Active and passive beamforming based on RWMMSE and Gradient Projection for IRS-aided MU-MISO Systems

Seraphin Kimaryo
Department of Electrical and Information Engineering
Seoul National University of Science and Technology
Seoul, Republic of Korea
email:seraphin@seoultech.ac.kr

Eduard E. Bahingayi
Biomedical Engineering Unit
Muhimbili University of Health and Allied Sciences
Dar es Salaam, Tanzania
email:eduard.bahingayi@muhas.ac.tz

Kyungchun Lee
Department of Electrical and Information Engineering and the Research Center for Electrical and Information Technology
Seoul National University of Science and Technology
Seoul, Republic of Korea
email:kclee@seoultech.ac.kr

***Abstract*—The intelligent reflecting surface (IRS), a planar array consisting of many low-cost configurable reflecting elements, offers many new possibilities, such as creating a virtual line of sight, interference cancellation, wireless power transfer, etc. To efficiently achieve these goals, the real-time reconfiguration of its elements, which is still challenging, needs to be as simple as possible. Prior works for downlink IRS-aided multi-user multiple-input single-output (MISO) systems, which are noted to have high complexity, are extended in this work by incorporating them with novel techniques that lower their complexity. In doing so, we create a weighted sum rate problem and adopt a reduced weighted minimum mean square error (RWMMSE) algorithm, which is an advanced low-complexity version of the weighted minimum mean square error (WMMSE) scheme to optimize the active precoders, whereas a gradient projection (GP) algorithm, whose procedure of choosing the step size is redesigned to avoid the high-complexity iterative processes, is adopted for optimization of passive beamformers. Finally, the presented numerical results demonstrate the effectiveness of our proposed scheme in attaining comparable performance to the benchmark schemes with a significant complexity reduction.**

***Index Terms*—Intelligent reflecting surface (IRS), multiple-input single-output (MISO), passive beamforming, active beamforming, reduced WMMSE (RWMMSE).**

I. Introduction

THE WIRELESS channel has been used for years as a medium for transmitting and receiving electromagnetic waves. Apart from its advantages and flexibilities over its counterpart, the wired medium, it also has several issues, such as fading, scattering, and shadowing [1], [2], which make transmitting and receiving the signals challenging. Moreover, its free-propagation nature not only makes the transmitted information vulnerable to eavesdropping, hence raising some security issues [2], [3], but also allows the transmitted signals to add together either constructively or destructively, thus making interference another bottleneck even within the same network operator.

Multiple-input multiple-output (MIMO) [4], [5], cell-free massive MIMO [6], [7], and ultra-dense networks (UDN) [8], [9] are among the few solutions that have been proposed to tackle the aforementioned issues. Although most of these solutions provide great relief to the challenges above, they are only channel-adaptive techniques, i.e., they work based on the nature of the current channel conditions.

In an effort to gain more control over the conditions of the wireless channel, scholars and industry developed artificial metamaterials [10], whose primary characteristic is the ability to be reconfigured in real-time. This property, when exploited by several metamaterial-derived technologies such as intelligent transmitting surfaces (ITS) [11], intelligent reflecting surfaces (IRS) [10], and dynamic metasurface antennas (DMA) [12], [13], can offer great flexibility on how to deal with various wireless channel conditions.

Interestingly, these technologies can be amalgamated with several others to create various combinations [11], [13], [14] that best serve a specific desired goal. Although several combinations exist, the most popular one is between IRS and systems with multiple antennas because of their various use cases, like improving spectral efficiency, facilitating interference cancellation, supporting secure communication [15], and assisting in simultaneous wireless information and power transfer [16]. Specifically, this work focuses on the IRS-aided downlink multi-user (MU) multiple-input single-output (MISO) system.

In a nutshell, IRS is a planar surface consisting of a large number of nearly passive reflecting metamaterial elements, which are usually sub-wavelength-spaced. In its common usage, it is normally deployed by being plated on the surfaces

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2022R1A2C1006566)

of user-surrounding structures like building facades, ceilings, furniture, etc., and near the receiver to facilitate the transmission and reception of the signals in a controlled manner [14]. When used with MIMO, this system presents a highly non-convex problem, i.e., the joint optimization of active and passive beamformers, which needs to be efficiently addressed in order to achieve its full potential. Some previous works like [17], [18] proposed notable techniques for solving this problem; however, the mechanics of most of these techniques accumulate a lot of computation complexity along the way.

The work [17], for example, solved the weighted sum rate (WSR) maximization problem for the IRS-aided MU downlink MISO system by alternatively deploying a weighted minimum mean square error (WMMSE) [19] algorithm to find the active precoders at the base station (BS) and either a Manifold optimization (MO) or block coordinate descent (BCD) method to solve for the IRS' passive beamformer. Although it is appreciatively noted that these techniques jointly offer remarkable performance in terms of the achieved WSR, their accrued complexity, which unfortunately scales with the number of BS transmit antennas, is prohibitively high.

To reduce complexity, the work [20] adopted gradient projection (GP) and cross-entropy (CE) to solve a similar IRS phase shift optimization problem for downlink MISO systems but with a different objective, i.e., spectral efficiency (SE) maximization of the user. While GP is essentially a gradient descent scheme, which needs, at every point, to decide the direction and step size to descend towards a global minimum, CE resembles a genetic ant colony algorithm that depends on many parameters to kickstart and continue with its optimization process. Although the internal structure of GP is simpler compared to that of CE, both approaches still have high complexity for practical systems.

Contributions: Due to the high complexity of these existing schemes, in this paper, we extend the prior efforts of using WMMSE and GP to alternately optimize the BS' precoders (active beamformers) and IRS' phase shifts (passive beamformers) of the IRS-aided MU downlink MISO system. Specifically, we propose using the RWMSSE algorithm [21] for optimizing active beamformers and the GP algorithm for optimizing passive beamformers.

Notations: Throughout this paper, $\mathbb{C}$ denotes the complex domain. A bold-face capital letter, $\boldsymbol{X}$, is used to denote a matrix, a bold-face lower-case letter, $\boldsymbol{x}$, denotes a vector, and a lower-case italic letter, x, denotes a scalar. The superscripts T, H, and $*$ denote the transpose, conjugate transpose, and the complex conjugate, respectively. We use $\boldsymbol{X}^{-1}$ and $\mathrm{Tr}\left(\boldsymbol{X}\right)$ to denote the inverse and trace of the matrix $\boldsymbol{X}$, respectively. The magnitude of the complex number x is denoted by $|x|$, whereas its argument is denoted as $\arg(x)$. For any vector $\boldsymbol{x}$, $[\boldsymbol{x}]_i$ is the i-th element of $\boldsymbol{x}$; $|\boldsymbol{x}|$ and $\|\boldsymbol{x}\|$ denote its absolute value and Euclidean norm, respectively. $\mathrm{diag}(a_1, \cdots, a_N)$ is an $N \times N$ diagonal matrix containing $a_1, \cdots, a_N$ as its diagonal elements.

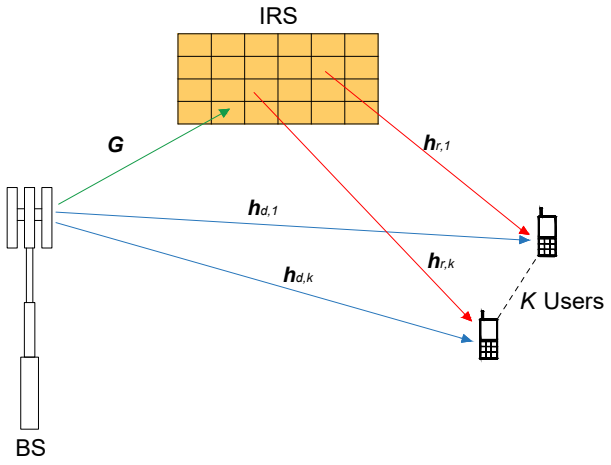

Fig. 1: System model

II. System Model and Problem Formulation

Fig. 1 shows a communication system considered in this work, where a BS equipped with M transmit antennas transmits K streams of data signals with transmit power P to K single-antenna users. This transmission is assisted by an IRS equipped with N reflecting elements arranged on a planar surface. The signal reaches the users through a set of baseband channels, which are all assumed to be flat-fading and known to the BS. Particularly, we denote the link between BS and IRS as $\boldsymbol{G} \in \mathbb{C}^{N\times M}$, the reflected link between IRS and user k as $\boldsymbol{h}_{r,k} \in \mathbb{C}^{N\times 1}$, and the direct link between BS and user k as $\boldsymbol{h}_{d,k} \in \mathbb{C}^{M\times 1}$. Moreover, the reflection coefficients of all IRS elements are arranged in a diagonal matrix $\boldsymbol{\Theta} = \mathrm{diag}\left(e^{j\theta_1}, e^{j\theta_2}, \cdots, e^{j\theta_N}\right) \in \mathbb{C}^{N\times N}$, where θ_i is the phase shift of the i-th element. Prior to transmission, all the unit-power symbols $s_1, s_2, \cdots, s_K$, where s_i is the symbol intended for user i, are precoded with their respective precoders $(\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_K) \in \mathbb{C}^{M\times 1}$, and added together to form a transmitted signal:

$$\boldsymbol{d} = \sum_{k=1}^{K} \boldsymbol{f}_k s_k. \tag{1}$$

Then, the received signal y_k at user k is given by

$$y_k = \left(\boldsymbol{h}_{d,k}^T + \boldsymbol{h}_{r,k}^T \boldsymbol{\Theta} \boldsymbol{G}\right) \boldsymbol{d} + z_k, \tag{2}$$

where z_k is additive complex white Gaussian noise with zero mean and variance σ^2. Let $\boldsymbol{h}_k = \boldsymbol{h}_{d,k}^T + \boldsymbol{\theta}^T \boldsymbol{E}_k$, denote the composite channel comprising of both the direct and reflect links between BS and user k, where $\boldsymbol{\theta} = \left[e^{j\theta_1}, e^{j\theta_2}, \cdots, e^{j\theta_N}\right]^T \in \mathbb{C}^{N\times 1}$ is the vectorized collection of IRS reflection coefficients and $\boldsymbol{E}_k = \mathrm{diag}\left(\boldsymbol{h}_{r,k}\right) \boldsymbol{G} \in \mathbb{C}^{N\times M}$ is used to exchange the shapes and positions of the variables $\boldsymbol{h}_{r,k}$ and $\boldsymbol{\Theta}$. With this transformation, the signal-

to-interference-plus-noise ratio (SINR) for user k can be compactly expressed as

$$SINR_k = \frac{|\boldsymbol{h}_k \boldsymbol{f}_k|^2}{\sum_{i\neq k}^{K} |\boldsymbol{h}_k \boldsymbol{f}_i|^2 + \sigma^2}. \tag{3}$$

Note that the presence of interference in (3) normally complicates the analysis of such systems. Alternatively, the WSR, which for our case is given by

$$R\left(\boldsymbol{F}, \boldsymbol{\theta}\right) = \sum_{k=1}^{K} \omega_k \log_2\left(1 + SINR_k\right), \tag{4}$$

where $\boldsymbol{F} = [\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_K] \in \mathbb{C}^{M\times K}$ and ω_k denotes the priority of user k, has been used successfully to deal with communication systems of this nature. In this work, we aim to find $\boldsymbol{\theta}$ and $\boldsymbol{F}$ that maximize the system's WSR. This problem is given by

$$\begin{aligned} \mathcal{P}(A): \quad \max_{\boldsymbol{F},\boldsymbol{\theta}} \quad & R\left(\boldsymbol{F}, \boldsymbol{\theta}\right) && \text{(5a)}\\ \text{s.t.} \quad & |[\boldsymbol{\theta}]_\ell| = 1 \quad \forall \ell = 1, 2, \cdots, N, && \text{(5b)}\\ & \sum_{k=1}^{K} \|\boldsymbol{f}_k\|^2 \leq P. && \text{(5c)} \end{aligned}$$

We notice that the objective function of $\mathcal{P}(A)$ is non-convex with respect to $\boldsymbol{\theta}$ due to its unit modulus constraint on each element. Therefore, designing a global optimal algorithm to solve $\mathcal{P}(A)$ is challenging. In the next section, we propose an efficient algorithm that generates a locally optimal solution to $\mathcal{P}(A)$.

III. Proposed Solution to Problem $\mathcal{P}(A)$

Apart from the non-convexity of $\mathcal{P}(A)$, the deep coupling of $\boldsymbol{F}$ and $\boldsymbol{\theta}$ makes the problem even harder to solve. Notably, AO is very efficient in solving these kinds of problems by alternatively fixing one parameter and optimizing the other in a repetitive manner until the objective function converges. Therefore, in this work, we adopt the same technique AO, where at each turn, when one optimization variable is fixed, we create and propose a solution to a subproblem of the other variable.

A. Optimization of Active Precoding

When $\boldsymbol{\theta}$ is fixed, a prior work [17] solved the resulting problem $\mathcal{P}(B)$ by using WMMSE technique [19].

$$\begin{aligned} \mathcal{P}(B): \quad \max_{\boldsymbol{F}} \quad & R_1\left(\boldsymbol{F}, \boldsymbol{\theta}\right) && \text{(6a)}\\ \text{s.t.} \quad & \sum_{k=1}^{K} \|\boldsymbol{f}_k\|^2 \leq P. && \text{(6b)} \end{aligned}$$

We note that this approach ends up having very high complexity as it needs to repetitively compute the inverse of a high-dimensional matrix and iteratively find the Lagrangian power constraint. To reduce this complexity, we adopt and present the RWMMSE technique in this subsection. A great detail of RWMMSE is given in [21]; however, for the sake of convenience, we briefly touch on some of its important concepts as we explain the solution to this subproblem. To begin, we define $\boldsymbol{H} = \left[\boldsymbol{h}_1^T, \boldsymbol{h}_2^T, \cdots, \boldsymbol{h}_K^T\right]^T \in \mathbb{C}^{K\times M}$ to be a collection of composite channels from all users to the BS. The key idea of RWMMSE lies in two novel observations of the nontrivial stationary points. The first observation is that any nontrivial stationary point $\boldsymbol{f}_k^\star$, which does not result in a zero WSR, must satisfy $\boldsymbol{f}_k^\star = \boldsymbol{H}^H \boldsymbol{x}_k$ for some $\boldsymbol{x}_k \in \mathbb{C}^{K\times 1}$. This immediately transforms $\mathcal{P}(B)$ to

$$\begin{aligned} \max_{\boldsymbol{X}} \quad & \sum_{k=1}^{K} \omega_k \log_2\left(1 + \frac{|\boldsymbol{h}_k \boldsymbol{H}^H \boldsymbol{x}_k|^2}{\sum_{i\neq k}^{K} |\boldsymbol{h}_k \boldsymbol{H}^H \boldsymbol{x}_i|^2 + \sigma^2}\right) && \text{(7a)}\\ \text{s.t.} \quad & \sum_{k=1}^{K} \|\boldsymbol{H}^H \boldsymbol{x}_k\|^2 \leq P, && \text{(7b)} \end{aligned}$$

where $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_K] \in \mathbb{C}^{K\times K}$. This shows that $\mathcal{P}(B)$ can be solved by equivalently solving (7). Notice that problem (7) has now much-reduced dimensions on the optimizable parameter $\boldsymbol{X}$ than that of $\boldsymbol{F}$. Nevertheless, this problem is still computationally expensive if WMMSE is applied because of the power constraint (7b), which needs an iterative line search process for its computation. This brings us to the second observation, which helps to circumvent this constraint; i.e., any nontrivial stationary point of $\mathcal{P}(B)$ must satisfy the power constraint (6b) with equality. Using this observation, we obtain the following unconstrained problem

$$\max_{\boldsymbol{X}} \quad \sum_{k=1}^{K} \omega_k \log_2\left(1 + \frac{\left|\bar{\boldsymbol{h}}_k \boldsymbol{x}_k\right|^2}{a_k}\right), \tag{8}$$

with the necessary modification of the relationship between $\boldsymbol{f}_k^\star$ and $\boldsymbol{x}_k$ given as $\boldsymbol{f}_k^\star = \sqrt{\beta}\boldsymbol{H}^H \boldsymbol{x}_k$, where $a_k = \sum_{j\neq k}^{K} \left|\bar{\boldsymbol{h}}_k \boldsymbol{x}_j\right|^2 + \frac{\sigma_k^2}{P}\sum_{i=1}^{K} \mathrm{Tr}\left(\bar{\boldsymbol{H}}\boldsymbol{x}_i \boldsymbol{x}_i^H\right)$ is used to simplify the notations, $\bar{\boldsymbol{H}} = \boldsymbol{H}\boldsymbol{H}^H \in \mathbb{C}^{K\times K}$, $\bar{\boldsymbol{h}}_k = \boldsymbol{h}_k \boldsymbol{H}^H \in \mathbb{C}^{1\times K}$, and $\beta = \frac{P}{\sum_{K=1}^{K} \mathrm{Tr}\left(\bar{\boldsymbol{H}}\boldsymbol{x}_k \boldsymbol{x}_k^H\right)}$ is the scaling factor. To solve (8) we adopt its equivalent unconstrained weighted sum-mean square error (MSE) problem $\mathcal{P}(C)$:

$$\begin{aligned} \mathcal{P}(C): \quad \min_{u_k, \chi_k, \boldsymbol{X}} \quad & \sum_{k=1}^{K} \omega_k \Big(\chi_k \left|1 - u_k^* \bar{\boldsymbol{h}}_k \boldsymbol{x}_k\right|^2 \\ & + \chi_k \left|u_k\right|^2 a_k - \log_2 \chi_k\Big), \end{aligned} \tag{9}$$

which is obtained by introducing two auxiliary variables u and χ, whose closed-form solutions obtained by the BCD method are given by

$$u_k = \frac{\bar{\boldsymbol{h}}_k \boldsymbol{x}_k}{a_k + \left|\bar{\boldsymbol{h}}_k \boldsymbol{x}_k\right|^2}. \tag{10}$$

$$\chi_k = \frac{1}{1 - u_k^* \bar{\boldsymbol{h}}_k \boldsymbol{x}_k}. \tag{11}$$

Similarly, the update rule for $\boldsymbol{X}$, which is also in a closed-form and given in (12), is obtained by solving $\mathcal{P}(C)$ when the blocks u and χ are fixed.

$$\boldsymbol{x}_k = \omega_k \chi_k u_k \left(\sum_{i=1}^{K} \frac{\sigma_i^2}{P} \omega_i \chi_i \left| u_i \right|^2 \bar{\boldsymbol{H}} + \sum_{j=1}^{K} \omega_j \chi_j \left| u_j \right|^2 \bar{\boldsymbol{h}}_j^H \bar{\boldsymbol{h}}_j \right)^{-1} \bar{\boldsymbol{h}}_k^H. \quad (12)$$

This process of finding the active beamformers is summarized in Algorithm 1. The algorithm starts by initializing the matrix $\boldsymbol{X}$ such that when applied to the system the total transmit power is not exceeded. Then, the auxiliary variables together with the $\boldsymbol{x}_k, \forall k$ are iteratively updated by their corresponding update rules until the objective function converges, upon which the active beamformers are obtained by $\boldsymbol{f}_k = \sqrt{\beta} \boldsymbol{H}^H \boldsymbol{x}_k$.

Algorithm 1: RWMMSE method to generate $\boldsymbol{F}$

Input : $\boldsymbol{X}$, $\boldsymbol{H}$, and ω_k, $\forall k$
Initialization: Generate $\bar{\boldsymbol{H}} = \boldsymbol{H}\boldsymbol{H}^H$.
Set $\chi_k = 1$
repeat
  $\chi_k^{'} = \chi_k$
  Update u_k, $\forall k$ by (10).
  Update χ_k , $\forall k$ by (11).
  Update x_k , $\forall k$ by (12).
until $\left| \sum_{j=1}^{K} \omega_k \log \chi_k - \sum_{j=1}^{K} \omega_k \log \chi_k^{'} \right| \leq \epsilon_f$
Output: $\boldsymbol{X}$ and $\boldsymbol{f}_k = \sqrt{\beta} \boldsymbol{H}^H \boldsymbol{x}_k, \forall k$ such that $\sum_{k=1}^{K} \| \boldsymbol{f}_k \|^2 \leq P$

B. Passive precoding

Next, we fix the active precoder $\boldsymbol{F}$ and optimize the passive precoder $\boldsymbol{\theta}$. The fixation of $\boldsymbol{F}$ transforms $\mathcal{P}(A)$ into

$$\mathcal{P}(D): \quad \max_{\boldsymbol{\theta}} \quad f(\boldsymbol{\theta}) \quad (13)$$

$$\text{s.t.} \quad |[\boldsymbol{\theta}]_\ell| = 1 \quad \forall \ell = 1, 2, \cdots, N, \quad (14)$$

where

$$f(\boldsymbol{\theta}) = \sum_{k=1}^{K} \omega_k \log_2 \left(1 + \frac{\left| \boldsymbol{\theta}^H \boldsymbol{e}_{k,k}^* + r_{k,k}^* \right|^2}{\sum_{i \neq k}^{K} \left| \boldsymbol{\theta}^H \boldsymbol{e}_{k,i}^* + r_{k,i}^* \right|^2 + \sigma^2} \right) \quad (15)$$

is the single-variable objective function with $\boldsymbol{e}_{k,i} = \boldsymbol{E}_k \boldsymbol{f}_i$ and $r_{k,i} = \boldsymbol{h}_{d,k}^T \boldsymbol{f}_i$. It is noted that $\mathcal{P}(D)$ is continuous and differentiable, thus it has attracted several gradient descent-based approaches for its optimization. In this study, we extend GP to the MU system and enhance it with robust procedures for choices of step sizes at each iteration.

The GP is basically a gradient descent method, in which at every given point $\boldsymbol{\theta}^i$ of iteration i, a search direction $\boldsymbol{\rho}^i$ is computed, and then we decide how far to move along that direction to obtain the next point $\boldsymbol{\theta}^{i+1}$. This can be mathematically represented as

$$\boldsymbol{\theta}^{i+1} = \boldsymbol{\theta}^i + \alpha^i \boldsymbol{\rho}^i,$$

where $\alpha^i > 0$ is called the step-size. While the search direction at each point is easily obtained by the negative of the gradient of the objective function at that particular point, which in our case is given by $\boldsymbol{\rho}^i = -\nabla f(\boldsymbol{\theta}^i)$[1] $= -2 \sum_{k=1}^{K} \omega_k g_k(\boldsymbol{\theta}^i)$ with

$$g_k(\boldsymbol{\theta}) = \frac{\sum_{i=1}^{K} \boldsymbol{e}_{k,i}^* \boldsymbol{e}_{k,i}^T \boldsymbol{\theta} + \sum_{i=1}^{K} \boldsymbol{e}_{k,i}^* r_{k,i}}{\sum_{i=1}^{K} \left| \boldsymbol{\theta}^H \boldsymbol{e}_{k,i}^* + r_{k,i}^* \right|^2 + \sigma^2} - \frac{\sum_{i \neq k}^{K} \boldsymbol{e}_{k,i}^* \boldsymbol{e}_{k,i}^T \boldsymbol{\theta} + \sum_{i \neq k}^{K} \boldsymbol{e}_{k,i}^* r_{k,i}}{\sum_{i \neq k}^{K} \left| \boldsymbol{\theta}^H \boldsymbol{e}_{k,i}^* + r_{k,i}^* \right|^2 + \sigma^2}$$

being the Euclidean gradient; the search of optimal α_i is usually very challenging. The work [20] proposes a procedure that involves the computation of the maximum eigenvalue λ_{max} of a high-order matrix $\boldsymbol{E}_k \boldsymbol{E}_k^H \in \mathbb{C}^{N \times N}$, which apart from being computationally expensive on its own, the extension to the MU system is even harder as it needs to compute λ_{max} for each user. To avoid this problem we adopt a two-point step size method, which is given as follows [22]

$$\alpha_i = \begin{cases} \frac{1}{c \| \nabla f(\boldsymbol{\theta}^i) \|}, & \text{if } i = 0 \\ \frac{\boldsymbol{\upsilon}^{i H} \boldsymbol{\xi}^i}{\| \boldsymbol{\xi}^i \|}, & \text{otherwise} \end{cases} \quad (16)$$

where $\boldsymbol{\xi}^i = \nabla f(\boldsymbol{\theta}^i) - \nabla f(\boldsymbol{\theta}^{i-1})$, $\boldsymbol{\upsilon}^i = \boldsymbol{\theta}^i - \boldsymbol{\theta}^{i-1}$, and $c \in (0, 1)$ is a scaling constant. The proposed GP method to solve $\mathcal{P}(D)$ is summarized in Algorithm 2.

Algorithm 2: GP-based method for solving $\mathcal{P}(D)$

Input : $\boldsymbol{\theta}$, $\boldsymbol{F}$, $\boldsymbol{E}_k$, ω_k, $\forall k$
Set $i = 0$ and $\boldsymbol{\theta}^i = \boldsymbol{\theta}$.
repeat
  Compute $\nabla f(\boldsymbol{\theta}^i)$.
  Generate α^i by using (16).
  $\boldsymbol{\varpi} \leftarrow \boldsymbol{\theta}^i - \alpha^i \nabla f(\boldsymbol{\theta}^i)$
  $\boldsymbol{\theta}^{i+1} \leftarrow e^{j \arg(\boldsymbol{\varpi})}$
  $i \leftarrow i + 1$
until $\left| f(\boldsymbol{\theta}^i) - f(\boldsymbol{\theta}^{i-1}) \right| \leq \epsilon_\theta$
Output: $\boldsymbol{\theta}$

The complete algorithm, which is henceforth referred to as "RWMMSE-GP", is presented in Algorithm 3.

IV. Simulation Results

In this section, we present the numerical results of our proposed algorithm. We compare its effectiveness with the benchmark schemes BCD and WMMSE-MO [17]. It should be noted that contrary to RWMMSE-GP, which uses RWMMSE

[1]Note that the gradient is computed using the natural logarithm version of (15) because of its analytical simplicity and the independency of the solution form the base of the logarithm.

Algorithm 3: Proposed RWMMSE-GP scheme

Input : $\boldsymbol{G}$, $\boldsymbol{h}_{d,k}$, $\boldsymbol{h}_{r,k}$, ω_k, $\forall k \in [1, 2, \cdots, K]$

Initialization: Randomly generate $\boldsymbol{\theta}$ and $\boldsymbol{X}$ such that $\sum_{j=1}^{K} \mathrm{Tr}(\bar{\boldsymbol{H}} \boldsymbol{x}_j \boldsymbol{x}_j^H) \leq P$.

Set $j = 0$.

Calculate WSR, C^j.

repeat

  Update $\boldsymbol{\theta}$ by Algorithm 2.

  Update $\boldsymbol{F}$ by Algorithm 1.

  $j \leftarrow j + 1$

  Calculate WSR, C^j.

until $\left| C^j - C^{j-1} \right| \leq \epsilon_C$

Output: $\boldsymbol{\theta}$, $\boldsymbol{F}$

and GP schemes to optimize the active and passive beamformers, the benchmark scheme WMMSE-MO uses the WMMSE and MO algorithms for the optimization of active and passive beamformers, respectively. We also include the performance of the random phase scheme, which randomly chooses the phases of the IRS, and the unassisted system (without IRS) to act as the performance lower bounds.

A. Simulation Settings

We adopt nearly similar simulation environments to those adopted in [17]. Unless otherwise stated, we consider a system in which a BS equipped with $M = 8$ antennas and located at $(0, 0)$ is used to serve $K = 4$ single-antenna users, which are located at $(205.65\text{m}, 34.48\text{m})$, $(193.47\text{m}, 30.24\text{m})$, $(198.30\text{m}, 22.40\text{m})$, and $(206.00\text{m}, 24.28\text{m})$. This communication is assisted by an IRS located at $(200\ \text{m}, 0)$. The channels $\boldsymbol{G}$, $\boldsymbol{h}_{d,k}$, and $\boldsymbol{h}_{r,k}$ are generated using mmWave clustered channel model [23] and then multiplied by the square root of the large-scale fading which is given by

$$PL\left(\partial, \eta\right)[\text{dB}] = \begin{cases} 35.6 + 10\eta \log\left(\partial\right), & \text{LoS case} \\ 32.6 + 10\eta \log\left(\partial\right), & \text{NLoS case,} \end{cases} \tag{17}$$

where ∂ is the distance in meters between BS and the user, and η denotes the path-loss exponent, which is given as $\eta = 2.2$ and $\eta = 3.67$ for LoS and NLoS cases, respectively. The inverse of the path-loss of the link between BS and each user is used as the priority of the corresponding user. These priorities are further normalized such that $\sum_{k=1}^{K} \omega_k = 1$. Finally, the transmit power at the BS is set to 20 dBm, the noise power at the user's receiver is set to $\sigma^2 = -170$ dBm/Hz, and the transmission bandwidth is set to 180 kHz. Control parameters of the proposed algorithm are set as follows $\epsilon_\theta = \epsilon_C = \epsilon_f = 10^{-5}$ and $c = \frac{1}{6}$.

B. WSR Versus the Number IRS elements:

We start by presenting in Fig. 2 the achievable WSR for various sizes of IRS. It is quickly noted that the proposed scheme, BCD, and WMMSE-MO all exhibit the expected behavior of performance improvement as the size of the IRS increases. In addition, it is also seen that the performance

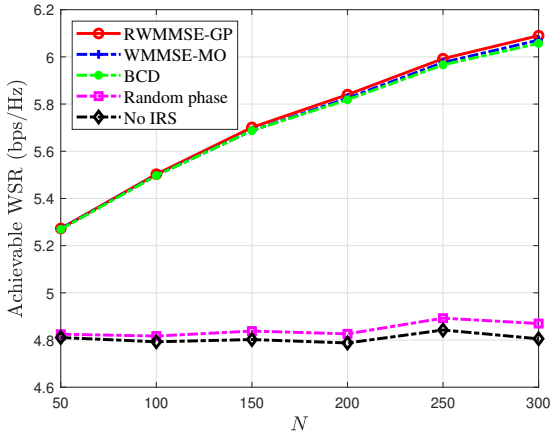

Fig. 2: WSR vs. N for $K = 4$, $M = 8$, and $P = 20$ dBm.

of the proposed scheme is almost the same as that of the benchmark schemes. As the size of the IRS increases, our proposed scheme performs marginally better than the benchmark schemes. This clearly demonstrates the effectiveness of our proposed scheme in attaining better performance with reduced complexity. Furthermore, we note from Fig. 2 that the performance of the randomly configured IRS-aided system is almost the same as that of the unassisted system. This observation highlights the necessity of a proper IRS configuration to reap its benefits fully.

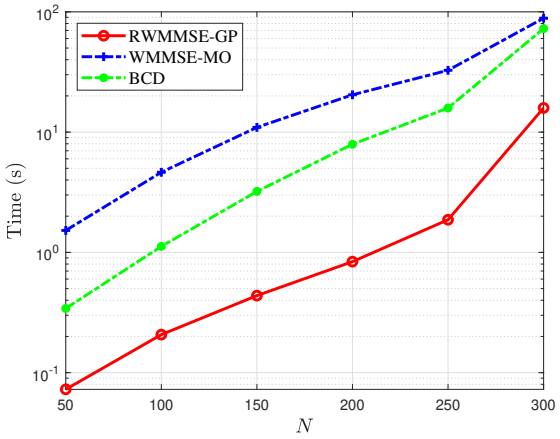

Fig. 3: Runtime vs. N for $K = 4$, $M = 8$, and $P = 20$ dBm.

C. Complexity comparison:

Next, we examine the accrued complexity of both the proposed and the benchmark schemes to produce the results in Fig. 2. This complexity, which is presented in Fig. 3 in terms of the running time versus the number of IRS elements, is obtained from a MATLAB R2023a running on a Windows 10 computer having a 12th Gen Intel(R) Core(TM) i7-12700 CPU @ 2.10 GHz processor and 32 GB of RAM. As expected, the running times of all schemes are observed to increase as

the size of the IRS increases. This is due to the increased computational processing in configuring the phase of each IRS element. Apart from that, we note from Fig. 3 that the proposed scheme attains the shortest running time for all sizes of the IRS. Quantitatively, note that when $N = 50$, the running times of the proposed scheme and that of BCD are 0.07 seconds and 0.34 seconds, respectively. Moreover, as N increases to 300, the corresponding running times of the proposed scheme and the BCD algorithm are 15.89 seconds and 72.53 seconds, respectively. This reduction in running time is attributed to the faster convergence of the proposed scheme.

V. Conclusion

This work investigated one potential way of reducing the complexity for optimization of the IRS phase shifts (passive beamformer) and transmit precoders (active beamformers) of the IRS-aided MISO systems. We formulated the WSR problem and adopted an AO scheme to decouple the optimizable parameters, i.e., passive and active beamformers, which are, after that, alternately optimized by complexity-efficient techniques. Specifically, the RWMMSE algorithm that reduces the size of the invertible matrices, avoids iterative processes, and circumvents the Lagrangian power constraint is adopted for active beamformer optimization. On the other hand, the passive beamformer is optimized by a GP algorithm that is enhanced with a BB step-size selection. Simulation results demonstrated the effectiveness of the proposed algorithm in attaining nearly the same performance as the benchmark scheme with a significant complexity reduction.

References

[1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[2] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[3] H.-M. Wang, T.-X. Zheng, J. Yuan, D. Towsley, and M. H. Lee, "Physical layer security in heterogeneous cellular networks," *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 1204–1219, 2016.

[4] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.

[5] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.

[6] H. A. Ammar, R. Adve, S. Shahbazpanahi, G. Boudreau, and K. V. Srinivas, "User-centric cell-free massive MIMO networks: A survey of opportunities, challenges and solutions," *IEEE Commun. Surv. Tutorials*, vol. 24, no. 1, pp. 611–652, 2022.

[7] G. Femenias and F. Riera-Palou, "Wideband cell-free mmWave massive MIMO-OFDM: Beam squint-aware channel covariance-based hybrid beamforming," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 4695–4710, 2022.

[8] H. Zhang, J. Lee, Q. T. Q. S., and C.-L. I., *Ultra-dense networks: Principles and applications*. Cambridge University Press, 2020.

[9] L.-P. David and M. Ding, *Fundamentals of ultra-dense Wireless Networks*. Cambridge University Press, 2022.

[10] 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, 2019.

[11] V. Jamali, A. M. Tulino, G. Fischer, R. R. Müller, and R. Schober, "Intelligent surface-aided transmitter architectures for millimeter-wave ultra massive MIMO systems," *IEEE open j. Commun. Soc.*, vol. 2, pp. 144–167, 2021.

[12] S. F. Kimaryo and K. Lee, "Downlink beamforming for dynamic metasurface antennas," *IEEE Trans. Wireless Commun.*, vol. 22, no. 7, pp. 4745–4755, 2023.

[13] N. Shlezinger, G. C. Alexandropoulos, M. F. Imani, Y. C. Eldar, and D. R. Smith, "Dynamic metasurface antennas for 6G extreme massive MIMO communications," *IEEE Wireless Commun.*, vol. 28, no. 2, pp. 106–113, 2021.

[14] 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, 2021.

[15] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1410–1414, 2019.

[16] C. Pan, H. Ren, K. Wang, *et al.*, "Intelligent reflecting surface aided MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1719–1734, 2020.

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

[18] W. Chen, X. Ma, Z. Li, and N. Kuang, "Sum-rate maximization for intelligent reflecting surface based terahertz communication systems," in *2019 IEEE/CIC International Conference on Communications Workshops in China (ICCC Workshops)*, 2019, pp. 153–157.

[19] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, 2011.

[20] J.-C. Chen, "Beamforming optimization for intelligent reflecting surface-aided MISO communication systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 504–513, 2020.

[21] X. Zhao, S. Lu, Q. Shi, and Z.-Q. Luo, "Rethinking wmmse: Can its complexity scale linearly with the number of BS antennas?" *arXiv preprint arXiv:2205.06225*, 2022.

[22] J. Barzilai and J. M. Borwein, "Two-Point Step Size Gradient Methods," *IMA J. Numer. Anal.*, vol. 8, no. 1, pp. 141–148, Jan. 1988.

[23] Y. Wang, W. Zou, and Y. Tao, "Analog precoding designs for millimeter wave communication systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11 733–11 745, 2018.