Achievable Rate Maximization for Underlay Spectrum Sharing MIMO System With Intelligent Reflecting Surface

Vaibhav Kumar, *Member, IEEE*, Mark F. Flanagan, *Senior Member, IEEE*, Rui Zhang, *Fellow, IEEE*, and Le-Nam Tran, *Senior Member, IEEE*

***Abstract*—In this letter, the achievable rate maximization problem is considered for intelligent reflecting surface (IRS) assisted multiple-input multiple-output (MIMO) systems in an underlay spectrum sharing scenario, subject to interference power constraints at the primary users. The formulated non-convex optimization problem is challenging to solve due to its non-convexity as well as coupling design variables in the constraints. Different from existing works that are mostly based on alternating optimization (AO), we propose a penalty dual decomposition based gradient projection (PDDGP) algorithm to solve this problem. We also provide a convergence proof and a complexity analysis for the proposed algorithm. We benchmark the proposed algorithm against two known solutions, namely a minimum mean-square error based AO algorithm and an inner approximation method with block coordinate descent. Specifically, the complexity of the proposed algorithm grows *linearly* with respect to the number of reflecting elements at the IRS, while that of the two benchmark methods grows with the third power of the number of IRS elements. Moreover, numerical results show that the proposed PDDGP algorithm yields considerably higher achievable rate than the benchmark solutions.**

***Index Terms*—Intelligent reflecting surface, underlay spectrum sharing, MIMO, penalty dual decomposition, gradient projection.**

I. Introduction

With the recent development of software-controlled metasurface technology, intelligent reflecting surfaces (IRSs) are being regarded as a promising technology to enhance the performance of next-generation wireless communications [1]. By virtue of flexible configuration of the radio propagation environment, IRSs are capable of increasing the spectral as well as the energy efficiency of wireless communication systems. On the other hand, spectrum sharing has great potential to support massive connectivity in the congested radio spectrum for beyond-fifth-generation (B5G) communications standards [2], [3]. Therefore, the design and analysis of IRS-assisted spectrum sharing to enhance the achievable rate of wireless communication systems have gained considerable attention in the recent years.

Manuscript received 10 April 2022; revised 26 May 2022; accepted 1 June 2022. Date of publication 8 June 2022; date of current version 9 August 2022. This work was supported in part by the Science Foundation Ireland under Grant 17/CDA/4786. The associate editor coordinating the review of this article and approving it for publication was M. Hasna. *(Corresponding author: Vaibhav Kumar.)*

Vaibhav Kumar, Mark F. Flanagan, and Le-Nam Tran are with the School of Electrical and Electronic Engineering, University College Dublin, Dublin 4, D04 V1W8 Ireland (e-mail: vaibhav.kumar@ieee.org; mark.flanagan@ieee.org; nam.tran@ucd.ie).

Rui Zhang is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore (e-mail: elezhang@nus.edu.sg).

Digital Object Identifier 10.1109/LWC.2022.3180988

In this context, the problem of optimal active and passive beamforming for secondary user (SU) rate maximization in an IRS-assisted single-input single-output (SISO) and multiple-input single-output (MISO) spectrum sharing systems were considered in [4]–[8]. However, studies on SU rate maximization for the more challenging IRS-assisted multiple-input multiple-output (MIMO) underlay spectrum sharing systems is still limited. In order to maximize the weighted sum rate of the secondary network in an IRS-assisted MIMO system in an underlay spectrum sharing scenario with multiple secondary receivers (SRs) and a single primary receiver (PR), a block coordinate descent algorithm along with an inner approximation method (BCD-IAM) was proposed in [9]. Similarly, in [10], the authors considered the problem of weighted sum rate maximization of the secondary network for an IRS-assisted MIMO spectrum sharing system consisting of multiple SRs as well as multiple PRs, where they adopted a weighted minimum mean-square error (WMMSE) method and an alternating optimization (AO)-based algorithm to find a suboptimal solution. On the other hand, a suboptimal solution to the problem of achievable secrecy rate maximization for an IRS-assisted MIMO system using AO, barrier method, SDR and minorization-maximization (MM), was presented in [11].

It is noteworthy that the methods proposed to maximize the SU rate for MISO systems in [4]–[8] are not applicable to MIMO systems, unless each receive antenna in the MIMO system is treated as a separate user. However, this reduces the beamforming/multiplexing gains which are major benefits of MIMO systems. The method presented in [9] for MIMO systems is proposed only for the case of a single PR. Similarly, the system considered in [11], when omitting the eavesdropper, reduces to the conventional IRS-assisted MIMO spectrum sharing system. However, the solution proposed in [11] is limited to the scenario where the direct links between the (secondary) transmitter and (primary and secondary) receivers are absent. On the other hand, the AO-based solution proposed in [10] can be inefficient due to the possibly high coupling of design variables, especially when the SU rate performance is dominated by the interference constraints at the PRs. Motivated by the drawbacks of existing methods, in this letter we propose a low-complexity but efficient algorithm to maximize the achievable rate of the SU in an IRS-assisted MIMO underlay spectrum sharing system consisting of multiple PRs, where direct links between the (secondary) transmitter and (primary and secondary) receivers are also present. Our main contributions in this letter are as follows: (i) we propose a penalty dual decomposition based gradient projection (PDDGP) algorithm to maximize the achievable rate; (ii) we provide a convergence proof and complexity analysis of the proposed algorithm; and (iii) we provide

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

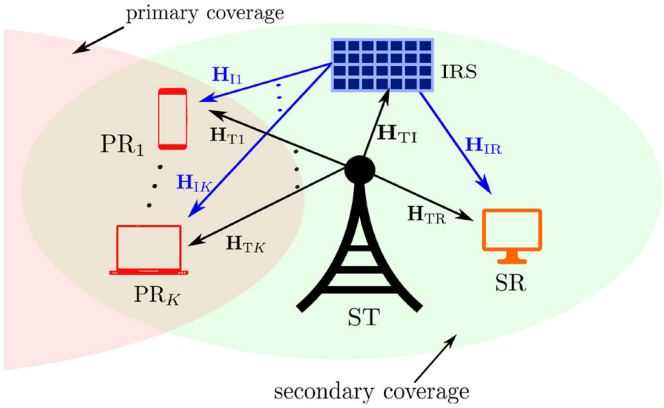

Fig. 1. System model for IRS-assisted MIMO underlay spectrum sharing.

extensive numerical results to demonstrate the superiority of the proposed solution over the benchmark solutions provided in [9] and [10], and to show the effect of different system parameters on the achievable rate.

Notations: Bold lowercase and uppercase letters denote vectors and matrices, respectively. $\mathbf{X}^*$, $\mathbf{X}^{\mathsf{T}}$, $\mathbf{X}^{\dagger}$, $|\mathbf{X}|$, $\|\mathbf{X}\|$ and $\mathrm{Tr}(\mathbf{X})$ denote the complex conjugate, (ordinary) transpose, Hermitian transpose, determinant, Frobenius norm and trace of a matrix $\mathbf{X}$, respectively. The vector space of complex matrices of size $M \times N$ is denoted by $\mathbb{C}^{M\times N}$. $\mathrm{diag}(\mathbf{x})$ denotes a square diagonal matrix where the elements of $\mathbf{x}$ comprise the main diagonal, $\mathrm{vec}_{\mathrm{d}}(\mathbf{X})$ denotes the column vector containing the elements of the main diagonal of the matrix $\mathbf{X}$, and $\nabla_{\mathbf{X}} f(\cdot)$ denotes the complex-valued gradient of $f(\cdot)$ with respect to (w.r.t.) $\mathbf{X}^*$. $\mathbb{E}\{\cdot\}$ denotes the expectation operator. $\mathbf{X} \succeq \mathbf{Y}$ represents that $\mathbf{X} - \mathbf{Y}$ is positive semidefinite, and $\Pi_{\mathcal{X}}(\mathbf{x})$ denotes the Euclidean projection of the point $\mathbf{x}$ onto the set $\mathcal{X}$, i.e., $\Pi_{\mathcal{X}}(\mathbf{x}) \triangleq \mathrm{argmin}_{\hat{\mathbf{x}}\in\mathcal{X}} \|\hat{\mathbf{x}} - \mathbf{x}\|$.

II. System Model and Problem Formulation

Let us first consider an IRS-assisted MIMO system in an underlay spectrum sharing scenario shown in Fig. 1, consisting of one secondary transmitter (ST), one SR, and K PRs. The number of antennas at the ST and SR are denoted by N_{T} and N_{R}, respectively, while the K PRs are each equipped with N_{P} antennas. The communication between the ST and SR is assisted by an IRS consisting of N_{I} low-cost passive reflecting elements. The ST-SR, ST-IRS, and IRS-SR channels are denoted by $\mathbf{H}_{\mathrm{TR}} \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{T}}}$, $\mathbf{H}_{\mathrm{TI}} \in \mathbb{C}^{N_{\mathrm{I}}\times N_{\mathrm{T}}}$, and $\mathbf{H}_{\mathrm{IR}} \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{I}}}$, respectively. On the other hand, the channel between the ST and the k-th PR ($k \in \mathcal{K} \triangleq \{1, 2, \ldots, K\}$) is denoted by $\mathbf{H}_{\mathrm{T}\,k} \in \mathbb{C}^{N_{\mathrm{P}}\times N_{\mathrm{T}}}$, while the channel between the IRS and the k-th PR is denoted by $\mathbf{H}_{\mathrm{I}\,k} \in \mathbb{C}^{N_{\mathrm{P}}\times N_{\mathrm{I}}}$. Let $\boldsymbol{\theta} \triangleq [\theta_1\ \theta_2\ \ldots\ \theta_{N_{\mathrm{I}}}]^{\mathsf{T}} \in \mathbb{C}^{N_{\mathrm{I}}\times 1}$ denote the phase-shift vector at the IRS, where $\theta_l = e^{j\phi_l}$, $l \in \mathcal{N}_{\mathrm{I}} \triangleq \{1, 2, \ldots, N_{\mathrm{I}}\}$, and $\phi_l \in [0, 2\pi)$. Here the ST, SR and IRS constitute the secondary network, and in order to communicate with the SR, the ST uses the spectrum band whose license is owned by a primary network. However, with the help of properly designed beamformers at the ST and IRS, it is ensured that the interference inflicted by the secondary network on the PRs is below a predefined threshold.[1]

The signal received at the SR is given by

$$\mathbf{y}_{\mathrm{R}} = (\mathbf{H}_{\mathrm{IR}}\,\boldsymbol{\Theta}\mathbf{H}_{\mathrm{TI}} + \mathbf{H}_{\mathrm{TR}})\mathbf{x} + \mathbf{n}_{\mathrm{R}}, \tag{1}$$

where $\boldsymbol{\Theta} \triangleq \mathrm{diag}(\boldsymbol{\theta})$, $\mathbf{x} \in \mathbb{C}^{N_{\mathrm{T}}\times 1}$ is the transmitted signal vector from the ST (which is intended for the SR), $\mathbf{n}_{\mathrm{R}} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$ denotes the complex additive white Gaussian noise (AWGN) vector at the SR, and σ^2 is the noise power. Without loss of generality, we assume that the noise at all of the PRs is also distributed as $\mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$. With a slight abuse of notation, in the rest of letter, we make the replacements $\mathbf{H}_{\mathrm{TI}} \leftarrow \mathbf{H}_{\mathrm{TI}}/\sigma$, $\mathbf{H}_{\mathrm{TR}} \leftarrow \mathbf{H}_{\mathrm{TR}}/\sigma$ and $\mathbf{H}_{\mathrm{T}\,k} \leftarrow \mathbf{H}_{\mathrm{T}\,k}/\sigma$. This normalization step will also mitigate potential numerical issues since we avoid dealing with extremely small quantities.

Problem Formulation: Denote the transmit covariance matrix at the ST by $\mathbf{X} \triangleq \mathbb{E}\{\mathbf{x}\mathbf{x}^{\dagger}\}$, the maximum transmit power budget at the ST by $P_{\max}$, and the maximum tolerable interference power (normalized to the noise power) at the k-th PR by P_k. Then, the achievable rate maximization problem for the secondary network can be formulated as follows:

$$\begin{aligned}
\underset{\mathbf{X}\succeq\mathbf{0},\boldsymbol{\theta}}{\text{maximize}} \quad & R(\mathbf{X}, \boldsymbol{\theta}) = \ln|\mathbf{I} + \mathbf{Z}_{\mathrm{R}}\mathbf{X}\mathbf{Z}_{\mathrm{R}}^{\dagger}|, && \text{(2a)}\\
\text{subject to} \quad & \mathrm{Tr}(\mathbf{X}) \le P_{\max}, && \text{(2b)}\\
& \mathrm{Tr}(\mathbf{Z}_k\mathbf{X}\mathbf{Z}_k^{\dagger}) \le P_k, \forall k \in \mathcal{K}, && \text{(2c)}\\
& |\theta_l| = 1, \forall l \in \mathcal{N}_{\mathrm{I}}, && \text{(2d)}
\end{aligned}$$

where $\mathbf{Z}_{\mathrm{R}} \triangleq \mathbf{H}_{\mathrm{IR}}\,\boldsymbol{\Theta}\mathbf{H}_{\mathrm{TI}} + \mathbf{H}_{\mathrm{TR}}$, and $\mathbf{Z}_k \triangleq \mathbf{H}_{\mathrm{I}\,k}\boldsymbol{\Theta}\mathbf{H}_{\mathrm{TI}} + \mathbf{H}_{\mathrm{T}\,k}$. Note that the constraints in (2b)–(2d) represent the transmit power constraint at the ST, the interference power constraints (IPCs) at the PRs and the unit-modulus constraints (UMCs) at the IRS, respectively. Furthermore, the feasible sets for $\mathbf{X}$ and $\boldsymbol{\theta}$ are defined as $\mathcal{X} \triangleq \{\mathbf{X} \in \mathbb{C}^{N_{\mathrm{T}}\times N_{\mathrm{T}}} : \mathrm{Tr}(\mathbf{X}) \le P_{\max}; \mathbf{X} \succeq \mathbf{0}\}$ and $\vartheta \triangleq \{\boldsymbol{\theta} \in \mathbb{C}^{N_{\mathrm{I}}\times 1} : |\theta_l| = 1, l \in \mathcal{N}_{\mathrm{I}}\}$, respectively.

III. Proposed Solution

A. PDDGP Algorithm

It is easy to see that problem (2) is non-convex due to the non-convexity of the objective as well as constraints (2c) and (2d). If the IPCs were not present, (2) would reduce to the conventional MIMO capacity problem for which the gradient projection algorithm in [13] can be applied to obtain a stationary solution. In fact, it is the coupling of $\mathbf{X}$ and $\boldsymbol{\theta}$ due to the IPCs in (2c) that makes the problem much more difficult to solve and the gradient projection algorithm proposed in [13] is no longer applicable. The main idea of the existing optimization methods in [9], [10] for solving (2) is to alternately optimize $\mathbf{X}$ and $\boldsymbol{\theta}$. However, due to the coupling of $\mathbf{X}$ and $\boldsymbol{\theta}$, these methods are not guaranteed to return a stationary solution. Indeed we numerically demonstrate in Section IV that these methods are inferior to the PDDGP algorithm that we propose in the following.

Specifically, in order to cope with the IPCs, we apply the penalty dual decomposition (PDD) method recently

[1]Following the arguments in the related works of [6], [7], [9], [11], [12], we assume that the secondary network is deployed outside the coverage of the primary transmitter, and that the distance between the primary transmitter and secondary receiver is large enough for the primary-to-secondary interference to be negligible. However, it is noteworthy that the proposed PDDGP algorithm is also applicable when the secondary network is deployed completely or partially inside the coverage of the primary transmitter.

presented in [14]. First, let $g_k(\mathbf{X},\boldsymbol{\theta},P_k,s_k) \triangleq \mathrm{Tr}(\mathbf{Z}_k\mathbf{X}\mathbf{Z}_k^\dagger)+ s_k - P_k$. Then it is easy to see that (2c) is equivalent to $g_k(\mathbf{X},\boldsymbol{\theta},P_k,s_k)=0, s_k \geq 0, \forall k \in \mathcal{K}$. Next, the augmented Lagrangian function can be written as follows:

$$\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X},\boldsymbol{\theta},\mathbf{s}) \triangleq \ln|\mathbf{I}+\mathbf{Z}_{\mathrm{R}}\mathbf{X}\mathbf{Z}_{\mathrm{R}}^\dagger| - \sum_{k\in\mathcal{K}} \upsilon_k g_k(\mathbf{X},\boldsymbol{\theta},P_k,s_k) - \frac{1}{2\rho}\sum_{k\in\mathcal{K}} g_k^2(\mathbf{X},\boldsymbol{\theta},P_k,s_k),$$

where $\mathbf{s} \triangleq [s_1,s_2,\ldots,s_K]^\mathsf{T}$, $\boldsymbol{\upsilon} \triangleq [\upsilon_1,\upsilon_2,\ldots,\upsilon_K]^\mathsf{T}$, υ_k is the Lagrangian multiplier corresponding to the constraint $g_k(\mathbf{X},\boldsymbol{\theta},P_k,s_k)=0$ and ρ is the *penalty parameter*. Hence, for a given $(\boldsymbol{\upsilon},\rho)$ an equivalent optimization problem can be formulated as

$$\underset{\mathbf{X},\boldsymbol{\theta},\mathbf{s}}{\text{maximize}} \quad \hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X},\boldsymbol{\theta},\mathbf{s}), \tag{3a}$$

$$\text{subject to} \quad \mathbf{s}\geq\mathbf{0}, (2\text{b}), (2\text{d}). \tag{3b}$$

Since the IPCs (and thus the coupling of $\mathbf{X}$ and $\boldsymbol{\theta}$) is now included in the augmented objective, the optimization variables in (3) are decoupled. As a result, a simple but efficient method can be applied to find a stationary solution to (3) which is based on the alternating projected gradient method (APGM). The APGM is motivated by the fact that the feasible set with respect to individual variables is simple in the sense that the projection onto this set can be expressed in closed form.

First, a projected gradient step for $\boldsymbol{\theta}$ is performed while other variables are held fixed. In this regard, the gradient of $\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X},\boldsymbol{\theta},\mathbf{s})$ w.r.t. $\boldsymbol{\theta}$ is given by

$$\nabla_{\boldsymbol{\theta}}\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X},\boldsymbol{\theta},\mathbf{s}) = \nabla_{\boldsymbol{\theta}}\ln|\mathbf{I}+\mathbf{Z}_{\mathrm{R}}\mathbf{X}\mathbf{Z}_{\mathrm{R}}^\dagger| - \sum_{k\in\mathcal{K}}\left[\upsilon_k+\frac{1}{\rho}g_k(\mathbf{X},\boldsymbol{\theta},P_k,s_k)\right]\nabla_{\boldsymbol{\theta}}g_k(\mathbf{X},\boldsymbol{\theta},P_k,s_k). \tag{4}$$

Next using (4), [13, eqn. (17a)] and [15, Table 4.3 and eqn. (6.153)], $\nabla_{\boldsymbol{\theta}}\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X},\boldsymbol{\theta},\mathbf{s})$ is given by (5), as shown at the bottom of the page. Suppose that, from the current point $\boldsymbol{\theta}$, after moving along $\nabla_{\boldsymbol{\theta}}\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X},\boldsymbol{\theta},\mathbf{s})$ with some step size one arrives at a point $\hat{\boldsymbol{\theta}}=[\hat{\theta}_1,\hat{\theta}_2,\ldots,\hat{\theta}_{N_{\mathrm{I}}}]^\mathsf{T}\in\mathbb{C}^{N_{\mathrm{I}}\times 1}$. To obtain the next iterate, the projection of $\hat{\boldsymbol{\theta}}$ onto the set ϑ is needed. It is easy to see that this projection is given by $\Pi_\vartheta(\hat{\boldsymbol{\theta}}) = [\bar{\theta}_1,\bar{\theta}_2,\ldots,\bar{\theta}_{N_{\mathrm{I}}}]^\mathsf{T}$, where for each $l\in\mathcal{N}_{\mathrm{I}}$

$$\bar{\theta}_l = \begin{cases} \hat{\theta}_l/|\hat{\theta}_l|, & \text{if } \hat{\theta}_l \neq 0 \\ \exp(j\phi), \phi\in[0,2\pi), & \text{otherwise.}\end{cases} \tag{6}$$

Next, a projected gradient step for $\mathbf{X}$ is carried out. Following a similar line of argument and using [15, eqns. (6.207), (6.195) and Table 4.3], a closed-form expression for the gradient of $\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X},\boldsymbol{\theta},\mathbf{s})$ w.r.t. $\mathbf{X}$ can be obtained in (7), as shown at the bottom of the page. Moreover, the projection of a given point $\mathbf{W}\in\mathbb{C}^{N_{\mathrm{T}}\times N_{\mathrm{T}}}$ onto the feasible set $\mathcal{X}$ can be shown to admit a water-filling solution. Due to space constraints, the details regarding the projection of $\mathbf{W}$ onto $\mathcal{X}$ are omitted, while the interested reader may refer to [13, Sec. III-C] for details.

Algorithm 1: Gradient Projection Algorithm to Solve (3) for Fixed $\boldsymbol{\upsilon}$ and ρ

Input: $\boldsymbol{\theta}_0$, $\mathbf{X}_0$, $\mathbf{s}_0$, $\boldsymbol{\upsilon}$, μ_0, α_0, ρ
Output: $\boldsymbol{\theta}_n, \mathbf{X}_n$
$n\leftarrow 1$
repeat
  $\boldsymbol{\theta}_n = \Pi_\vartheta(\boldsymbol{\theta}_{n-1}+\mu_n\nabla_{\boldsymbol{\theta}}\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}_{n-1},\boldsymbol{\theta}_{n-1},\mathbf{s}_{n-1}))$;
  $\mathbf{X}_n = \Pi_{\mathcal{X}}(\mathbf{X}_{n-1}+\alpha_n\nabla_{\mathbf{X}}\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}_{n-1},\boldsymbol{\theta}_n,\mathbf{s}_{n-1}))$;
  $s_{n,k}=\max\{0,P_k-\mathrm{Tr}(\mathbf{Z}_k\mathbf{X}_n\mathbf{Z}_k^\dagger)\},\forall k\in\mathcal{K}$
  where $\mathbf{Z}_k = \mathbf{H}_{\mathrm{I}\,k}\boldsymbol{\Theta}_n\mathbf{H}_{\mathrm{TI}}+\mathbf{H}_{\mathrm{T}\,k}$;
  $n\leftarrow n+1$;
until *convergence*;

Algorithm 2: The PDDGP Algorithm

Input: $\boldsymbol{\theta}_0$, $\mathbf{X}_0$, $\mathbf{s}_0$, $\boldsymbol{\upsilon}_0$, μ_0, α_0, ρ, $\kappa<1$
Output: $\boldsymbol{\theta}^\star, \mathbf{X}^\star$
repeat
  Solve problem (3) using **Algorithm** 1;
  $\mathbf{X}^\star\leftarrow\mathbf{X}_n$, $\boldsymbol{\theta}^\star\leftarrow\boldsymbol{\theta}_n$, $\mathbf{s}^\star\leftarrow\mathbf{s}_n$;
  $\upsilon_k\leftarrow\upsilon_k+\frac{1}{\rho}g_k(\mathbf{X}^\star,\boldsymbol{\theta}^\star,P_k,s_k^\star),\ \forall k\in\mathcal{K}$;
  $\rho\leftarrow\kappa\rho$;
until *convergence*;

We now turn to the optimization over $\mathbf{s}$. It is easy to see that the optimal solution to (3), when $\boldsymbol{\theta}$ and $\mathbf{X}$ are fixed, is simply given by

$$s_{n,k}=\max\{0,P_k-\mathrm{Tr}(\mathbf{Z}_k\mathbf{X}_n\mathbf{Z}_k^\dagger)\},\forall k\in\mathcal{K}$$

and thus a projection step for $\mathbf{s}$ is not necessary. The proposed APGM to find a stationary solution to (3) for fixed $\boldsymbol{\upsilon}$ and ρ is summarized in Algorithm 1. Note that in Algorithm 1, μ_n and α_n denote the step size corresponding to the gradient step for the $\boldsymbol{\theta}$ and $\mathbf{X}$ update in iteration *n*, respectively. Appropriate values of μ_n and α_n can be obtained by a backtracking*line search. Specifically,* we can set $\mu_n=\mu_{n-1}\gamma^{i_n}$ where i_n is the smallest positive integer such that

$$\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}_{n-1},\boldsymbol{\theta}_n,\mathbf{s}_{n-1}) \geq \hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}_{n-1},\boldsymbol{\theta}_{n-1},\mathbf{s}_{n-1}) + \langle\nabla_{\boldsymbol{\theta}}\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}_{n-1},\boldsymbol{\theta}_{n-1},\mathbf{s}_n),\boldsymbol{\theta}_n-\boldsymbol{\theta}_{n-1}\rangle - \frac{1}{\mu_{n-1}\gamma^{i_n}}\|\boldsymbol{\theta}_n-\boldsymbol{\theta}_{n-1}\|^2, \tag{8}$$

where $\langle\mathbf{x},\mathbf{y}\rangle\triangleq 2\Re(\mathbf{x}^\dagger\mathbf{y})$ and $\gamma<1$. A similar routine can be followed to obtain α_n.

After solving (3) for given $(\boldsymbol{\upsilon},\rho)$, the penalty parameter, ρ, is decreased and the Lagrange multipliers are updated. The overall description of the proposed PDDGP algorithm to find a stationary solution to (3) is outlined in Algorithm 2.

B. Extension to Multiple SRs

We have presented the system model where there is a single SR for the sake of simplicity. However, we remark that the extension to deal with multiple SRs is straightforward. In this case, we can consider the maximization of the weighted sum

$$\nabla_{\boldsymbol{\theta}}\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X},\boldsymbol{\theta},\mathbf{s}) = \mathrm{vec}_{\mathrm{d}}\left(\mathbf{H}_{\mathrm{IR}}^\dagger\left(\mathbf{I}+\mathbf{Z}_{\mathrm{R}}\mathbf{X}\mathbf{Z}_{\mathrm{R}}^\dagger\right)^{-1}\mathbf{Z}_{\mathrm{R}}\mathbf{X}\mathbf{H}_{\mathrm{TI}}^\dagger\right) - \sum_{k\in\mathcal{K}}\left[\upsilon_k+\tfrac{1}{\rho}\{\mathrm{Tr}(\mathbf{Z}_k\mathbf{X}\mathbf{Z}_k^\dagger)+s_k-P_k\}\right]\mathrm{vec}_{\mathrm{d}}(\mathbf{H}_{\mathrm{I}\,k}^\dagger\mathbf{Z}_k\mathbf{X}\mathbf{H}_{\mathrm{TI}}^\dagger) \tag{5}$$

$$\nabla_{\mathbf{X}}\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X},\boldsymbol{\theta},\mathbf{s}) = \mathbf{Z}_{\mathrm{R}}^\dagger(\mathbf{I}+\mathbf{Z}_{\mathrm{R}}\mathbf{X}\mathbf{Z}_{\mathrm{R}}^\dagger)^{-1}\mathbf{Z}_{\mathrm{R}} - \sum_{k\in\mathcal{K}}\left[\upsilon_k+\tfrac{1}{\rho}\{\mathrm{Tr}(\mathbf{Z}_k\mathbf{X}\mathbf{Z}_k^\dagger)+s_k-P_k\}\right]\mathbf{Z}_k^\dagger\mathbf{Z}_k \tag{7}$$

rate (WSR) of all SRs as in [9], [10] and the transmit power constraint (2b) becomes the sum power constraint. It is easy to check that the projected gradient step for $\boldsymbol{\theta}$ requires minimal modifications. Also, the projection onto the set defined by the sum power constraint still admits a water-filling solution. Thus, it is trivial to see that Algorithm 2 can be easily modified to solve the resulting WSR maximization problem. The detailed derivations are skipped here due to space limitation and only the numerical results are presented for the multiple SR case in the next section.

C. Proof of Convergence

The convergence proof of Algorithm 2 follows the same arguments as given in [14], and thus we only summarize the main points here. First, for given $(\boldsymbol{\upsilon}, \rho)$, Algorithm 1 generates a strictly increasing objective sequence, $\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}, \boldsymbol{\theta}, \mathbf{s})$. Since the feasible set is bounded, the iterates returned by Algorithm 1 converge to a limit point of (3a) to some accuracy ϵ_k. Next, it can be shown that the sequence $\|\boldsymbol{\upsilon}_{k+1} - \boldsymbol{\upsilon}_k\|$ is bounded. Thus, from the dual update it holds that $\sqrt{\sum_{k\in\mathcal{K}} g_k^2(\mathbf{X}, \boldsymbol{\theta}, P_k, s_k)} = \rho\|\boldsymbol{\upsilon}_{k+1} - \boldsymbol{\upsilon}_k\| \to 0$ as $\rho \to 0$.

D. Complexity Analysis

It is obvious that the complexity of the proposed PDDGP algorithm is dominated by that of Algorithm 1. Thus the complexity of Algorithm 1 is obtained by counting the required number of complex multiplications using the big-O notation. First, the computational complexity associated with $\nabla_{\boldsymbol{\theta}} \hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}, \boldsymbol{\theta}, \mathbf{s})$ in (5) is as follows. The complexity for computing the terms $\mathbf{Z}_{\mathrm{R}}$, $\mathbf{Z}_{\mathrm{R}}\mathbf{X}\mathbf{Z}_{\mathrm{R}}^{\dagger}$, $(\mathbf{I} + \mathbf{Z}_{\mathrm{R}}\mathbf{X}\mathbf{Z}_{\mathrm{R}}^{\dagger})^{-1}\mathbf{Z}_{\mathrm{R}}$ ($\triangleq \boldsymbol{\Xi}$), $\boldsymbol{\Xi}\mathbf{X}$ and $\mathrm{vec_d}(\hat{\mathbf{H}}_{\mathrm{IR}}^{\dagger}\boldsymbol{\Xi}\mathbf{X}\mathbf{H}_{\mathrm{TI}}^{\dagger})$ is respectively $O(N_{\mathrm{R}}N_{\mathrm{I}}N_{\mathrm{T}})$, $O(N_{\mathrm{R}}N_{\mathrm{T}}^2 + N_{\mathrm{R}}^2N_{\mathrm{T}})$, $O(N_{\mathrm{R}}^3 + N_{\mathrm{R}}^2N_{\mathrm{T}})$, $O(N_{\mathrm{R}}N_{\mathrm{T}}^2)$ and $O(N_{\mathrm{I}}N_{\mathrm{T}}N_{\mathrm{R}})$. Similarly, the complexity for obtaining $\mathrm{Tr}(\mathbf{Z}_k\mathbf{X}\mathbf{Z}_k^{\dagger})$ and $\mathrm{vec_d}(\mathbf{H}_{\mathrm{I}\,k}^{\dagger}\mathbf{Z}_k\mathbf{X}\mathbf{H}_{\mathrm{TI}}^{\dagger})$ is respectively $O(N_{\mathrm{P}}N_{\mathrm{T}}(N_{\mathrm{I}} + N_{\mathrm{T}}))$ and $O(N_{\mathrm{T}}^2N_{\mathrm{I}} + N_{\mathrm{P}}N_{\mathrm{T}}N_{\mathrm{I}})$. Therefore, the total complexity for calculating $\nabla_{\boldsymbol{\theta}} \hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}, \boldsymbol{\theta}, \mathbf{s})$ is $O(2N_{\mathrm{R}}N_{\mathrm{I}}N_{\mathrm{T}} + 2N_{\mathrm{R}}N_{\mathrm{T}}^2 + 2N_{\mathrm{R}}^2N_{\mathrm{T}} + N_{\mathrm{R}}^3 + (K+1)N_{\mathrm{P}}N_{\mathrm{T}}N_{\mathrm{I}} + KN_{\mathrm{T}}^2N_{\mathrm{I}} + KN_{\mathrm{P}}N_{\mathrm{T}}^2)$. We remark that, since the projection onto ϑ has negligible complexity, the additional complexity to find an appropriate value of μ_n using the backtracking line search in (8) can be ignored.

Next, the complexity to compute $\nabla_{\mathbf{X}} \hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}, \boldsymbol{\theta}, \mathbf{s})$ is $O(N_{\mathrm{T}}^2(N_{\mathrm{R}} + KN_{\mathrm{P}}))$. Additionally, the complexity for projecting $\mathbf{X}_{n-1} + \alpha_{n-1}\nabla_{\mathbf{X}} \hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}_{n-1}, \boldsymbol{\theta}_n, \mathbf{s}_{n-1})$ onto the feasible set $\mathcal{X}$ is $O(N_{\mathrm{T}}^3)$. Therefore, the computational complexity for the projected gradient step for $\mathbf{X}$ is $O(N_{\mathrm{R}}N_{\mathrm{T}}^2 + KN_{\mathrm{P}}N_{\mathrm{T}}^2 + N_{\mathrm{T}}^3)$.

Moreover, it is straightforward to observe that computing $\mathbf{s}_n$ requires $O(K(N_{\mathrm{P}}N_{\mathrm{I}}N_{\mathrm{T}} + N_{\mathrm{P}}N_{\mathrm{T}}^2))$ complex multiplications. Therefore, the per-iteration complexity of Algorithm 1 is $O(N_{\mathrm{R}}^3 + N_{\mathrm{T}}^3 + 2N_{\mathrm{R}}^2N_{\mathrm{T}} + N_{\mathrm{T}}^2(3N_{\mathrm{R}} + KN_{\mathrm{I}} + 3KN_{\mathrm{P}}) + N_{\mathrm{T}}N_{\mathrm{I}}(2N_{\mathrm{R}} + 2KN_{\mathrm{P}}))$.

It is noteworthy that in a practical IRS-assisted MIMO underlay spectrum sharing system, $N_{\mathrm{I}} \gg \max\{N_{\mathrm{T}}, N_{\mathrm{R}}, N_{\mathrm{P}}\}$, and therefore, the per-iteration complexity of Algorithm 1 can be approximated by $O(N_{\mathrm{I}}(KN_{\mathrm{T}}^2 + 2N_{\mathrm{R}}N_{\mathrm{T}} + 2KN_{\mathrm{P}}))$, which is linear w.r.t. N_{I}. Note that the complexity of the algorithms proposed in both [9] and [10] grows as $O(N_{\mathrm{I}}^3 + KN_{\mathrm{I}}^2\max\{N_{\mathrm{P}}, N_{\mathrm{T}}\})$ which is significantly higher than that of our proposed method.

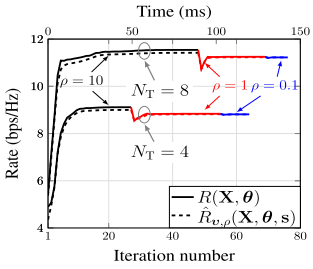

Fig. 2. Convergence results for $P_{\max} = 20$ dBm.

IV. Results and Discussion

In this section, the results of numerical experiments are presented to evaluate the performance of the system under consideration. We consider $K = 4$ PRs (unless stated otherwise), and the location of the ST, SR, IRS and the four PRs in Cartesian coordinates are respectively given by (300 m,0 m), (600 m, 0 m), (300 m, 30 m) and {(0 m, 0 m), (0 m, 5 m), (0 m, 10 m), (0 m, 15 m)}. The small-scale fading is assumed to be Rayleigh distributed, and that the path loss between two nodes is modeled as $\mathrm{PL} = (-30 - 10\xi\log_{10}(d/d_0))$ dB, where ξ is the path loss exponent, d is the distance between the nodes, and d_0 (= 1 m) is the reference distance. We assume that the path loss exponent for the ST-SR and ST-PR links is 3.75, whereas that for the ST-IRS, IRS-SR and IRS-PR links is 2.2 [9]. Furthermore, it is assumed that $N_{\mathrm{R}} = 4$, $N_{\mathrm{P}} = 4$, $P_k = 10^{-13}$ W ($\forall k \in \mathcal{K}$), the noise power spectral density $\mathcal{N}_0 = -174$ dBm/Hz and the system bandwidth is $B = 10$ MHz. In Figs. 3-5, the average achievable rate is computed over 100 channel realizations. To run Algorithm 2, we start with $\rho = 10$ and decrease it as $\rho \leftarrow \kappa\rho$ where $\kappa = 0.1$ when the relative objective progress of $\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}, \boldsymbol{\theta}, \mathbf{s})$ [14, Eqn. (48)] in Algorithm 1 is less than 10^{-5}. For Algorithm 1 we set $\gamma = 0.5$ in the backtracking line search. Algorithm 2 terminates when the difference between $R(\mathbf{X}, \boldsymbol{\theta})$ and $\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}, \boldsymbol{\theta}, \mathbf{s})$ is less than 10^{-5}.

In Fig. 2, the convergence of Algorithms 1 and 2 are shown for $N_{\mathrm{I}} = 64$ and different values of N_{T}. In the figure, each iteration is comprised of one update of $\mathbf{X}$, $\boldsymbol{\theta}$ and $\mathbf{s}$. Each color in Fig. 2 in fact represents the convergence of Algorithm 1 for a fixed ρ whose value is clearly indicated in the figure. We recall that Algorithm 1 aims to maximize the augmented objective $\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}, \boldsymbol{\theta}, \mathbf{s})$, and thus, for a given ρ, monotonic increase is only guaranteed for $\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}, \boldsymbol{\theta}, \mathbf{s})$, but not for the original objective $R(\mathbf{X}, \boldsymbol{\theta})$. This point is clearly seen in Fig. 2. We also observe that as the number of iterations increases, the difference between $R(\mathbf{X}, \boldsymbol{\theta})$ and the augmented objective $\hat{R}_{\boldsymbol{\upsilon},\rho}(\mathbf{X}, \boldsymbol{\theta}, \mathbf{s})$ decreases and approaches zero for sufficiently small ρ, in accordance with the principle of a penalty method. In the figure, we also show the run time for Algorithm 2, where the algorithm converges before 150 ms for $N_{\mathrm{T}} = 8$. Note that we have implemented the algorithm using Python v3.9.7 on a Linux PC with 7.5 GiB memory and Intel Core i5-7200U CPU.

In Fig. 3, we demonstrate the superiority of the proposed PDDGP algorithm over the BCD-IAM and WMMSE approaches proposed in [9] and [10], respectively, for $N_{\mathrm{I}} = 64$. Note that here only the first PR located at (0 m, 0 m) is considered since the algorithm in [9] was proposed for a single PR. It is evident from the figure that all the schemes result in a similar performance for small values of $P_{\max}$; however, the proposed PDDGP algorithm offers a significantly higher

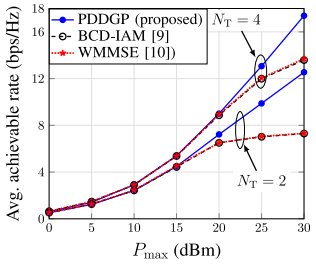

Fig. 3. Average achievable rate for a single PR.

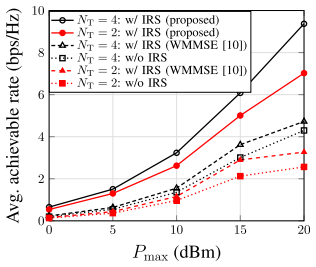

Fig. 4. Average achievable sum rate versus $P_{\max}$.

average achievable rate for large values of $P_{\max}$ compared to that achieved via the solutions proposed in [9] and [10]. Also, it is expected that the SU rate performance achieved via the BCD-IAM and WMMSE schemes is the same because both of these solutions are derived from similar principles. The main reason for the inferior performance of the BCD-IAM and WMMSE for large $P_{\max}$ is that as $P_{\max}$ increases, the SU rate performance is mostly dominated by the IPCs at the PRs. In other words, the IPCs become binding for large $P_{\max}$. Consequently, the coupling between $\mathbf{X}$ and $\boldsymbol{\theta}$ becomes stronger, and it is known that AO-based optimization cannot guarantee a stationary solution under such a scenario.

Next, in Fig. 4, the average achievable sum rate versus the maximum transmit power budget at the ST ($P_{\max}$) is shown for $N_{\mathrm{I}} = 64$. In particular, we compare the proposed PDDGP algorithm to the WMMSE-based AO algorithm introduced in [10] for the case of two SRs located at (600 m, 0 m) and (600 m, 5 m). In Fig. 4, it can be clearly seen that the proposed PDDGP always outperforms the WMMSE-based algorithm. Note that the BCD-IAM and the WMMSE-based AO are similar in principle and thus share the same drawbacks. This explains why the WMMSE-based AO in [10] is inferior to our proposed algorithm. However, note that different from the case of a single PR in Fig. 3, the PDDGP algorithm outperforms the WMMSE algorithm even in the lower $P_{\max}$ regime for the multiple PRs case, and the gains are more significant for larger $P_{\max}$. The benefit of an IRS-assisted MIMO system over that of the system without IRS can also be observed from Fig. 4, as the former results in a considerably higher achievable rate.

To investigate the effect of the number of reflecting elements at the IRS, in Fig. 5 we plot the average achievable rate achieved by the proposed PDDGP algorithm for one SR as a function of N_{I} for $N_{\mathrm{T}} = 16$, $N_{\mathrm{R}} = 4$, $N_{\mathrm{P}} = 2$ and $K = 4$. Interestingly, for an exponential increase in N_{I}, the achievable rate of the MIMO system also increases near-exponentially. This occurs because for large values of N_{I}, the IRS can create a highly-focused beam to increase the signal-to-noise ratio (SNR) at the SR while satisfying the interference constraints at the PRs, resulting in a higher achievable rate at the SR.

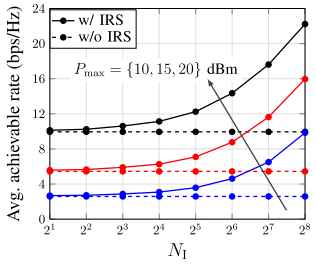

Fig. 5. Average achievable rate versus N_{I}.

V. Conclusion

In this letter, we have presented the optimization problem for achievable rate maximization in an IRS-assisted SU MIMO system in underlay spectrum sharing with multiple PRs, subject to a transmit power constraint at the ST, unit modulus constraints at the IRS and multiple interference power constraints at the PRs. A PDDGP algorithm has been proposed to jointly optimize the transmit beamforming and the IRS phase shifts. The numerical results established the superiority of the proposed algorithm over two existing approaches in terms of the achievable rate. We also showed that the complexity of the proposed PDDGP algorithm is considerably lower than that of the existing approaches.

References

[1] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.

[2] P. Yang, L. Kong, and G. Chen, "Spectrum sharing for 5G/6G URLLC: Research frontiers and standards," *IEEE Commun. Stand. Mag.*, vol. 5, no. 2, pp. 120–125, Jun. 2021.

[3] V. Kumar, Z. Ding, and M. F. Flanagan, "On the performance of downlink NOMA in underlay spectrum sharing," *IEEE Trans. Veh. Technol.*, vol. 70, no. 5, pp. 4523–4540, May 2021.

[4] X. Guan, Q. Wu, and R. Zhang, "Joint power control and passive beamforming in IRS-assisted spectrum sharing," *IEEE Commun. Lett.*, vol. 24, no. 7, pp. 1553–1557, Jul. 2020.

[5] D. Xu, X. Yu, and R. Schober, "Resource allocation for intelligent reflecting surface-assisted cognitive radio networks," in *Proc. IEEE SPAWC*, 2020, pp. 1–5.

[6] J. Yuan, Y.-C. Liang, J. Joung, G. Feng, and E. G. Larsson, "Intelligent reflecting surface-assisted cognitive radio system," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 675–687, Jan. 2021.

[7] L. Zhang, C. Pan, Y. Wang, H. Ren, and K. Wang, "Robust beamforming design for intelligent reflecting surface aided cognitive radio systems with imperfect cascaded CSI," *IEEE Trans. Cogn. Commun. Netw.*, vol. 8, no. 1, pp. 186–201, Mar. 2022.

[8] S. F. Zamanian, S. M. Razavizadeh, and Q. Wu, "Vertical beamforming in intelligent reflecting surface-aided cognitive radio networks," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 1919–1923, Sep. 2021.

[9] L. Zhang, Y. Wang, W. Tao, Z. Jia, T. Song, and C. Pan, "Intelligent reflecting surface aided MIMO cognitive radio systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11445–11457, Oct. 2020.

[10] W. Jiang, Y. Zhang, J. Zhao, Z. Xiong, and Z. Ding, "Joint transmit precoding and reflect beamforming design for IRS-assisted MIMO cognitive radio systems," *IEEE Trans. Wireless Commun.*, early access, Nov. 2, 2021, doi: 10.1109/TWC.2021.3122959.

[11] L. Dong, H.-M. Wang, H. Xiao, and J. Bai, "Secure intelligent reflecting surface assisted MIMO cognitive radio transmission," in *Proc. IEEE WCNC*, 2021, pp. 1–6.

[12] V. Kumar, M. F. Flanagan, D. B. da Costa, and L.-N. Tran, "On the secrecy rate of downlink NOMA in underlay spectrum sharing with imperfect CSI," in *Proc. 28th Int. Conf. Telecommun. (ICT)*, 2021, pp. 167–173.

[13] N. S. Perović, L.-N. Tran, M. Di Renzo, and M. F. Flanagan, "Achievable rate optimization for MIMO systems with reconfigurable intelligent surfaces," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3865–3882, Jun. 2021.

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

[15] A. Hjørungnes, *Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.