OPEN ACCESS

University of BRISTOL

Yang, Y., Dang, S., Wen, M., Ai, B., & Hu, R. Q. (2024). RIS-Assisted Mobile Millimeter Wave MIMO Communications: A Blockage-Aware Robust Beamforming Approach. In *ICC 2024 - IEEE International Conference on Communications* (pp. 3737-3742). IEEE Computer Society. https://doi.org/10.1109/ICC51166.2024.10622320

Peer reviewed version

License (if available):
CC BY

Link to published version (if available):
10.1109/ICC51166.2024.10622320

Link to publication record on the Bristol Research Portal
PDF-document

This is the accepted author manuscript (AAM) of the article which has been made Open Access under the University of Bristol's Scholarly Works Policy. The final published version (Version of Record) can be found on the publisher's website. The copyright of any third-party content, such as images, remains with the copyright holder.

University of Bristol – Bristol Research Portal

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/red/research-policy/pure/user-guides/brp-terms/

RIS-Assisted Mobile Millimeter Wave MIMO Communications: A Blockage-Aware Robust Beamforming Approach

Yan Yang*, Shuping Dang†, Miaowen Wen‡, Bo Ai*, and Rose Qingyang Hu§
*School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China
†School of Electrical, Electronic and Mechanical Engineering, University of Bristol, Bristol BS8 1UB, U.K
‡School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China
*School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China
§Department of Electrical and Computer Engineering, Utah State University, Logan, UT 84322 USA
*yyang@bjtu.edu.cn; †shuping.dang@bristol.ac.uk; ‡eemwwen@scut.edu.cn;*bai@bjtu.edu.cn;§rose.hu@usu.edu

***Abstract*—Millimeter wave (mmWave) communications are highly affected by blockage, whereas the emerging reconfigurable intelligent surface (RIS) has the potential to overcome this issue. This paper proposes a Neyman-Pearson (N-P) criterion-based blockage-aware algorithm to improve resilience to blockage in mobile mmWave multiple input multiple output (MIMO) systems. By virtue of this pragmatic blockage-aware technique, we further propose an outage-constrained beamforming design for RIS-assisted mmWave MIMO transmission to achieve outage probability minimization and achievable rate maximization. Specifically, we propose an accelerated projected gradient descent (PGD) algorithm to solve the computational challenge of high-dimensional RIS phase-shift matrix (PSM) optimization. Particularly, we formulate a new Nesterov momentum acceleration scheme to speed up the convergence rate. Extensive experiments confirm the effectiveness of the proposed blockage-aware approach and the proposed accelerated PGD algorithm outperforms a number of representative baseline algorithms in terms of the achievable rate performance.**

***Keywords*—Blockage detection, mmWave, reconfigurable intelligent surfaces, robust beamforming, projected gradient descent**

I. Introduction

Signal transmissions at mmWave bands are susceptible to an extremely high path and penetration loss. By virtue of large-scale antenna arrays, path loss at mmWave bands can be compensated by highly directional beamforming at both transmitter and receiver, in which a very narrow beam pattern is used to provide high-directional gain [1], [2]. However, the high-directional mmWave transmission experiences extreme sensitivity to the blockage. It normally incurs a rapid degradation of signal strength, resulting in a higher outage probability. To overcome such a challenge, the recent emergence of reconfigurable intelligent surfaces (RISs) provides a potential solution by utilizing their unnatural reflection property [3], [4]. Unlike traditional amplify-and-forward relays and coordinated multi-point (CoMP) transmission, RIS is almost passive. Such a deployment has the potential to give rise to a dramatic improvement in the overall network's signal-to-interference-plus-noise ratio (SINR) performance.

Note that the validity of a RIS substantially depends on whether the blockages can be effectively detected. Prior works in the literature have focused on blockage prediction, robust beamforming, and outage minimum to overcome such challenges [3], [4]. With the aid of CoMP transmission, a stochastic-learning approach has been proposed to capture crucial blockage patterns and a robust beamforming design to combat uncertain path blockages. Most recently, a stochastic learning-based robust beamforming design and a lightweight algorithm for blockage status sensing have been proposed for RIS-assisted mmWave systems in the presence of random blockages [5].

In this paper, we develop a pragmatic blockage-aware algorithm to substantially reap the benefits of the RIS, such that when the line-of-sight (LOS) link is blocked, the base station (BS) can immediately steer the dominant beam towards the RIS and create an alternative reflection link for the mobile station (MS). Furthermore, we propose a robust beamforming design for downlink mmWave MIMO transmission to achieve outage probability minimization and achievable rate maximization.

The rest of the paper is organized as follows. In Section II, we introduce the system model and elucidate the problems for the outage probability minimization and rate optimization. Section III, we provide a statistical decision-making framework for blockage detection. In Section IV, a robust beamformer is formulated, and a projected gradient descent (PGD) based algorithm is derived for achievable rate maximization. Simulation results are shown in Section V, and the paper draws conclusions in Section VI.

II. System Model and Problem Statement

A. System Model

Consider a RIS-assisted mmWave MIMO communication system in the downlink, as illustrated in Fig. 1, where the RIS, with $\mathcal{N}$ reflecting elements arranged, is utilized to provide a substitutable link in the presence of random link blockages. A BS with a HAD beamforming architecture uses a uniform linear array (ULA) of N_t

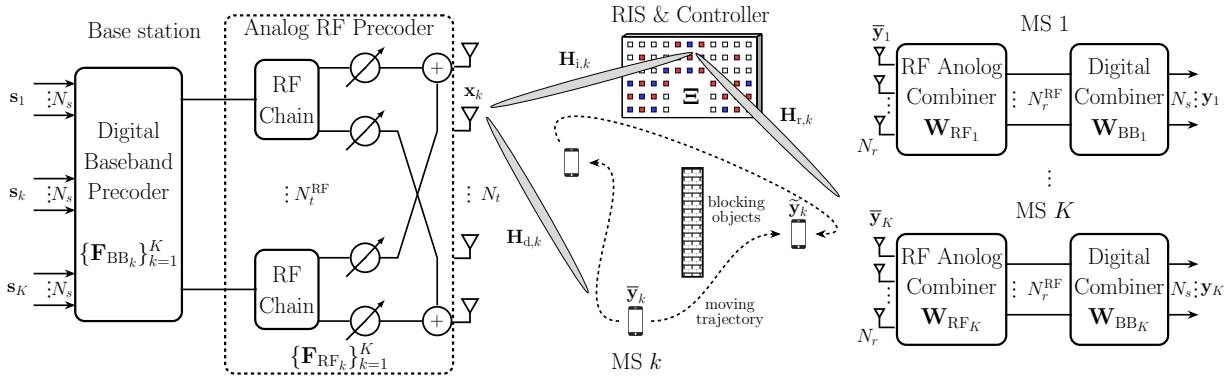

Fig. 1: RIS-assisted mmWave MIMO communication system: A downlink transmission subject to blockages.

transmit antennas and N_t^{RF} RF chains to serve K MSs. We assume that each MS is equipped with N_r-antennas ULA and N_r^{RF} RF chains, and $N_{t,k}^{\mathrm{RF}}$ RF chains can be configured for a specific MS k to transmit varying N_s data streams ($N_s \leq N_t^{\mathrm{RF}}$). During downlink transmission, the data stream $\mathbf{s} = \{\mathbf{s}_k\}_{k=1}^K, \mathbf{s}_k \in \mathbb{C}^{N_s \times 1}$ is processed by a set of digital baseband precoders $\{\mathbf{F}_{\mathrm{BB}_k}\}_{k=1}^K, \mathbf{F}_{\mathrm{BB}_k} \in \mathbb{C}^{N_{t,k}^{\mathrm{RF}} \times N_s}$ followed by a set of analog RF precoders $\{\mathbf{F}_{\mathrm{RF}_k}\}_{k=1}^K, \mathbf{F}_{\mathrm{RF}_k} \in \mathbb{C}^{N_{t,k} \times N_s}$. Likewise, a hybrid combiner is realized at the receiver by concatenation of an RF combiner $\mathbf{W}_{\mathrm{RF}_k} \in \mathbb{C}^{N_{r,k} \times N_r^{\mathrm{RF}}}$ and a digital baseband combiner $\mathbf{W}_{\mathrm{BB}_k} \in \mathbb{C}^{N_r^{\mathrm{RF}} \times N_s}$.

We define $\mathcal{K} = \{1, 2, ..., K\}$ as the set of K MSs. For an arbitrary MS $k \in \mathcal{K}$, when the direct link is blocked, we suppose that $\mathcal{N}_k$ element of the RIS can be partitioned as a subarray of size $\mathcal{I}_k \times \mathcal{J}_k$, given $\sum_{k=1}^K \mathcal{N}_k \leq \mathcal{N}$, $\mathcal{I}_k \leq \mathcal{I}$ and $\mathcal{J}_k \leq \mathcal{J}$. $\mathbf{\Xi}_k(\boldsymbol{\theta}) = \mathrm{diag}(\boldsymbol{\theta}) \in \mathbb{C}^{\mathcal{N}_k \times \mathcal{N}_k}$ is the phase-shift matrix (PSM), where $\boldsymbol{\theta} = \left[e^{j\theta_{1,k}}, \ldots, e^{j\theta_{\imath\jmath,k}}, \ldots\right]^T \in \mathbb{C}^{\mathcal{N}_k \times 1}$, $\{\theta_{\imath\jmath,k}\}_{\imath\jmath=1}^{\mathcal{N}_k}$ represents a set of phase shifts (PSs) induced by $\mathcal{N}_k$ RIS elements, and $\theta_{\imath\jmath,k}$ is the $(\imath,\jmath)$th entry.

On the downlink, the transmitted signal to the kth MS in a generic HAD mmWave system is given by

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{s}_k, \quad k \in \mathcal{K}, \tag{1}$$

where $\mathbf{F}_k = \mathbf{F}_{\mathrm{RF}_k}\mathbf{F}_{\mathrm{BB}_k} \in \mathbb{C}^{N_r \times N_{t,k}}$ is the hybrid precoder, and can be naturally different for individual MS; $\mathbf{s}_k \in \mathbb{C}^{N_s \times 1}$ is a normalized transmit signal vector, i.e., $\mathbb{E}\{||\mathbf{s}_k||^2\} = 1$.

With the Rician fading channel model, both the direct path and the reflection path are considered, such that the channel matrix between the BS and the kth MS can be described by

$$\overline{\mathbf{H}}_k(\boldsymbol{\theta}) = \varepsilon_k \sqrt{\frac{\kappa \overline{\varrho}_{\mathrm{d},k}}{\kappa+1}} \mathbf{H}_{\mathrm{d},k} + \sqrt{\frac{\overline{\varrho}_{\mathrm{c},k}}{\kappa+1}} \mathbf{H}_{\mathrm{c},k}(\boldsymbol{\theta}), \; k \in \mathcal{K}^{\mathrm{LOS}} \tag{2a}$$

$$\widetilde{\mathbf{H}}_k(\boldsymbol{\theta}) = \varepsilon_k \sqrt{\frac{\widetilde{\varrho}_{\mathrm{d},k}}{\kappa+1}} \mathbf{H}_{\mathrm{d},k} + \sqrt{\frac{\kappa \widetilde{\varrho}_{\mathrm{c},k}}{\kappa+1}} \mathbf{H}_{\mathrm{c},k}(\boldsymbol{\theta}), \; k \in \mathcal{K}^{\mathrm{NLOS}} \tag{2b}$$

where $\mathbf{H}_{\mathrm{d},k} \in \mathbb{C}^{N_r \times N_{t,k}}$ represents the direct BS-MS channel, and $\mathbf{H}_{\mathrm{c},k}(\boldsymbol{\theta}) = \mathbf{H}_{\mathrm{r},k}^H \mathbf{\Xi}_k(\boldsymbol{\theta}) \mathbf{H}_{\mathrm{i},k} \in \mathbb{C}^{N_r \times N_{t,k}}$ represents the cascaded BS-RIS-MS channel, both associated with the MS k. Here $\mathbf{H}_{\mathrm{r},k} \in \mathbb{C}^{\mathcal{N}_k \times N_t}$ and $\mathbf{H}_{\mathrm{i},k} \in \mathbb{C}^{\mathcal{N}_r \times \mathcal{N}_k}$ are the indirect channels between the BS and the RIS (BS-RIS) and the RIS and MS (RIS-MS), respectively. In addition, $\kappa \in [0, \infty)$ is the Rician K factor; $\varepsilon_k \in [0, 1]$ is a normalized random variable that represents the link loss rate caused by the blocker, and $\varepsilon_k = 0$ associates to worst-case channel condition, that is, the case that the kth MS is fully blocked. $\varrho_{\mathrm{d},k}$ and $\varrho_{\mathrm{r},k}$ are the fading coefficients of the direct and indirect links, respectively. In order to unburden the notation, we denote $\upsilon_0 = \sqrt{\frac{\kappa \varrho_{\mathrm{d},k}}{\kappa+1}}$ and $\upsilon_1 = \sqrt{\frac{\varrho_{\mathrm{c},k}}{\kappa+1}}$ by the two random variables accounting for the effects of both the path loss and Rician K-factor in sparse scattering environments.

Accordingly, the $N_r \times 1$ received signal vector at the antenna array of the kth MS, $\mathbf{y}_k$ can be expressed as

$$\overline{\mathbf{y}}_k = \underbrace{\varepsilon_k \overline{\upsilon}_0 \mathbf{W}_k^H \mathbf{H}_{\mathrm{d},k} \mathbf{x}_k}_{\text{dominant signal: } \overline{\mathbf{y}}_{\mathrm{d},k}} + \underbrace{\overline{\upsilon}_1 \mathbf{W}_k^H \mathbf{H}_{\mathrm{r},k}^H \mathbf{\Xi}_k(\boldsymbol{\theta}) \mathbf{H}_{\mathrm{i},k} \mathbf{x}_k}_{\text{reflected auxiliary signal: } \overline{\mathbf{y}}_{\mathrm{r},k}} + \mathbf{n}_k, \tag{3a}$$

$$\widetilde{\mathbf{y}}_k = \underbrace{\varepsilon_k \widetilde{\upsilon}_0 \mathbf{W}_k^H \mathbf{H}_{\mathrm{d},k} \mathbf{x}_k}_{\text{auxiliary signal: } \widetilde{\mathbf{y}}_{\mathrm{d},k}} + \underbrace{\widetilde{\upsilon}_1 \mathbf{W}_k^H \mathbf{H}_{\mathrm{r},k}^H \mathbf{\Xi}_k(\boldsymbol{\theta}) \mathbf{H}_{\mathrm{i},k} \mathbf{x}_k}_{\text{reflected dominant signal: } \widetilde{\mathbf{y}}_{\mathrm{r},k}} + \mathbf{n}_k, \tag{3b}$$

where $\mathbf{W}_k^H = \mathbf{W}_{\mathrm{BB}_k}^H \mathbf{W}_{\mathrm{RF}_k}$; $\mathbf{n}_k \sim \mathcal{CN}\left\{0, \sigma_k^2 \mathbf{I}\right\}$ is the circularly symmetric complex additive white Gaussian noise (AWGN) at the kth MS; σ_k^2 represents the noise variance associated with the kth MS; $\mathbf{I} \in \mathbb{C}^{N_r \times N_r}$ is the identity matrix.

For convenience, we denote $\mathbf{U}_k = \mathbf{F}_k^H \mathbf{F}_k$ as the transmit beamforming gain, $\mathbf{V}_k = \mathbf{W}_k^H \mathbf{W}_k$ as the receive beamforming gain, and $\mathbf{Z}_k(\boldsymbol{\theta}) = \mathbf{H}_{\mathrm{r},k}^H \mathbf{\Xi}_k(\boldsymbol{\theta}) \mathbf{H}_{\mathrm{i},k}$. The total beamforming gain is given by $\mathbf{Q} = \mathbf{V}_k \mathbf{U}_k$. The SINR on the kth MS can now be defined in a closed form as

$$\overline{\Gamma}_k(\boldsymbol{\theta}, \mathbf{Q}) = \frac{\varepsilon_k^2 \overline{\upsilon}_0^2}{P_0} \left|\mathbf{H}_{\mathrm{d},k} \mathbf{Q} \mathbf{H}_{\mathrm{d},k}^H\right| + \frac{\overline{\upsilon}_1^2}{P_0} \left|\mathbf{Z}_k(\boldsymbol{\theta}) \mathbf{Q} \mathbf{Z}_k^H(\boldsymbol{\theta})\right|, \tag{4a}$$

$$\widetilde{\Gamma}_k(\boldsymbol{\theta}, \mathbf{Q}) = \frac{\varepsilon_k^2 \widetilde{\upsilon}_0^2}{P_0} \left|\mathbf{H}_{\mathrm{d},k} \mathbf{Q} \mathbf{H}_{\mathrm{d},k}^H\right| + \frac{\widetilde{\upsilon}_1^2}{P_0} \left|\mathbf{Z}_k(\boldsymbol{\theta}) \mathbf{Q} \mathbf{Z}_k^H(\boldsymbol{\theta})\right|, \tag{4b}$$

where $P_0 = N_{t,k}\sigma_k^2 + \sum_{j \neq k} g_j$ is the total signal power received from all the other MSs, and s a measure of the k MS noise. The achievable rate of the kth user, $\mathcal{R}_k$,

measured in bits-per-second/Hz, can be calculated by

$$\overline{\mathcal{R}}_k(\boldsymbol{\theta},\mathbf{Q}) = \log_2 \det\left(\mathbf{I}_k + \frac{1}{P_0}\overline{\mathbf{H}}_k(\boldsymbol{\theta})\mathbf{Q}\overline{\mathbf{H}}_k^H(\boldsymbol{\theta})\right), \tag{5a}$$

$$\widetilde{\mathcal{R}}_k(\boldsymbol{\theta},\mathbf{Q}) = \log_2 \det\left(\mathbf{I}_k + \frac{1}{P_0}\widetilde{\mathbf{H}}_k(\boldsymbol{\theta})\mathbf{Q}\widetilde{\mathbf{H}}_k^H(\boldsymbol{\theta})\right). \tag{5b}$$

At the kth MS, suppose that the numbers of effective channel clusters for the BS-MS links, the BS-RIS links, and the RIS-MS links are $L_{\mathrm{d},k}$, $L_{\mathrm{r},k}$ and $L_{\mathrm{i},k}$, respectively, the channel matrices in (3) can be represented as

$$\mathbf{H}_{\mathrm{d},k} = \frac{1}{\sqrt{L_{\mathrm{d},k}}}\sum_{\ell=1}^{L_{\mathrm{d},k}} g_{\mathrm{d},k}^{\ell}\mathbf{a}_{\mathrm{MS}}\left(\psi_{\mathrm{d},k}^{\mathrm{AOA}_\ell}\right)\mathbf{a}_{\mathrm{BS}}\left(\phi_{\mathrm{d},k}^{\mathrm{AOD}_\ell}\right)^H, \tag{6}$$

$$\mathbf{H}_{\mathrm{r},k} = \frac{1}{\sqrt{L_{\mathrm{r},k}}}\sum_{\ell=1}^{L_{\mathrm{r},k}} g_{\mathrm{r},k}^{\ell}\mathbf{a}_{\mathrm{MS}}\left(\psi_{\mathrm{r},k}^{\mathrm{AOA}_\ell}\right)\mathbf{a}_{\mathrm{RIS}}\left(\phi_{\mathrm{r},k}^{\mathrm{AOD}_\ell},\varphi_{\mathrm{r},k}^{\mathrm{AOD}_\ell}\right)^H, \tag{7}$$

$$\mathbf{H}_{\mathrm{i},k} = \frac{1}{\sqrt{L_{\mathrm{i},k}}}\sum_{\ell=1}^{L_{\mathrm{i},k}} g_{\mathrm{i},k}^{\ell}\mathbf{a}_{\mathrm{RIS}}\left(\phi_{\mathrm{i},k}^{\mathrm{AOA}_\ell},\varphi_{\mathrm{i},k}^{\mathrm{AOA}_\ell}\right)\mathbf{a}_{\mathrm{BS}}\left(\phi_{\mathrm{d},k}^{\mathrm{AOD}_\ell}\right)^H, \tag{8}$$

where the angle-of-arrival (AOA) and angle-of-departure (AOD) are angle-domain representations, $g_{\mathrm{d},k}^{\ell}$, $g_{\mathrm{r},k}^{\ell}$ and $g_{\mathrm{i},k}^{\ell}$ are the complex gain of ℓth tap of the corresponding links, respectively; $\left\{\psi_{\mathrm{d},k}^{\mathrm{AOA}_\ell},\phi_{\mathrm{d},k}^{\mathrm{AOD}_\ell}\right\}$, $\left\{\psi_{\mathrm{r},k}^{\mathrm{AOA}_\ell},\left\{\phi_{\mathrm{r},k}^{\mathrm{AOD}_\ell},\varphi_{\mathrm{r},k}^{\mathrm{AOD}_\ell}\right\}\right\}$ and $\left\{\left\{\phi_{\mathrm{i},k}^{\mathrm{AOA}_\ell},\varphi_{\mathrm{i},k}^{\mathrm{AOA}_\ell}\right\},\phi_{\mathrm{d},k}^{\mathrm{AOD}_\ell}\right\}$ are the azimuth (horizontal) AOA and AOD pairs of ℓth tap for the BS-MS links, the RIS-MS links, and the BS-RIS links, respectively. Moreover, $\mathbf{a}_{\mathrm{BS}}(\phi)$, $\mathbf{a}_{\mathrm{MS}}(\phi)$ and $\mathbf{a}_{\mathrm{RIS}}(\phi,\varphi)$ represent the transmit (BS), receive (MS) and reflect (RIS) array responses (i.e., steering vectors), respectively, given by

$$\begin{aligned}
\mathbf{a}_{\mathrm{BS}}(\phi) &= \frac{1}{\sqrt{N_{t,k}}}\left[1, e^{j\cos(\phi)},\ldots,e^{j(N_{t,k}-1)\cos(\phi)}\right]^T,\\
\mathbf{a}_{\mathrm{MS}}(\psi) &= \frac{1}{\sqrt{N_r}}\left[1, e^{j\cos(\psi)},\ldots,e^{j(N_r-1)\cos(\psi)}\right]^T,\\
\mathbf{a}_{\mathrm{RIS}}(\phi,\varphi) &= \frac{1}{\sqrt{\mathcal{N}_k}}\left[1, e^{j\xi_1(\phi,\varphi)},\ldots,e^{j\xi_{\imath\jmath}(\phi,\varphi)}\right]^T,
\end{aligned} \tag{9}$$

where $\xi_{\imath\jmath}(\phi,\varphi) = (\imath-1)\sin\varphi\sin\phi + (\jmath-1)\cos\phi$ is the phase of an incoming plane wave at the $\imath\jmath$th RIS element, the geometric factor of the RIS, λ is the wavelength and d is the inter-element distance of the RIS.

B. Problem Statement

We first consider the minimization of outage probability for arbitrary MS k under pure non-line-of-sight (NLOS) conditions, neglecting the case of any LOS. Let $\Pr\left\{\widetilde{\Gamma}_k(\boldsymbol{\theta},\mathbf{Q}) \le \gamma\right\}$ be the cumulative outage probability of the MS k. Mathematically, this minimization problem can be formulated as

$$\begin{aligned}
\mathcal{P}_1: \min_{\boldsymbol{\theta},\mathbf{Q}}\ & \Pr\left(\widetilde{\Gamma}_k(\boldsymbol{\theta},\mathbf{Q}) \le \gamma_{\mathrm{th}}\right), \quad k \in \mathcal{K}^{\mathrm{NLOS}},\\
\text{s.t.}\ & \mathrm{Tr}(\mathbf{Q}) \le \Omega_{\mathrm{total}},
\end{aligned} \tag{10}$$

where γ_{th} are the required SINR thresholds, corresponding to minimum acceptable rates. By the definition of the outage probability, an outage could take place when the SINR falls below a threshold μ_{th}.

Next, we aim to find a set of optimum $\mathbf{U}_k$, $\mathbf{V}_k$ and $\boldsymbol{\theta}_k$ that maximize the achievable rate under the total power constraint Ω_{total}. To be mathematically precise, this can be cast as the following optimization problem:

$$\begin{aligned}
\mathcal{P}_2: \max_{\boldsymbol{\theta},\mathbf{Q}}\ & \widetilde{\mathcal{R}}_k(\boldsymbol{\theta},\mathbf{Q}), \quad k \in \mathcal{K}^{\mathrm{NLOS}}\\
\text{s.t.}\ & \mathrm{Tr}(\mathbf{Q}) \le \Omega_{\mathrm{total}},\ \widetilde{\Gamma}_k(\boldsymbol{\theta},\mathbf{Q}) > \gamma_{\mathrm{th}},
\end{aligned} \tag{11}$$

where Ω_{total} the total transmit power constraint. Unless stated otherwise, we assume that the objective function $\mathcal{R}_k(\boldsymbol{\theta},\mathbf{Q})$ per se is convex and smooth.

III. Binary Hypothesis Testing for Blockage Detection

From the perspective of the statistical decision theory, we formulate the blockages detection as a binary hypothesis testing problem, whose process can be described by a 3-tuple $(\mathcal{H},\mathcal{Y}_\varepsilon,\mathscr{P}_\varepsilon)$. $\mathcal{H} = \left\{\mathcal{H}_0^{\mathrm{LOS}},\mathcal{H}_1^{\mathrm{NLOS}}\right\}$ is a set that consists of two hypotheses: $\mathcal{H}_0^{\mathrm{LOS}}$ is the null hypothesis of no blockage and $\mathcal{H}_1^{\mathrm{NLOS}}$ is the alternative hypothesis of blockage; $\mathcal{Y}_\varepsilon = \left\{\mathcal{Y}_0^{\mathrm{LOS}},\mathcal{Y}_1^{\mathrm{NLOS}}\right\}$ is the decision region with parameter ε, representing to the disjoint sample sets $\mathcal{Y}_0^{\mathrm{LOS}}$ and $\mathcal{Y}_1^{\mathrm{NLOS}}$ that the decide LOS or NLOS channel condition between the BS and the MS.

Without loss of generality, we can rewrite (3a) as a time-domain noisy observation $\mathbf{y}_q(\varepsilon)$ at time slot t_q with the unknown parameter ε and path delay τ, which is given by

$$\mathbf{y}_q(\varepsilon) = \mathbf{y}_k(t_q) = \overline{\mathbf{y}}_{\mathrm{d},k}(t_q) + \overline{\mathbf{y}}_{\mathrm{r},k}(t_q-\tau) + \mathbf{n}_k(t_q). \tag{12}$$

Our goal here is to preemptively detect whether the strong LOS signal $\overline{\mathbf{y}}_{\mathrm{d},k}$ of interest is present. Formally, this can come down to distinguishing between the following two simple hypotheses:

$$\begin{aligned}
\mathcal{H}_0^{\mathrm{LOS}}&: \mathbf{y}_q(\varepsilon) = \overline{\mathbf{y}}_{\mathrm{d},k} + \overline{\mathbf{y}}_{\mathrm{r},k} + \mathbf{n}_k, \quad \textit{if}\ \ \varepsilon > \varepsilon_{\mathrm{th}},\\
\mathcal{H}_1^{\mathrm{NLOS}}&: \mathbf{y}_q(\varepsilon) = \overline{\mathbf{y}}_{\mathrm{r},k} + \mathbf{n}_k, \quad \textit{if}\ \ \varepsilon < \varepsilon_{\mathrm{th}},
\end{aligned} \tag{13}$$

where $\varepsilon_{\mathrm{th}}$ is a pre-defined threshold value, identifying the absence or presence of the blockage.

Formally, the probability mass function (PMF) $f_{\mathbf{Y}_q}(\cdot)$ of this binary hypothesis testing problem under each hypothesis can be obtained by the following form

$$\begin{aligned}
\mathcal{H}_0^{\mathrm{LOS}}&: \mathcal{Y}_\varepsilon \sim f_{\mathbf{Y}_q}\left(\varepsilon > \varepsilon_{\mathrm{th}}|\mathcal{H}_0^{\mathrm{LOS}}\right),\\
\mathcal{H}_1^{\mathrm{NLOS}}&: \mathcal{Y}_\varepsilon \sim f_{\mathbf{Y}_q}\left(\varepsilon < \varepsilon_{\mathrm{th}}|\mathcal{H}_1^{\mathrm{NLOS}}\right).
\end{aligned} \tag{14}$$

Upon using the Neyman-Pearson (N-P) Lemma, an alternative decision rule is to solving the following optimization problem:

$$\begin{aligned}
\delta_{\text{N-P}}(\mathbf{y}_q(\varepsilon)) = \arg\max_{\varepsilon}\ & P_{\mathrm{D}}^{\mathrm{NLOS}}(\varepsilon_{\mathrm{th}}),\\
\text{s.t.}\ & P_{\mathrm{FA}}^{\mathrm{LOS}}(\varepsilon_{\mathrm{th}}) \le \alpha,
\end{aligned} \tag{15}$$

where $P_{\mathrm{D}}^{\mathrm{NLOS}}(\varepsilon_{\mathrm{th}})$ is the probability of correct blockage detection; $P_{\mathrm{FA}}^{\mathrm{LOS}}(\varepsilon_{\mathrm{th}})$ is the probability of false alarm,

corresponding to the detection threshold ε_{th} respectively. In this way, we can model the observed data by a discrete observation set $\mathbf{Y} = \{\mathbf{Y}_1, \ldots, \mathbf{Y}_q, \ldots\}$, where $\mathbf{Y}_q = \left[\mathbf{y}_q^{(0)}, \ldots, \mathbf{y}_q^{(m)}, \ldots\right]^T \in \mathbb{C}^{N_{t,k} \times M}$, and $\mathbf{y}_q^{(m)}$ represents for each possible observation.

By convention, the probability of detection $P_{\text{D}}^{\text{NLOS}}(\varepsilon_{\text{th}})$ and the probability of false alarm $P_{\text{FA}}^{\text{LOS}}(\varepsilon_{\text{th}})$ can be respectively calculated by

$$\begin{aligned} P_{\text{D}}^{\text{NLOS}}(\varepsilon_{\text{th}}) &= P\left(\mathcal{H}_1^{\text{NLOS}} | \mathcal{H}_1^{\text{NLOS}}\right) \\ &= \prod_{l:\ \mathcal{L}_l(\varepsilon_{\text{th}}) > \Lambda} \int_0^{\varepsilon_{\text{th}}} f_{\mathbf{y}_q^{(m)}}(\varepsilon)\, d\varepsilon = \beta, \end{aligned} \tag{16}$$

and

$$\begin{aligned} P_{\text{FA}}^{\text{LOS}}(\varepsilon_{\text{th}}) &= P\left(\mathcal{H}_1^{\text{NLOS}} | \mathcal{H}_0^{\text{LOS}}\right) \\ &= \prod_{l:\ \mathcal{L}_l(\varepsilon_{\text{th}}) < \Lambda} \int_{\varepsilon_{\text{th}}}^{1} f_{\mathbf{y}_q^{(m)}}(\varepsilon) d\varepsilon \leq \alpha, \end{aligned} \tag{17}$$

where $\alpha \in [0, 1]$ is the *significance level* of the test of user k. It may be viewed as a transition probability of going from state LOS to NLOS. For given $F_{\text{FA}} = \alpha$, N-P's optimal solution for achieving maximum P_{D}, can be determined by a likelihood ratio test (LRT), that is, it follows that the LRT can be expressed as

$$\begin{aligned} \mathcal{L}(\varepsilon_{\text{th}}) &= \frac{\displaystyle\int_0^{\varepsilon_{\text{th}}} f_{\mathbf{Y}_q}\left(\varepsilon | \mathcal{H}_1^{\text{NLOS}}\right)\, d\varepsilon}{\displaystyle\int_{\varepsilon_{\text{th}}}^{1} f_{\mathbf{Y}_q}\left(\varepsilon | \mathcal{H}_0^{\text{LOS}}\right)\, d\varepsilon} \\ &= \prod_{l=1}^{\mathcal{M}} \frac{\displaystyle\int_0^{\varepsilon_{\text{th}}} f_{\mathbf{y}_q^{(m)}}\left(\varepsilon | \mathcal{H}_1^{\text{NLOS}}\right)\, d\varepsilon}{\displaystyle\int_{\varepsilon_{\text{th}}}^{1} f_{\mathbf{y}_q^{(m)}}\left(\varepsilon | \mathcal{H}_0^{\text{LOS}}\right)\, d\varepsilon} \underset{\mathcal{H}_0^{\text{LOS}}}{\overset{\mathcal{H}_1^{\text{NLOS}}}{\gtrless}} \Lambda. \end{aligned} \tag{18}$$

Mathematically, a deterministic decision rule $\delta_{\text{N-P}} : \mathbf{y}_k \to \{1, 0\}$ can be reformulated as

$$\delta_{\text{N-P}}(\mathbf{y}_q(\varepsilon)) = \begin{cases} 1; & \mathcal{L}(\varepsilon_{\text{th}}) > \Lambda \\ \varepsilon_{\text{th}}; & \mathcal{L}(\varepsilon_{\text{th}}) = \Lambda \\ 0; & \mathcal{L}(\varepsilon_{\text{th}}) < \Lambda \end{cases}, \tag{19}$$

where $\delta_{\text{N-P}}(\cdot)$ is the N-P criterion-based statistic test operator. At this point, we have solved the statistical decision problem raised at the beginning, which follows that

$$\begin{aligned} \mathcal{Y}_0^{\text{LOS}} &\triangleq \{\mathbf{y}_q | \delta_{\text{N-P}}(\mathbf{y}_q(\varepsilon)) = 0\}, \\ \mathcal{Y}_1^{\text{NLOS}} &\triangleq \{\mathbf{y}_q | \delta_{\text{N-P}}(\mathbf{y}_q(\varepsilon)) = 1\}. \end{aligned} \tag{20}$$

IV. Outage Minimization and Rate Maximization

A. Outage Minimization with Robust Beamforming

Motivated by the above analysis, we propose a beamwidth-variant beamforming solution enabling flexible beamwidth adaptation. Normally, the BS-RIS beam is designed to be the narrowest beamwidth, such that the optimization of the transmit precoder $\mathbf{F}_k$ and the PSM $\boldsymbol{\Xi}(\boldsymbol{\theta})$ can be decoupled. As a consequence, the solution of the outage-constrained problem $\mathcal{P}_1$ can be divided into the following two steps. The narrowest BS-MIS beam is designed first. Then, assuming a local optimal $\mathbf{F}_k^*$, an adaptive beamwidth control method can be applied to further decrease the outage probability.

The first step of the algorithm considers the design of hybrid precoder $\mathbf{F}_{\text{BB}_k}$ and $\mathbf{F}_{\text{RF}_k}$, of the purpose to minimize the beamwidth of the main lobe. Most commonly, the beamwidth can be defined by the half-power beamwidth (HPBW), equivalent to the angular width of the radiation pattern that is at 3-dB down from the maximum gain of the beam (beam peak). The design problem of the narrowest transmit beam can be written as

$$\underset{\mathbf{F}_{\text{BB}_k}, \mathbf{F}_{\text{RF}_k}}{\arg\min}\ \Phi_{\text{BS}}(\mathbf{F}_{\text{BB}_k}, \mathbf{F}_{\text{RF}_k}) := |\phi_{\text{3-dB}}^{+} - \phi_{\text{3-dB}}^{-}|. \tag{21}$$

In general, the narrowest possible beamwidth $\Phi_{\text{BS}}^{\min}(\mathbf{F}_{\text{BB}_k}, \mathbf{F}_{\text{RF}_k})$ can be directly determined by the optimal beamforming vector $\mathbf{F}_{\text{BB}_k}$ and $\mathbf{F}_{\text{RF}_k}$ in a hierarchical multiple-resolution codebook.

Thus, in the follow-up passive beamforming design step, we formulate an adjustable beamwidth control strategy by scaling the number of the RIS reflective elements. In a more compact form, (10) can be explicitly rewritten as

$$\begin{aligned} \min_{\Phi_{\text{RIS}}(\boldsymbol{\theta}, \mathcal{N}_k),\, \mathbf{Q}}\ & \Pr\left(\widetilde{\Gamma}_k(\boldsymbol{\theta}, \mathbf{Q}) \leq \gamma_{\text{th}} | _{\Phi_{\text{RIS}}(\boldsymbol{\theta}, \mathcal{N}_k)}\right),\ k \in \mathcal{K}^{\text{NLOS}}, \\ \text{s.t.}\quad & \text{Tr}(\mathbf{Q}) \leq \Omega_{\text{total}}, \quad \text{MSE}_{\boldsymbol{\theta}} < \zeta_{\boldsymbol{\theta}}^{\max}, \end{aligned} \tag{22}$$

where $\Phi_{\text{RIS}}(\boldsymbol{\theta}, \mathcal{N}_k)$ is a function of $\boldsymbol{\theta}$, $\mathcal{N}_k$, controlling the beamwidth in the RIS; $\zeta_{\boldsymbol{\theta}}^{\max}$ is the maximum tolerable level of the inaccurate PSs estimates. $\Phi_{\text{RIS}}(\boldsymbol{\theta}, \mathcal{N}_k)$ is inversely proportional to the distance between the RIS and the MS.

Compared to a constant-beamwidth strategy, the adjustable beamwidth has more design flexibility by scaling $\mathcal{N}_k$, and there exists an optimal beamwidth that significantly suppresses the outage probability under different circumstances. This enables us to carry out an adaptive beamwidth control procedure that provides more resilient connectivity.

B. Rate Maximization with Projected Gradient Descent

We now consider the globally optimal solution for problem $\mathcal{P}_2$. Following the famous Nesterov momentum acceleration method, we devise an accelerated PGD algorithm.

To solve problem $\mathcal{P}_2$ in a standard gradient descent fashion, we can reformulate, in terms of the feasible set $\mathcal{F} = \{\Theta, \mathcal{Q}\}$, the optimization problem in a more compact form

$$\underset{\boldsymbol{\theta} \in \Theta,\ \mathbf{Q} \in \mathcal{Q}}{\arg\min}\quad \widetilde{\mathcal{R}}_k^{-1}(\boldsymbol{\theta}, \mathbf{Q}), \quad k \in \mathcal{K}^{\text{NLOS}}. \tag{23}$$

Similarly, the objective function $\widetilde{\mathcal{R}}_k^{-1}$ holds continuously differentiable and has a known gradient Lipschitz constant (e.g., L-Lipschitz continuous). Our main focus of this subsection is to find the global minima of the function $\mathcal{R}_k^{-1}$, which is equivalent to excluding all local minima.

It is convenient to define $\mathbf{P}_{\mathcal{F}}(\mathbf{u}) = \frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{v}||^2} \mathbf{v}$ as the

Euclidean projection from a vector of $\mathbf{u}$ onto an individual vector $\mathbf{v}$ in a feasible set $\mathcal{F}$. Exactly speaking, it refers to a function from $\mathbb{R}^n$ to $\mathbb{R}^n$, which in our case is to simply find $\boldsymbol{\theta}$ and $\mathbf{Q}$ that is closest to feasible sets Θ and $\mathcal{Q}$, respectively. Mathematically, the projection operation onto the constraint set Θ and $\mathcal{Q}$ can be cast as the following optimization problems:

$$\mathbf{P}_{\Theta}(\boldsymbol{\theta}) = \arg\min_{\boldsymbol{\vartheta}\in\Theta} \ ||\boldsymbol{\vartheta} - \boldsymbol{\theta}||_2, \tag{24a}$$

$$\mathbf{P}_{\mathcal{Q}}(\mathbf{Q}) = \arg\min_{\boldsymbol{\mathcal{Q}}\in\mathcal{Q}} \ ||\boldsymbol{\mathcal{Q}} - \mathbf{Q}||_2. \tag{24b}$$

In standard Nesterov's accelerated gradient method, an extra momentum term is first introduced to gradient descent on $\widetilde{\mathcal{R}}_k^{-1}$ at each iteration. After that, the parameters $\boldsymbol{\theta}$ and $\mathbf{Q}$ are iteratively updated by re-projecting each of them onto the constraint set Θ and $\mathcal{Q}$. The updating process of Nesterov's momentum and parameters over two successive steps is summarized through the following recursive formulas

$$\begin{aligned} \boldsymbol{\eta}^{(n+1)} &= \boldsymbol{\nu}\boldsymbol{\eta}^{(n)} - \mu_{\boldsymbol{\theta}}^{(n)} \nabla_{\boldsymbol{\theta}} \widetilde{\mathcal{R}}_k^{-1} \left(\boldsymbol{\theta}^{(n)}, \mathbf{Q}^{(n)}\right), \\ \boldsymbol{\theta}^{(n+1)} &= \mathbf{P}_{\Theta}\left(\boldsymbol{\theta}^{(n)} + \boldsymbol{\eta}^{(n+1)}\right), \end{aligned} \tag{25}$$

and

$$\begin{aligned} \boldsymbol{\chi}^{(n+1)} &= \boldsymbol{\Delta}\boldsymbol{\chi}^{(n)} - \mu_{\mathbf{Q}}^{(n)} \nabla_{\mathbf{Q}} \widetilde{\mathcal{R}}_k^{-1} \left(\boldsymbol{\theta}^{(n)}, \mathbf{Q}^{(n)}\right), \\ \mathbf{Q}^{(n+1)} &= \mathbf{P}_{\mathcal{Q}}\left(\mathbf{Q}^{(n)} + \boldsymbol{\chi}^{(n+1)}\right), \end{aligned} \tag{26}$$

where $\boldsymbol{\eta}^{(n+1)}$ and $\boldsymbol{\chi}^{(n+1)}$ are the accumulated momentum terms at the $(n+1)$th iteration; The hyperparameter $\boldsymbol{\nu} \in [0,1]$ and $\boldsymbol{\Delta} \in [0,1]$ represent the level of inertia in the descent direction (a.k.a momentum coefficient), which can be usually chosen by trial and/or with a model selection criterion [6]; $\mu_{\boldsymbol{\theta}}^{(n)} \in (0,1)$ and $\mu_{\mathbf{Q}}^{(n)} \in (0,1)$ are the step sizes used in nth iteration.

Following directly from the derivation in [7], the complex-value gradient of $\widetilde{\mathcal{R}}_k^{-1}$ with respect to $\boldsymbol{\theta}$ and $\mathbf{Q}$ can be explicitly formulated as follows:

$$\nabla_{\boldsymbol{\theta}} \widetilde{\mathcal{R}}_k^{-1}(\boldsymbol{\theta}, \mathbf{Q}) = \mathcal{V}_d\left(\mathbf{H}_{\text{RU},k}^H \boldsymbol{\Sigma}^{-1} \mathbf{Q}(\boldsymbol{\theta}) \mathbf{H}_{\text{BR},k}\right), \tag{27a}$$

$$\nabla_{\mathbf{Q}} \widetilde{\mathcal{R}}_k^{-1}(\boldsymbol{\theta}, \mathbf{Q}) = \mathbf{Q}^H(\boldsymbol{\theta}) \boldsymbol{\Sigma}^{-1} \mathbf{Q}(\boldsymbol{\theta}), \tag{27b}$$

where $\mathcal{V}(\cdot)$ denotes the vectorization operator that stacks the columns to create a single long column vector, as a straightforward calculation technique. The detailed calculation procedure of the complex-value gradient is rigorously derived in [7], as a straightforward technique.

V. Numerical Results

In this section, we evaluate the outage probability and achievable rate with the detailed Monte Carlo simulation. We consider a rectangle RIS comprising 1024 elements arranged in a 64×64 matrix. The distance between BS and RIS $d_{\text{BS-RIS}}$ is set to 100 meters, and the corresponding angle between BS and RIS is fixed at 45 degrees. We adopt the total free space path loss (FSPL) ratio model proposed by [7]. The system is assumed to operate at 28 GHz carrier frequency.

Fig. 2 plots the detection probability versus false alarm probability with a given SNR value, and demonstrates the effectiveness of the proposed blockage detection algorithm. It can be observed from Fig. 4 that with the SNR progressively increasing, the false alarm probability $P_{\text{FA}}^{\text{LOS}}$ is significantly reduced. For instance, assuming the optimal threshold to achieve a probability of detection $P_{\text{FA}}^{\text{LOS}}$ of 0.9, the probability of false alarm $P_{\text{FA}}^{\text{LOS}}$ drops from 0.1 (SNR=2 dB) to less 0.001 (SNR=10 dB). Moreover, we observe that a higher SNR can substantively improve the detection threshold, and SNR=10 dB is generally acceptable for blockage detection.

Next, Fig. 3 shows the outage probability as a function of the transmit power at the BS. We compare the outage performance of the proposed adaptive beamwidth control scheme with the fixed beamwidth solution given in [8]. In our experiment, the reliability of estimate s is set to be 0.95, 0.85, and 0.5, so as to evaluate the impact of the channel estimation error. From Fig. 3, one can observe that outage performance with the proposed approach significantly outperforms the method in [8] under various CSI uncertainty settings. It can be explained as a wider reflecting beam is capable of providing a larger range of coverage, thereby providing resilient connectivity over the indirect RIS-MS link.

In Fig. 4, we analyze two cases for the achievable rate $\widetilde{\mathcal{R}}_k$ and for the downlink sum-rate $\mathcal{R}_{\text{sys}}$. The performance is compared with the reference paper [9] and [10], with consideration of the impact of the number of the RIS elements $\mathcal{N}_k$ on the achievable rate, where $\varsigma_k \in [0,1]$ represents the reliability of the channel estimate. For the following analysis, we further investigate the effectiveness of the proposed PGD-based optimization method for RIS-assisted transmission. By examining Fig. 4, we clearly see that the proposed PGD approach significantly boosts the achievable rate performance with respect to the CSSCA-based and AO-based schemes obtained in [9] and [10]. We observe that by increasing the number of PSs $\mathcal{N}_k$, the achievable rate increases.

As a final numerical experiment, we investigate the convergence behavior of the proposed PGD-based algorithms compared with the aforementioned baseline algorithm. In Fig. 5, our numerical results demonstrate that the proposed PGD-based algorithm with momentum acceleration, compares favorably with other baseline algorithms, i.e., CSSCA-based and AO-based algorithms. Evidently, we see from Fig. 5 that the momentum-accelerated PGD algorithm is the fastest algorithm. Particularly, by using momentum, the perturbations are significantly reduced. This agrees with the intuition that the momentum term allows for finer convergence to take place, helpful in finding a global minimizer.

VI. Conclusion

In this paper, a new way of alleviating the effect of the blockages in the RIS-assisted mmWave MIMO communication system is proposed. Our analysis highlights the importance of the N-P criterion-based detector in substantially alleviating the impact of the blockages. Preliminary results have proven that with this integrated blockage sensing capability, the detection of mmWave link availability could be tractable. We further confirm that by

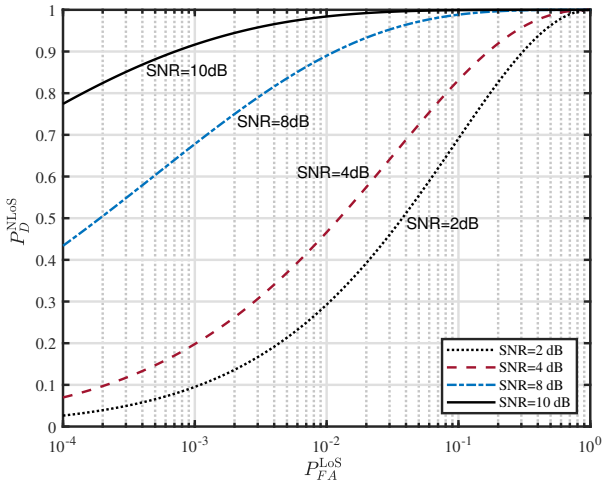

Fig. 2: Blockage detection performance.

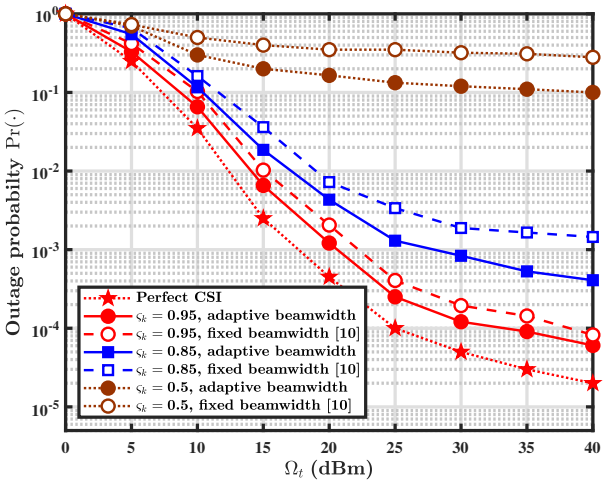

Fig. 4: Achievable rate performance.

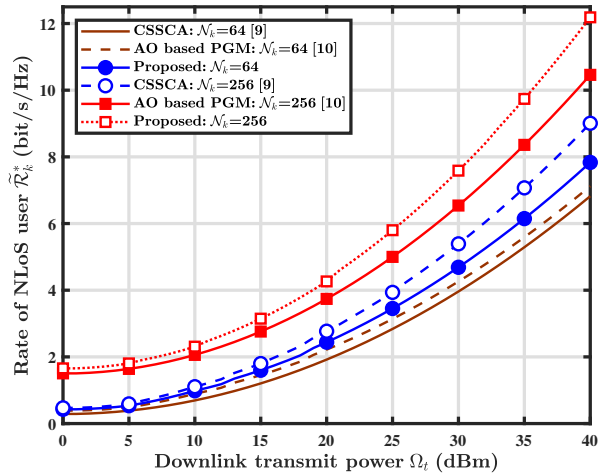

Fig. 3: Outage performance.

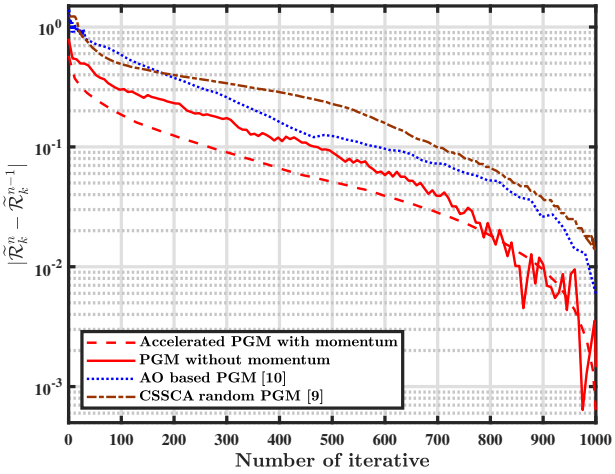

Fig. 5: Convergence rates.

leveraging the momentum-based optimization algorithms, the proposed PGD algorithm significantly outperforms the existing baseline algorithms and provably converges to the global optimum.

REFERENCES

[1] I. K. Jain, R. Kumar, and S. S. Panwar. The impact of mobile blockers on millimeter wave cellular systems. *IEEE J. Sel. Areas Commun.*, 37(4):854–868, April 2019.

[2] V. Raghavan et al. Hand and body blockage measurements with form-factor user equipment at 28 ghz. *IEEE Trans. Antennas Propag.*, 70(1):607–620, January 2022.

[3] Q. Wu and R. Zhang. Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming. *IEEE Commun. Mag.*, 18(11):5394–5409, November 2019.

[4] M. D. Renzo, A. Zappone, M. Debbah, M. Alouini, C. Yuen, J. Rosny, and S. Tretyakov. Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and road ahead. *IEEE J. Sel. Areas Commun.*, 38(11):2450–2525, November 2020.

[5] L. Jiao, P. Wang, A. Alipour-Fanid, H. Zeng, and K. Zeng. Enabling efficient blockage-aware handover in RIS-assisted mmwave cellular networks. *IEEE Trans. Wireless Commun.*, 21(4):2243–2257, April 2022.

[6] Y. Bengio. Gradient-based optimization of hyperparameters. *Neural Computation*, 12(8):1889–1900, August 2000.

[7] N. S. Perović, L. Tran, M. Di Renzo, and M. F. Flanagan. Achievable rate optimization for MIMO systems with reconfigurable intelligent surfaces. *IEEE Trans. Wireless Commun.*, 20(6):3865–3882, June 2021.

[8] M. Zhao, A. Liu, and R. Zhang. Outage-constrained robust beamforming for intelligent reflecting surface aided wireless communication. *IEEE Trans. Signal Process.*, 69.

[9] M. Shi, X. Li, T. Fan, J. Liu, and S. Lv. Intelligent reflecting surface-assisted millimeter wave communications: Joint active and passive precoding design. *IEEE Trans. Wireless Commun.*, 11(10):2215–2219, April 2022.

[10] S. Zhang and R. Zhang. Capacity characterization for intelligent reflecting surface aided mimo communication. *IEEE J. Sel. Areas Commun.*, 38(8):1823–1838, August 2020.