mode, this is equivalent to $\angle v_m^t - \angle v_m^r \in \{\pi/2\,(\mathrm{mod}2\pi), -\pi/2\,(\mathrm{mod}2\pi)\}$.

Different types of STAR-RIS are the results of mix and match of above constraints, and they are summarized in Table I. Notice that MS with coupled-phase is just the basic MS STAR-RIS since MS requires $\left|v_m^t\right|, |v_m^r| \in \{0,1\}$. Together with the lossless power constraint $\left|v_m^t\right|^2 + |v_m^r|^2 = 1$, we must have either $\left|v_m^t\right| = 0$ or $|v_m^r| = 0$. This makes the coupled-phase constraint $\left|v_m^t\right| |v_m^r| \cos\left(\angle v_m^t - \angle v_m^r\right) = 0$ automatically satisfied. For the TS mode, since only reflection phase or transmission phase of each STAR-RIS element is used in a certain time period, coupled-phase constraint could not exist. Therefore, MS and TS STAR-RIS with coupled-phase (with or without discrete phase constraint) are not feasible and we do not list them in Table I.

Let $\boldsymbol{z}$ denote other communication resources to be optimized and $\boldsymbol{v}^t = \left[v_1^t, \cdots, v_M^t\right]^T$, $\boldsymbol{v}^r = \left[v_1^r, \cdots, v_M^r\right]^T$ be the shorthand notations for the collection of $\left\{v_m^t\right\}_{m=1}^M$ and $\{v_m^r\}_{m=1}^M$ respectively, a general optimization problem involving STAR-RIS can be formulated as

$$\min_{\{\boldsymbol{z},\boldsymbol{v}^t,\boldsymbol{v}^r,\lambda^t,\lambda^r\}} \quad \mathscr{F}\left(\boldsymbol{z}, \boldsymbol{v}^t, \boldsymbol{v}^r, \lambda^t, \lambda^r\right) \tag{1a}$$

$$s.t. \quad \lambda^t \left|v_m^t\right|^2 + \lambda^r |v_m^r|^2 = 1, \tag{1b}$$

$$\begin{cases} \left|v_m^t\right| = |v_m^r| = 1, \left\{\lambda^t, \lambda^r\right\} \geq 0, & \text{if TS} \\ \left|v_m^t\right|, |v_m^r| \in \{0,1\}, \lambda^t = \lambda^r = 1, & \text{if MS} \\ \angle v_m^t - \angle v_m^r \in \{\pi/2\,(\mathrm{mod}2\pi), \\ \quad -\pi/2\,(\mathrm{mod}2\pi)\}, \lambda^t = \lambda^r = 1, & \text{if ES} \end{cases} \tag{1c}$$

$$\angle v_m^t, \angle v_m^r \in \{0, 2\pi/L, \cdots, 2\pi\,(L-1)/L\}, \tag{1d}$$

$$\left(\boldsymbol{z}, \boldsymbol{v}^t, \boldsymbol{v}^r, \lambda^t, \lambda^r\right) \in \boldsymbol{\Omega}. \tag{1e}$$

where $\mathscr{F}$ is the objective function and is assumed to be bounded from below, which is a trivial assumption since the problem in (1) is a minimization problem. The objective function in (1) can represent different forms of system per-

formance. For example, it can be power consumption [15], [19], or mean square error function [26]. On the other hand, the sum-rate [14], [20], spectral efficiency [29] and secrecy capacity [30], [34] are also frequently used optimization objectives, but they need a negative sign if they are used in (1) since these criteria should be maximized instead of minimized. $\boldsymbol{\Omega}$ is the coupled constraint set of $\boldsymbol{z}$, $\boldsymbol{v}^t$, $\boldsymbol{v}^r$, λ^t, λ^r. Constraint (1b) is a general expression covering all three modes of STAR-RIS. In particular, when $\left|v_m^t\right| = |v_m^r| = 1$, (1b) reduces to time allocation constraint $\lambda^t + \lambda^r = 1$ in the TS mode. On the other hand, in ES or MS mode, they do not involve the allocated time variables λ^t, λ^r and therefore setting $\lambda^t = \lambda^r = 1$ make (1b) reduces to the lossless constraint $\left|v_m^t\right|^2 + |v_m^r|^2 = 1$. Notice that the ES mode constraint $\left|v_m^t\right|$, $|v_m^r| \in [0,1]$ is implicitly included in (1b), therefore it is not listed in (1). The discrete phase constraint (1d) are compatible with the coupled-phase constraint (third line of (1c)) if L is an even number greater than 2. For example, when $L = 4$, $\angle\varphi_m^t \in \{0, \pi/2, \pi, 3\pi/2\}$ and the coupled constraint would make the reflection phase $\angle\varphi_m^r = \angle\varphi_m^t \pm \pi/2$, which is still in $\{0, \pi/2, \pi, 3\pi/2\}$. Similar observations can be made as long as $L > 2$ and is an even number. This condition can be easily satisfied when the number of information bit for phase control is larger than 1.

Existing works only solve special cases of (1). For example, when only (1b) and the second constraint in (1c) are included, this corresponds to the typical MS model, and a penalty term $\left|v_m^t\right|^2 - \left|v_m^t\right|$ is commonly introduced to relax the $\{0,1\}$ constraint into $[0,1]$ [14], [24], [27]. On the other hand, when only (1b) with $\lambda^t = \lambda^r = 1$ and (1d) are included, (1) becomes the discrete ES STAR-RIS, and a common approach is to relax the discrete phase temporarily and then quantizing the continuous-valued result into discrete phase [31]. Furthermore, with (1b) and the first case of (1c), the general formulation (1) becomes the TS STAR-RIS optimization problem. Since the transmission and reflection coefficients are not coupled in TS STAR-RIS and the time allocation constraint is convex, techniques such as semidefinite relaxation (SDR) [15] and gradient descent (GD) method [25] for traditional reflection only RIS can be adopted to handle the phase optimization. Recently, a coupled-phase STAR-RIS model was considered in [19], [20], where (1b) and the third line of (1c) are included. To overcome the difficulty introduced by the elementwise coupled-phase constraint, [19] proposes an elementwise-AO method, and [20] puts forward a penalty-based algorithm by moving both (1b) and the second line of (1c) into a penalty term. Also, [34] uses the same penalty form to solve the secrecy beamforming problem under the coupled-phase STAR-RIS.

Although there exists studies handling different special cases of problem (1) by invoking different optimization schemes, a unified framework for solving the general problem including diverse types of STAR-RIS in (1) is missing. Especially, the discrete phase-shift constraint (1d) is dominantly handled by quantization, which may introduce noticeable performance loss when the number of allowable phases is small (e.g., $L = 2$ or 4), not to mention the lack of solution quality guarantee by such approach. Coming up with a unified framework not only saves the effort of finding optimization algorithms for various special cases falling into the form of (1), but also facilitates the comparison among various STAR-RIS models in a particular communication scenario.

B. A Penalty-based Reformulation of (1)

Notice that problem (1) is challenging to solve for STAR-RIS coefficients $\boldsymbol{v}^t$ and $\boldsymbol{v}^r$ since the constraint (1c) contains binary selection. More specifically, the second constraint and the third constraint in (1c) are the binary selection for the amplitude and the difference of the two phases, respectively. Moreover, possible occurance of the discrete phases (1d) also makes (1) a mixed integer optimization problem.

To tackle this problem, we propose to employ auxiliary vectors $\boldsymbol{\varphi}^t$, $\boldsymbol{\varphi}^r \in \mathbb{C}^{M\times 1}$ together with a penalty term to handle constraints (1c) and (1d). The reformulated problem is written as

$$\min_{\left\{\substack{\boldsymbol{z},\boldsymbol{v}^t,\boldsymbol{v}^r,\boldsymbol{\varphi}^t,\\ \boldsymbol{\varphi}^r,\lambda^t,\lambda^r}\right\}} \mathscr{F}\left(\boldsymbol{z},\boldsymbol{v}^t,\boldsymbol{v}^r,\lambda^t,\lambda^r\right) + \frac{\gamma}{2}\sum_{p=t,r}|\boldsymbol{v}^p - \boldsymbol{\varphi}^p|_2^2 \tag{2a}$$

$$s.t. \quad \begin{cases} \left|\varphi_m^t\right| = |\varphi_m^r| = 1, \left\{\lambda^t,\lambda^r\right\} \geq 0, & \text{if TS} \\ \left|\varphi_m^t\right|, |\varphi_m^r| \in \{0,1\}, \lambda^t = \lambda^r = 1, & \text{if MS} \\ \angle\varphi_m^t - \angle\varphi_m^r \in \{\pi/2\,(\mathrm{mod}2\pi), \\ \quad -\pi/2\,(\mathrm{mod}2\pi)\}, \lambda^t = \lambda^r = 1, & \text{if ES} \end{cases} \tag{2b}$$

$$\angle\varphi_m^t, \angle\varphi_m^r \in \{0, 2\pi/L, \cdots, 2\pi\,(L-1)/L\}, \tag{2c}$$

$$(1b), (1e)$$

where γ is the penalty coefficient. When the penalty coefficient increases, the RIS coefficient vectors $\boldsymbol{v}^t$ and $\boldsymbol{v}^r$ will be forced to take the same values as the auxiliary vectors $\boldsymbol{\varphi}^t$ and $\boldsymbol{\varphi}^r$, respectively, which makes $\boldsymbol{v}^t$ and $\boldsymbol{v}^r$ satisfy the constraints (1c) and (1d).

Recognizing that the constraints for $\left\{\boldsymbol{\varphi}^t,\boldsymbol{\varphi}^r\right\}$ and $\left\{\boldsymbol{z},\boldsymbol{v}^t,\boldsymbol{v}^r,\lambda^t,\lambda^r\right\}$ in (2) are not coupled, BCD framework can be adopted to handle this problem, which involves solving the following two subproblems alternatively:

$$\mathrm{P1}: \min_{\{\boldsymbol{\varphi}^t,\boldsymbol{\varphi}^r\}} \left|\boldsymbol{v}^t - \boldsymbol{\varphi}^t\right|_2^2 + |\boldsymbol{v}^r - \boldsymbol{\varphi}^r|_2^2$$

$$s.t. \quad \begin{cases} \left|\varphi_m^t\right| = |\varphi_m^r| = 1, & \text{if TS} \\ \left|\varphi_m^t\right|, |\varphi_m^r| \in \{0,1\}, & \text{if MS} \\ \angle\varphi_m^t - \angle\varphi_m^r \in \{\pi/2\,(\mathrm{mod}2\pi), \\ \quad -\pi/2\,(\mathrm{mod}2\pi)\}, & \text{if ES} \end{cases}$$

$$\angle\varphi_m^t, \angle\varphi_m^r \in \{0, 2\pi/L, \cdots, 2\pi\,(L-1)/L\}.$$

$$\mathrm{P2}: \min_{\left\{\substack{\boldsymbol{z},\boldsymbol{v}^t,\boldsymbol{v}^r,\\ \lambda^t,\lambda^r}\right\}} \mathscr{F}\left(\boldsymbol{z},\boldsymbol{v}^t,\boldsymbol{v}^r,\lambda^t,\lambda^r\right) + \frac{\gamma}{2}\sum_{p=t,r}|\boldsymbol{v}^p - \boldsymbol{\varphi}^p|_2^2$$

$$s.t. \quad \begin{cases} \lambda^t + \lambda^r = 1, \left\{\lambda^t,\lambda^r\right\} \geq 0, & \text{if TS} \\ \left|v_m^t\right|^2 + |v_m^r|^2 = 1, \\ \lambda^t = \lambda^r = 1, & \text{if MS/ES} \end{cases}$$

$$\left(\boldsymbol{z},\boldsymbol{v}^t,\boldsymbol{v}^r,\lambda^t,\lambda^r\right) \in \boldsymbol{\Omega}$$

The advantage of solving subproblems P1 and P2 iteratively is that it separates the constraints due to discrete selection (i.e., STAR-RIS coefficients of (2b) and (2c)) from the objective function $\mathscr{F}$. This means that when we solve subproblem P2, we do not need to consider discrete selection constraints, and when we solve P1, we do not need to consider $\mathscr{F}$ and other communication resources. Together with an increasing weight of the penalty term, the overall algorithm for solving (1) is summarized in **Algorithm 1**.

Notice that P2 appears as an optimization problem similar to that in many conventional wireless systems. In fact, the proposed framework intentionally handles the discrete constraints arising from STAR-RIS in P1, so that P2 can be tackled by traditional optimization techniques. Therefore, the proposed framework covers any wireless system and resource allocation objective, with the variation of specific details only reflected in P2 but not P1. This judicious design allows easy incorporation of STAR-RIS to any wireless system, without complicating the corresponding optimization problem.

One may doubt the necessity of introducing auxiliary variables in the proposed framework, as setting $\gamma = 0$ also break the original problem (1) into two subproblems, with the first subproblem being a projection onto discrete constraint, while the second subproblem handling other continuous constraints. However, such approach would lead to both subproblems containing constraints of v_m^t and v_m^r. Then the projection operation in the first subproblem would lead to violation of the constraints in another subproblem, resulting in infeasible solution for the overall problem. In contrast, the proposed framework uses the auxiliary variables $\{\boldsymbol{\varphi}^t, \boldsymbol{\varphi}^r\}$ to satisfy the STAR-RIS coefficient constraints of P1 and $\{\boldsymbol{v}^t, \boldsymbol{v}^r\}$ to satisfy the lossless power constraints $|v_m^t|^2 + |v_m^r|^2 = 1$ of P2. Then by using penalty method, $\boldsymbol{v}^t$ and $\boldsymbol{v}^r$ are forced to be close enough to $\boldsymbol{\varphi}^t$ and $\boldsymbol{\varphi}^r$. Consequently, $\boldsymbol{v}^t$ and $\boldsymbol{v}^r$ will satisfy all the STAR-RIS coefficient constraints in the proposed framework.

In a recent work dealing with ES STAR-RIS with coupled-phase [20], the idea of penalty is also employed. The difference in the proposed framework is that we cover different STAR-RIS types (shown in Table I) while [20] is only tailored for ES STAR-RIS with coupled-phase. Furthermore, we include discrete phase constraint in the penalty while [20] did not, making the proposed framework more general. More importantly, we retain the constraint (1b) in the subproblem P2 with respect to the original optimization variables, while [20] enforces (1b) using auxiliary variables. Although the last point seems to be a unremarkable difference, this subtle change leads to significant consequences due to the following reasons:

1) For the subproblem P1, closed-form global optimal solutions for all STAR-RIS models in Table I can be obtained, even with the presence of discrete phase constraint (1d). Details will be presented in the next section. In contrast, in the corresponding subproblem of [20], it requires alternatively solving for the amplitude and phase of $\boldsymbol{\varphi}^t$ and $\boldsymbol{\varphi}^r$, which slow down the convergence of the algorithm at the subproblem level.
2) Under the special case of ES STAR-RIS with coupled-phase in (2), the split of the constraints (1b) and (2b)-(2c) means that we are moving the solution of basic ES STAR-RIS (corresponding to subproblem P2) toward the solution of the coupled-phase ES STAR-RIS. However, the penalty method in [20] is to move the unconstrained STAR-RIS (with no operating mode information) solution to approach the coupled-phase ES STAR-RIS solution. In this case, the penalty weight required by the framework in [20] would be larger than that of the proposed framework to reach convergence. Since the penalty loop is the outer layer of the algorithm, a small penalty ratio for reaching convergence reduces the computational time of the proposed algorithm. This point will be further illustrated in the simulation results section.

Algorithm 1 Penalty-based BCD algorithm

Input:

Initialize γ, increasing ratio of the penalty $c > 1$, and the penalty fulfillment threshold δ.

General step:

Initialize a feasible starting point for $\boldsymbol{z}$, $\boldsymbol{v}^t$, $\boldsymbol{v}^r$, $\boldsymbol{\varphi}^t$, $\boldsymbol{\varphi}^r$, λ^t, λ^r .

Do

For $n = 0, 1, 2, ...$ execute the following steps:

optimize $\{\boldsymbol{z}, \boldsymbol{v}^t, \boldsymbol{v}^r, \lambda^t, \lambda^r\}$ by solving P2.

optimize $\{\boldsymbol{\varphi}^t, \boldsymbol{\varphi}^r\}$ by solving P1.

until the objective function (2a) converge.

$\gamma \leftarrow c\gamma$.

until $\max_m |v_m^t - \varphi_m^t| \leq \delta$ and $\max_m |v_m^r - \varphi_m^r| \leq \delta$.

Before we present the solution to the subproblem P1 in the next section, we reveal the following proposition about the solution quality of the **Algorithm 1**.

Proposition 1. *Suppose the solution from solving* P2 *takes finite value for all* n*. Also assume that a stationary solution of* P2 *and the global optimal solution of* P1 *can be obtained. Then we have*

1) *The limit point generated by the BCD loop in* ***Algorithm 1*** *is a stationary solution of* (2) *under any fixed penalty* γ*.*
2) *As the penalty weight* γ *increase to* ∞*, the limit point generated by* ***Algorithm 1*** *is a stationary point of the original problem* (1)*.*

Proof. See Appendix A in supplementary material.

Proposition 1 gives the legitimacy of using BCD under a fixed penalty parameter, and the penalty framework to enforce the STAR-RIS constraints in (1). Notice that convergence of solution also guarantees the convergence of the objective function value. Therefore, **Proposition 1** is stronger than just the objective function value convergence. To facilitate the understanding of the proposed penalty framework and the solution quality guarantee by **Proposition 1**, Fig. 3 summarizes various key problem formulations and the conditions of convergence.

Although penalty method and the BCD method are commonly used in multivariate optimization, convergence are only thoroughly studied in cases with smooth objective function and

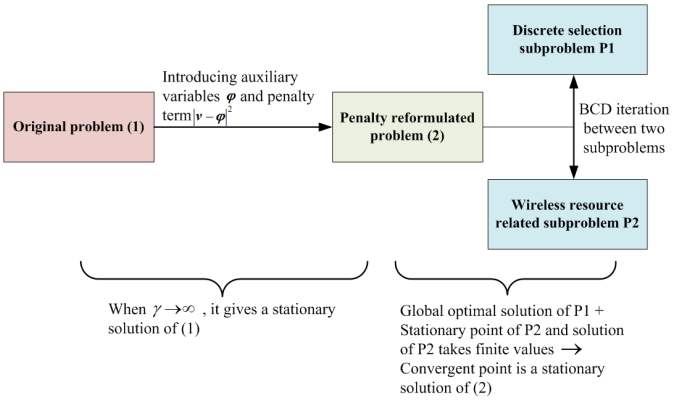


Fig. 3: Summary of the proposed framework and its solution quality

constraints [35], [36]. However, the problem in this paper is a mixed discrete-continuous optimization which may contains discrete phase and 0-1 amplitude of the STAR-RIS coefficient. Current research of BCD and penalty method in discrete optimization are mainly for linear and polynomial optimization [37], [38]. Even there exists a recent work focusing on nonlinear mixed discrete programming [39], many strict assumptions, which are hard to be verified, are imposed to ensure the solution quality. Therefore, even though we employ penalty and BCD methods for optimization, due to the mixed discrete-continuous variables, the solution quality and convergence guarantee cannot be covered by existing theory. This challenge is resolved in this paper through the convergence guarantee from **Proposition 1**, which enable other researchers to be free from the burden of analyzing the convergence themselves.

Regarding the assumption in **Proposition 1**, the first assumption is trivial and can be easily satisfied since the wireless resources are limited and their corresponding optimization variables should not grow to infinity. For the requirement of the stationary solution of P2, since no discrete selection constraint is involved, it is easily satisfied from solutions obtained by successive convex approximation (SCA) technique [40], [41] or first-order optimization methods [42], [43] in many communication problems. On the other hand, it seems quite daunting at the first sight as the global optimal solution of P1 is required. We will reveal in the next section that this is possible.

III. Optimizing Auxiliary Variables $\boldsymbol{\varphi}$ in P1

Notice that the elements of $\boldsymbol{\varphi}^t$ and $\boldsymbol{\varphi}^r$ are not coupled in P1. Therefore, P1 can be parallelized into M subproblems with the m^{th} subproblem given by

$$\begin{aligned}\min_{\varphi_m^t,\varphi_m^r}\quad & \left|v_m^t-\varphi_m^t\right|^2+\left|v_m^r-\varphi_m^r\right|^2\\ s.t.\quad & \begin{cases}\left|\varphi_m^t\right|=\left|\varphi_m^r\right|=1, & \text{if TS}\\ \left|\varphi_m^t\right|,\left|\varphi_m^r\right|\in\{0,1\}, & \text{if MS}\\ \angle\varphi_m^t-\angle\varphi_m^r\in\{\pi/2\,(\mathrm{mod}2\pi),\\ \quad -\pi/2\,(\mathrm{mod}2\pi)\}, & \text{if ES}\end{cases}\\ & \angle\varphi_m^t,\angle\varphi_m^r\in\{0,2\pi/L,\cdots,2\pi\,(L-1)/L\}.\end{aligned}\tag{3}$$

As shown in Table I, different combinations of the constraints in (3) result in different types of STAR-RIS. Below, we divide the discussion in two cases. First, we consider the cases without coupled-phase constraint, which corresponds to cases 1-6 in Table I. Then we discuss coupled-phase cases 7 and 8 in Table I.

STAR-RISs without coupled-phase: The resulting subproblem is

$$\min_{\varphi_m^t,\varphi_m^r}\quad \left|v_m^t-\varphi_m^t\right|^2+\left|v_m^r-\varphi_m^r\right|^2 \tag{4a}$$

$$s.t.\quad \begin{cases}\left|\varphi_m^t\right|=\left|\varphi_m^r\right|=1, & \text{if TS}\\ \left|\varphi_m^t\right|,\left|\varphi_m^r\right|=\{0,1\}\ \text{or}\ \{1,0\}, & \text{if MS}\\ \varphi_m^t,\varphi_m^r\in\mathbb{C}, & \text{if ES}\end{cases}\tag{4b}$$

$$\angle\varphi_m^t,\angle\varphi_m^r\in\{0,2\pi/L,\cdots,2\pi\,(L-1)/L\}. \tag{4c}$$

Generally, this is a problem containing interger variables since the phase is discrete, which is hard to solve. However, since the amplitude and phase constraints are not coupled in (4), we can solve them separately and obtain closed-form solution of (4) given by the following lemma.

Lemma 1. *Define* $\alpha_m^t=\mathrm{Proj}_{\boldsymbol{\Theta}}\left(\angle v_m^t\right)$, $\alpha_m^r=\mathrm{Proj}_{\boldsymbol{\Theta}}\left(\angle v_m^r\right)$, $\beta_m^t=\left|v_m^t\right|\cos\left(\alpha_m^t-\angle v_m^t\right)$ *and* $\beta_m^r=\left|v_m^r\right|\cos\left(\alpha_m^r-\angle v_m^r\right)$. *The optimal solution of* (4) *is*

$$\begin{aligned}&\varphi_m^t=e^{j\alpha_m^t},\varphi_m^r=e^{j\alpha_m^r}, && \text{if TS}\\ &\varphi_m^t=\tfrac{1+\mathrm{sgn}\left(\beta_m^t-\beta_m^r\right)}{2}e^{j\alpha_m^t},\varphi_m^r=\tfrac{1+\mathrm{sgn}\left(\beta_m^r-\beta_m^t\right)}{2}e^{j\alpha_m^r}, && \text{if MS}\\ &\varphi_m^t=\beta_m^te^{j\alpha_m^t},\varphi_m^r=\beta_m^re^{j\alpha_m^r} && \text{if ES}\end{aligned}\tag{5}$$

where $\boldsymbol{\Theta}=\{0,2\pi/L,\cdots,2\pi\,(L-1)/L\}$, $\mathrm{Proj}_{\mathscr{A}}(b)$ *is to project* b *to the set* $\mathscr{A}$, *and* sgn *is the sign function.*

Proof. See Appendix B in supplementary material.

STAR-RISs with coupled-phase: Noting that coupled-phase only occurs in ES mode, leading to the subproblem P1 reduces to

$$\min_{\varphi_m^t,\varphi_m^r}\quad \left|v_m^t-\varphi_m^t\right|^2+\left|v_m^r-\varphi_m^r\right|^2 \tag{6a}$$

$$s.t.\quad \angle\varphi_m^t-\angle\varphi_m^r\in\{\pi/2\,(\mathrm{mod}2\pi),-\pi/2\,(\mathrm{mod}2\pi)\}, \tag{6b}$$

$$\angle\varphi_m^t,\angle\varphi_m^r\in\{0,2\pi/L,\cdots,2\pi\,(L-1)/L\}. \tag{6c}$$

As illustrated below (1), when $L>2$ and is an even number, constraint (6b) and (6c) are compatible. In this case, the global optimal solution of (6) is given in **Lemma 2**.

Lemma 2. *If* $L>2$ *and is an even number, the optimal solution of* (6) *is*

$$\begin{aligned}&\varphi_m^t=\left|v_m^t\right|\cos\left(\theta_m^t-\angle v_m^t\right)e^{j\theta_m^t},\\ &\varphi_m^r=\left|v_m^r\right|\left|\sin\left(\theta_m^t-\angle v_m^r\right)\right|e^{j\left(\theta_m^t-\frac{\pi}{2}\mathrm{sgn}\left(\sin\left(\theta_m^t-\angle v_m^r\right)\right)\right)},\end{aligned}\tag{7}$$

where $\theta_m^t=\mathrm{Proj}_{\boldsymbol{\Theta}}\left(\angle v_m^t-b_m/2+\pi/2\right)$ *and* $b_m=-j\ln\left(\frac{\left[\left|v_m^r\right|^2\cos\left(2\angle v_m^t-2\angle v_m^r\right)+\left|v_m^t\right|^2\right]+j\left|v_m^r\right|^2\sin\left(2\angle v_m^t-2\angle v_m^r\right)}{\sqrt{\left[\left|v_m^r\right|^2\cos(2\angle v_m^t-2\angle v_m^r)+\left|v_m^t\right|^2\right]^2+\left[\left|v_m^r\right|^2\sin(2\angle v_m^t-2\angle v_m^r)\right]^2}}\right)$.

Proof. See Appendix C in supplementary material.

Lemmas 1 and **2** cover both discrete and continuous phase STAR-RIS. For the latter case, it is equivalent to taking $L \to \infty$, and the projection functions in **Lemmas 1** and **2** can be skipped. It is worth noting that for the simplest ES STAR-RIS (the fifth case in Table I), since there is no phase and amplitude constraint involved in (3), the optimal solution of the auxiliary variables $\boldsymbol{\varphi}^t$ and $\boldsymbol{\varphi}^r$ will always be equal to $\boldsymbol{v}^t$ and $\boldsymbol{v}^r$, respectively. Hence, the penalty loop would only be executed once, which reduces to the conventional non-penalty design in many existing works [28], [44].

Since the solutions of P1 cover all existing type of STAR-RISs, once P2 is solved, one can easily compare the system performance of all STAR-RIS types. This is the first time such comparison is enabled in a unified way. This is important since if we optimize the resource allocation problems (1) under different STAR-RIS modes independently, the difference in performance may not only come from the difference in operating modes, but also the difference in algorithms. Hence, the proposed framework provides a unified way for communication system researchers to decide which types of STAR-RIS is the best for a particular scenario or application. An example of such comparison will be provided in Section V.

IV. A CASE STUDY OF P2 ON DOWNLINK STAR-RIS ASSISTED TRANSMISSION SYSTEM

The proposed framework in Section II is based on judiciously decomposing the original problem into two subproblems: P1 for the discrete constraints arising from STAR-RIS; and P2 for the optimization of other wireless resources. With the closed-form solution of P1 derived for various types of STAR-RIS in Section III, researchers only need to focus on solving P2, which is typical in wireless communication resource allocation. This section provides an illustrative example on downlink STAR-RIS assisted communication system.

Assume that the BS has N antennas and the STAR-RIS comprises M elements. Furthermore, there are K^r users and K^t users with single antenna in the reflection region and the transmission region, respectively. In this paper, we assume that channel state information (CSI) is available at the BS. Let $s_i \in \mathbb{C}$ with normalized power represents the information symbol targeted to the i^{th} user ($i = 1, ..., K^r + K^t$). With the individual beamforming vectors $\boldsymbol{w}_i \in \mathbb{C}^{N\times 1}$ applied at the BS, the signal transmitted from the BS is $\Sigma_{i=1}^{K^r+K^t} \boldsymbol{w}_i s_i$. Define the reflection user set and transmission user set as $\mathscr{K}^r = \{1, \cdots, K^r\}$ and $\mathscr{K}^t = \left\{K^r+1, \cdots, K^r+K^t\right\}$ respectively, the received signal of the l^{th} user is

$$y_l = \begin{cases} \sum\limits_{i=1}^{K^r+K^t} \left(\boldsymbol{h}_l^T \mathrm{diag}\left(\boldsymbol{v}^r\right)\boldsymbol{G} + \boldsymbol{d}_l^T\right)\boldsymbol{w}_i s_i + n_l, & l \in \mathscr{K}^r \\ \sum\limits_{i=1}^{K^r+K^t} \left(\boldsymbol{h}_l^T \mathrm{diag}\left(\boldsymbol{v}^t\right)\boldsymbol{G} + \boldsymbol{d}_l^T\right)\boldsymbol{w}_i s_i + n_l, & l \in \mathscr{K}^t \end{cases}, \tag{8}$$

where $\boldsymbol{G} \in \mathbb{C}^{M\times N}$ and $\boldsymbol{h}_l \in \mathbb{C}^{M\times 1}$ are the channels from the BS to the STAR-RIS and the STAR-RIS to the l^{th} user, respectively. $\boldsymbol{d}_l \in \mathbb{C}^{N\times 1}$ is the direct link channel from the BS to the l^{th} user and $n_l \in \mathbb{C}$ is Gaussian noise with zero mean and variance σ_l^2.

In this section, we consider the downlink sum-rate maximization problem. Since the information symbols of different users and noise are uncorrelated, the sum-rate of the whole system is [23]

$$\begin{aligned} &\mathscr{R}\left(\boldsymbol{w}, \boldsymbol{v}^t, \boldsymbol{v}^r, \lambda^t, \lambda^r\right) = \\ &\lambda^r \sum_{l=1}^{K^r} \log\left(1 + \frac{\left|\boldsymbol{a}_l^T \boldsymbol{w}_l\right|^2 \Big/ \lambda^r}{\sum_{i=1,\neq l}^{K^r+K^t} \left|\boldsymbol{a}_l^T \boldsymbol{w}_i\right|^2 \Big/ \lambda^r + \sigma_l^2}\right) \\ &+ \lambda^t \sum_{l=K^r+1}^{K^r+K^t} \log\left(1 + \frac{\left|\boldsymbol{a}_l^T \boldsymbol{w}_l\right|^2 \Big/ \lambda^t}{\sum_{i=1,\neq l}^{K^r+K^t} \left|\boldsymbol{a}_l^T \boldsymbol{w}_i\right|^2 \Big/ \lambda^t + \sigma_l^2}\right), \end{aligned} \tag{9}$$

where $\boldsymbol{a}_l = \begin{cases} \boldsymbol{G}^T \mathrm{diag}\left(\boldsymbol{v}^r\right)\boldsymbol{h}_l + \boldsymbol{d}_l, & l \in \mathscr{K}^r \\ \boldsymbol{G}^T \mathrm{diag}\left(\boldsymbol{v}^t\right)\boldsymbol{h}_l + \boldsymbol{d}_l, & l \in \mathscr{K}^t \end{cases}$ and $\boldsymbol{w} = \left\{\boldsymbol{w}_l\right\}_{l=1}^{K^r+K^t}$. In (9), the term $\left|\boldsymbol{a}_l^T \boldsymbol{w}_i\right|^2$ is divided by λ^r or λ^t because the signal is recieved only in a fraction (λ^r or λ^t) of the total communication time [15]. This often occurs in system model involving time allocation [45], [46]. Furthermore, for MS and ES modes, λ^t and λ^r are set to 1, and therefore (9) reduces to traditional sum-rate expression.

In (9), the set of beamforming vectors $\boldsymbol{w}$ and the RIS coefficients $\boldsymbol{v}^t$, $\boldsymbol{v}^r$ are nonlinearly coupled in both numerator and denominator. Besides, there is a summation of various data rates, which makes this objective function hard to tackle. If there is only one user, optimizing the data rate is equivalent to maximizing the signal-to-interference plus noise ratio (SINR) and hence quadratic transform [47] can be adopted to handle this single fraction. On the other hand, due to the presence of multiple users, we need the following equivalent sum-rate function for (9).

Lemma 3. *The sum-rate function* $\mathscr{R}\left(\boldsymbol{w}, \boldsymbol{v}^t, \boldsymbol{v}^r, \lambda^t, \lambda^r\right) = \max\limits_{\boldsymbol{\rho},\boldsymbol{x}} \mathscr{F}_1\left(\boldsymbol{w}, \boldsymbol{\rho}, \boldsymbol{x}, \boldsymbol{v}^t, \boldsymbol{v}^r, \lambda^t, \lambda^r\right)$*, where* $\mathscr{F}_1$ *is shown in* (10) *on the top of next page,* $\boldsymbol{\rho} = \{\rho_1, \cdots, \rho_{K^r+K^t}\}$ *and* $\boldsymbol{x} = \{x_1, \cdots, x_{K^r+K^t}\}$ *are sets of auxiliary variables.*

Proof. See Appendix D in supplementary material.

According to the proposed framework, since the objective function is only involved in P2, we set the sum-rate function in **Lemma 3** as the $\mathscr{F}$ in P2. Further recognizing that $\boldsymbol{z}$ in P2 corresponds to $\{\boldsymbol{w}, \boldsymbol{\rho}, \boldsymbol{x}\}$ in this particular application, P2 can be written as

$$\begin{aligned} \min_{\left\{\substack{\boldsymbol{w},\boldsymbol{\rho},\boldsymbol{x},\\ \boldsymbol{v}^t,\boldsymbol{v}^r,\\ \lambda^t,\lambda^r}\right\}} \; & -\mathscr{F}_1\left(\boldsymbol{w}, \boldsymbol{\rho}, \boldsymbol{x}, \boldsymbol{v}^t, \boldsymbol{v}^r, \lambda^t, \lambda^r\right) + \frac{\gamma}{2}\sum_{p=t,r} |\boldsymbol{v}^p - \boldsymbol{\varphi}^p|_2^2 \\ s.t. \; & \begin{cases} \lambda^t + \lambda^r = 1, \left\{\lambda^t, \lambda^r\right\} \geq 0, & \text{if TS} \\ \left|v_m^t\right|^2 + \left|v_m^r\right|^2 = 1, & \text{if MS/ES} \end{cases} \\ & \sum_{l=1}^{K^r+K^t} |\boldsymbol{w}_l|_2^2 \leq P_{BS}. \end{aligned} \tag{11}$$

To solve (11), recognizing that the constraints for $\boldsymbol{w}$, $\boldsymbol{\rho}$, $\boldsymbol{x}$, $\left\{\boldsymbol{v}^t, \boldsymbol{v}^r\right\}$ and $\left\{\lambda^t, \lambda^r\right\}$ are separable, BCD algorithm can be adopted, and the details of solving different subproblems are given below.

$$\begin{aligned}\mathscr{F}_1\left(\boldsymbol{w},\boldsymbol{\rho},\boldsymbol{x},\boldsymbol{v}^t,\boldsymbol{v}^r,\lambda^t,\lambda^r\right)=&\lambda^r\sum_{l=1}^{K^r}\left[\log\left(1+\rho_l\right)-\rho_l\right]+\lambda^t\sum_{l=K^r+1}^{K^r+K^t}\left[\log\left(1+\rho_l\right)-\rho_l\right]\\&+\lambda^r\sum_{l=1}^{K^r}\left\{2\left(1+\rho_l\right)\mathrm{Re}\left[\overline{x_l}\boldsymbol{a}_l^T\boldsymbol{w}_l\right]-\left(1+\rho_l\right)\left|x_l\right|^2\left(\sum_{i=1}^{K^r+K^t}\left|\boldsymbol{a}_l^T\boldsymbol{w}_i\right|^2+\lambda^r\sigma_l^2\right)\right\}\\&+\lambda^t\sum_{l=K^r+1}^{K^r+K^t}\left\{2\left(1+\rho_l\right)\mathrm{Re}\left[\overline{x_l}\boldsymbol{a}_l^T\boldsymbol{w}_l\right]-\left(1+\rho_l\right)\left|x_l\right|^2\left(\sum_{i=1}^{K^r+K^t}\left|\boldsymbol{a}_l^T\boldsymbol{w}_i\right|^2+\lambda^t\sigma_l^2\right)\right\},\end{aligned}\tag{10}$$

A. Optimizing $\boldsymbol{x}$, $\boldsymbol{\rho}$, $\boldsymbol{w}$, λ^t and λ^r

Under the BCD formulation, the subproblems with respect to the auxiliary variables $\boldsymbol{x}$, $\boldsymbol{\rho}$, the beamforming vector $\boldsymbol{w}$ and time allocation variables $\left\{\lambda^t,\lambda^r\right\}$ are convex problems, and they are relatively easy to handle.

Optimizing auxiliary variable $\boldsymbol{x}$: Noting that $\mathscr{F}_1$ is a convex function of $\boldsymbol{x}$ with each element $\{x_l\}_{l=1}^{K^r+K^t}$ being separable, we can take derivative of $\mathscr{F}_1$ with respect to each x_l and set them to zero, yielding the following the optimal solution

$$x_l=\begin{cases}\frac{\boldsymbol{a}_l^T\boldsymbol{w}_l}{\Sigma_{i=1}^{K^r+K^t}\left|\boldsymbol{a}_l^T\boldsymbol{w}_i\right|^2+\lambda^r\sigma_l^2}, & l\in\mathscr{K}^r,\\ \frac{\boldsymbol{a}_l^T\boldsymbol{w}_l}{\Sigma_{i=1}^{K^r+K^t}\left|\boldsymbol{a}_l^T\boldsymbol{w}_i\right|^2+\lambda^t\sigma_l^2}, & l\in\mathscr{K}^t.\end{cases}\tag{12}$$

Optimizing auxiliary variable $\boldsymbol{\rho}$: Since each element in $\boldsymbol{\rho}$ is also separable, by taking the derivative of $\mathscr{F}_1$ with respect to each ρ_l ($l=1,2,...,K^r+K^t$) and set them to zero, we obtain

$$\rho_l=\begin{cases}\frac{\left|\boldsymbol{a}_l^T\boldsymbol{w}_l\right|^2}{\Sigma_{i=1,\neq l}^{K^r+K^t}\left|\boldsymbol{a}_l^T\boldsymbol{w}_i\right|^2+\lambda^r\sigma_l^2}, & l\in\mathscr{K}^r,\\ \frac{\left|\boldsymbol{a}_l^T\boldsymbol{w}_l\right|^2}{\Sigma_{i=1,\neq l}^{K^r+K^t}\left|\boldsymbol{a}_l^T\boldsymbol{w}_i\right|^2+\lambda^t\sigma_l^2}, & l\in\mathscr{K}^t.\end{cases}\tag{13}$$

Optimizing beamforming vector $\boldsymbol{w}$: Focusing on the components related to $\boldsymbol{w}$ in (11), the following subproblem of $\boldsymbol{w}$ is obtained and it is a convex quadratic problem:

$$\min_{\boldsymbol{w}}\quad\sum_{l=1}^{K^r+K^t}\left\{\boldsymbol{w}_l^H\boldsymbol{\Xi}\boldsymbol{w}_l-2\mathrm{Re}\left[\boldsymbol{q}_l^H\boldsymbol{w}_l\right]\right\}\tag{14a}$$

$$s.t.\quad\sum_{l=1}^{K^r+K^t}\boldsymbol{w}_l^H\boldsymbol{w}_l\le P_{BS},\tag{14b}$$

where $\boldsymbol{q}_l=\begin{cases}\lambda^r\left(1+\rho_l\right)x_l\overline{\boldsymbol{a}_l}, & l\in K^r\\ \lambda^t\left(1+\rho_l\right)x_l\overline{\boldsymbol{a}_l}, & l\in K^t\end{cases}$ and $\boldsymbol{\Xi}=\lambda^r\Sigma_{i=1}^{K^r}(1+\rho_i)\left|x_i\right|^2\overline{\boldsymbol{a}_i}\boldsymbol{a}_i^T+\lambda^t\Sigma_{i=K^r+1}^{K^r+K^t}(1+\rho_i)\left|x_i\right|^2\overline{\boldsymbol{a}_i}\boldsymbol{a}_i^T$ with $\overline{\boldsymbol{a}_i}$ denoting the conjugate of $\boldsymbol{a}_i$. Notice that problem (14) is a convex quadratically constrained quadratic program (QCQP), one popular option is to employ CVX to tackle this problem.

However, CVX tool does not take the advantage of the structure of (14). Since this problem has only one constraint, the following lemma with bisection method can be adopted to solve this problem optimally.

Lemma 4. *Denote the eigenvalue decomposition of $\boldsymbol{\Xi}$ as $\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H$, and $\boldsymbol{B}=\Sigma_{l=1}^{K^r+K^t}\boldsymbol{U}^H\boldsymbol{q}_l\boldsymbol{q}_l^H\boldsymbol{U}$, the optimal beamforming vector of* (14) *is $\boldsymbol{w}_l=\boldsymbol{U}(\boldsymbol{\Lambda}+\mu\boldsymbol{I})^{-1}\boldsymbol{U}^H\boldsymbol{q}_l$, where*

$$\mu=\begin{cases}0, & \text{if } \mathrm{Tr}\left(\boldsymbol{\Lambda}^{-2}\boldsymbol{B}\right)\le P_{BS}\\ \text{solution of } \mathrm{Tr}\left(\left(\boldsymbol{\Lambda}+\mu\boldsymbol{I}\right)^{-2}\boldsymbol{B}\right)=P_{BS}, & \text{otherwise}\end{cases}.\tag{15}$$

Since $\mathrm{Tr}\left(\left(\boldsymbol{\Lambda}+\mu\boldsymbol{I}\right)^{-2}\boldsymbol{B}\right)$ is a strictly decreasing function of μ, the solution of the second case in (15) *can be found by bisection method from interval $\left[0,\sqrt{\mathrm{Tr}\left(\boldsymbol{B}\right)/P_{BS}}\right]$.*

Proof. See Appendix E in supplementary material.

Optimizing time allocation variables $\left\{\lambda^t,\lambda^r\right\}$: For ES and MS mode, we have $\lambda^t=\lambda^r=1$. Hence, only λ^t and λ^r in the TS mode need to be optimized. Focusing on the components related to λ^t and λ^r in (11), the following subproblem is obtained,

$$\min_{\lambda^t,\lambda^r}\quad\left(\lambda^t\right)^2\psi^t+\left(\lambda^r\right)^2\psi^r+\lambda^t\sum_{l=K^r+1}^{K^r+K^t}\eta_l^t+\lambda^r\sum_{l=1}^{K^r}\eta_l^r\tag{16a}$$

$$s.t.\ \ \lambda^t+\lambda^r=1,\left\{\lambda^t,\lambda^r\right\}\ge 0,\tag{16b}$$

where $\eta_l^r=\left(1+\rho_l\right)\left[\left|x_l\right|^2\sum_{i=1}^{K^r+K^t}\left|\boldsymbol{a}_l^T\boldsymbol{w}_i\right|^2-2\mathrm{Re}\left(\overline{x_l}\boldsymbol{a}_l^T\boldsymbol{w}_l\right)\right]-\log\left(1+\rho_l\right)+\rho_l$, $\eta_l^t=\left(1+\rho_l\right)\left[\left|x_l\right|^2\sum_{i=1}^{K^r+K^t}\left|\boldsymbol{a}_l^T\boldsymbol{w}_i\right|^2-2\mathrm{Re}\left(\overline{x_l}\boldsymbol{a}_l^T\boldsymbol{w}_l\right)\right]-\log\left(1+\rho_l\right)+\rho_l$, $\psi^r=\Sigma_{l=1}^{K^r}\sigma_l^2$ and $\psi^t=\Sigma_{l=K^r+1}^{K^r+K^t}\sigma_l^2$. Substituting $\lambda^r=1-\lambda^t$ in (16a), (16) is equivalent to

$$\begin{aligned}\min_{\lambda^t}\quad&\left(\psi^t+\psi^r\right)\left(\lambda^t-\frac{2\psi^r-\eta^t+\eta^r}{2\psi^t+2\psi^r}\right)^2\\ s.t.\ \ &\lambda^t\in[0,1],\end{aligned}\tag{17}$$

where $\eta^t=\sum_{l=K^r+1}^{K^r+K^t}\eta_l^t$ and $\eta^r=\sum_{l=1}^{K^r}\eta_l^r$. Therefore, the optimal solution is $\lambda^t=\mathrm{Proj}_{[0,1]}\left(\frac{2\psi^r-\eta^t+\eta^r}{2\psi^t+2\psi^r}\right)$ and $\lambda^r=1-\mathrm{Proj}_{[0,1]}\left(\frac{2\psi^r-\eta^t+\eta^r}{2\psi^t+2\psi^r}\right)$.

B. Optimizing $\boldsymbol{v}^t$ and $\boldsymbol{v}^r$

To better illustrate the optimization subproblem with respect to $\boldsymbol{v}^t$ and $\boldsymbol{v}^r$, we define two shorthand notations:

$$\boldsymbol{A}_l = (1+\rho_l)\left|x_l\right|^2 \mathrm{diag}\left(\boldsymbol{h}_l^H\right)\overline{\boldsymbol{G}}\left(\sum_{i=1}^{K^r+K^t}\overline{\boldsymbol{w}_i}\boldsymbol{w}_i^T\right)\boldsymbol{G}^T\mathrm{diag}\left(\boldsymbol{h}_l\right), \tag{18}$$

$$\boldsymbol{b}_l = 2(1+\rho_l)\,\mathrm{diag}\left(\boldsymbol{h}_l^H\right)\overline{\boldsymbol{G}}\left[\left|x_l\right|^2\left(\sum_{i=1}^{K^r+K^t}\overline{\boldsymbol{w}_i}\boldsymbol{w}_i^T\right)\boldsymbol{d}_l - x_l\overline{\boldsymbol{w}_l}\right]. \tag{19}$$

Focusing on the terms related to $\boldsymbol{v}^t$ and $\boldsymbol{v}^r$ in (11), the subproblem becomes

$$\min_{\boldsymbol{v}^t,\boldsymbol{v}^r}\ \sum_{p=t,r}(\boldsymbol{v}^p)^H\boldsymbol{A}^p\boldsymbol{v}^p + \mathrm{Re}\left[(\boldsymbol{b}^p)^H\boldsymbol{v}^p\right] \tag{20a}$$

$$s.t.\ \left|v_m^t\right|^2 + \left|v_m^r\right|^2 = 1,\ \ m\in\{1,2,...,M\},\quad \text{if ES/MS} \tag{20b}$$

where $\boldsymbol{A}^r = \gamma\boldsymbol{I}_M/2 + \lambda^r\Sigma_{l=1}^{K^r}\boldsymbol{A}_l$, $\boldsymbol{A}^t = \gamma\boldsymbol{I}_M/2 + \lambda^t\Sigma_{l=K^r+1}^{K^r+K^t}\boldsymbol{A}_l$, $\boldsymbol{b}^r = -\gamma\boldsymbol{\varphi}^r + \lambda^r\Sigma_{l=1}^{K^r}\boldsymbol{b}_l$ and $\boldsymbol{b}^t = -\gamma\boldsymbol{\varphi}^t + \lambda^t\Sigma_{l=K^r+1}^{K^r+K^t}\boldsymbol{b}_l$.

For the TS mode, we do not have the constraint (20b). This is an unconstrained quadratic optimization and the closed-form solution is obtained by

$$\boldsymbol{v}^t = -\frac{1}{2}\left(\boldsymbol{A}^t\right)^{-1}\boldsymbol{b}^t\ ;\ \boldsymbol{v}^r = -\frac{1}{2}(\boldsymbol{A}^r)^{-1}\boldsymbol{b}^r. \tag{21}$$

For the ES and MS modes, recognizing that (20) has M non-convex and quadratic equality constraints, SCA is not suitable since the equality constraints cannot be approximated. On the other hand, noticing that the constraints (20b) is separable with different index m, the coefficient pair $\left\{v_m^t, v_m^r\right\}$ can be sequentially updated under the BCD framework from $m=1$ to $m=M$ with the rest coefficient pairs fixed. Hence, the subproblem with respect to $\left\{v_m^t, v_m^r\right\}$ is

$$\begin{aligned}\min_{v_m^t, v_m^r}\ & \sum_{p=t,r}\boldsymbol{A}_{m,m}^p|v_m^p|^2 + \mathrm{Re}\left\{\left(\overline{c_m^p}\right)v_m^p\right\}\\ s.t.\ & \left|v_m^t\right|^2 + \left|v_m^r\right|^2 = 1,\end{aligned} \tag{22}$$

where $c_m^t = b_m^t + 2\Sigma_{j\neq m}\boldsymbol{A}_{m,j}^t v_j^t$ and $c_m^r = b_m^r + 2\Sigma_{j\neq m}\boldsymbol{A}_{m,j}^r v_j^r$, with $\boldsymbol{A}_{i,j}^t$ and $\boldsymbol{A}_{i,j}^r$ denote the $(i,j)^{th}$ element of $\boldsymbol{A}^t$ and $\boldsymbol{A}^r$, respectively.

Expressing $v_m^t = \left|v_m^t\right|e^{j\angle v_m^t}$ and $v_m^r = \left|v_m^r\right|e^{j\angle v_m^r}$, and noticing that the phases $\angle v_m^t$ and $\angle v_m^r$ only affect the value of $\mathrm{Re}\left\{\left(\overline{c_m^t}\right)v_m^t + \left(\overline{c_m^r}\right)v_m^r\right\}$, the phases $\angle v_m^t$ and $\angle v_m^r$ that minimize (22) should be chosen as

$$\angle v_m^t = \pi + \angle c_m^t\ \ ;\ \ \angle v_m^r = \pi + \angle c_m^r. \tag{23}$$

Putting this result to (22) and it reduces to

$$\begin{aligned}\min_{|v_m^t|,|v_m^r|}\ & \sum_{p=t,r}\boldsymbol{A}_{m,m}^p|v_m^p|^2 - |c_m^p|\,|v_m^p|\\ s.t.\ & \left|v_m^t\right|^2 + \left|v_m^r\right|^2 = 1.\end{aligned} \tag{24}$$

Problem (24) is a real-valued optimization problem with a circular constraint. This problem would be easier to solve if we define $|v_m^r| = \cos(\phi_m)$ and $\left|v_m^t\right| = \sin(\phi_m)$ and transfer the unknown to $\phi_m\in[0,\pi/2]$ with the constraint of (24) guaranteed to satisfy. Then, the resulting one dimension unconstrained problem becomes

$$\begin{aligned}\min_{\phi_m\in\left[0,\frac{\pi}{2}\right]} f(\phi_m) = & \left(\boldsymbol{A}_{m,m}^r - \boldsymbol{A}_{m,m}^t\right)\cos^2(\phi_m)\\ & - |c_m^r|\cos(\phi_m) - \left|c_m^t\right|\sin(\phi_m).\end{aligned} \tag{25}$$

If we compute the gradient $\nabla f(\phi_m)$, it is found that

$$\begin{aligned}\nabla f(\phi_m) = & \sin(\phi_m)\cos(\phi_m)\\ & \times\left[\frac{|c_m^r|}{\cos(\phi_m)} - \frac{\left|c_m^t\right|}{\sin(\phi_m)} - 2\left(\boldsymbol{A}_{m,m}^r - \boldsymbol{A}_{m,m}^t\right)\right].\end{aligned} \tag{26}$$

Since the function in the square bracket is a monotonic increasing function taking values from $(-\infty,\infty)$, the zero point of the gradient can be obtained by the bisection method from the interval $[0,\pi/2]$. Denoting the estimated ϕ_m of (25) by bisection method as $\hat{\phi}_m$, the solution of (22) is then given by

$$v_m^t = \sin\left(\hat{\phi}_m\right)e^{j\left(\pi+\angle c_m^t\right)}\ \ ;\ \ v_m^r = \cos\left(\hat{\phi}_m\right)e^{j(\pi+\angle c_m^r)}. \tag{27}$$

C. Summary and Time Complexity of the Proposed Algorithm

The proposed algorithms for sum-rate maximization under different STAR-RISs are summarized in Table II. From Table II, we can see the operating mode constraint and discrete phase constraint are handled using equations (5) and (7). Hence, algorithm designers only need to focus on the subproblem P2. Since P2 under ES mode and MS mode are the same, there are only two variations (ES/MS and TS) for three kinds of STAR-RIS modes. Furthermore, from Table II, we can see that there are many common equations in algorithms for different STAR-RISs. This is how the proposed framework facilitates the design of algorithms under different types and operating modes of STAR-RIS.

Since individual variable in (11) is updated with the optimal closed-form solution under the BCD framework, the solution of (11) is a stationary point [48], [49]. Together with the global optimal solution of P1 and **Proposition 1**, the solution generated by the **Algorithm 1** is at least a stationary point of the original STAR-RIS assisted downlink sum-rate maximization problem.

For the complexity of **Algorithm 1** in the context of STAR-RIS assisted downlink transmission, updating $\boldsymbol{x}$ and $\boldsymbol{\rho}$ with closed-form expressions (12), (13) takes $\mathcal{O}\left(\left(K^r+K^t\right)N/P\right)$, where P is the parallel processing factor since all these variables can be updated in parallel. For updating $\boldsymbol{w}$, according to **Lemma 4**, its complexity mainly come from the eigenvalue decomposition of $\boldsymbol{\Xi}$ and the bisection method of finding μ, Therefore, its complexity is $\mathcal{O}\left(N^3\right) + \mathcal{O}\left(\log\left(\sqrt{\mathrm{Tr}\left(\boldsymbol{B}\right)/\varepsilon^2 P_{BS}}\right)\right)$, where ε is the accuracy of bisection search. It is shown in Appendix F of

TABLE II: Equations related to different types of STAR-RIS

Operating mode	Coupled-phase constraint	Discrete phase constraint	Equations related to subproblem P1	Time allocation constraint	Equations related to subproblem P2
TS	✘	✘	First line of (5)	✔	(12),(13),(15),(21)
	✘	✔		✔	
MS	✘	✘	Second line of (5)	✘	(12),(13),(15),(27)
	✘	✔		✘	
ES	✘	✘	Third line of (5)	✘	
	✘	✔		✘	
	✔	✘	(7)	✘	
	✔	✔		✘	

the supplementary material that $\mathrm{Tr}(\boldsymbol{B})$ scales linearly with the number of users in the system.

For the update of RIS coefficient $\boldsymbol{v}^t$ and $\boldsymbol{v}^r$, since we sequentially update M pairs of $\left\{v_m^t, v_m^r\right\}_{m=1}^M$, and each pair involves a bisection search with range from 0 to $\pi/2$, the corresponding complexity is $\mathcal{O}\left(M\log\left(\pi/2\varsigma\right)\right)$, where ς is the accuracy of bisection search. For the auxiliary variables $\boldsymbol{\varphi}^t$ and $\boldsymbol{\varphi}^r$, it can be parallelized and updated with closed-form according to **Lemmas 1** and **2**, with the complexity of this update being $\mathcal{O}\left(M/P\right)$.

In total, the complexity of the penalty based BCD algorithm for this sum-rate maximization is $I_{pen}I_{BCD}\big(\mathcal{O}\left(\left(K^r+K^t\right)N/P+N^3\right)+\mathcal{O}\left(M/P\right)+\mathcal{O}\left(M\log\left(\pi/2\varsigma\right)\right)+\mathcal{O}\left(\log\left(\sqrt{\mathrm{Tr}\left(\boldsymbol{B}\right)/\varepsilon^2P_{BS}}\right)\right)\big)$, where I_{Pen} and I_{BCD} are the number of penalty iterations and BCD iterations, respectively.

V. Simulation Results and Discussions

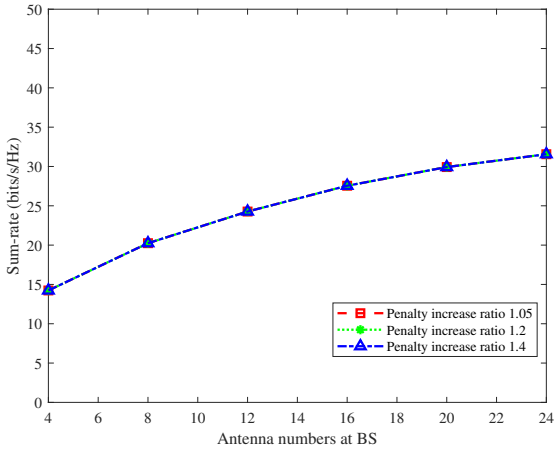
In this section, we evaluate the downlink sum-rate transmission performance through simulation. All problem instances are simulated using Matlab-R2023a on a Windows x64 desktop with 2.8 GHz CPU and 16 GB RAM, and the simulation results are obtained via averaging over 100 simulation trials. The parallel processing index $P=4$ is used in the simulation. In the simulations, the BS and STAR-RIS are located at the coordinate (0,20m) and (40m,0), respectively. The STAR-RIS is placed along the y-axis and perpendicular to the ground. The reflection and transmission users are uniformly located within 8m of two sides of the STAR-RIS. The parameters $\{\boldsymbol{G},\boldsymbol{h}_l,\boldsymbol{d}_l\}$ are modeled as the Rician fading channels, which contain both the line-of-sight (LoS) and non-LoS (NLoS) components [50]. Take $\boldsymbol{h}_l$ as an example, the Rician fading channel model is $\boldsymbol{h}_l=\sqrt{\nu_{\boldsymbol{h}_l}/(\kappa_{\boldsymbol{h}}+1)}\left(\sqrt{\kappa_{\boldsymbol{h}}}\boldsymbol{h}_l^{LoS}+\boldsymbol{h}_l^{NLoS}\right)$ and each parameter is specified below.

1) $\nu_{\boldsymbol{h}_l}=L_0(\mathscr{d}_{\boldsymbol{h}_l}/\mathscr{d}_0)^{-\alpha_{\boldsymbol{h}}}$ is the distance dependent path-loss from RIS to the l^{th} user, where $L_0=-30$dB denotes the path loss at the reference distance $\mathscr{d}_0=1$m. $\mathscr{d}_{\boldsymbol{h}_l}$ is the distance between STAR-RIS and the l^{th} user. $\alpha_{\boldsymbol{h}}=2.2$ denotes the path-loss exponent for the RIS-user link. Correspondingly, the path-loss exponents for the BS-RIS link and the BS-user link are 2.2 and 3.6, respectively.
2) $\kappa_{\boldsymbol{h}}$ is the Rician factor for the RIS-user link. A higher value of Rician factor means stronger LoS component. When the Rician factor is 0, it means there is no LoS signal and the channel reduces to the Rayleigh fading. In the simulations, the Rician factors for the RIS-user link $\kappa_{\boldsymbol{h}}$ is set to 5. Correspondingly, the Rician factor of the BS-RIS link and the BS-user link are 5 and 0, respectively.
3) The LoS component $\boldsymbol{h}_l^{LoS}$ is modeled as the steering vector of the array responses. Hence the m^{th} element $\left[\boldsymbol{h}_l^{LoS}\right]_m=e^{j2\pi(m-1)d_A\sin(\omega)/\lambda}$, where ω denotes the angle-of-arrival (AoA) or angle-of-departure (AoD) of the array. In the simulation, $d_A/\lambda=1/2$ and the ω is modeled as uniform distributed in $[0,2\pi)$. On the other hand, $\boldsymbol{h}_l^{NLoS}$ denotes the NLoS component signal with each element obeying the normalized complex Gaussian distribution.

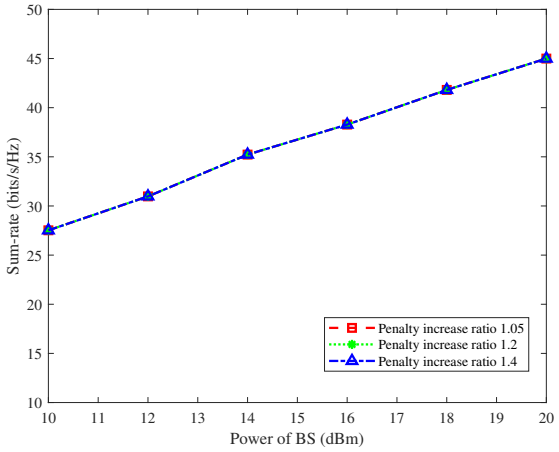
To avoid repeating figure descriptions, the settings for (M, N, K^r, K^t, L, P_{BS}, σ_l^2) are provided in the caption of each figure.

For implementation, we employ exponential increase of γ, $\gamma\leftarrow\gamma\times1.2$ after each iteration, which is widely adopted in existing works involving penalty method [51]–[54]. The penalty parameter is increased until the distance between $\left\{\boldsymbol{v}^t,\boldsymbol{v}^r\right\}$ and $\left\{\boldsymbol{\varphi}^t,\boldsymbol{\varphi}^r\right\}$ of alternating optimization falls below a predefined tolerance. It is known that as long as the increase of penalty parameter is within a moderate range in each iteration, the performance would remain the same. This is shown by simulations in Figs. 4(a) and 4(b) for the ES STAR-RIS and MS STAR-RIS, respectively.

Fig. 5 compares the sum-rate performance among the eight types of STAR-RIS using the proposed general optimization framework. Firstly, we can see that additional discrete phase constraint only slightly reduce the network throughput (2.94% for two discrete phases in ES, 6.16% and 8.31% for two discrete phases in MS and TS, respectively), which is different from the finding in [55] that sparse phase (lower than 8 phases) will significantly affect the performance. The key reason is that the conclusion in [55] is based on quantization from continuous phase solution to its nearest discrete phase. Under the very few allowable discrete phases (e.g., in the simulation only 2 phases are allowed), the continuous phase and its closest discrete value are very different. Since the quantization is applied independently in each RIS element, the cumulated performance degradation will be significant compared to the continuous optimal solution. In contrast, the auxiliary variables $\boldsymbol{\varphi}^t$, $\boldsymbol{\varphi}^r$ in the proposed algorithm are only affected by the phase constraints, which can be regarded as the phase corrector



(a) ES STAR-RIS at $M = 30$, $K^r = K^t = 4$, $P_{BS} = 20$dBm and $\sigma_l^2 = -80$dBm



(b) MS STAR-RIS at $M = 20$, $N = 16$, $K^r = K^t = 5$ and $\sigma_l^2 = -80$dBm

Fig. 4: Performance comparison under different penalty increase ratio

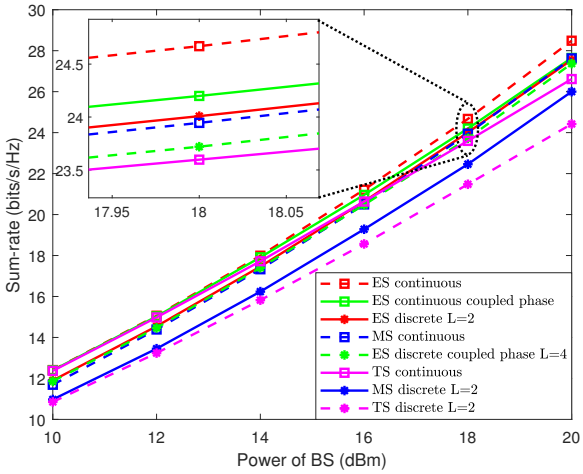
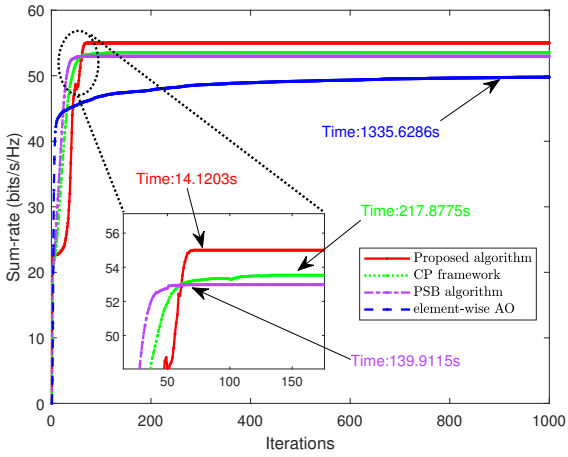


Fig. 5: Sum-rate comparison of the eight types of STAR-RIS at $M = 30$, $N = 16$, $K^r = K^t = 4$ and $\sigma_l^2 = -80$dBm

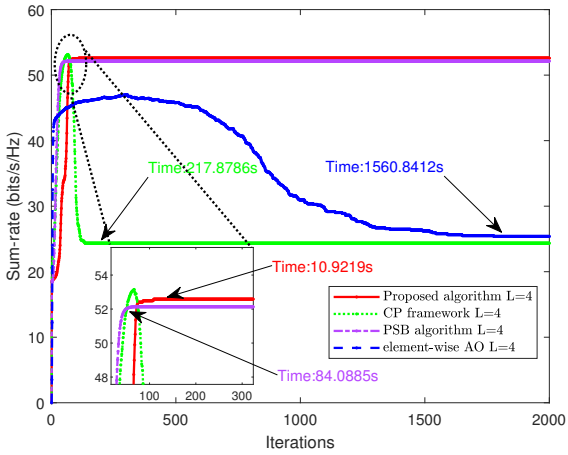
for $\boldsymbol{v}^t$, $\boldsymbol{v}^r$. Instead of directly quantizing the $\boldsymbol{v}^t$, $\boldsymbol{v}^r$ to $\boldsymbol{\varphi}^t$, $\boldsymbol{\varphi}^r$, a penalty term is introduced to enforce an indirect quantization and it allows more freedom for $\boldsymbol{v}$ and $\boldsymbol{\varphi}$ to search for better solution. This avoids the significant performance loss and shows that quantization is a more influential factor than discrete phase that leads to performance degradation. Secondly, it is notice that the discrete phase constraint in MS and TS modes lead to more performance loss than that in the ES mode. This is probably because of the additional amplitude constraint of (1c) imposed on MS and TS STAR-RIS coefficients. Without the amplitude constraint, ES STAR-RIS can optimize the amplitude to compensate the loss brought by the discrete phase constraint. Thirdly, from the magnified part of Fig. 5, we notice that the continuous coupled-phase constraint in STAR-RIS almost have no influence (only 0.89% loss) on the system throughput compared to noncoupled-phase case, which is consistent with the conclusion in [20]. In contrast, the throughput degradation introduced by the 0-1 amplitude constraint in the MS mode is more obvious. This shows that amplitude constraint affects system performance more than the phase constraint. This insight is revealed for the first time due to the ease of comparison using the proposed unified penalty framework.

Figs. 6(a) and 6(b) focus on the ES STAR-RIS with coupled-phase (i.e., cases 7 and 8 in Table I) and compare the convergence behavior of the proposed algorithm with the elementwise-AO [19], the coupled-phase STAR-RIS framework (named as CP framework) [20] and penalty-based secrecy beamforming (PSB) algorithm [34]. Although the original PSB algorithm was derived for secrecy sum-rate, we can modify it to handle sum-rate maximation. For the continuous coupled-phase case in Fig. 6(a), the proposed algorithm converges to the highest sum-rate with the shortest execution time. PSB algorithm and CP framework are slightly worse and the elementwise-AO performs the worst. As explained before **Proposition 1** (Section IIB), the penalty weight required in the CP and PSB algorithms would be larger than the proposed algorithm and slows down their convergence speeds. On the other hand, elementwise-AO needs to solve nonconvex subproblems on STAR-RIS element-wise level, which are handled by SCA and CVX, and thus incurs heavy computations.

To further illustrate the complexity of the proposed and compared algorithms, Table III summarizes the complexities of various algorithms, where I_{AO} is the number of alternative optimization (AO) iteration. Notice that ς in the proposed framework is the accuracy of bisection search. From Table III, ignoring the terms common to all methods, it can be seen that the proposed framework has a complexity order scales linearly with respect to M in each iteration, i.e., $\mathcal{O}(M)$. In contrast, the CP framework and PSB algorithm take at least $\mathcal{O}(M^2)$ in each iteration. As the number of STAR-RIS element M is usually larger than the number of BS antennas N and number of users K^r and K^t, the complexities of the CP framework and the PSB algorithms would be much larger than that of the proposed framework. Although the element-wise AO does not have outer penalty loop, the complexity order contains a term



(a) Coupled-phase STAR-RIS



(b) Discrete coupled-phase STAR-RIS

Fig. 6: Convergence behaviour with $M = 30$, $N = 16$, $K^r = K^t = 4$, $P_{BS} = 20\text{dBm}$ and $\sigma_l^2 = -100\text{dBm}$

$I_{AO}\mathcal{O}\left(M^{3.5}\right)$, which heavily slows down the whole execution time. Due to the low complexity of the proposed method in each iteration, as shown in Fig. 6(a), the proposed algorithm reduces almost a factor of 10 in computation time compared to the CP framework and the PSB algorithm, and almost a factor of 100 compared to the elementwise AO algorithm. Notice that since the CP framework, PSB algorithm and the proposed algorithm introduce penalty terms, the sum-rate may not be monotonic with respect to iterations as the monotonic property only holds for objective function including the penalty term.

For the discrete coupled-phase case in Fig. 6(b), the CP framework and elementwise-AO have noticeable performance decline when the number of iteration increases. This is due to the quantization to the continuous phase solution. This phenomenon coincides with the general statement from [55] that quantization in sparse phase (lower than 3 or 4 information bits) will strongly affect the performance. The PSB algorithm performs better than the elementwise-AO and CP framework because it employs quantization at RIS coefficient subproblem level. Compared with only one time quantization on the final continuous phase solution, repeated quantization at the level of

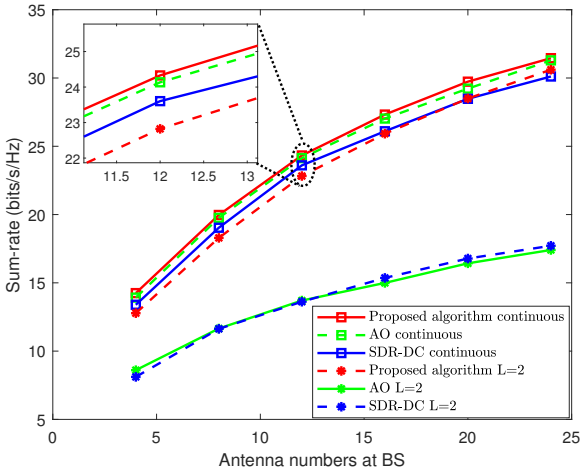


Fig. 7: Sum-rate of ES STAR-RIS with $M = 30$, $K^r = K^t = 4$, $P_{BS} = 20\text{dBm}$ and $\sigma_l^2 = -80\text{dBm}$

RIS-subproblems avoids a sharp performance loss. Different from the above three methods, the proposed algorithm finds the global optimal closed-form solution at the RIS coefficient level subproblem even under the discrete phase constraint, which leads to the best throughput performance for the discrete and coupled-phase STAR-RIS in Fig. 6(b).

Although the execution times reported in Fig. 6 are in the order of second and are much longer than the typical wireless channel coherence time, these numbers are obtained by Matlab running on a consumer grade computer. The purpose of these numbers are for comparing the speeds of different algorithms under the same hardware platform. It does not necessarily mean this is the time needed if the algorithm is implemented in a product. For example, in actual deployment, we can use C++, instead of Matlab, to shorten its execution time. Furthermore, we can use better hardware, instead of desktop computer, for implementation. Moreover, we can run the proposed framework offline to generate many training samples and use them to train a deep neural network. This would enable a super fast inference time.

Fig. 7 compares the sum-rate performance between the proposed algorithm, the AO algorithm [14] and SDR-DC method [15] for the basic ES STAR-RIS (i.e., the fifth and sixth cases in Table I). Under continuous phase shift, the proposed algorithm performs the best and SDR-DC performs the worst, but the performance gap is not significant. However, under 2 discrete phases, the proposed algorithm performs only slight worse than continuous phase case (only 3.81% degradation), and significantly outperforms the other two algorithms. This again verifies that discrete phase is not the major reason for performance degradation. It is just that the widely used quantization is not a good strategy for ES STAR-RIS under small number of phases.

Fig. 8 compares the sum-rate performance in TS STAR-RIS system (i.e., first two cases in Table I) among the proposed algorithm, the SDR algorithm [15] with Gaussian random, and gradient descent (GD) method [25], [56]. From Fig. 8, we can see that under continuous phase shift, the proposed algorithm

TABLE III: Computational complexities of various algorithms

Algorithms	Computational complexity
CP framework	$I_{pen}I_{BCD}\left(\mathcal{O}\left(\left(K^r+K^t\right)N/P+N^3\right)+\mathcal{O}\left(M^2\right)+\mathcal{O}\left(\log\left(\sqrt{\mathrm{Tr}\left(\boldsymbol{B}\right)/\varepsilon^2 P_{BS}}\right)\right)+I_{AO}\mathcal{O}\left(M/P\right)\right)$
PSB algorithm	$I_{pen}I_{BCD}\left(\mathcal{O}\left(\left(K^r+K^t\right)N/P+N^3\right)+\mathcal{O}\left(M^2\right)+\mathcal{O}\left(\log\left(\sqrt{\mathrm{Tr}\left(\boldsymbol{B}\right)/\varepsilon^2 P_{BS}}\right)\right)+\mathcal{O}\left(2M/P\right)\right)$
Elementwise-AO	$I_{BCD}\left(\mathcal{O}\left(\left(K^r+K^t\right)N/P+N^3\right)+\mathcal{O}\left(\log\left(\sqrt{\mathrm{Tr}\left(\boldsymbol{B}\right)/\varepsilon^2 P_{BS}}\right)\right)+I_{AO}\mathcal{O}\left(M^{3.5}\right)\right)$
Proposed framework	$I_{pen}I_{BCD}\left(\mathcal{O}\left(\left(K^r+K^t\right)N/P+N^3\right)+\mathcal{O}\left(M\log\left(\pi/2\varsigma\right)\right)\right)+\mathcal{O}\left(\log\left(\sqrt{\mathrm{Tr}\left(\boldsymbol{B}\right)/\varepsilon^2 P_{BS}}\right)\right)+\mathcal{O}\left(M/P\right)\right)$

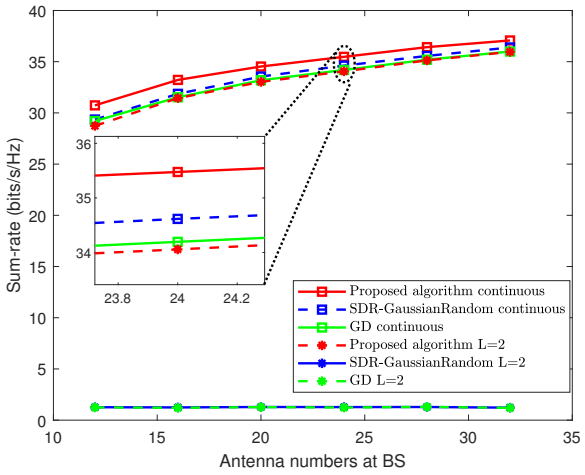


Fig. 8: Sum-rate of TS STAR-RIS with $M=40$, $K^r=K^t=5$, $P_{BS}=20\text{dBm}$ and $\sigma_l^2=-90\text{dBm}$

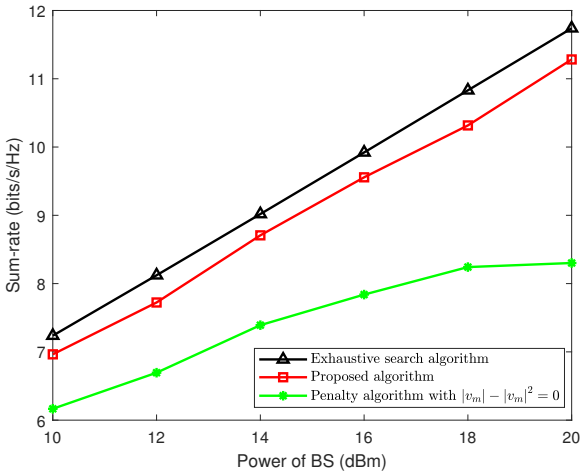


Fig. 10: Sum-rate of MS STAR-RIS with $L=2$, $M=6$, $N=4$, $K^r=K^t=1$ and $\sigma_l^2=-80\text{dBm}$

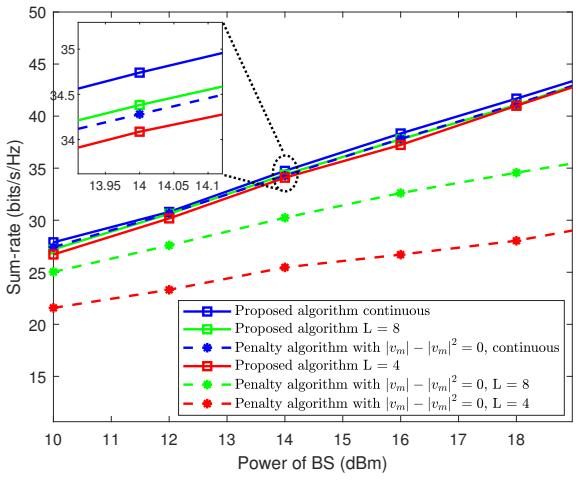


Fig. 9: Sum-rate of MS STAR-RIS with $M=20$, $K^r=K^t=5$, and $\sigma_l^2=-80\text{dBm}$

performs the best, and then followed closely by SDR algorithm and GD method. However, under 2 discrete phase case, the GD method and SDR algorithm fail completely.

Fig. 9 compares the sum-rate performance between the proposed algorithm and the direct penalty algorithm with the widely used penalty term $\left|v_m^{\hbar}\right|-\left|v_m^{\hbar}\right|^2=0, \forall \hbar\in\{t,r\}, m\in\{1,\cdots,M\}$ [15], [34] for MS STAR-RIS (i.e., the third and fourth cases in Table I). Since no existing literature studies the MS STAR-RIS under discrete phase, we add a quantization step after the direct penalty algorithm to make the discrete phase constraint satisfied. From Fig. 9, we observe that when there is no discrete phase constraint, the proposed auxiliary variable based penalty method outperforms the direct penalty method, and the performance of direct penalty method is even worse than the proposed algorithm with discrete phase $L=8$. This shows the possibility of exploring better penalty term to improve the performance of continuous phase MS STAR-RIS. In additional, the discrete phase constraint just slightly degrades the performance (2.28% for $L=4$ and 1.27% for $L=8$) if the proposed algorithm is employed, which coincides with the conclusion in ES STAR-RIS that the discrete phase is not the major reason for performance degradation. In contrast, the quantization adopted in discrete MS STAR-RIS heavily impairs the performance (28.3% degradation at $L=4$ and 14.1% degradation at $L=8$). Besides, quantization in MS STAR-RIS makes the sum-rate grows slowly with the power of BS. This all shows that quantization is not a good option for MS STAR-RIS under discrete phase.

Next, we evaluate the performance of the proposed algorithm when compared to exhaustive search solution. In particular, we consider the discrete phase MS STAR-RIS case, in which both amplitude and phase are discrete variables. For a MS STAR-RIS with M elements and L discrete phase, it will have $(2L)^M$ different combinations, which is exponential with M. Therefore, exhaustive search algorithm is only possible for small M and L. Fig. 10 shows the results of $L=2$ and

$M = 6$. It can be seen that the performance degradation of the proposed method from the optimal solution using exhaustive searching is only 2.4%∼4.3%, and their sum-rates grow at the same rate as the power of BS increases. However, for the direct penalty algorithm with $\left|v_m^{\mathscr{k}}\right| - \left|v_m^{\mathscr{k}}\right|^2 = 0, \forall \mathscr{k} \in \{t, r\}$, the degradation is more significant (14.7%∼29.2%) and shows a slower rate of sum-rate increase with power of BS.

Finally, to show the applicability of the proposed framework to other resource allocation problems, we consider the energy efficiency maximization problem, which has the same constraints as the sum-rate maximization problem (11) but the objective function $\mathscr{F}_1$ is divided by

$$\frac{1}{\eta} \sum_{l=1}^{K^r+K^t} |\boldsymbol{w}_l|^2 + P_a + \left(K^r + K^t\right) P_c + MP_s, \qquad (28)$$

where η is the power amplifier efficiency, P_a, P_c and P_s are the hardware-dissipated power at the BS, the circuit power at each user, and the hardware-dissipated power at each STAR-RIS element, respectively [8], [57]. In this simulation, these parameters are set as $(\eta, P_a, P_c, P_s) = (0.311, 39\text{dBm}, 20\text{dBm}, 10\text{dBm})$ [8], [58]. For the optimization algorithm, we need Dinkelbatch algorithm as an extra outer layer to handle the fractional form of the objective function compared to the sum-rate optimization algorithm. It can be seen from Fig. 11 that the ES STAR-RIS with coupled phase performs the best. Then, it is followed by MS STAR-RIS and ES STAR-RIS, while TS STAR-RIS performs much worse than other modes. Compared to Fig. 5, it is observed that the relative performance order of different operating modes in energy efficiency maximization is different from that of the sum-rate maximization. Besides, the influence of the discrete phase constraint is even smaller in energy efficiency problem than in sum-rate maximization. Fig. 11 not only shows the versatile applicability of the proposed framework to different objective functions, but also demonstrates the importance of having a convenient way of comparing the performance of different operating modes of STAR-RIS, as their relative performance would highly depend on the objective of the optimization problem.

Notice that although TS STAR-RIS performs the worst among different STAR-RIS operating modes, the proposed algorithm still outperforms an existing approach: complex circle manifold (CCM) approach [59], [60], which projects the STAR-RIS coefficient onto the circle manifold constraints $\left\{\boldsymbol{v}^t : \left|v_m^t\right| = 1, m = 1, 2, \cdots, M\right\}$ and $\{\boldsymbol{v}^r : |v_m^r| = 1, m = 1, 2, \cdots, M\}$. From Fig. 12, it can be seen that the proposed framework performs better than CCM algorithm for both continuous phase and discrete phase STAR-RIS, but the advantage of the proposed framework is more pronounced for the case of two allowable discrete phases.

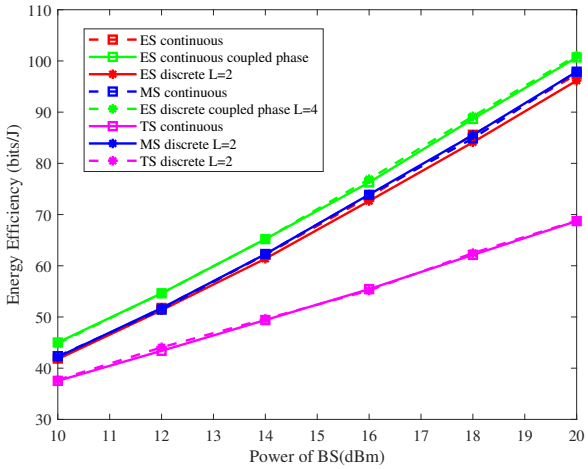


Fig. 11: Energy efficiency comparison of various operating modes at $M = 30$, $N = 16$, $K^r = K^t = 4$ and $\sigma_l^2 = -80\text{dBm}$

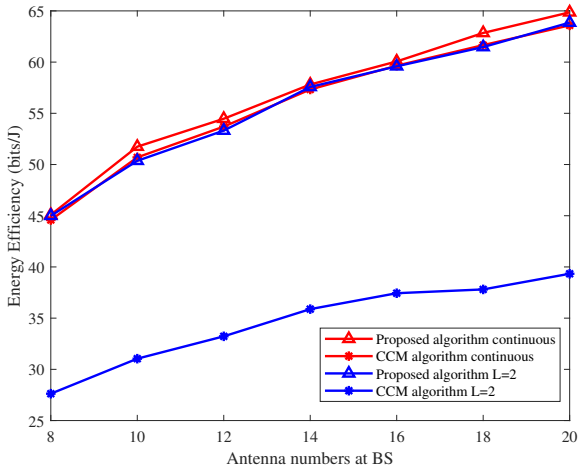


Fig. 12: Energy efficiency comparison in TS STAR-RIS at $M = 30$, $K^r = K^t = 4$, $P_{BS} = 20\text{dBm}$ and $\sigma_l^2 = -80\text{dBm}$

VI. Conclusions

This paper proposed a unified framework to efficiently handle the constraints introduced by various kinds of STAR-RISs, even with discrete and coupled phase constraints. The proposed unified framework introduces auxiliary variables for the STAR-RIS phases such that closed-form global optimal solution is possible at the subproblem level. In addition to unifying the algorithm derivations for systems involving various kinds of STAR-RISs and lowering the computational complexity, convergence to a stationary point under such framework was established theoretically. As an illustrated example, a downlink STAR-RIS assisted transmission system was investigated under the proposed framework. Simulation results showed that the proposed framework outperforms other existing state-of-the-art methods, and revealed for the first time that discrete phase may not cause significant performance degradation.

References

[1] Y. Zhou, L. Liu, L. Wang, N. Hui, X. Cui, J. Wu, Y. Peng, Y. Qi, and C. Xing, "Service-aware 6G: An intelligent and open network based on the convergence of communication, computing and caching," *Digit. Commun. Netw.*, vol. 6, no. 3, pp. 253–260, Aug 2020.

[2] Y. Zhou, L. Tian, L. Liu, and Y. Qi, "Fog computing enabled future mobile communication networks: A convergence of communication and computing," *IEEE Commun Mag*, vol. 57, no. 5, pp. 20–27, May 2019.

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

[4] D. Kudathanthirige, D. Gunasinghe, and G. Amarasuriya, "Performance analysis of intelligent reflective surfaces for wireless communication," in *ICC 2020*, July. 2020, pp. 1–6.

[5] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, Aug. 2019.

[6] J. Chen, Y.-C. Liang, Y. Pei, and H. Guo, "Intelligent reflecting surface: A programmable wireless environment for physical layer security," *IEEE Access*, vol. 7, pp. 82 599–82 612, June. 2019.

[7] M. Renzo, M. Debbah, D. Phan-Huy, and et al, "Smart radio enviroments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *J Wireless Com Network*, no. 129, pp. 129–149, May. 2019.

[8] Z. Li, S. Wang, M. Wen, and Y.-C. Wu, "Secure multicast energy-efficiency maximization with massive RISs and uncertain CSI: First-order algorithms and convergence analysis," *IEEE Trans. Wirel. Commun.*, vol. 21, no. 9, pp. 6818–6833, Sept. 2022.

[9] Y. Yang, Y. Gong, and Y.-C. Wu, "Intelligent-Reflecting-Surface-aided mobile edge computing with binary offloading: Energy minimization for IoT devices," *IEEE Internet Things J.*, vol. 9, no. 15, pp. 12 973–12 983, May. 2022.

[10] J. Xu, Y. Liu, X. Mu, and O. A. Dobre, "STAR-RISs: Simultaneous transmitting and reflecting reconfigurable intelligent surfaces," *IEEE Commun. Lett.*, vol. 25, no. 9, pp. 3134–3138, May. 2021.

[11] C. Wu, Y. Liu, X. Mu, X. Gu, and O. A. Dobre, "Coverage characterization of STAR-RIS networks: NOMA and OMA," *IEEE Commun. Lett.*, vol. 25, no. 9, pp. 3036–3040, Sept. 2021.

[12] H. Niu, Z. Chu, F. Zhou, and Z. Zhu, "Simultaneous transmission and reflection reconfigurable intelligent surface assisted secrecy MISO networks," *IEEE Commun. Lett.*, vol. 25, no. 11, pp. 3498–3502, Aug. 2021.

[13] M. Z. Alom, B. Van Essen, A. T. Moody, D. P. Widemann, and T. M. Taha, "Quadratic unconstrained binary optimization (QUBO) on neuromorphic computing system," in *IJCNN-2017*, July. 2017, pp. 3922–3929.

[14] P. P. Perera, V. G. Warnasooriya, D. Kudathanthirige, and H. A. Suraweera, "Sum rate maximization in STAR-RIS assisted full-duplex communication systems," in *ICC 2022*, Aug. 2022, pp. 3281–3286.

[15] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Simultaneously transmitting and reflecting (STAR) RIS aided wireless communications," *IEEE Trans. Wirel. Commun.*, vol. 21, no. 5, pp. 3083–3098, May. 2022.

[16] X. Ju, S. Gong, N. Zhao, C. Xing, A. Nallanathan, and D. Niyato, "A framework on complex matrix derivatives with special structure constraints for wireless systems," *IEEE Trans Commun*, pp. 1–1, 2024.

[17] Y. Liu, J. Xu, and X. Mu, "Fast beam splitting technique for STAR-RISs with coupled T&R phase shifts," in *AT-AP-RASC 2022*, May. 2022, pp. 1–3.

[18] R. Zhong, Y. Liu, X. Mu, Y. Chen, X. Wang, and L. Hanzo, "Hybrid reinforcement learning for STAR-RISs: A coupled phase-shift model based beamformer," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 9, pp. 2556–2569, Sept. 2022.

[19] Y. Liu, X. Mu, R. Schober, and H. V. Poor, "Simultaneously transmitting and reflecting (STAR)-RISs: A coupled phase-shift model," in *ICC 2022*, May. 2022, pp. 2840–2845.

[20] Z. Wang, X. Mu, Y. Liu, and R. Schober, "Coupled phase-shift STAR-RISs: A general optimization framework," *IEEE Wireless Commun. Lett.*, vol. 12, no. 2, pp. 207–211, Nov. 2023.

[21] Y. Chen, B. Ai, H. Zhang, Y. Niu, L. Song, Z. Han, and H. Vincent Poor, "Reconfigurable intelligent surface assisted device-to-device communications," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 5, pp. 2792–2804, May. 2021.

[22] H. Gao, K. Cui, C. Huang, and C. Yuen, "Robust beamforming for RIS-assisted wireless communications with discrete phase shifts," *IEEE Wireless Commun. Lett.*, vol. 10, no. 12, pp. 2619–2623, Aug. 2021.

[23] M. Katwe, K. Singh, B. Clerckx, and C.-P. Li, "Improved spectral efficiency in STAR-RIS aided uplink communication using rate splitting multiple access," *IEEE Trans. Wirel. Commun.*, pp. 1–1, Jan. 2023.

[24] X. Qin, Z. Song, T. Hou, W. Yu, J. Wang, and X. Sun, "Joint resource allocation and configuration design for STAR-RIS-enhanced wireless-powered MEC," *IEEE Trans Commun*, vol. 71, no. 4, pp. 2381–2395, Jan. 2023.

[25] W. Du, Z. Chu, G. Chen, P. Xiao, Y. Xiao, X. Wu, and W. Hao, "STAR-RIS assisted wireless powered IoT networks," *IEEE Trans. Veh. Technol.*, pp. 1–15, Mar. 2023.

[26] C. Wu, C. You, Y. Liu, X. Gu, and Y. Cai, "Channel estimation for STAR-RIS-aided wireless communication," *IEEE Commun. Lett.*, vol. 26, no. 3, pp. 652–656, Dec. 2022.

[27] W. Ni, Y. Liu, Y. C. Eldar, Z. Yang, and H. Tian, "STAR-RIS integrated nonorthogonal multiple access and over-the-air federated learning: Framework, analysis, and optimization," *IEEE Internet Things J.*, July. 2022.

[28] Q. Zhang, Y. Zhao, H. Li, S. Hou, and Z. Song, "Joint optimization of STAR-RIS assisted UAV communication systems," *IEEE Wireless Commun. Lett.*, vol. 11, no. 11, pp. 2390–2394, Sept. 2022.

[29] F. Fang, B. Wu, S. Fu, Z. Ding, and X. Wang, "Energy-Efficient design of STAR-RIS aided MIMO-NOMA networks," *IEEE Trans Commun*, vol. 71, no. 1, pp. 498–511, Nov. 2023.

[30] Z. Zhang, J. Chen, Y. Liu, Q. Wu, B. He, and L. Yang, "On the secrecy design of STAR-RIS assisted uplink NOMA networks," *IEEE Trans. Wirel. Commun.*, vol. 21, no. 12, pp. 11 207–11 221, Dec. 2022.

[31] M. F. U. Abrar, M. Talha, R. I. Ansari, S. A. Hassan, and H. Jung, "Optimization of STAR-RIS-assisted hybrid NOMA mmWave communication," *IEEE Trans. Veh. Technol.*, pp. 1–16, Mar. 2023.

[32] J. Zhao, Y. Zhu, X. Mu, K. Cai, Y. Liu, and L. Hanzo, "Simultaneously transmitting and reflecting reconfigurable intelligent surface (STAR-RIS) assisted UAV communications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 10, pp. 3041–3056, Aug. 2022.

[33] X. Zhai, G. Han, Y. Cai, Y. Liu, and L. Hanzo, "Simultaneously transmitting and reflecting (STAR) RIS assisted over-the-air computation systems," *IEEE Trans Commun*, vol. 71, no. 3, pp. 1309–1322, Mar. 2023.

[34] Z. Zhang, Z. Wang, Y. Liu, B. He, L. Lv, and J. Chen, "Security enhancement for coupled phase-shift STAR-RIS networks," *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 8210–8215, Feb. 2023.

[35] A. Beck and L. Tetruashvili, "On the convergence of block coordinate descent type methods," *SIAM J. Optim.*, vol. 23, no. 4, pp. 2037–2060, 2013.

[36] Q. Lin, R. Ma, and Y. Xu, "Complexity of an inexact proximal-point penalty method for constrained smooth non-convex optimization," *Comput Optim Appl*, vol. 82, pp. 175–224, Mar 2022.

[37] J. Sven and S. Anita, "The blockwise coordinate descent method for integer programs," *Math. Methods Oper. Res.*, vol. 91, pp. 357–381, Jun 2019.

[38] N. Gusmeroli and A. Wiegele, "EXPEDIS: An exact penalty method over discrete sets," *Discrete Optim.*, vol. 44, p. 100622, May 2022.

[39] C. Yu, K. L. Teo, and Y. Bai, "An exact penalty function method for nonlinear mixed discrete programming problems," *Optim. Lett.*, vol. 7, pp. 23–38, Aug 2013.

[40] A. S. Bedi, K. Rajawat, V. Aggarwal, and A. Koppel, "Escaping saddle points for successive convex approximation," *IEEE Trans. Signal Process.*, vol. 70, pp. 307–321, Dec. 2022.

[41] Y. Li, M. Xia, and Y.-C. Wu, "Energy-Efficient precoding for non-orthogonal multicast and unicast transmission via first-order algorithm," *IEEE Trans. Wirel. Commun.*, vol. 18, no. 9, pp. 4590–4604, July. 2019.

[42] A. Beck, *First-Order methods in optimization*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2017. [Online]. Available: https://epubs.siam.org/doi/abs/10.1137/1.9781611974997

[43] Z. Li, M. Xia, M. Wen, and Y.-C. Wu, "Massive access in secure NOMA under imperfect CSI: Security guaranteed sum-rate maximization with first-order algorithm," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 998–1014, Apr. 2021.

[44] T. Hou, J. Wang, Y. Liu, X. Sun, A. Li, and B. Ai, "A joint design for STAR-RIS enhanced NOMA-CoMP networks: A simultaneous-signal-enhancement-and-cancellation-based (SSECB) design," *IEEE Trans. Veh. Technol.*, vol. 71, no. 1, pp. 1043–1048, Jan 2022.

[45] H. Lee, K.-J. Lee, H. Kim, B. Clerckx, and I. Lee, "Resource allocation techniques for wireless powered communication betworks with energy storage constraint," *IEEE Trans. Wirel. Commun.*, vol. 15, no. 4, pp. 2619–2628, Apr. 2016.

[46] H. Ju and R. Zhang, "Optimal resource allocation in full-duplex wireless-powered communication network," *IEEE Trans Commun*, vol. 62, no. 10, pp. 3528–3540, Oct. 2014.

[47] K. Shen and W. Yu, "Fractional programming for communication systems—Part II: Uplink scheduling via matching," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2631–2644, Mar. 2018.

[48] Y. Yang, M. Pesavento, Z.-Q. Luo, and B. Ottersten, "Inexact block coordinate descent algorithms for nonsmooth nonconvex optimization," *IEEE Trans. Signal Process.*, vol. 68, pp. 947–961, Dec. 2020.

[49] Y. Xu and W. Yin, "A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion," *SIAM J. Imaging Sci.*, vol. 6, no. 3, pp. 1758–1789, 2013.

[50] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[51] Q. Lin, Y. Li, and Y.-C. Wu, "Sparsity constrained joint activity and data detection for massive access: A difference-of-norms penalty framework," *IEEE Trans. Wirel. Commun.*, vol. 22, no. 3, pp. 1480–1494, Mar 2023.

[52] Q. Shi and M. Hong, "Penalty dual decomposition method for non-smooth nonconvex optimization—Part i: Algorithms and convergence analysis," *IEEE Trans. Signal Process.*, vol. 68, pp. 4108–4122, Jun 2020.

[53] H. Li, C. Fang, W. Yin, and Z. Lin, "Decentralized accelerated gradient methods with increasing penalty parameters," *IEEE Trans. Signal Process.*, vol. 68, pp. 4855–4870, Aug 2020.

[54] Z. J. Towfic and A. H. Sayed, "Adaptive penalty-based distributed stochastic convex optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3924–3938, Aug 2014.

[55] C. Wu, X. Mu, Y. Liu, X. Gu, and X. Wang, "Resource allocation in STAR-RIS-aided networks: OMA and NOMA," *IEEE Trans. Wirel. Commun.*, vol. 21, no. 9, pp. 7653–7667, Mar. 2022.

[56] Z. Liu, Z. Li, M. Wen, Y. Gong, and Y.-C. Wu, "STAR-RIS aided mobile edge computing: Computation rate maximization with binary amplitude coefficients," *IEEE Trans Commun*, vol. 71, no. 7, pp. 4313–4327, July. 2023.

[57] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wirel. Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.

[58] L. You, J. Xiong, Y. Huang, D. W. K. Ng, C. Pan, W. Wang, and X. Gao, "Reconfigurable intelligent surfaces-assisted multiuser MIMO uplink transmission with partial CSI," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 9, pp. 5613–5627, Apr 2021.

[59] C. Feng, W. Shen, J. An, and L. Hanzo, "Joint hybrid and passive RIS-assisted beamforming for mmwave MIMO systems relying on dynamically configured subarrays," *IEEE Internet Things J.*, vol. 9, no. 15, pp. 13 913–13 926, Jan 2022.

[60] C. Pan, H. Ren, K. Wang, W. Xu, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug 2020.

[61] R. G. Bartle and D. R. Sherbert, *Introduction to real analysis*. John Wiley & Sons, Inc., 2000.

[62] K. C. Kiwiel, "Convergence and efficiency of subgradient methods for quasiconvex minimization," *Math Program*, vol. 90, pp. 1–25, Mar. 2001.

[63] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization for nonconvex and nonsmooth problems," *Math Program*, vol. 146, no. 1-2, pp. 459–494, July. 2014.

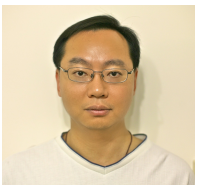
[64] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 5, pp. 3064–3076, May. 2020.

Hancheng Zhu received the B.Eng. degree from the Faculty of Computer Science and Technology, Nanjing Tech University, Nanjing, China, and the M.Eng. degree from the Faculty of Information Science and Engineering, Southeast University, Nanjing, China, in 2015 and 2018, respectively. He is currently working toward the Ph.D. degree with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong. His research interests include first-order optimization, and wireless communication.

Yuanwei Liu (S'13-M'16-SM'19-F'24, http://www.eecs.qmul.ac.uk/ yuanwei) is a Senior Lecturer (Associate Professor) with the School of Electronic Engineering and Computer Science, Queen Mary University of London. His research interests include non-orthogonal multiple access, reconfigurable intelligent surface, near field communications, integrated sensing and communications, and machine learning.

Yuanwei Liu is a Fellow of the IEEE, a Fellow of AAIA, a Web of Science Highly Cited Researcher, an IEEE Communication Society Distinguished Lecturer, an IEEE Vehicular Technology Society Distinguished Lecturer, the rapporteur of ETSI Industry Specification Group on Reconfigurable Intelligent Surfaces on work item of "Multi-functional Reconfigurable Intelligent Surfaces (RIS): Modelling, Optimisation, and Operation", and the UK representative for the URSI Commission C on "Radio communication Systems and Signal Processing". He was listed as one of 35 Innovators Under 35 China in 2022 by MIT Technology Review. He received IEEE ComSoc Outstanding Young Researcher Award for EMEA in 2020. He received the 2020 IEEE Signal Processing and Computing for Communications (SPCC) Technical Committee Early Achievement Award, IEEE Communication Theory Technical Committee (CTTC) 2021 Early Achievement Award. He received IEEE ComSoc Outstanding Nominee for Best Young Professionals Award in 2021. He is the co-recipient of the 2024 IEEE Communications Society Heinrich Hertz Award, and 5 IEEE conference best paper Awards. He serves as the Co-Editor-in-Chief of IEEE ComSoc TC Newsletter, an Area Editor of IEEE CL, an Editor of IEEE COMST/TWC/TVT/TNSE/TCCN. He serves as the (leading) Guest Editor for Proceedings of the IEEE and IEEE JSAC/JSTSP/Network. He serves the academic Chair for the Next Generation Multiple Access Emerging Technology Initiative, vice chair of SPCC and Technical Committee on Cognitive Networks (TCCN).

Yik-Chung Wu received the B.Eng. (EEE) and M.Phil. degrees from The University of Hong Kong (HKU), in 1998 and 2001, respectively, and the Ph.D. degree from Texas A&M University, College Station, in 2005. From 2005 to 2006, he was with the Thomson Corporate Research, Princeton, NJ, as a member of Technical Staff. Since 2006, he has been with HKU, currently as an Associate Professor. He was a Visiting Scholar at Princeton University in 2015 and 2017. His research interests include general areas of communication systems, signal processing, and machine learning. He received four best paper awards in international conferences, with the most recent one from IEEE International Conference on Communications (ICC) 2020. He served as an Editor for the IEEE COMMUNICATIONS LETTERS and the IEEE TRANSACTIONS ON COMMUNICATIONS. He is currently a Senior Area Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING, an Associate Editor for IEEE WIRELESS COMMUNICATION LETTERS (where he was elected the Best Editor of the Year 2023) and an Editor of Journal of Communications and Networks.



Vincent obtained B.Eng (Distinction 1st Hons) from the University of Hong Kong (1989-1992) and Ph.D. from the Cambridge University (1995-1997). He joined Bell Labs from 1997-2004 and the Department of ECE, Hong Kong University of Science and Technology (HKUST) in 2004. He is currently a Chair Professor and the Founding Director of Huawei-HKUST Joint Innovation Lab at HKUST. His current research focus includes Stochastic Optimization and Analysis for wireless systems, Massive MIMO Systems, Sparse Recovery, Bayesian Machine Learning, Mission-Critical IoT as well as PHY Caching for Wireless Networks.

SUPPLEMENTARY MATERIAL

VII. APPENDIX

A. Proof of Proposition 1

Let $\boldsymbol{p}_\gamma^n = \left[\boldsymbol{z}_\gamma^n, \lambda_\gamma^{t,n}, \lambda_\gamma^{r,n}\right]$, $\boldsymbol{v}_\gamma^n = \left[\boldsymbol{v}_\gamma^{t,n}, \boldsymbol{v}_\gamma^{r,n}\right]$ and $\boldsymbol{\varphi}_\gamma^n = \left[\boldsymbol{\varphi}_\gamma^{t,n}, \boldsymbol{\varphi}_\gamma^{r,n}\right]$ as the solutions at the n^{th} BCD iteration under the penalty γ. Using the above compact notations, the penalty term at the n^{th} BCD iteration under the penalty γ is reformulated as $\mathscr{G}\left(\boldsymbol{v}_\gamma^n, \boldsymbol{\varphi}_\gamma^n\right) = \frac{\gamma}{2}\left|\boldsymbol{v}_\gamma^n - \boldsymbol{\varphi}_\gamma^n\right|_2^2$. Since the BCD iteration is monotonic under any fixed γ, we have

$$\mathscr{G}\left(\boldsymbol{v}_\gamma^n, \boldsymbol{\varphi}_\gamma^n\right) \leq \mathscr{F}\left(\boldsymbol{p}_\gamma^0, \boldsymbol{v}_\gamma^0\right) + \mathscr{G}\left(\boldsymbol{v}_\gamma^0, \boldsymbol{\varphi}_\gamma^0\right) - \mathscr{F}\left(\boldsymbol{p}_\gamma^n, \boldsymbol{v}_\gamma^n\right). \quad (29)$$

With $\left\{\boldsymbol{p}_\gamma^0, \boldsymbol{v}_\gamma^0, \boldsymbol{\varphi}_\gamma^0\right\}$ being the initial point of the BCD iteration, we can always select them so that the first two terms of the right hand side of (29) are bounded. Besides, $\mathscr{F}$ is bounded from below in order to make its minimization meaningful. This makes $-\mathscr{F}\left(\boldsymbol{p}_\gamma^n, \boldsymbol{v}_\gamma^n\right)$ bounded from above, so does the right hand side of (29). Therefore, the left hand side of (29) is bounded from above for all n. That is, $\frac{\gamma}{2}\left|\boldsymbol{v}_\gamma^n - \boldsymbol{\varphi}_\gamma^n\right|_2^2$ is bounded from above for all n.

On the other hand, since the **Proposition 1** assumes that the $\left\{\boldsymbol{p}_\gamma^n, \boldsymbol{v}_\gamma^n\right\}$ of P2 does not contains infinite value, there exists a sufficient large D so that $\left|\left[\boldsymbol{p}_\gamma^n, \boldsymbol{v}_\gamma^n\right]\right| \leq D$ for all n (otherwise, $\left|\left[\boldsymbol{p}_\gamma^n, \boldsymbol{v}_\gamma^n\right]\right| \to \infty$ for some n and it contains infinite point, which violates the assumption in **Proposition 1**). Therefore, we obtain $\boldsymbol{\varphi}_\gamma^n$ is bounded for all n. Notice that the solution $\boldsymbol{p}_\gamma^n, \boldsymbol{v}_\gamma^n$ are also bounded by D, there must exist a subsequence $\left\{\boldsymbol{p}_\gamma^{n_j}, \boldsymbol{v}_\gamma^{n_j}, \boldsymbol{\varphi}_\gamma^{n_j}\right\}_{j\in\mathbb{N}}$ converges to some limit point $\left\{\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}_\gamma^*\right\}$ based on Bolzano-Weierstrass Theorem [61].

For the ease of representation, the constraints of P2 are denoted using an indicator function $\mathbb{I}_1(\boldsymbol{p}, \boldsymbol{v}) = \begin{cases} 0, & \text{if the constraints in P2 hold} \\ \infty, & \text{otherwise} \end{cases}$. By assumption in **Proposition 1** that the solution $\boldsymbol{p}_\gamma^{n_j}$, $\boldsymbol{v}_\gamma^{n_j}$ of P2 is a stationary point, the first-order optimality condition holds [42], which is shown in (30), where $\partial_{\boldsymbol{p}}\mathbb{I}_1(\boldsymbol{p}, \boldsymbol{v})$ and $\partial_{\boldsymbol{v}}\mathbb{I}_1(\boldsymbol{p}, \boldsymbol{v})$ are the limiting subdifferential of the non-smooth function $\mathbb{I}_1(\boldsymbol{p}, \boldsymbol{v})$ with respect to $\boldsymbol{p}$ and $\boldsymbol{v}$, respectively.

Since $\left\{\boldsymbol{p}_\gamma^{n_j}, \boldsymbol{v}_\gamma^{n_j}\right\}_{j\in\mathbb{N}}$ and $\left\{\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right\}$ are feasible points of P2, we have $\lim_{j\to\infty} \partial_{\boldsymbol{p}}\mathbb{I}_1\left(\boldsymbol{p}_\gamma^{n_j}, \boldsymbol{v}_\gamma^{n_j}\right) = \partial_{\boldsymbol{p}}\mathbb{I}_1\left(\lim_{j\to\infty} \boldsymbol{p}_\gamma^{n_j}, \boldsymbol{v}_\gamma^{n_j}\right) = \partial_{\boldsymbol{p}}\mathbb{I}_1\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right)$ and $\lim_{j\to\infty} \partial_{\boldsymbol{p}}\mathbb{I}_1\left(\boldsymbol{p}_\gamma^{n_j}, \boldsymbol{v}_\gamma^{n_j}\right) = \partial_{\boldsymbol{p}}\mathbb{I}_1\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right)$. Then, taking $j \to \infty$ in (30), we obtain (31), which can be written as

$$\mathbf{0} \in \begin{bmatrix} \nabla_{\boldsymbol{p}}\mathscr{F}\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) + \partial_{\boldsymbol{p}}\mathbb{I}_1\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) \\ \nabla_{\boldsymbol{v}}\mathscr{F}\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) + \nabla_{\boldsymbol{v}}\mathscr{G}\left(\boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}_\gamma^*\right) + \partial_{\boldsymbol{v}}\mathbb{I}_1\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) \end{bmatrix}. \quad (32)$$

On the other hand, since by the assumption in **Proposition 1** that the global optimal solution of P1 is obtained, we have

$$\mathscr{G}\left(\boldsymbol{v}_\gamma^{n_j}, \boldsymbol{\varphi}\right) + \mathbb{I}_2(\boldsymbol{\varphi}) \geq \mathscr{G}\left(\boldsymbol{v}_\gamma^{n_j}, \boldsymbol{\varphi}_\gamma^{n_j}\right) + \mathbb{I}_2\left(\boldsymbol{\varphi}_\gamma^{n_j}\right), \ \forall \boldsymbol{\varphi}, \quad (33)$$

where $\mathbb{I}_2(\boldsymbol{\varphi}) = \begin{cases} 0, & \text{if the constraints in P1 hold} \\ \infty, & \text{otherwise} \end{cases}$. Recognizing the constraints in P1 are compact constraint sets, $\mathbb{I}_2(\boldsymbol{\varphi})$ is a lower semi-continuous function [62], and consequently we have

$$\liminf_{j\to\infty} \mathbb{I}_2\left(\boldsymbol{\varphi}_\gamma^{n_j}\right) \geq \mathbb{I}_2\left(\boldsymbol{\varphi}_\gamma^*\right). \quad (34)$$

Taking $j \to \infty$ in (33) and applying (34), we obtain

$$\mathscr{G}\left(\boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}\right) + \mathbb{I}_2(\boldsymbol{\varphi}) \geq \mathscr{G}\left(\boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}_\gamma^*\right) + \mathbb{I}_2\left(\boldsymbol{\varphi}_\gamma^*\right), \ \forall \boldsymbol{\varphi}, \quad (35)$$

which guarantees that

$$\mathbf{0} \in \nabla_{\boldsymbol{\varphi}}\mathscr{G}\left(\boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}_\gamma^*\right) + \partial_{\boldsymbol{\varphi}}\mathbb{I}_2\left(\boldsymbol{\varphi}_\gamma^*\right). \quad (36)$$

Since $\mathbb{I}_1(\boldsymbol{p}, \boldsymbol{v})$ does not depend on $\boldsymbol{\varphi}$, we have $\mathbf{0} \in \partial_{\boldsymbol{\varphi}}\mathbb{I}_1\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right)$. Using similar arguments, we also have $\mathbf{0} \in \nabla_{\boldsymbol{\varphi}}\mathscr{F}\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right)$, $\mathbf{0} \in \nabla_{\boldsymbol{p}}\mathscr{G}\left(\boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}_\gamma^*\right)$, $\mathbf{0} \in \partial_{\boldsymbol{p}}\mathbb{I}_2\left(\boldsymbol{\varphi}_\gamma^*\right)$ and $\mathbf{0} \in \partial_{\boldsymbol{v}}\mathbb{I}_2\left(\boldsymbol{\varphi}_\gamma^*\right)$. Based on these and combine with (32) and (36), we obtain (37). Therefore, this limit point $\left\{\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}_\gamma^*\right\}$ is a stationary point of (2) [63]. Part 1) of **Proposition 1** is thus proved.

With $\lim_{j\to\infty}\left(\boldsymbol{p}_\gamma^{n_j}, \boldsymbol{v}_\gamma^{n_j}, \boldsymbol{\varphi}_\gamma^{n_j}\right) = \left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}_\gamma^*\right)$, and the sequence of solutions $\left\{\boldsymbol{p}_\gamma^{n_j}, \boldsymbol{v}_\gamma^{n_j}, \boldsymbol{\varphi}_\gamma^{n_j}\right\}_{j\in\mathbb{N}}$ are bounded for any γ, we know $\left\{\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}_\gamma^*\right\}$ is also bounded for any γ. Then there must exist a subsequence $\left\{\boldsymbol{p}_{\gamma_l}^*, \boldsymbol{v}_{\gamma_l}^*, \boldsymbol{\varphi}_{\gamma_l}^*\right\}_{l\in\mathbb{N}}$ converges to the limit point $\left\{\boldsymbol{p}_\infty^*, \boldsymbol{v}_\infty^*, \boldsymbol{\varphi}_\infty^*\right\}$ based on Bolzano-Weierstrass Theorem [61].

Suppose that $\boldsymbol{v}_\infty^* \neq \boldsymbol{\varphi}_\infty^*$, we would have $\left|\boldsymbol{v}_{\gamma_l}^* - \boldsymbol{\varphi}_{\gamma_l}^*\right|_2^2 > c_0$ for a positive c_0 and sufficiently large l. That yields $\lim_{l\to\infty} \mathscr{G}\left(\boldsymbol{v}_{\gamma_l}^*, \boldsymbol{\varphi}_{\gamma_l}^*\right) = \lim_{l\to\infty} \frac{\gamma_l}{2}\left|\boldsymbol{v}_{\gamma_l}^* - \boldsymbol{\varphi}_{\gamma_l}^*\right|_2^2 = \infty$, which violates the bounded property obtained below (29). Hence, by contradiction, we must have $\boldsymbol{v}_\infty^* = \boldsymbol{\varphi}_\infty^*$. According to part 1) of the **Proposition 1**, $\left\{\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}_\gamma^*\right\}$ is a stationary point of (2) for any γ. Thus we have $\left\{\boldsymbol{p}_\infty^*, \boldsymbol{v}_\infty^*, \boldsymbol{\varphi}_\infty^*\right\}$ is a stationary point of (2). Together with $\boldsymbol{v}_\infty^* = \boldsymbol{\varphi}_\infty^*$, which means that (2) is equivalent to (1), we obtain that $\left\{\boldsymbol{p}_\infty^*, \boldsymbol{v}_\infty^*\right\}$ is a stationary point of (1). This completed the proof of part 2).

B. Proof of Lemma 1

Firstly, we discuss the ES and TS STAR-RIS cases. Since φ_m^t and φ_m^r are separable in both the objective function and constraints, we can consider one φ_m^t or φ_m^r at a time. Taking φ_m^t as an example, the optimization problem (4) with respect to φ_m^t is

$$\min_{\varphi_m^t} \ \left|\varphi_m^t\right|^2 - 2\left|v_m^t\right|\left|\varphi_m^t\right|\cos\left(\angle\varphi_m^t - \angle v_m^t\right) \quad (38a)$$

$$s.t. \ \ \left|\varphi_m^t\right| = 1, \quad \text{if TS} \quad (38b)$$

$$\angle\varphi_m^t \in \{0, 2\pi/L, \cdots, 2\pi(L-1)/L\}, \quad (38c)$$

where we have expanded $\left|v_m^t - \varphi_m^t\right|^2$ and removed terms not related to φ_m^t. To minimize (38a), $\cos\left(\angle\varphi_m^t - \angle v_m^t\right)$ should be maximized. Taking the consideration of (38c), the optimal phase of $\angle\varphi_m^t$ is equal to $\text{Proj}_{\boldsymbol{\Theta}}\left(\angle v_m^t\right)$. Since TS mode restrict the amplitude of φ_m^t, $\text{Proj}_{\boldsymbol{\Theta}}\left(\angle v_m^t\right)$ is the solution of the TS mode. For the ES STAR-RIS, we put this optimal phase back into the objective function (38a), the resulting problem with respect to $\left|\varphi_m^t\right|$ is an unconstrained quadratic optimization problem and the closed-form solution is shown in **Lemma 1**.

$$\left\langle \left[\begin{array}{c} \nabla_{\boldsymbol{p}} \mathscr{F}\left(\boldsymbol{p}_\gamma^{n_j}, \boldsymbol{v}_\gamma^{n_j}\right) + \partial_{\boldsymbol{p}} \mathbb{I}_1\left(\boldsymbol{p}_\gamma^{n_j}, \boldsymbol{v}_\gamma^{n_j}\right) \\ \nabla_{\boldsymbol{v}} \mathscr{F}\left(\boldsymbol{p}_\gamma^{n_j}, \boldsymbol{v}_\gamma^{n_j}\right) + \nabla_{\boldsymbol{v}} \mathscr{G}\left(\boldsymbol{v}_\gamma^{n_j}, \boldsymbol{\varphi}_\gamma^{n_j-1}\right) + \partial_{\boldsymbol{v}} \mathbb{I}_1\left(\boldsymbol{p}_\gamma^{n_j}, \boldsymbol{v}_\gamma^{n_j}\right) \end{array} \right], \left[\begin{array}{c} \boldsymbol{p} - \boldsymbol{p}_\gamma^{n_j} \\ \boldsymbol{v} - \boldsymbol{v}_\gamma^{n_j} \end{array} \right] \right\rangle \geq 0, \forall \boldsymbol{p}, \boldsymbol{v}, \tag{30}$$

$$\left\langle \left[\begin{array}{c} \nabla_{\boldsymbol{p}} \mathscr{F}\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) + \partial_{\boldsymbol{p}} \mathbb{I}_1\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) \\ \nabla_{\boldsymbol{v}} \mathscr{F}\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) + \nabla_{\boldsymbol{v}} \mathscr{G}\left(\boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}_\gamma^*\right) + \partial_{\boldsymbol{v}} \mathbb{I}_1\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) \end{array} \right], \left[\begin{array}{c} \boldsymbol{p} - \boldsymbol{p}_\gamma^* \\ \boldsymbol{v} - \boldsymbol{v}_\gamma^* \end{array} \right] \right\rangle \geq 0, \ \ \forall \boldsymbol{p}, \boldsymbol{v}, \tag{31}$$

$$\mathbf{0} \in \left[\begin{array}{c} \nabla_{\boldsymbol{p}} \mathscr{F}\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) + \nabla_{\boldsymbol{p}} \mathscr{G}\left(\boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}_\gamma^*\right) + \partial_{\boldsymbol{p}} \mathbb{I}_1\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) + \partial_{\boldsymbol{p}} \mathbb{I}_2\left(\boldsymbol{\varphi}_\gamma^*\right) \\ \nabla_{\boldsymbol{v}} \mathscr{F}\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) + \nabla_{\boldsymbol{v}} \mathscr{G}\left(\boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}_\gamma^*\right) + \partial_{\boldsymbol{v}} \mathbb{I}_1\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) + \partial_{\boldsymbol{v}} \mathbb{I}_2\left(\boldsymbol{\varphi}_\gamma^*\right) \\ \nabla_{\boldsymbol{\varphi}} \mathscr{F}\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) + \nabla_{\boldsymbol{\varphi}} \mathscr{G}\left(\boldsymbol{v}_\gamma^*, \boldsymbol{\varphi}_\gamma^*\right) + \partial_{\boldsymbol{\varphi}} \mathbb{I}_1\left(\boldsymbol{p}_\gamma^*, \boldsymbol{v}_\gamma^*\right) + \partial_{\boldsymbol{\varphi}} \mathbb{I}_2\left(\boldsymbol{\varphi}_\gamma^*\right) \end{array} \right]. \tag{37}$$

For the MS STAR-RIS, there are two possibilities for the amplitude variables.

1) $\left|\varphi_m^t\right| = 0$ and $|\varphi_m^r| = 1$. Through expanding the objective function in (4a) and removing the terms unrelated to φ_m^r, the following optimization problem can be obtained from (4):

$$\begin{aligned} \min_{\angle\varphi_m^r} \ & -2\left|v_m^r\right| \cos\left(\angle\varphi_m^r - \angle v_m^r\right) \\ s.t. \ & \angle\varphi_m^r \in \{0, 2\pi/L, \cdots, 2\pi\left(L-1\right)/L\}. \end{aligned} \tag{39}$$

Therefore, the optimal solution of $\angle\varphi_m^r$ is $\mathrm{Proj}_{\boldsymbol{\Theta}}\left(\angle v_m^r\right)$, and the minimal value is $-2\left|v_m^r\right| \cos\left(\mathrm{Proj}_{\boldsymbol{\Theta}}\left(\angle v_m^r\right) - \angle v_m^r\right) = -2\beta_m^r$.

2) $\left|\varphi_m^t\right| = 1$ and $|\varphi_m^r| = 0$. With similar derivations to the above case, the optimal solution of $\angle\varphi_m^t$ is $\mathrm{Proj}_{\boldsymbol{\Theta}}\left(\angle v_m^t\right)$. Hence, the minimal value is $-2\left|v_m^t\right| \cos\left(\mathrm{Proj}_{\boldsymbol{\Theta}}\left(\angle v_m^t\right) - \angle v_m^t\right) = -2\beta_m^t$.

Finally, the optimal solution can be obtained by choosing the minimal value of the above two cases and the result is expressed in **Lemma 1** using $\mathrm{sgn}\left(\cdot\right)$ function.

C. Proof of Lemma 2

Let $\varphi_m^t = a_m^t e^{j\theta_m^t}$, where $a_m^t \in \mathbb{R}$ and $\theta_m^t \in \{0, 2\pi/L, \cdots, 2\pi\left(L-1\right)/L\}$. By expanding (6a) and removing the terms irrelvant to φ_m^t and φ_m^r, we have the optimization problem (40).

To minimize (40), $\angle\varphi_m^r$ should be chosen to maximize the term $\mathrm{Re}\left\{e^{j\left(\angle\varphi_m^r - \angle v_m^r\right)}\right\}$. On the other hand, from the constraint in (40), once θ_m^t is obtained, there are only two possible values for $\angle\varphi_m^r$ to choose from. Hence we have

$$\angle\varphi_m^r = \left\{ \begin{array}{l} \theta_m^t + \pi/2, \ \ \text{if} \ \ \sin\left(\theta_m^t - \angle v_m^r\right) \leq 0, \\ \theta_m^t - \pi/2, \ \ \text{otherwise.} \end{array} \right. \tag{41}$$

Putting (41) into (40), the problem is reduced to (42). This is an unconstrained quadratic function for a_m^t and $|\varphi_m^r|$, hence their optimal solutions are

$$\left\{ \begin{array}{l} a_m^t = \left|v_m^t\right| \cos\left(\theta_m^t - \angle v_m^t\right), \\ |\varphi_m^r| = |v_m^r| \left|\sin\left(\theta_m^t - \angle v_m^r\right)\right|. \end{array} \right. \tag{43}$$

With the second line of (43) and using (41), we obtain the solution of φ_m^r.

Finally, substituting (43) back to (42), the resulting problem with respect to θ_m^t is

$$\begin{aligned} \min_{\theta_m^t} \ & -\left|v_m^t\right|^2 \cos^2\left(\theta_m^t - \angle v_m^t\right) - |v_m^r|^2 \sin^2\left(\theta_m^t - \angle v_m^r\right). \\ s.t. \ & \theta_m^t \in \{0, 2\pi/L, \cdots, 2\pi\left(L-1\right)/L\}. \end{aligned} \tag{44}$$

Since $\cos^2 u = \left(1 + \cos 2u\right)/2$ and $\sin^2 u = \left(1 - \cos 2u\right)/2$, optimizing (44) is equivalent to

$$\begin{aligned} \min_{\theta_m^t} \ & |v_m^r|^2 \cos\left(2\theta_m^t - 2\angle v_m^r\right) - \left|v_m^t\right|^2 \cos\left(2\theta_m^t - 2\angle v_m^t\right) \\ s.t. \ & \theta_m^t \in \{0, 2\pi/L, \cdots, 2\pi\left(L-1\right)/L\}. \end{aligned} \tag{45}$$

Using sum-difference-product formula for trigonometric functions, the objective function in (45) is equal to

$$\sqrt{\chi_m} \cos\left(2\theta_m^t - 2\angle v_m^t + b_m\right), \tag{46}$$

where $\chi_m = \left[|v_m^r|^2 \cos\left(2\angle v_m^t - 2\angle v_m^r\right) + \left|v_m^t\right|^2\right]^2 + \left[|v_m^r|^2 \sin\left(2\angle v_m^t - 2\angle v_m^r\right)\right]^2$ and b_m is defined in **Lemma 2**. Since χ_m is unrelated to θ_m^t, (46) is equivalent to

$$\begin{aligned} \min_{\theta_m^t} \ & \cos\left(2\theta_m^t - 2\angle v_m^t + b_m\right) \\ s.t. \ & \theta_m^t \in \{0, 2\pi/L, \cdots, 2\pi\left(L-1\right)/L\}. \end{aligned} \tag{47}$$

Hence, the optimal solution of θ_m^t is $\mathrm{Proj}_{\boldsymbol{\Theta}}\left(\angle v_m^t - b_m/2 + \pi/2\right)$ and the optimal solution of (6) shown in **Lemma 2**.

D. Proof of Lemma 3

Firstly, to deal with the sum-of-logarithms-of-ratio objective function in (9), the closed-form fractional programming (FP) approach in [47], [64] is introduced, which has two key steps.

1) *Lagrangian Dual Transform:* The logarithm function can be represented with an auxiliary variable ρ as

$$\log\left(1+\gamma\right) = \max_{\rho} \log\left(1+\rho\right) - \rho + \frac{\left(1+\rho\right)\gamma}{1+\gamma}, \tag{48}$$

$$\begin{aligned}\min_{a_m^t,\theta_m^t,\varphi_m^r}\ & \left(a_m^t\right)^2-2a_m^t\left|v_m^t\right|\cos\left(\theta_m^t-\angle v_m^t\right)+\left|\varphi_m^r\right|^2-2\left|v_m^r\right|\left|\varphi_m^r\right|\mathrm{Re}\left\{e^{j\left(\angle\varphi_m^r-\angle v_m^r\right)}\right\}\\ s.t.\ & \angle\varphi_m^r=\theta_m^t\pm\pi/2\left(\mathrm{mod}2\pi\right),\\ & \theta_m^t\in\left\{0,2\pi/L,\cdots,2\pi\left(L-1\right)/L\right\}.\end{aligned}\tag{40}$$

$$\begin{aligned}\min_{a_m^t,\theta_m^t,|\varphi_m^r|}\ & \left(a_m^t\right)^2-2a_m^t\left|v_m^t\right|\cos\left(\theta_m^t-\angle v_m^t\right)+\left|\varphi_m^r\right|^2-2\left|v_m^r\right|\left|\varphi_m^r\right|\left|\sin\left(\theta_m^t-\angle v_m^r\right)\right|\\ s.t.\ & \theta_m^t\in\left\{0,2\pi/L,\cdots,2\pi\left(L-1\right)/L\right\}.\end{aligned}\tag{42}$$

where the optimal solution occurs at $\rho=\gamma$. Based on (48), the objective function (9) can be written as

$$\begin{aligned}&\mathscr{R}\left(\boldsymbol{w},\boldsymbol{v}^t,\boldsymbol{v}^r,\lambda^t,\lambda^r\right)=\\ &\max_{\rho}\ \lambda^r\sum_{l=1}^{K^r}\log\left(1+\rho_l\right)-\rho_l+\frac{\left(1+\rho_l\right)\left|\boldsymbol{a}_l^T\boldsymbol{w}_l\right|^2}{\sum_{i=1}^{K^r+K^t}\left|\boldsymbol{a}_l^T\boldsymbol{w}_i\right|^2+\lambda^r\sigma_l^2}\\ &+\lambda^t\sum_{l=K^r+1}^{K^r+K^t}\log\left(1+\rho_l\right)-\rho_l+\frac{\left(1+\rho_l\right)\left|\boldsymbol{a}_l^T\boldsymbol{w}_l\right|^2}{\sum_{i=1}^{K^r+K^t}\left|\boldsymbol{a}_l^T\boldsymbol{w}_i\right|^2+\lambda^t\sigma_l^2}.\end{aligned}\tag{49}$$

The second step is to tackle the fractional term in (49).

2) *Quadratic transform:* By introducing the auxiliary variable x, the following equation holds.

$$\frac{\left|A\left(\boldsymbol{u}\right)\right|^2}{B\left(\boldsymbol{u}\right)}=2\mathrm{Re}\left\{\overline{x}A\left(\boldsymbol{u}\right)\right\}-\left|x\right|^2B\left(\boldsymbol{u}\right).\tag{50}$$

The equivalence can be proved by substituting $x=A\left(\boldsymbol{u}\right)/B\left(\boldsymbol{u}\right)$. With the transformation in (50), (49) can be converted into

$$\mathscr{R}\left(\boldsymbol{w},\boldsymbol{v}^t,\boldsymbol{v}^r,\lambda^t,\lambda^r\right)=\max_{\boldsymbol{\rho},\boldsymbol{x}}\mathscr{F}_1\left(\boldsymbol{w},\boldsymbol{\rho},\boldsymbol{x},\boldsymbol{v}^t,\boldsymbol{v}^r,\lambda^t,\lambda^r\right),\tag{51}$$

where $\mathscr{F}_1$ is defined in (10).

E. Proof of Lemma 4

Introducing the dual variable μ to the constraint (14b), the problem (14) is equivalent to

$$\max_{\mu\geq0}\min_{\boldsymbol{w}}\ \sum_{l=1}^{K^r+K^t}\boldsymbol{w}_l^H\boldsymbol{\Xi}\boldsymbol{w}_l-2\mathrm{Re}\left[\boldsymbol{q}_l^H\boldsymbol{w}_l\right]+\mu\left(\boldsymbol{w}_l^H\boldsymbol{w}_l-P_{BS}\right).\tag{52}$$

For each fixed μ, this problem is a quadratic function for all $\left\{\boldsymbol{w}_l\right\}_{l=1}^{K^r+K^t}$. Therefore, the optimal solution of $\boldsymbol{w}_l$ is $\left(\boldsymbol{\Xi}+\mu\boldsymbol{I}\right)^{-1}\boldsymbol{q}_l$. Through eigenvalue decomposition $\boldsymbol{\Xi}=\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H$, the optimal $\boldsymbol{w}_l$ can be rewritten as

$$\boldsymbol{w}_l=\boldsymbol{U}(\boldsymbol{\Lambda}+\mu\boldsymbol{I})^{-1}\boldsymbol{U}^H\boldsymbol{q}_l.\tag{53}$$

According to the complementary slackness, there are two cases for μ:

1) $\mu=0$. That means the constraint (14b) is inactive. This happens when $\Sigma_{l=1}^{K^r+K^t}\boldsymbol{w}_l^H\boldsymbol{w}_l\leq P_{BS}$ with $\boldsymbol{w}_l=\boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^H\boldsymbol{q}_l$, which is equivalent to $\Sigma_{l=1}^{K^r+K^t}\boldsymbol{q}_l^H\boldsymbol{U}\boldsymbol{\Lambda}^{-2}\boldsymbol{U}^H\boldsymbol{q}_l\leq P_{BS}$. Recall the definition of $\boldsymbol{B}$ in **Lemma 4**, $\Sigma_{l=1}^{K^r+K^t}\boldsymbol{q}_l^H\boldsymbol{U}\boldsymbol{\Lambda}^{-2}\boldsymbol{U}^H\boldsymbol{q}_l=\mathrm{Tr}\left(\boldsymbol{\Lambda}^{-2}\boldsymbol{B}\right)$. Hence, this case only occurs when $\mathrm{Tr}\left(\boldsymbol{\Lambda}^{-2}\boldsymbol{B}\right)\leq P_{BS}$.

2) $\mu>0$. That means the constraint (14b) is active. Therefore, we need to solve the equation $\Sigma_{l=1}^{K^r+K^t}\boldsymbol{w}_l^H\boldsymbol{w}_l=P_{BS}$ with each $\boldsymbol{w}_l$ given by (53). Putting (53) into the constraint $\Sigma_{l=1}^{K^r+K^t}\boldsymbol{w}_l^H\boldsymbol{w}_l=P_{BS}$ and rearrange the terms with the trace operator, it can be shown that

$$\mathrm{Tr}\left(\left(\boldsymbol{\Lambda}+\mu\boldsymbol{I}\right)^{-2}\boldsymbol{B}\right)=P_{BS}.\tag{54}$$

Notice that the left hand side expression is a monotonic decreasing function of μ. The bisection method with searching interval $[0,\mu_{\max}]$ can be adopted to find this solution, where $\mu_{\max}$ should satisfies

$$\mathrm{Tr}\left(\left(\boldsymbol{\Lambda}+\mu_{\max}\boldsymbol{I}\right)^{-2}\boldsymbol{B}\right)\leq P_{BS}.\tag{55}$$

Since $\boldsymbol{\Xi}$ is a positive semidefinite matrix, the diagonal element of $\boldsymbol{\Lambda}$ is non-negative. Hence, if $\mathrm{Tr}\left(\left(\mu_{\max}\boldsymbol{I}\right)^{-2}\boldsymbol{B}\right)\leq P_{BS}$ holds, we would also have (55) holds. From $\mathrm{Tr}\left(\left(\mu_{\max}\boldsymbol{I}\right)^{-2}\boldsymbol{B}\right)\leq P_{BS}$, we can set $\mu_{\max}=\sqrt{\mathrm{Tr}\left(\boldsymbol{B}\right)/P_{BS}}$.

F. Further analysis of $\mathrm{Tr}\left(\boldsymbol{B}\right)$

Since $\boldsymbol{B}=\Sigma_{l=1}^{K^r+K^t}\boldsymbol{U}^H\boldsymbol{q}_l\boldsymbol{q}_l^H\boldsymbol{U}$ in **Lemma 4** and $\boldsymbol{U}$ is a unitary matrix, we have $\mathrm{Tr}\left(\boldsymbol{B}\right)=\sum_{l=1}^{K^r+K^t}\mathrm{Tr}\left(\boldsymbol{U}^H\boldsymbol{q}_l\boldsymbol{q}_l^H\boldsymbol{U}\right)=\sum_{l=1}^{K^r+K^t}\left|\boldsymbol{q}_l\right|^2$. With the definition of $\boldsymbol{q}_l$ under (14), and noticing that the time allocation variables $\left\{\lambda^r,\lambda^t\right\}$ and the auxiliary variables $\left\{\rho_l,x_l\right\}_{l=1}^{K^r+K^t}$ are all bounded, we have

$$\begin{aligned}&\mathrm{Tr}\left(\boldsymbol{B}\right)\propto\sum_{l=1}^{K^r+K^t}\left|\boldsymbol{a}_l\right|^2=\\ &\sum_{l=1}^{K^r}\left|\boldsymbol{G}^T\mathrm{diag}\left(\boldsymbol{v}^r\right)\boldsymbol{h}_l+\boldsymbol{d}_l\right|^2+\sum_{l=K^r+1}^{K^r+K^t}\left|\boldsymbol{G}^T\mathrm{diag}\left(\boldsymbol{v}^t\right)\boldsymbol{h}_l+\boldsymbol{d}_l\right|^2,\end{aligned}\tag{56}$$

where the equality is based on the definition of $\boldsymbol{a}_l$ under (9). Now, we focus on the first term:

$$
\begin{aligned}
&\sum_{l=1}^{K^r} \left| \boldsymbol{G}^T \operatorname{diag}\left(\boldsymbol{v}^r\right) \boldsymbol{h}_l + \boldsymbol{d}_l \right|^2 \\
&\leq \sum_{l=1}^{K^r} \left(\left| \boldsymbol{G}^T \operatorname{diag}\left(\boldsymbol{v}^r\right) \boldsymbol{h}_l \right| + |\boldsymbol{d}_l| \right)^2 \\
&\leq 2 \sum_{l=1}^{K^r} \left(\left| \boldsymbol{G}^T \operatorname{diag}\left(\boldsymbol{v}^r\right) \boldsymbol{h}_l \right|^2 + |\boldsymbol{d}_l|^2 \right) \qquad (57) \\
&\leq 2 \sum_{l=1}^{K^r} \left(|\boldsymbol{G}|^2 |\operatorname{diag}\left(\boldsymbol{v}^r\right)|^2 |\boldsymbol{h}_l|^2 + |\boldsymbol{d}_l|^2 \right) \\
&\leq 2 \sum_{l=1}^{K^r} \left(|\boldsymbol{G}|^2 |\boldsymbol{h}_l|^2 + |\boldsymbol{d}_l|^2 \right).
\end{aligned}
$$

The first inequality is due to the triangle inequality of the 2-Norm. The second inequality holds due to the basic inequality $(a+b)^2 \leq 2a^2 + 2b^2$. The third inequality is due to the compatibility of 2-Norm. The last inequality holds since $\operatorname{diag}\left(\boldsymbol{v}^r\right)$ is a diagonal matrix with every diagonal element bounded by 1.

Similarly, the second term is bounded by $2 \sum_{l=1+K^r}^{K^r+K^t} \left(|\boldsymbol{G}|^2 |\boldsymbol{h}_l|^2 + |\boldsymbol{d}_l|^2 \right)$. Therefore, $\operatorname{Tr}\left(\boldsymbol{B}\right)$ is bounded by $2 \sum_{l=1}^{K^r+K^t} \left(|\boldsymbol{G}|^2 |\boldsymbol{h}_l|^2 + |\boldsymbol{d}_l|^2 \right)$. Notice that $\boldsymbol{G}$, $\boldsymbol{h}_l$ and $\boldsymbol{d}_l$ are the all channel coefficients, $\operatorname{Tr}\left(\boldsymbol{B}\right)$ is related to the strength of the channel coefficients and only grows linearly with the number of users.