Stochastic Learning-Based Robust Beamforming Design for RIS-Aided Millimeter-Wave Systems in the Presence of Random Blockages

Gui Zhou, Cunhua Pan, Hong Ren, Kezhi Wang, Maged Elkashlan, and Marco Di Renzo, *Fellow, IEEE*

Abstract

A fundamental challenge for millimeter wave (mmWave) communications lies in its sensitivity to the presence of blockages, which impact the connectivity of the communication links and ultimately the reliability of the entire network. In this paper, we analyze a reconfigurable intelligent surface (RIS)-aided mmWave communication system for enhancing the network reliability and connectivity in the presence of random blockages. To enhance the robustness of hybrid analog-digital beamforming in the presence of random blockages, we formulate a stochastic optimization problem based on the minimization of the sum outage probability. To tackle the proposed optimization problem, we introduce a low-complexity algorithm based on the stochastic block gradient descent method, which learns sensible blockage patterns without searching for all combinations of potentially blocked links. Numerical results confirm the performance benefits of the proposed algorithm in terms of outage probability and effective data rate.

Index Terms

Reconfigurable intelligent surface (RIS), intelligent reflecting surface (IRS), millimeter wave communications, stochastic optimization, robust beamforming design.

I. Introduction

Due to the abundance of available frequency bandwidth, millimeter wave (mmWave) communication is envisioned to be a potential technology to meet the high data rate demand of current and future

G. Zhou, C. Pan, H. Ren and M. Elkashlan are with the School of Electronic Engineering and Computer Science at Queen Mary University of London, London E1 4NS, U.K. (e-mail: g.zhou, c.pan, h.ren, maged.elkashlan@qmul.ac.uk). K. Wang is with Department of Computer and Information Sciences, Northumbria University, UK. (e-mail: kezhi.wang@northumbria.ac.uk). M. Di Renzo is with Université Paris-Saclay, CNRS and CentraleSupélec, Laboratoire des Signaux et Systèmes, Gif-sur-Yvette, France. (e-mail: marco.direnzo@centralesupelec.fr). (Corresponding author: Cunhua Pan)

wireless networks. In addition, due to the small wavelength, a large number of antenna elements can be packed in antenna arrays of reasonable size, which can compensate for the severe path loss caused by the high transmission frequency and can mitigate the inter-user interference by capitalizing on the design of high-directional beams [1]. Furthermore, the use of hybrid analog-digital beamforming reduces the high cost and power consumption of using several radio frequency (RF) chains and full-digital signal processing baseband units at mmWave base stations (BSs) [2].

However, a major challenge in mmWave communications lies in the high penetration loss and low diffraction of the signals [3]. This makes mmWave communications highly sensible to the presence of spatial blockages (e.g., buildings, human beings, etc.) and degrades the reliability of the communication links and the connectivity of the entire network. To address this difficulty, new robust beamforming designs, tackling the channel uncertainties due to the presence of random blockages, have been investigated in recent literature [4]–[6]. In particular, [4] demonstrated the feasibility of obtaining the blockage probability by learning the spatiotemporal image sensing side-information. By capitalizing on the predicted blockage probability, [5] proposed a worst-case robust coordinated multipoint (CoMP) beamforming design by considering all possible combinations of blockage patterns. To reduce the complexity and improve the robustness of mmWave communication that suffers from the presence of random blockages, an outage-minimum strategy based on a stochastic optimization method was proposed in [6].

However, the above mentioned methods are only suitable for CoMP scenarios, in which the space diversity gain brought by multiple BSs compensates for the link outage loss caused by the presence of random blockages. Deploying multiple BSs, however, increases the hardware cost and power consumption. To overcome these challenges, the emerging technology of reconfigurable intelligent surfaces (RISs) was recently proposed as a promising solution for establishing alternative communication routes at a low cost, high energy efficiency, and high reliability [7]–[12]. An RIS is a thin surface made of nearly-passive and reconfigurable scatterers, which are capable of modifying the incident radio waves, so as to enhance the received signals at some specified locations [13]–[18].

In this paper, motivated by these considerations, we study an mmWave RIS-aided communication system, and optimize the hybrid analog-digital beamforming at the BS and the passive beamforming at the RIS by explicitly taking into account the presence of random spatial blockages. To enhance the robustness of the considered system, we study a sum-outage probability minimization problem that minimizes the rate at which the system quality of service (QoS), which is quantified in terms of outage probability, is not fulfilled. The formulated stochastic optimization problem with highly coupled variables is solved by using the block stochastic gradient descent (BSGD) method. The obtained

simulation results show the performance benefits of the proposed scheme in terms of outage probability and effective data rate.

II. System Model

A. Signal Model

We consider the downlink of an mmWave system in which a BS equipped with a uniform linear array (ULA) with N antennas and N_{RF} RF chains serves K single-antenna users (denoted by $\mathcal{K} \triangleq \{1, ..., K\}$) in the presence of random blockages, where $K \leq N_{RF} \ll N$. We assume that an RIS, e.g., deployed on the facade of a building, has the capability of passively reflecting the signals transmitted from the BS to the users. The RIS is made of M passive reflecting elements that are arranged in a uniform planar array (UPA). It is assumed that the phase shifts of the RIS are calculated by the BS and are then sent back to the RIS controller through dedicated control channels [13], [14]. The BS adopts a hybrid precoding architecture, in which each RF chain is connected to all antennas, and transmits Gaussian data symbols $\mathbf{s} = [s_1, \cdots, s_K]^{\mathrm{T}} \in \mathbb{C}^{K\times 1}$ to the users through a digital precoding matrix $\mathbf{D} = [\mathbf{d}_1, \cdots, \mathbf{d}_K] \in \mathbb{C}^{N_{RF}\times K}$ and an analog precoding matrix $\mathbf{A} \in \mathbb{C}^{N\times N_{RF}}$. The transmit power of the BS fulfills the constraint $||\mathbf{AD}||_F^2 \leq P_{max}$, where P_{max} is the total transmit power limit. Each entry of $\mathbf{A}$ is constrained to have a unit modulus, i.e., $\mathbf{A} \in \mathcal{S}_A$ where $\mathcal{S}_A \triangleq \{\mathbf{A} || [\mathbf{A}]_{m,n}|^2 = 1, \forall m, n\}$, and the symbol $[\cdot]_{m,n}$ denotes the (m, n)-th element of a matrix. The received signal intended to the k-th user is expressed as

$$y_k = (\mathbf{h}_{\mathrm{b},k}^{\mathrm{H}} + \mathbf{h}_{\mathrm{i},k}^{\mathrm{H}} \mathbf{E} \mathbf{H}_{\mathrm{bi}}) \mathbf{A} \mathbf{D} \mathbf{s} + n_k, \tag{1}$$

where $\mathbf{E} = \mathrm{diag}([e_1, \ldots, e_M])$ is the reflection coefficient matrix (also known as the passive beamforming matrix) of the RIS and $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ is the additive white Gaussian noise (AWGN). The channels of the BS-user k, BS-RIS, and RIS-user k links are denoted by $\mathbf{h}_{\mathrm{b},k} \in \mathbb{C}^{N\times 1}$, $\mathbf{H}_{\mathrm{bi}} \in \mathbb{C}^{M\times N}$ and $\mathbf{h}_{\mathrm{i},k} \in \mathbb{C}^{M\times 1}$, respectively.

Furthermore, denoting by $\mathbf{H}_k = \begin{bmatrix} \mathrm{diag}(\mathbf{h}_{\mathrm{i},k}^{\mathrm{H}}) \mathbf{H}_{\mathrm{bi}} \\ \mathbf{h}_{\mathrm{b},k}^{\mathrm{H}} \end{bmatrix} \in \mathbb{C}^{(M+1)\times N}$ the equivalent channel that accounts for the BS to the user k and by $\mathbf{e} = [e_1, \ldots e_M, 1]^{\mathrm{T}} \in \mathbb{C}^{(M+1)\times 1}$ the equivalent reflection coefficient vector that belongs to the set $\mathcal{S}_e = \{\mathbf{e} \mid |e_m|^2 = 1, 1 \leq m \leq M, e_{M+1} = 1\}$, (1) can be rewritten as $y_k = \mathbf{e}^{\mathrm{H}} \mathbf{H}_k \mathbf{A} \mathbf{D} \mathbf{s} + n_k, \forall k \in \mathcal{K}$, and the corresponding achievable signal-to-interference-plus-noise ratio (SINR), $\Gamma_k(\mathbf{D}, \mathbf{A}, \mathbf{e})$, can be written as

$$\Gamma_k(\mathbf{D}, \mathbf{A}, \mathbf{e}) = \frac{|\mathbf{e}^{\mathrm{H}} \mathbf{H}_k \mathbf{A} \mathbf{d}_k|^2}{\sum_{i\neq k}^{K} |\mathbf{e}^{\mathrm{H}} \mathbf{H}_k \mathbf{A} \mathbf{d}_i|^2 + \sigma_k^2}. \tag{2}$$

B. Channel Model

Based on [2], we assume that the mmWave channel is characterized by a limited scattering and that a geometric channel model can be used. Assume that there are L_{BU}, L_{BI} and L_{IU} propagation paths for the BS-user links, the BS-RIS links and the RIS-user links respectively, we have

$$\mathbf{h}_{\mathrm{b},k} = \sqrt{\frac{1}{L_{BU}}} \sum_{l=1}^{L_{BU}} g_{k,l}^{\mathrm{b}} \mathbf{a}_L \left(\theta_{k,l}^{\mathrm{b},t}\right), \forall k \in \mathcal{K}, \tag{3}$$

$$\mathbf{h}_{\mathrm{i},k} = \sqrt{\frac{1}{L_{IU}}} \sum_{l=1}^{L_{IU}} g_{k,l}^{\mathrm{i}} \mathbf{a}_P \left(\theta_{k,l}^{\mathrm{i},t}, \phi_{k,l}^{\mathrm{i},t}\right), \forall k \in \mathcal{K}, \tag{4}$$

$$\mathbf{H}_{\mathrm{bi}} = \sqrt{\frac{1}{L_{BI}}} \sum_{l=1}^{L_{BI}} g_{l}^{\mathrm{bi}} \mathbf{a}_P \left(\theta_{l}^{\mathrm{i},r}, \phi_{l}^{\mathrm{i},r}\right) \mathbf{a}_L \left(\theta_{l}^{\mathrm{b},t}\right)^{\mathrm{H}}. \tag{5}$$

In (3)-(5), $\{g_{k,l}^{\mathrm{b}}, g_{k,l}^{\mathrm{i}}, g_{l}^{\mathrm{bi}}\}$ denote the large-scale fading coefficients. Define $g \in \{g_{k,l}^{\mathrm{b}}$ for $\forall k, g_{k,l}^{\mathrm{i}}$ for $\forall k, g_{l}^{\mathrm{bi}}\}$, then g has distribution $\mathcal{CN}(0, 10^{\frac{\mathrm{PL}}{10}})$, where $\mathrm{PL} = -C_0 - 10\alpha \log_{10}(D)$ dB, $C_0 = 30$ dB is the path loss at a reference distance of one meter, D is the link distance (in meters) and α is the pathloss exponent. The steering vector of the ULA is given by $\mathbf{a}_L(\theta) = [1, e^{j\frac{2\pi}{\lambda_c} d \sin(\theta)}, \ldots, e^{j\frac{2\pi}{\lambda_c} d(N-1)\sin(\theta)}]^{\mathrm{H}}$, where θ stands for the azimuth angle-of-departure (AoD) $\theta_{k,l}^{\mathrm{b},t}$ and $\theta_{l}^{\mathrm{b},t}$, d is the antenna inter-distance, and λ_c is the carrier wavelength. The steering vector of the UPA is $\mathbf{a}_P(\theta, \phi) = [1, \ldots, e^{j\frac{2\pi}{\lambda_c} d(p \sin(\theta)\sin(\phi) + q\cos(\theta)\sin(\phi))}, \ldots]^{\mathrm{H}}$, where $0 \leq \{p, q\} \leq \sqrt{M} - 1$, $\theta(\phi)$ is the azimuth (elevation) AoD $\theta_{k,l}^{\mathrm{i},t}(\phi_{k,l}^{\mathrm{i},t})$ and the angle-of-arrival (AoA) $\theta_{l}^{\mathrm{i},r}(\phi_{l}^{\mathrm{i},r})$.

According to [3], [5], [6], the BS-user links may experience the presence of a random blockage with 20%-60% probability, while the RIS-aided links can be assumed not to be affected by blockages, since the locations of the RISs can be appropriately optimized in order to ensure line of sight transmission. Due to the fact that the communication links in the mmWave frequency band are sensitive to the presence of blockages, the achievable data rate can be significantly reduced. In order to investigate the impact of the channel uncertainties caused by the presence of random blockages, we adopt a recently proposed probabilistic model for the BS-user links [6]. In particular, the channels between the BS and the users are modeled as

$$\mathbf{h}_{\mathrm{b},k} = \sqrt{\frac{1}{L_{BU}}} \sum_{l=1}^{L_{BU}} \omega_{k,l} g_{k,l}^{\mathrm{b}} \mathbf{a}_L \left(\theta_{k,l}^{\mathrm{b},t}\right), \forall k \in \mathcal{K}, \tag{6}$$

where the random variable $\omega_{k,l} \in \{0, 1\}$ is a blockage parameter that is distributed according to a Bernoulli distribution. In particular, the corresponding blockage probability is denoted by $p_{k,l}$. In this work, we assume that the blockage probability is known at the BS and that it is used for robust beamforming design.

C. Problem Formulation

In the presence of random blockages, we aim to propose a robust mmWave beamforming scheme with the objective of minimizing the sum outage probability [6]. The formulated optimization problem is given by

$$\min_{\mathbf{D},\mathbf{A},\mathbf{e}} \ \sum_{k\in\mathcal{K}} \Pr\{\Gamma_k\left(\mathbf{D},\mathbf{A},\mathbf{e}\right) \leq \gamma_k\} \tag{7a}$$

$$\text{s.t.} \ \ ||\mathbf{AD}||_F^2 \leq P_{max} \tag{7b}$$

$$\mathbf{A} \in \mathcal{S}_A \tag{7c}$$

$$\mathbf{e} \in \mathcal{S}_e, \tag{7d}$$

where $\gamma_k > 0$ is the SINR reliability threshold of user k and $\Pr\{\Gamma_k\left(\mathbf{D},\mathbf{A},\mathbf{e}\right) \leq \gamma_k\}$ denotes the probability that the required SINR cannot be satisfied.

Note that in contrast to the traditional deterministic formulation which might not always have feasible solutions for the QoS constraints, Problem (7) has always a feasible solution that yields the QoS target minimum outage probability.

III. Beamforming design

Problem (7) is challenging to solve due to the absence of a closed-form expression for the objective function (7a), the non-convex unit-modulus constraints (7c) and (7d), and that fact that the optimization variables are tightly coupled. In the following, we propose a robust beamforming design algorithm under the stochastic-learning-based alternating optimization (AO) framework.

A. Problem Transformation

To start with, we rewrite $\Pr\{\Gamma_k\left(\mathbf{D},\mathbf{A},\mathbf{e}\right) \leq \gamma_k\}$ as $\mathbb{E}_{\mathbf{H}_k}[\mathbb{I}_{\Gamma_k\leq\gamma_k}]$, where $\mathbb{I}_{\Gamma_k\leq\gamma_k}$ is the step function. Since the step function is non-differentiable, we approximate it with the following generalized smooth hinge surrogate function [19]

$$u_k\left(\mathbf{X}\right) = \begin{cases} 0 & \text{if } 1-\frac{\Gamma_k(\mathbf{X})}{\gamma_k} < 0 \\ \frac{1}{2\epsilon}\left(1-\frac{\Gamma_k(\mathbf{X})}{\gamma_k}\right)^2 & \text{otherwise} \\ 1-\frac{\Gamma_k(\mathbf{X})}{\gamma_k}-\frac{\epsilon}{2} & \text{if } 1-\frac{\Gamma_k(\mathbf{X})}{\gamma_k} > \epsilon, \end{cases}$$

where $0 < \epsilon \ll 1$ and $\mathbf{X} \triangleq \{\mathbf{D}, \mathbf{A}, \mathbf{e}\}$ denotes the collection of all variables. By replacing the step function $\mathbb{I}_{\Gamma_k \leq \gamma_k}$ with its smooth approximation $u_k(\mathbf{X})$, we obtain the following approximated reformulation for Problem (7)

$$\min_{\mathbf{X}} \ \sum_{k \in \mathcal{K}} \mathbb{E}_{\mathbf{H}_k}[u_k(\mathbf{X})] \tag{8a}$$

$$\text{s.t. } (7b), (7c), (7d). \tag{8b}$$

Problem (8) can be viewed as a risk minimization problem [20], which was studied thoroughly in several key application areas such as wireless resource optimization, compressive sensing, machine learning, etc. It can be solved efficiently by using stochastic approximation methods, which are widely adopted due to their easy-to-implement feature, although they may requite a high complexity and a large memory.

Since the BS is assumed to know the channel state information (CSI), namely the channel gain and AoDs, the expectation in (8a) is computed with respect to the blockage parameter $\omega_{k,l}$ in (6). By assuming that the blockage probability can be predicted with accuracy, we can generate $\omega_{k,l}$ randomly and can construct the training data sample set $\mathcal{H} = \{\mathbf{H}_k^{(t)}, \forall k \in \mathcal{K}\}_{t=1}^{T}$ that is utilized for stochastic optimization. Accordingly, an appropriate surrogate for the risk function $\mathbb{E}_{\mathbf{H}_k}[u_k(\mathbf{X})]$ is often assumed to be $\frac{1}{T}\sum_{t=1}^{T} u_k(\mathbf{X}, \mathbf{H}_k^{(t)})$, which is known as the empirical risk function [20]. Therefore, the resulting empirical risk minimization (ERM) problem can be formulated as follows

$$\min_{\mathbf{X}} \ \frac{1}{T}\sum_{t=1}^{T}\sum_{k \in \mathcal{K}} u_k(\mathbf{X}, \mathbf{H}_k^{(t)}) \tag{9a}$$

$$\text{s.t. } (7b), (7c), (7d). \tag{9b}$$

A popular method to solve ERM problems is the stochastic gradient descent (SGD) [21], which we leverage to solve Problem (9), by alternately optimizing one of the block variables $\{\mathbf{D}, \mathbf{A}, \mathbf{e}\}$ while keeping the others fixed. This algorithm is known as the block stochastic gradient descent (BSGD) method [22]. In particular, the solution at each iteration is updated based on the recursion formulas

$$\mathbf{D}^{(t)} = \mathbf{D}^{(t-1)} - \alpha_t \sum_{k \in \mathcal{K}} \nabla_{\mathbf{D}} u_k\left(\mathbf{X}, \mathbf{H}_k^{(t)}\right), \tag{10a}$$

$$\mathbf{A}^{(t)} = \mathbf{A}^{(t-1)} - \alpha_t \sum_{k \in \mathcal{K}} \nabla_{\mathbf{A}} u_k\left(\mathbf{X}, \mathbf{H}_k^{(t)}\right), \tag{10b}$$

$$\mathbf{e}^{(t)} = \mathbf{e}^{(t-1)} - \alpha_t \sum_{k \in \mathcal{K}} \nabla_{\mathbf{e}} u_k\left(\mathbf{X}, \mathbf{H}_k^{(t)}\right), \tag{10c}$$

where $\nabla_{\mathbf{x}}$ for $\mathbf{x} \in \{\mathbf{D}, \mathbf{A}, \mathbf{e}\}$ denotes the stochastic gradient, and the step size $\{\alpha_t \in (0,1]\}$ is a decreasing sequence satisfying $\alpha_t \to 0$, $\sum_{t=1}^{\infty} \alpha_t \to \infty$ and $\sum_{t=1}^{\infty} \alpha_t^2 \to \infty$. It is worth noting that

the normalization factor $1/N$ in (10) is omitted, since the stochastic gradient $\nabla_{\mathbf{x}} u_k(\mathbf{X}, \mathbf{H}_k^{(t)})$ is an unbiased estimate of the batch gradient $\frac{1}{T}\sum_{t=1}^{T} \nabla_{\mathbf{x}} u_k(\mathbf{X}, \mathbf{H}_k^{(t)})$ [20], i.e., $\mathbb{E}_{\mathbf{H}_k}[\nabla_{\mathbf{x}} u_k(\mathbf{X}, \mathbf{H}_k^{(t)})|\mathbf{x}^{(t-1)}] = \frac{1}{T}\sum_{t=1}^{T} \nabla_{\mathbf{x}} u_k(\mathbf{X}, \mathbf{H}_k^{(t)})$.

The partial gradient of $u_k\left(\mathbf{X}, \mathbf{H}_k^{(t)}\right)$ with respect to $\mathbf{x}$ is

$$\nabla_{\mathbf{x}} u_k(\mathbf{X}, \mathbf{H}_k^{(t)}) = \begin{cases} 0 & \text{if } 1 - \frac{\Gamma_k\left(\mathbf{X}, \mathbf{H}_k^{(t)}\right)}{\gamma_k} < 0 \\ \frac{\frac{\Gamma_k\left(\mathbf{X}, \mathbf{H}_k^{(t)}\right)}{\gamma_k} - 1}{\epsilon} \frac{\nabla_{\mathbf{x}} \Gamma_k\left(\mathbf{X}, \mathbf{H}_k^{(t)}\right)}{\gamma_k} & \text{otherwise} \\ -\frac{\nabla_{\mathbf{x}} \Gamma_k\left(\mathbf{X}, \mathbf{H}_k^{(t)}\right)}{\gamma_k} & \text{if } 1 - \frac{\Gamma_k\left(\mathbf{X}, \mathbf{H}_k^{(t)}\right)}{\gamma_k} > \epsilon. \end{cases} \tag{11}$$

In order to compute the partial gradient $\nabla \Gamma_k\left(\mathbf{x}, \mathbf{H}_k^{(t)}\right)$ with respect to any $\mathbf{x} \in \{\mathbf{D}, \mathbf{A}, \mathbf{e}\}$, it is necessary to apply some mathematical transformations to the SINR in (2). In particular, we have

$$\Gamma_k\left(\mathbf{X}, \widetilde{\mathbf{h}}_k^t\right) = \frac{\mathrm{vec}(\mathbf{D})^{\mathrm{H}} \mathbf{Q}_{\mathbf{D},k}^{(t)} \mathrm{vec}(\mathbf{D})}{\mathrm{vec}(\mathbf{D})^{\mathrm{H}} \overline{\mathbf{Q}}_{\mathbf{D},k}^{(t)} \mathrm{vec}(\mathbf{D}) + \sigma_k^2} \tag{12a}$$

$$= \frac{\mathrm{vec}(\mathbf{A})^{\mathrm{H}} \mathbf{Q}_{\mathbf{A},k}^{(t)} \mathrm{vec}(\mathbf{A})}{\mathrm{vec}(\mathbf{A})^{\mathrm{H}} \overline{\mathbf{Q}}_{\mathbf{A},k}^{(t)} \mathrm{vec}(\mathbf{A}) + \sigma_k^2} \tag{12b}$$

$$= \frac{\mathbf{e}^{\mathrm{H}} \mathbf{Q}_{\mathbf{e},k}^{(t)} \mathbf{e}}{\mathbf{e}^{\mathrm{H}} \overline{\mathbf{Q}}_{\mathbf{e},k}^{(t)} \mathbf{e} + \sigma_k^2} \tag{12c}$$

with

$$\begin{aligned} \mathbf{Q}_{\mathbf{D},k}^{(t)} &= \mathrm{diag}(\mathbf{i}_k) \otimes \mathbf{A}^{\mathrm{H}} \mathbf{H}_k^{(t),\mathrm{H}} \mathbf{e}\mathbf{e}^{\mathrm{H}} \mathbf{H}_k^{(t)} \mathbf{A}, \\ \overline{\mathbf{Q}}_{\mathbf{D},k}^{(t)} &= \mathrm{diag}(\bar{\mathbf{i}}_k) \otimes \mathbf{A}^{\mathrm{H}} \mathbf{H}_k^{(t),\mathrm{H}} \mathbf{e}\mathbf{e}^{\mathrm{H}} \mathbf{H}_k^{(t)} \mathbf{A}, \\ \mathbf{Q}_{\mathbf{A},k}^{(t)} &= \mathbf{d}_k \mathbf{d}_k^{\mathrm{H}} \otimes \mathbf{H}_k^{(t),\mathrm{H}} \mathbf{e}\mathbf{e}^{\mathrm{H}} \mathbf{H}_k^{(t)}, \\ \overline{\mathbf{Q}}_{\mathbf{A},k}^{(t)} &= \mathbf{D}_{-k} \mathbf{D}_{-k}^{\mathrm{H}} \otimes \mathbf{H}_k^{(t),\mathrm{H}} \mathbf{e}\mathbf{e}^{\mathrm{H}} \mathbf{H}_k^{(t)}, \\ \mathbf{Q}_{\mathbf{e},k}^{(t)} &= \mathbf{H}_k^{(t)} \mathbf{A} \mathbf{d}_k \mathbf{d}_k^{\mathrm{H}} \mathbf{A}^{\mathrm{H}} \mathbf{H}_k^{(t),\mathrm{H}}, \\ \overline{\mathbf{Q}}_{\mathbf{e},k}^{(t)} &= \mathbf{H}_k^{(t)} \mathbf{A} \sum_{i \neq k}^{K} \mathbf{d}_i \mathbf{d}_i^{\mathrm{H}} \mathbf{A}^{\mathrm{H}} \mathbf{H}_k^{(t),\mathrm{H}}, \end{aligned}$$

where $\mathbf{i}_k$ is the k-th column of the $K \times K$ identity matrix $\mathbf{I}_K$, $\bar{\mathbf{i}}_k$ denotes the one's complement of $\mathbf{i}_k$, and $\otimes$ is the Kronecker product.

Since equations (12a)-(12c) have the same mathematical structure, the unified partial gradient for any $\mathbf{x}$ is given by

$$\nabla_{\mathbf{x}} \Gamma_k\left(\mathbf{X}, \widetilde{\mathbf{h}}_k^t\right) = \frac{\mathbf{Q}_{\mathbf{x},k}^t \mathbf{x}}{t} - \frac{\mathbf{x}^{\mathrm{H}} \mathbf{Q}_{\mathbf{x},k}^t \mathbf{x}}{t^2} \overline{\mathbf{Q}}_{\mathbf{x},k}^t \mathbf{x},$$

where $t = \mathbf{x}^{\mathrm{H}}\overline{\mathbf{Q}}_{\mathbf{x},k}^{t}\mathbf{x} + \sigma_k^2$.

B. Algorithm Description

Algorithm 1 summarizes the proposed BSGD-based outage minimum robust hybrid beamforming design for RIS-aided mmWave systems in which the BS-user links undergo random blockages. The proposed algorithm is referred to as BSGD-OutMin.

Algorithm 1 BSGD-OutMin Algorithm

Initialize: Initialize $\mathbf{D}^{(0)}$, $\mathbf{A}^{(0)}$, $\mathbf{e}^{(0)}$, and the data set $\mathcal{H}$. Set $t = 1$ and $T_{max} = 10^5$.

1: **repeat**
2: Sample the data $\mathbf{H}_k^{(t)}$ $\forall k \in \mathcal{K}$ from $\mathcal{H}$.
3: Compute $\nabla_{\mathbf{x}} u_k(\mathbf{X}, \mathbf{H}_k^{(t)})$ $\forall \mathbf{x} \in \{\mathbf{D}, \mathbf{A}, \mathbf{e}\}$ from (11).
4: Update $\mathbf{D}^{(t)}$ according to (10a).
5: Project $\mathbf{D}^{(t)}$ onto the feasible set in (7b) as $\mathbf{D}^{(t)} = \frac{\left(\mathbf{D}^{(t)}\right)}{||\mathbf{A}^{(t-1)}\mathbf{D}^{(t)}||_F}\sqrt{P_{max}}$.
6: Update $\mathbf{A}^{(t)}$ according to (10b).
7: Project $\mathbf{A}^{(t)}$ onto the feasible set $\mathcal{S}_A$ as $\mathbf{A}^{(t)} = \exp\left\{j\angle\mathbf{A}^{(t)}\right\}$, where $j \triangleq \sqrt{-1}$ is the imaginary unit and $\angle(\cdot)$ denote the angle of a complex number.
8: Update $\mathbf{e}^{(t)}$ according to (10c).
9: Project $\mathbf{e}^{(t)}$ onto the feasible set $\mathcal{S}_e$ as $\mathbf{e}^{(t)} = \exp\left\{j\angle\left(\mathbf{e}^{(t)}/\left[\mathbf{e}^{(t)}\right]_{M+1}\right)\right\}$.
10: $t = t + 1$.
11: **until** The objective value in (9a) converges.

a) Convergence analysis: The projections in Step 5, Step 7 and Step 9 pose challenges to the convergence analysis of the BSGD-OutMin algorithm. According to [22], the algorithm can be guaranteed to converge to a stationary point when the step size is set as $\alpha_t \in (0, 1/L]$ where L is the Lipschitz constant. However, it is difficult to find the Lipschitz constants corresponding to the variables $\{\mathbf{D}, \mathbf{A}, \mathbf{e}\}$ due to the complex objective in (9a). By using a sufficiently small step size, however, our simulation results show that Algorithm 1 is convergent.

b) Initial point: Problem (7) has, in general, multiple local minima points due to the non-convex unit-modulus constraints $\mathbf{A} \in \mathcal{S}_A$ and $\mathbf{e} \in \mathcal{S}_e$. The accurate selection of the initial points in Algorithm 1 plays an important role for the convergence speed and the optimality of the obtained local solution. To

that end, we first initialize $\mathbf{e}$ to maximize the equivalent total channel gain, resulting in the following optimization problem

$$\max_{\mathbf{e}\in\mathcal{S}_e} \sum_{k\in\mathcal{K}} ||\mathbf{e}^{\mathrm{H}}\mathbf{H}_k||_2^2. \tag{13a}$$

The objective function in (13) is convex and can be linearized by using its first-order Taylor approximation, i.e., $\mathrm{Re}\{\mathbf{e}^{[n],\mathrm{H}}\mathbf{H}\mathbf{H}^{\mathrm{H}}\mathbf{e}\}-\mathbf{e}^{[n],\mathrm{H}}\mathbf{H}\mathbf{H}^{\mathrm{H}}\mathbf{e}^{[n]}$, where $\mathbf{H}=[\mathbf{H}_1,\ldots,\mathbf{H}_K]$ and $\mathbf{e}^{[n]}$ is the optimal solution obtained at the n-th iteration, and $\mathrm{Re}\{\cdot\}$ denotes the real part of a complex number. Therefore, the optimal solution to Problem (13) at the $(n+1)$-th iteration is

$$\mathbf{e}^{[n+1]} = \exp\left\{j\angle\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}\mathbf{e}^{[n]}/\left[\mathbf{H}\mathbf{H}^{\mathrm{H}}\mathbf{e}^{[n]}\right]_{M+1}\right)\right\}.$$

Furthermore, $\mathbf{A}$ is initialized to align the phases of the equivalent channel to the mmWave BS, i.e.,

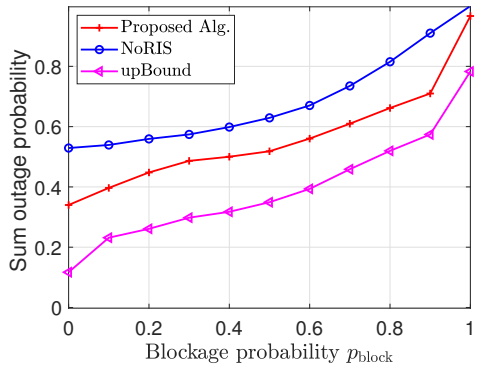
$$\mathbf{A}^{(0)} = \exp\left\{j\angle\left(\mathbf{H}(\mathbf{I}_K\otimes\mathbf{e}^{(0)})\right)\right\}.$$

IV. Numerical results and discussion

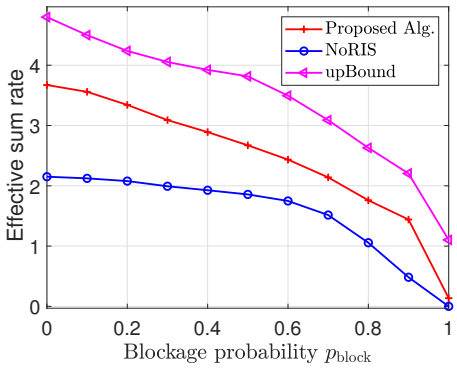
In this section, numerical results are illustrated to evaluate the performance of the proposed algorithm. All results are obtained by averaging over 500 channel realizations. Unless stated otherwise, we assume $N=32$, $K=N_{RF}=2$, $M=64$, $L_{BU}=L_{IU}=L_{BI}=5$, and the BS and the RIS are located at (0 m, 0 m) and (40 m, 10 m), respectively, and the users are assumed to be randomly distributed in a circle centered at (50 m, 0 m) with radius 5 m. The path-loss exponents of the three types of links are $\alpha_{BU}=3.6$ and $\alpha_{BI}=\alpha_{IU}=2.2$ [23]. Moreover, it is assumed that the active/passive antenna spacing at the BS/RIS is $d=\lambda_c/2$. The transmit power limit of the BS is $P_{max}=5$ W and the noise at each user is $\sigma_1^2=,\ldots,\sigma_K^2=-100$ dBm. For simplicity, it is assumed that the blockage probabilities are equal $p_{k,l}=p_{\mathrm{block}},\forall k,l$, and that the target SINR of all users is $\gamma_1=\ldots=\gamma_K=\gamma$, leading to the minimum target rate $R_{\mathrm{targ}}=\log_2(1+\gamma)$. To evaluate the performance of the proposed BSGD-OutMin algorithm, we consider two benchmark schemes: 1) upBound: full-digital precoding at the BS, given by $\mathbf{F}\in\mathbb{C}^{N\times K}$ with $||\mathbf{F}||_F^2\leq P_{max}$, and 2) NoRIS: there is no RIS in the considered system.

In order to demonstrate the robustness of the proposed algorithm, we consider two performance metrics: the sum outage probability and the effective sum rate. In particular, the sum outage probability is the sum of the achievable outage probabilities for all users, as formulated in Problem (7a). Then, the corresponding effective rate of user k is defined as $R_{\mathrm{eff},k}\triangleq\mathbb{E}[\log_2(1+\Gamma_k(\mathbf{D},\mathbf{A},\mathbf{e}))]$ if $\Gamma_k(\mathbf{D},\mathbf{A},\mathbf{e})\geq\gamma_k$ and $R_{\mathrm{eff},k}\triangleq 0$ otherwise.

Fig. 1 illustrates the impact of the blockage probability p_{block} on the system performance by assuming a target rate of $R_{\mathrm{targ}}=2$ bit/s/Hz. It is observed that an mmWave system in the absence of an RIS



(a) Outage probability



(b) Effective rate

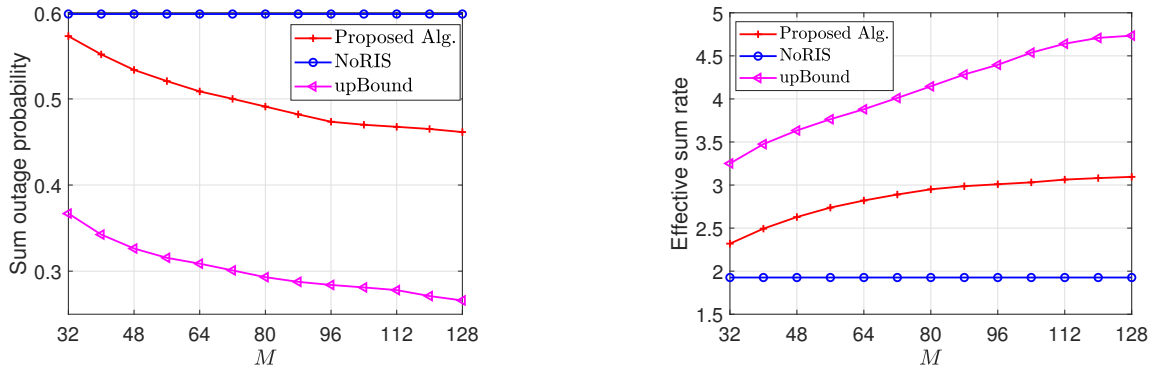
Fig. 1: Comparison of outage probability and effective rate as a function of the blockage probability p_{block} for $N = 32$, $M = 64$, $K = N_{RF} = 2$ and $R_{\text{targ}} = 2$ bit/s/Hz.

is highly sensitive to the presence of blockages, which results in the worst sum outage probability and effective sum rate. On the other hand, the proposed scheme provides a high data rate and a low outage probability over a wide range of blockage probabilities (i.e., $0\% \sim 40\%$), which substantiates the robustness of the proposed algorithm to the presence of random blockages. As expected, the performance of the adopted hybrid beamforming scheme is inferior to that of the upBound scheme that is obtained by using full-digital precoding. This is, however, obtained at the cost of increasing the power consumption and the hardware cost.

Fig. 2 illustrates the impact of the number of reflecting elements at the RIS in order to compensate for the performance loss caused by the hybrid beamforming design and the presence of random blockages. By increasing M, in particular, the sum outage probability decreases and the sum effective rate increases. Thus, the RIS plays an important role for improving the robustness and performance of mmWave systems in the presence of random blockages.

V. Conclusions

This work introduced a stochastic-learning-based robust beamforming design for RIS-aided mmWave systems which is aimed to combat the channel uncertainties caused by the presence of random blockages. The formulated stochastic optimization problem was solved by using the stochastic block gradient descent algorithm. Simulation results demonstrated the robustness of the proposed hybrid beamforming design in the presence of random blockages and confirmed its superior performance in terms of outage probability and effective data rate compared with baseline schemes.



(a) Outage probability

(b) Effective rate

Fig. 2: Comparison of outage probability and effective rate as a function of M for $N = 32$, $M = 64$, $K = N_{RF} = 2$, $R_{\text{targ}} = 2$ bit/s/Hz and $p_{\text{block}} = 40\%$.

REFERENCES

[1] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: it will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[2] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Top. Sign. Proces.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[3] V. Raghavan *et al.*, "Statistical blockage modeling and robustness of beamforming in millimeter-wave systems," *IEEE Trans. Microwave Theory Tech.*, vol. 67, no. 7, pp. 3010–3024, Jul. 2019.

[4] T. Nishio *et al.*, "Proactive received power prediction using machine learning and depth images for mmWave networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2413–2427, Nov. 2019.

[5] D. Kumar, J. Kaleva, and A. Tolli, "Rate and reliability trade-Off for mmWave communication via multi-point connectivity," in *IEEE GLOBECOM*, 2019, pp. 1–6.

[6] H. Iimori *et al.*, "Stochastic learning robust beamforming for millimeter-wave systems with path blockage," *IEEE Wireless Commun. Lett.*, pp. 1–1, 2020.

[7] M. Di Renzo *et al.*, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 798–807, 2020.

[8] E. Basar *et al.*, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019.

[9] M. Di Renzo *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *J. Wireless Commun. Netw.*, 2019,129(2019).

[10] ——, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and road ahead," *IEEE J. Sel. Areas Commun., early access*, 2020.

[11] N. S. Perović, M. Di Renzo, and M. F. Flanagan, "Channel capacity optimization using reconfigurable intelligent surfaces in indoor mmWave environments," in *IEEE ICC (2020)*, pp. 1–7.

[12] C. Huang *et al.*, "Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends," *IEEE Wireless Commun.*, pp. 1–8, Jul. 2020.

[13] 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.*, 2020.

[14] C. Pan, H. Ren, K. Wang *et al.*, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2020.

[15] T. Bai, C. Pan, Y. Deng *et al.*, "Latency minimization for intelligent reflecting surface aided mobile edge computing," *IEEE J. Sel. Areas Commun., early access*, 2020.

[16] A. Zappone, M. Di Renzo, F. Shams, X. Qian, and M. Debbah, "Overhead-aware design of reconfigurable intelligent surfaces in smart radio environments," *IEEE Trans. Wireless Commun. (to appear)*.

[17] G. Zhou, C. Pan, H. Ren *et al.*, "Robust beamforming design for intelligent reflecting surface aided MISO communication systems," *IEEE Wireless Commun. Lett.*, pp. 1–1, 2020.

[18] T. A. Le, T. Van Chien, and M. Di Renzo, "Robust probabilistic-constrained optimization for IRS-aided MISO communication systems," *IEEE Wireless Commun. Lett., early access*, 2020.

[19] P. L. Bartlett and M. H. Wegkamp, "Classification with a reject option using a hinge loss," *J. Machine Learning Research*, vol. 9, pp. 1823–1840, Jun. 2008.

[20] R. Xin, S. Kar, and U. A. Khan, "Decentralized stochastic optimization and machine learning: a unified variance-reduction framework for robust performance and fast convergence," *IEEE Signal Processing Mag.*, vol. 37, no. 3, pp. 102–113, 2020.

[21] L. Bottou, *Online Learning and Stochastic Approximations*. Cambridge Univ. Press, 1998.

[22] Y. Xu and W. Yin, "Block stochastic gradient iteration for convex and nonconvex optimization," *SIAM J. Optim.*, vol. 25, no. 3, pp. 1686–1716, 2015.

[23] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Intelligent reflecting surface aided multigroup multicast MISO communication systems," *IEEE Trans. Signal Process.*, vol. 68, pp. 3236–3251, 2020.