Fairness-Oriented Multiple RIS-Aided mmWave Transmission: Stochastic Optimization Methods

Gui Zhou, Cunhua Pan, Hong Ren, Kezhi Wang, Marco Di Renzo

Fairness-Oriented Multiple RIS-Aided mmWave Transmission: Stochastic Optimization Methods

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

Abstract

In millimeter wave (mmWave) systems, it is challenging to ensure the reliable connectivity of communication due to the high sensitivity to the presence of blockages. In order to improve the robustness of mmWave systems under the presence of random blockages, multiple reconfigurable intelligent surfaces (RISs) are deployed to enhance the spatial diversity gain, and robust beamforming is then designed based on a stochastic optimization for minimizing the maximum outage probability among multiple users to ensure fairness. Under the stochastic optimization framework, we adopt the stochastic majorization–minimization (SMM) method and the stochastic successive convex approximation (SSCA) method to construct deterministic surrogate problems at each iteration for new channel realizations, and to obtain closed-form solutions of the precoding matrix at the base station (BS) and the passive beamforming vectors at the RISs. Both stochastic optimization methods are proved to converge to the set of stationary points of the original stochastic problems. Finally, simulation results show that the proposed robust beamforming in RIS-aided systems can effectively compensate for the performance loss caused by the presence of random blockages, especially when the blockage probability is high, compared with benchmark solutions.

Index Terms

(Corresponding author: Cunhua Pan) G. Zhou and C. Pan 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@qmul.ac.uk). H. Ren is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China. (hren@seu.edu.cn). 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 UniversitÃ© Paris-Saclay, CNRS and CentraleSupÃ©lec, Laboratoire des Signaux et SystÃšmes, Gif-sur-Yvette, France. (e-mail: marco.di-renzo@universite-paris-saclay.fr).

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

I. Introduction

Millimeter wave (mmWave) communication is expected to be a promising technology to meet the growing demand for data rate in current and future wireless networks due to the abundant spectrum available at high frequencies. High frequency communication inevitably has severe attenuation, but this can be compensated by an array of reasonable size containing a large number of antennas due to the small wavelength [1]. In addition, the high-directional beams of large arrays are capable of mitigating the inter-user interference. However, mmWave communication systems suffer from high penetration loss [2]–[4]. Hence, mmWave systems are much more susceptible to the presence of spatial blockages than sub-6 GHz systems, and the reliability of the communication links for the entire network cannot always be guaranteed [2]–[4].

In particular, spatial blockages can be divided into static blockages (e.g., buildings and other immobile fixtures), dynamic blockages (e.g., humans, vehicles, or moving obstructions) and self-blockages (e.g., hand blocking of the user itself and blockage from other body parts). Some statistical models are established to characterize the properties of the random dynamic blockages and self-blockages [3]–[5]. The authors of [6] have developed a distance-dependent blockage probability model, in which the probability of link blocking increases exponentially with the link length. Furthermore, [7] and [8] proposed a mechanism to predict the blockage probability via machine learning. When the blockage probability is obtained, some robust beamforming design strategies have been proposed in the recent literature [9], [10] to address channel uncertainties caused by the presence of random blockages. Specifically, the authors of [9] proposed a worst-case robust beamforming design for application to coordinated multipoint (CoMP) systems in which all possible combinations of blockage patterns were considered. Due to the high computational complexity of the ergodic method in [9], an outage-minimum strategy based on a stochastic optimization method was proposed for CoMP systems in [10] to improve the robustness of mmWave systems. The idea of adopting multiple base stations (BSs) in CoMP systems is effective to compensate for the performance loss caused by the presence of random blockages by exploiting spatial diversity gains. However, this will incur excessive hardware cost and power consumption. Another promising scheme proposed in [11] is to deploy cost-

efficient reconfigurable intelligent surfaces (RISs) in mmWave systems to create an alternative communication link via the RISs.

Due to the promising advantages in terms of providing energy- and spectral-efficient communications, RISs have attracted extensive research attention from both academia and industry [12]–[15]. An RIS is a thin surface consisting of nearly-passive and reconfigurable reflecting elements, which reflects the impinging radio waves without adopting radio frequency (RF) chains. The passive elements on the RIS can be tuned to alter their electromagnetic response such that the signals reflected from an RIS can be constructively superimposed to enhance the signal power at the intended receiver or destructively combined to avoid the information leakage to undesired receivers. These characteristics make RISs appealing to be applied in various communication systems as shown in [16]. For instance, RISs can be applied in single-cell multiple-input and multiple-output (MIMO) systems [17]–[21], multicell MIMO communications [22], simultaneous wireless information and power transfer (SWIPT) systems [17], [23], secure communications [24], mmWave systems [25]–[28] and THz systems [29], [30].

Although the performance advantages of deploying an RIS in mmWave systems have been demonstrated in recent contributions, there still exist major open problems to solve. The authors of [26] only considered the BS-RIS-user channels and assumed that the direct BS-user communication links were completely blocked by obstacles. However, this assumption only applies to the case of static blockages with outage probability of 1, but not to the case of dynamic blockage with outage probability varying between 0 and 1 [2]–[4]. The numerical results in [27] showed that the gain from additional reflection channels could compensate for the performance loss caused by the presence of random blockages, but the impact of blockages was not considered in the beamforming design. Most recently, we have considered the robust beamforming design for RIS-aided mmWave communication systems in [11] by taking the random blockages into consideration. However, the objective function therein is to minimize the sum outage probability, which cannot ensure the fairness for all the users.

A. Novelty and contributions

Against the above background, this paper proposes a robust transmission strategy in an RIS-aided mmWave communication system to deal with the channel uncertainties caused by the random blockages while ensuring the fairness among the users. Some common approaches to handle the imperfect and partial CSI problem are the outage constrained robust optimization

and the worst-case robust optimization techniques [31]. However, both of them still need the estimation of the instantaneous CSI, and the worst-case robust method is conservative and hence suboptimal due to the low probability of occurrence for the worst case. An alternative approach is to design the robust beamforming by optimizing the statistical performance under a stochastic optimization framework in which only the long-term CSI is required. In what follows, we propose a maximum outage probability minimization problem to be solved by using the stochastic optimization framework.

Specifically, the main contributions of this work are summarized as follows:

- To the best of our knowledge, this is the first work exploring the robust beamforming design for an RIS-aided downlink multiuser mmWave system with the knowledge of long-term CSI and blockage probability. Specifically, our optimization objective in this work is to minimize the maximum outage probability of all the users. Different from the sum outage probability minimization problem in [10], [11], the min-max outage probability objective can ensure the quality of service (QoS) performance for the worst-case user in the downlink multiuser system. Due to the non-differentiable objective function, the stochastic gradient descent (SGD) method adopted in [10], [11] cannot be directly applied. To resolve this problem, two more general and powerful stochastic optimization frameworks are adopted to jointly optimize the active precoding at the BS and the passive beamforming at the RISs.
- The single user case is firstly considered to obtain insights for the robust beamforming design by minimizing the outage probability in the presence of long-term CSI and random blockages. Due to the non-deterministic expression of the probability function in the objective function, we replace it with an expectation function over the long-term CSI and blockage probability. Then, the inner step function in the expectation is approximated by a smooth function which is twice differentiable for the precoding at the BS and the passive beamforming at the RISs, respectively. The resulting expectation optimization problem is solved by adopting the stochastic majorization–minimization (SMM) method in which an upper bound surrogate function of the original differentiable function is constructed for a new channel realization at each iteration. The constructed surrogate problem has closed-form solutions and is computationally efficient. We prove that the proposed SMM method is guaranteed to converge to the set of stationary points of the original expectation minimization problem.
- The robust beamforming for a more general multiuser case is designed by solving a maxi-

mum outage probability minimization problem. To tackle the non-differentiability of the max objective function, we firstly replace it with its log-sum-exp upper bound. The stochastic successive convex approximation (SSCA) method is then adopted, which has more flexibility in the choice of the surrogate function than the SMM method, and generates closed-form solutions at each iteration. Furthermore, we prove that the final solution generated by the iterative algorithm is guaranteed to converge to the stationary point of the original expectation minimization problem.

- We demonstrate through numerical results that the proposed robust beamforming algorithm outperforms its non-robust counterpart and the robust beamforming in conventional RIS-free systems both in terms of maximum outage probability and minimum effective rate if the blockage probability is high. Moreover, deploying multiple small-size RISs is shown to be more effective than deploying a single large-size RIS in terms of improving the performance of the worst-case user.

The remainder of this paper is organized as follows. Section II introduces the system model. The outage probability minimization problem is formulated for a single-user case in Section III. Section IV further investigates the min-max outage probability problem for the multiuser system. Finally, Sections VI and VII report the numerical results and conclusions, respectively.

Notations: The following mathematical notations and symbols are used throughout this paper. Vectors and matrices are denoted by boldface lowercase letters and boldface uppercase letters, respectively. The symbols $\mathbf{X}^*$, $\mathbf{X}^{\mathrm{T}}$, $\mathbf{X}^{\mathrm{H}}$, and $||\mathbf{X}||_F$ denote the conjugate, transpose, Hermitian (conjugate transpose), Frobenius norm of matrix $\mathbf{X}$, respectively. The symbol $||\mathbf{x}||_2$ denotes 2-norm of vector $\mathbf{x}$. The symbols $\mathrm{Tr}\{\cdot\}$, $\mathrm{Re}\{\cdot\}$, $|\cdot|$, $\lambda(\cdot)$, and $\angle(\cdot)$ denote the trace, real part, modulus, eigenvalue, and angle of a complex number, respectively. $\mathrm{diag}(\mathbf{x})$ is a diagonal matrix with the entries of $\mathbf{x}$ on its main diagonal. $[\mathbf{x}]_m$ means the m-th element of the vector $\mathbf{x}$. The Kronecker product between two matrices $\mathbf{X}$ and $\mathbf{Y}$ is denoted by $\mathbf{X} \otimes \mathbf{Y}$. $\mathbf{X} \succeq \mathbf{Y}$means that $\mathbf{X} - \mathbf{Y}$ is positive semidefinite. Additionally, the symbol $\mathbb{C}$ denotes complex field, $\mathbb{R}$ represents real field, and $\mathrm{j} \triangleq \sqrt{-1}$ is the imaginary unit. The inner product $\langle \bullet, \bullet \rangle : \mathbb{C}^{M \times N} \times \mathbb{C}^{M \times N} \to \mathbb{R}$ is defined as $\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbb{R}\{\mathrm{Tr}\{\mathbf{X}^{\mathrm{H}}\mathbf{Y}\}\}$.

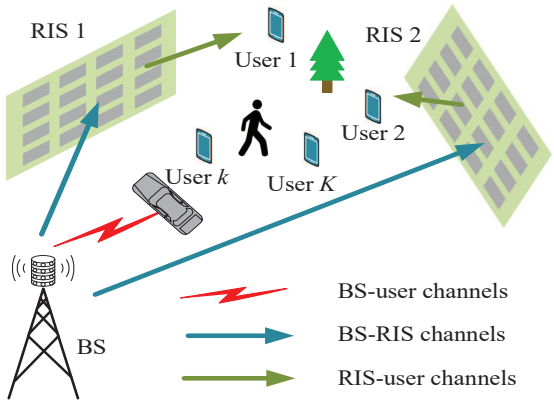


Fig. 1: Multiple RIS-aided mmWave communication system.

II. System Model

A. Signal Model

As shown in Fig. 1, we consider an RIS-aided downlink mmWave communication system. In order to ensure high QoS for the users in the presence of random blockages, U RISs, each of which has M passive antennas, are deployed to assist the communication from the BS equipped with N active antennas to K single-antenna users (denoted by $\mathcal{K} \triangleq \{1, \ldots, K\}$). The RISs are assumed to be connected to controllers that exchange control information with the BS through dedicated channels [17], [22]. The baseband transmitted signal at the BS is $\mathbf{x} = \mathbf{F}\mathbf{s}$, where $\mathbf{s} \in \mathbb{C}^{K\times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ is the Gaussian data symbol vector and $\mathbf{F} = [\mathbf{f}_1, \ldots, \mathbf{f}_K] \in \mathbb{C}^{N\times K}$ denotes the full-digital beamforming matrix. The baseband transmit power is limited to the total transmit power P_{max}. Hence, $\mathbf{F}$ belongs to the set $\mathcal{S}_f = \{\mathbf{F} \mid ||\mathbf{F}||_F^2 \leq P_{max}\}$.

Let $\mathbf{h}_{\mathrm{b},k} \in \mathbb{C}^{N\times 1}$, $\mathbf{H}_u \in \mathbb{C}^{M\times N}$ and $\mathbf{h}_{u,k} \in \mathbb{C}^{M\times 1}$ denote the channels of the BS-user k, BS-RIS u, and RIS u-user k links, respectively. Then, the received signal intended to the k-th user is expressed as

$$y_k = \left(\mathbf{h}_{\mathrm{b},k}^{\mathrm{H}} + \sum_{u=1}^{U} \mathbf{h}_{u,k}^{\mathrm{H}} \mathbf{E}_u \mathbf{H}_u\right) \mathbf{x} + n_k, \tag{1}$$

where $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ is the additive white Gaussian noise (AWGN), and $\mathbf{E}_u = \zeta \mathrm{diag}([e_{(u-1)M+1}, \ldots, e_{uM}])$ is the reflection coefficient matrix (also known as the passive beamforming matrix) of the u-th RIS. Element $e_{(u-1)M+m}$ is the unit modulus coefficient of the m-th phase shift at the u-th RIS, and $\zeta \in [0, 1]$ indicates the reflection efficiency. Here, it is assumed that $\zeta = 1$ for investigation of the performance upper bound of RIS.

Furthermore, after defining the compact matrices $\mathbf{h}_k = [\mathbf{h}_{1,k}^{\mathrm{H}}, \ldots, \mathbf{h}_{U,k}^{\mathrm{H}}]^{\mathrm{H}}$ and $\mathbf{H} = [\mathbf{H}_1^{\mathrm{H}}, \ldots, \mathbf{H}_U^{\mathrm{H}}]^{\mathrm{H}}$, we obtain the equivalent channel $\mathbf{G}_k = \begin{bmatrix} \mathrm{diag}(\mathbf{h}_k^{\mathrm{H}})\mathbf{H} \\ \mathbf{h}_{\mathrm{b},k}^{\mathrm{H}} \end{bmatrix} \in \mathbb{C}^{(UM+1)\times N}$ between the BS and the k-th user. The corresponding equivalent reflection coefficient vector is given by $\mathbf{e} = [e_1, \ldots, e_{UM+1}]^{\mathrm{T}} \in \mathbb{C}^{(UM+1)\times 1}$ which belongs to the set $\mathcal{S}_e = \{\mathbf{e} \mid |e_m|^2 = 1, 1 \le m \le UM, e_{UM+1} = 1\}$. Then, (1) can be rewritten into a compact form as

$$y_k = \mathbf{e}^{\mathrm{H}}\mathbf{G}_k\mathbf{F}\mathbf{s} + n_k, \tag{2}$$

and the corresponding achievable signal-to-interference-plus-noise ratio (SINR) is

$$\Gamma_k(\mathbf{F}, \mathbf{e}) = \frac{|\mathbf{e}^{\mathrm{H}}\mathbf{G}_k\mathbf{f}_k|^2}{\sum_{i\neq k}^{K}|\mathbf{e}^{\mathrm{H}}\mathbf{G}_k\mathbf{f}_i|^2 + \sigma_k^2}. \tag{3}$$

B. Channel Model

It is important to mention that the perfect instantaneous CSI of a mmWave system is difficult to obtain due to the passive implementation of the RIS reflecting elements. More specifically, estimating the full instantaneous CSI for all links would inevitably require a large training overhead. In this work, since the channel modeling of the RIS-related links is the same as the traditional wireless channel modeling [32], we adopt a Saleh-Valenzuela (SV) channel model [33] to characterize the mmWave channels where only the large-scale fading characteristics is required during the transmission design. In particular, it is assumed that the uniform planar array (UPA) is deployed on the BS and each RIS. The steering vector of the UPA is $\mathbf{a}(\varphi, \phi)$ in which $\varphi(\phi)$ denotes the azimuth (elevation) AoD and the angle-of-arrival (AoA) of the transceivers. We assume that there are $L_{\mathrm{b},k}$, $L_{u,k}$ and $L_{\mathrm{b},u}$ sparse scatterers for the BS-user k, the RIS u-user k and the BS-RIS u links, respectively, and each of the scatterers comprises I subpaths. Thus, the mmWave channels can be expressed as

$$\mathbf{h}_{\mathrm{b},k} = g_0^{\mathrm{b},k}\mathbf{a}\left(\varphi_{\mathrm{b},k,0}^{\mathrm{AoD}}, \phi_{\mathrm{b},k,0}^{\mathrm{AoD}}\right) + \sqrt{\frac{1}{IL_{\mathrm{b},k}}}\sum_{l=1}^{L_{\mathrm{b},k}}\sum_{i=1}^{I} g_{l,i}^{\mathrm{b},k}\mathbf{a}\left(\varphi_{\mathrm{b},k,l,i}^{\mathrm{AoD}}, \phi_{\mathrm{b},k,l,i}^{\mathrm{AoD}}\right), \forall k, \tag{4}$$

$$\mathbf{h}_{u,k} = g_0^{u,k}\mathbf{a}\left(\varphi_{u,k,0}^{\mathrm{AoD}}, \phi_{u,k,0}^{\mathrm{AoD}}\right) + \sqrt{\frac{1}{IL_{u,k}}}\sum_{l=1}^{L_{u,k}}\sum_{i=1}^{I} g_{l,i}^{u,k}\mathbf{a}\left(\varphi_{u,k,l,i}^{\mathrm{AoD}}, \phi_{u,k,l,i}^{\mathrm{AoD}}\right), \forall k, \forall u, \tag{5}$$

$$\begin{aligned}\mathbf{H}_u = {} & g_0^{\mathrm{b},u}\mathbf{a}\left(\varphi_{u,0}^{\mathrm{AoA}}, \phi_{u,0}^{\mathrm{AoA}}\right)\mathbf{a}\left(\varphi_{\mathrm{b},0}^{\mathrm{AoD}}, \phi_{\mathrm{b},0}^{\mathrm{AoD}}\right)^{\mathrm{H}} \\ & + \sqrt{\frac{1}{IL_{\mathrm{b},u}}}\sum_{l=1}^{L_{\mathrm{b},u}}\sum_{i=1}^{I} g_{l,i}^{\mathrm{b},u}\mathbf{a}\left(\varphi_{u,l,i}^{\mathrm{AoA}}, \phi_{u,l,i}^{\mathrm{AoA}}\right)\mathbf{a}\left(\varphi_{\mathrm{b},l,i}^{\mathrm{AoD}}, \phi_{\mathrm{b},l,i}^{\mathrm{AoD}}\right)^{\mathrm{H}}, \forall u,\end{aligned} \tag{6}$$

where, by denoting an arbitrary element $q \in \{(\mathrm{b},k),(u,k),(\mathrm{b},u)\}_{\forall k,\forall u}$, then $g_0^q\mathbf{a}\left(\varphi_{q,0}^{\mathrm{AoD}},\phi_{q,0}^{\mathrm{AoD}}\right)$ is the line-of-sight (LoS) component with fading coefficient following the distribution $g_0^q \sim \mathcal{CN}(0,\zeta_0^q 10^{\frac{\mathrm{PL}}{10}})$, where $\zeta_0^q = \frac{\kappa}{1+\kappa}$ is the power fraction that corresponds to the Rician factor κ, and PL is the distance-dependent large-scale fading model. The remaining paths are the line-of-sight (NLoS) components whose fading coefficients follow the distribution $g_{l,i}^q \sim \mathcal{CN}(0,\zeta_l^q 10^{\frac{\mathrm{PL}}{10}})$ with power fraction $\zeta_l^q = \frac{1}{(L_q-1)(1+\kappa)}$.

We assume that the user's locations are quasi-static over milliseconds or even seconds. Therefore, the large-scale fading characteristic parameters, such as the distance-dependent path loss, the numbers of clusters, their power fraction, the cluster central angles and angular beamspreads, change relatively slowly [34] and can be perfectly known by the BS. However, the instantaneous CSI, given by $\{\mathbf{h}_{\mathrm{b},k},\mathbf{h}_{u,k},\mathbf{H}_u\}$[1], vary during the transmission due to the rapidly varying small-scale fading coefficients $\{g_0^q, g_{l,i}^q\}$, AoDs and AoAs according to an ergodic stationary process. These AoDs and AoAs can be generated according to a Gaussian distribution, whose expectation values are the cluster central angles and variance values are the angular spreads [34].

Based on the fact that the communication links in the mmWave frequency band are sensitive to the presence of blockages, most existing contributions considered a worst-case scenario where the BS-user links are completely blocked by the obstacles during the whole transmission, while the RIS-related links are not affected by blockages since the locations of the RISs can be appropriately chosen to ensure line-of-sight transmission. However, this assumption may be impractical in many scenarios. Traditionally, blockage effects are incorporated into the shadowing model, along with reflections, scattering, and diffraction [5]. In contrast, we adopt a recently proposed probabilistic model [10] to characterize the channel uncertainties caused by the presence of random blockages. In particular, by introducing the blockage parameters $\omega_{k,l} \in \{0,1\}, 0 \leq l \leq L_{\mathrm{b},k}, \forall k \in \mathcal{K}$, which are random variables following a Bernoulli distribution with a blockage probability of p_k, the channels in (4) for the BS-user links are modified to

$$\mathbf{h}_{\mathrm{b},k} = \omega_{k,0} g_0^{\mathrm{b},k}\mathbf{a}\left(\varphi_{\mathrm{b},k,0}^{\mathrm{AoD}},\phi_{\mathrm{b},k,0}^{\mathrm{AoD}}\right) + \sqrt{\frac{1}{I}}\sum_{l=1}^{L_{\mathrm{b},k}}\omega_{k,l}\sum_{i=1}^{I} g_{l,i}^{\mathrm{b},k}\mathbf{a}\left(\varphi_{\mathrm{b},k,l,i}^{\mathrm{AoD}},\phi_{\mathrm{b},k,l,i}^{\mathrm{AoD}}\right), \forall k, \tag{7}$$

Analytical results show that the probability of link blocking increases exponentially with the link length [5], thus we adopt a distance-dependent blockage probability model given by $p_k(d_k) =$

[1]With the knowledge of long-term large-scale fading characteristic parameters, each sample of $\{\mathbf{h}_{\mathrm{b},k},\mathbf{h}_{u,k},\mathbf{H}_u\}$ can be obtained by generating $\{g_0^q, g_{l,i}^q\}$, AoDs and AoAs according to their Gaussian distribution.

$\max(0, 1 - e^{-a_{out}d_k + b_{out}})$, where physical distance parameter d_k between the BS and user k can be extracted from the estimated CSI by the BS, and parameters a_{out} and b_{out} are fit from the data [6].

C. Problem Formulation

To account for the presence of random blockages and to enhance the QoS of the users, we design a robust beamforming by minimizing the maximum outage probability among all users over the random CSI. The formulated optimization problem is given by

$$\min_{\mathbf{F},\mathbf{e}} \quad \max_{k\in\mathcal{K}} \Pr\{\Gamma_k(\mathbf{F},\mathbf{e}) \le \gamma_k\} \tag{8a}$$

$$\text{s.t.} \quad \mathbf{F} \in \mathcal{S}_f \tag{8b}$$

$$\mathbf{e} \in \mathcal{S}_e, \tag{8c}$$

where the outage probability $\Pr\{\Gamma_k(\mathbf{F},\mathbf{e}) \le \gamma_k\}$ is the probability that the SINR $\Gamma_k(\mathbf{F},\mathbf{e})$ of the user k is less than the SINR reliability threshold γ_k for all possible realizations of the random channel $\mathbf{G} = [\mathbf{G}_1, \ldots, \mathbf{G}_K]$.

Compared with the sum outage probability minimization problem in [11], the objective function in (8) can ensure the fairness among the users. However, due to the min-max function, the objective function is not smooth or differentiable, and the algorithms in [11] cannot be directly applied.

III. Single-User System

In this section, we consider a single-user system model in order to obtain some design insights. By setting $K = 1$ and dropping the user index, Problem (8) is simplified to

$$\min_{\mathbf{f},\mathbf{e}} \quad \Pr\{\Gamma(\mathbf{f},\mathbf{e}) \le \gamma\} \tag{9a}$$

$$\text{s.t.} \quad \mathbf{f} \in \mathcal{S}_f \tag{9b}$$

$$\mathbf{e} \in \mathcal{S}_e. \tag{9c}$$

A. Problem Reformulation

The probability $\Pr\{\Gamma(\mathbf{f},\mathbf{e}) \le \gamma\}$ has no closed-form expression and thus the problem in (9) is prohibitively challenging to be solved directly. An alternative solution is to replace the

probability function with an equivalent expectation function, i.e., $\Pr\{\Gamma(\mathbf{f},\mathbf{e}) \leq \gamma\} = \mathbb{E}_{\mathbf{G}}[\mathbb{I}_{\Gamma\leq\gamma}]$ where $\mathbb{I}_{\Gamma\leq\gamma}$ is the step function of the event $\Gamma \leq \gamma$. By doing so, various stochastic programming techniques can be used. However, the step function is discontinuous, and the existing stochastic programming methods cannot be directly applied.

To resolve this issue, we approximate the step function with a smooth approximating function

$$u\left(x\right) = \frac{1}{1+e^{-\theta x}}, \tag{10}$$

where $x = \gamma - \Gamma$ and θ is the smooth parameter controlling the approximation error.

By defining $f\left(\mathbf{f},\mathbf{e}|\mathbf{G}\right) = u\left(\gamma\sigma^2 - |\mathbf{e}^{\mathrm{H}}\mathbf{G}\mathbf{f}|^2\right)$, a convenient approximation of Problem (9) is

$$\min_{\mathbf{f}\in\mathcal{S}_f,\mathbf{e}\in\mathcal{S}_e} \; g(\mathbf{f},\mathbf{e}|\mathbf{G}) = \mathbb{E}\left[f\left(\mathbf{f},\mathbf{e}|\mathbf{G}\right)\right]. \tag{11}$$

B. Stochastic Majorization-Minimization Method

A simple approach for solving the above problem is the sample average approximation (SAA) method. However, the SAA method is computationally prohibitive since it requires large-sized memory storage due to the fact that the solution obtained at each iteration is calculated by averaging over a large number of channel realizations. To overcome these difficulties, we adopt the widely used SMM [35] (also known as stochastic successive minimization [36]) method in which a well-chosen upper bound approximation of the function $f(\mathbf{f},\mathbf{e}|\mathbf{G})$ for a new channel sample is constructed at each iteration and the solution is obtained on the average of the new channel sample and the others generated in previous iterations.

The standard of constructing the upper bound approximation function of $f(\mathbf{f},\mathbf{e}|\mathbf{G})$ is to make the corresponding surrogate problem easy to solve and even obtain the closed-form solutions. In particular, denote $\mathbf{x} \in \{\mathbf{f},\mathbf{e}\}$ that belongs to $\mathcal{S}_x \in \{\mathcal{S}_f,\mathcal{S}_e\}$, the surrogate function $\hat{f}(\mathbf{x},\mathbf{x}^{i-1}|\mathbf{G})$ of $f(\mathbf{x}|\mathbf{G})$ around any feasible point $\mathbf{x}^{i-1}$ needs to satisfy the following assumptions [36].

Assumption A

(A1) : $\hat{f}\left(\mathbf{x},\mathbf{x}^{i-1}|\mathbf{G}\right)$ is continuous in $\mathbf{x}$ for $\forall\mathbf{x}^{i-1} \in \mathcal{S}_x$.

(A2) : $\hat{f}(\mathbf{x}^{i-1},\mathbf{x}^{i-1}|\mathbf{G}) = f(\mathbf{x}^{i-1}|\mathbf{G}), \forall\mathbf{x}^{i-1} \in \mathcal{S}_x$.

(A3) : $\hat{f}(\mathbf{x},\mathbf{x}^{i-1}|\mathbf{G}) \geq f(\mathbf{x}|\mathbf{G}), \forall\mathbf{x},\mathbf{x}^{i-1} \in \mathcal{S}_x$.

(A4) : $\hat{f}'(\mathbf{x}^{i-1}|\mathbf{G};\mathbf{d}) = f'(\mathbf{x}^{i-1}|\mathbf{G};\mathbf{d})$, for all $\mathbf{x}^{i-1} \in \mathcal{S}_x$ and feasible directions $\forall\mathbf{d}$ at $\mathbf{x}^{i-1}$.

$f^{'}(\mathbf{x}^{i-1}|\mathbf{G};\mathbf{d})$ defines the directional derivative of $f(\mathbf{x}^{i-1}|\mathbf{G})$ in the direction $\mathbf{d}$ and is given by

$$f^{'}(\mathbf{x}^{i-1}|\mathbf{G};\mathbf{d}) = \lim_{\lambda \to 0} \frac{f(\mathbf{x}^{i-1} + \lambda \mathbf{d}|\mathbf{G}) - f(\mathbf{x}^{i-1}|\mathbf{G})}{\lambda}.$$

The Assumptions (A2)-(A3) indicate that the surrogate function $\hat{f}(\mathbf{x}, \mathbf{x}^{i-1}|\mathbf{G})$ is a locally upper bound of the original function $f(\mathbf{x}|\mathbf{G})$ around the feasible point $\mathbf{x}^{i-1}$. Assumption (A4) is a derivative consistency condition. To ensure the convergence of the SMM algorithm, we further make the following assumptions [36].

Assumption B

$(\mathrm{B1})$: The feasible set $\mathcal{S}_x$ and channel realizations are bounded.

$(\mathrm{B2})$: The functions $\hat{f}(\mathbf{x}, \mathbf{x}^{i-1}|\mathbf{G})$ and $f(\mathbf{x}|\mathbf{G})$, their derivatives, and their second-order derivatives are uniformly bounded.

Since the variables $\mathbf{f}$ and $\mathbf{e}$ are highly coupled with each other, we adopt an alternating optimization (AO) method to update them. Based on the above assumptions, the variables $\mathbf{f}$ and $\mathbf{e}$ are updated by solving the following two SMM subproblems:

$$\mathbf{f}^n = \arg\min_{\mathbf{f} \in \mathcal{S}_f} \ \frac{1}{n} \sum_{i=1}^{n} \hat{f}\left(\mathbf{f}, \mathbf{f}^{i-1}|\mathbf{G}^i\right), \tag{12}$$

and

$$\mathbf{e}^n = \arg\min_{\mathbf{e} \in \mathcal{S}_e} \ \frac{1}{n} \sum_{i=1}^{n} \hat{f}\left(\mathbf{e}, \mathbf{e}^{i-1}|\mathbf{G}^i\right). \tag{13}$$

Here, $\mathbf{G}^1, \mathbf{G}^2, ...$ are some independent samples of the random equivalent channel $\mathbf{G}$. $\hat{f}\left(\mathbf{f}, \mathbf{f}^{i-1}|\mathbf{G}^i\right)$ is a surrogate function corresponding to $\mathbf{f}$ when $\mathbf{e}$ is given, while $\hat{f}\left(\mathbf{e}, \mathbf{e}^{i-1}|\mathbf{G}^i\right)$ is the corresponding surrogate function of $\mathbf{e}$ with given $\mathbf{f}$.

1) Optimizing $\mathbf{f}$*:* First, we construct $\hat{f}\left(\mathbf{f}, \mathbf{f}^{i-1}|\mathbf{G}^i\right)$ satisfying Assumptions A and B, which is shown in the following lemma.

Lemma 1 *For the twice differentiable function* $f(\mathbf{f}|\mathbf{G}^i)$*, we construct its second-order upper bound approximation around any fixed* $\mathbf{f}^{i-1}$*, which is given by*

$$\hat{f}(\mathbf{f}, \mathbf{f}^{i-1}|\mathbf{G}^i) = 2\mathrm{Re}\left\{\mathbf{d}_f^{i,\mathrm{H}}\mathbf{f}\right\} + \alpha_f^i||\mathbf{f}||_2^2 + const_f^i, \tag{14}$$

where

$$\mathbf{d}_f^i = \mathbf{m}_f^i - \alpha_f^i \mathbf{f}^{i-1}, \tag{15a}$$

$$\mathbf{m}_f^i = \frac{-\theta e^{-\theta x^i}}{\left(1+e^{-\theta x^i}\right)^2}\mathbf{G}^{i,\mathrm{H}}\mathbf{e}^{i-1}\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}^i\mathbf{f}^{i-1}, \tag{15b}$$

$$x^i = \gamma\sigma^2 - |\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}^i\mathbf{f}^{i-1}|^2, \tag{15c}$$

$$\alpha_f^i = \frac{\theta^2}{2}P_{max}|\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}^i\mathbf{G}^{i,\mathrm{H}}\mathbf{e}^{i-1}|^2, \tag{15d}$$

$$const_f^i = f(\mathbf{f}^{i-1}|\mathbf{G}^i) + \alpha_f^i||\mathbf{f}^{i-1}||_2^2 - 2\mathrm{Re}\left\{\mathbf{m}_f^{i,\mathrm{H}}\mathbf{f}^{i-1}\right\}. \tag{15e}$$

Proof: See Appendix A. ■

By using (14) and ignoring constants, the subproblem in (12) for updating $\mathbf{f}$ is formulated as

$$\min_{\mathbf{f}\in\mathcal{S}_f} \; 2\mathrm{Re}\left\{\frac{1}{n}\sum_{i=1}^n \mathbf{d}_f^{i,\mathrm{H}}\mathbf{f}\right\} + \frac{1}{n}\sum_{i=1}^n \alpha_f^i||\mathbf{f}||_2^2. \tag{16}$$

Problem (16) is convex and can be solved by computing its Lagrange function given by

$$\mathcal{L}(\mathbf{f},\kappa) = 2\mathrm{Re}\left\{\frac{1}{n}\sum_{i=1}^n \mathbf{d}_f^{i,\mathrm{H}}\mathbf{f}\right\} + \frac{1}{n}\sum_{i=1}^n \alpha_f^i||\mathbf{f}||_2^2 + \kappa\left(||\mathbf{f}||_2^2 - P_{max}\right), \tag{17}$$

where $\kappa \geq 0$ is a Lagrange multiplier associated with the power constraint. By setting $\partial\mathcal{L}(\mathbf{f})/\partial\mathbf{f}^* = \mathbf{0}$, the globally optimal solution of $\mathbf{f}$ at the n-th iteration is derived as

$$\mathbf{f}^n = \frac{-1}{\kappa + \frac{1}{n}\sum_{i=1}^n \alpha_f^i}\frac{1}{n}\sum_{i=1}^n \mathbf{d}_f^i. \tag{18}$$

Also, (18) must satisfy the power constraint, which yields

$$\frac{||\frac{1}{n}\sum_{i=1}^n \mathbf{d}_f^i||_2^2}{(\kappa + \frac{1}{n}\sum_{i=1}^n \alpha_f^i)^2} \leq P_{max}. \tag{19}$$

Based on the fact that the left hand side of (19) is a decreasing function of κ, we obtain the following closed-form solution

$$\mathbf{f}^n = \begin{cases} \frac{-1}{\sum_{i=1}^n \alpha_f^i}\sum_{i=1}^n \mathbf{d}_f^i, & \text{if } \frac{||\sum_{i=1}^n \mathbf{d}_f^i||_2^2}{(\sum_{i=1}^n \alpha_f^i)^2} \leq P_{max}, \\ -\sqrt{\frac{P_{max}}{||\sum_{i=1}^n \mathbf{d}_f^i||_2^2}}\sum_{i=1}^n \mathbf{d}_f^i, & \text{otherwise.} \end{cases} \tag{20}$$

The first option in (20) is based on $\kappa = 0$. The second option is due to the fact that there must exist a $\kappa > 0$ that (19) holds with equality.

2) Optimizing $\mathbf{e}$*:* As for the update of $\mathbf{e}$ with given $\mathbf{f}$, we first construct a surrogate function corresponding to $\mathbf{e}$ in the following lemma.

Lemma 2 *For the twice differentiable function* $f(\mathbf{e}|\mathbf{G}^i)$*, we construct its second-order upper bound approximation around any feasible* $\mathbf{e}^{i-1}$*, which is given by*

$$\hat{f}(\mathbf{e}, \mathbf{e}^{i-1}|\mathbf{G}^i) = 2\mathrm{Re}\left\{\mathbf{d}_e^{i,\mathrm{H}}\mathbf{e}\right\} + const_e^i, \tag{21}$$

where

$$\mathbf{d}_e^i = \mathbf{m}_e^i - \alpha_e^i \mathbf{e}^{i-1}, \tag{22a}$$

$$\mathbf{m}_e^i = \frac{-\theta e^{-\theta x^i}}{\left(1+e^{-\theta x^i}\right)^2} \mathbf{G}^i \mathbf{f}^{i-1} \mathbf{f}^{i-1,\mathrm{H}} \mathbf{G}^{i,\mathrm{H}} \mathbf{e}^{i-1}, \tag{22b}$$

$$\alpha_e^i = \frac{\theta^2}{2}(UM+1)|\mathbf{f}^{i-1,\mathrm{H}} \mathbf{G}^{i,\mathrm{H}} \mathbf{G}^i \mathbf{f}^{i-1}|^2, \tag{22c}$$

$$const_e^i = f(\mathbf{e}^{i-1}|\mathbf{G}^i) + 2(UM+1)\alpha_e^i - 2\mathrm{Re}\left\{\mathbf{m}_e^{i,\mathrm{H}} \mathbf{e}^{i-1}\right\}. \tag{22d}$$

Proof: The proof of Lemma 2 is similar to that of Lemma 1 and hence omitted for brevity. ■

By substituting (21) into the objective function of subproblem (13) and ignoring constants, we obtain:

$$\min_{\mathbf{e}\in\mathcal{S}_e} \; 2\mathrm{Re}\left\{\frac{1}{n}\sum_{i=1}^{n} \mathbf{d}_e^{i,\mathrm{H}} \mathbf{e}\right\}. \tag{23}$$

The globally optimal solution of the above problem is given by

$$\mathbf{e}^n = \exp\left\{\mathrm{j}\angle\left(\left(\sum_{i=1}^{n}\mathbf{d}_e^i\right) / \left[\sum_{i=1}^{n}\mathbf{d}_e^i\right]_{UM+1}\right)\right\}, \tag{24}$$

where $[\cdot]_m$ means the m-th element of vector, $\mathrm{j} \triangleq \sqrt{-1}$ is the imaginary unit, $\angle(\cdot)$ denotes the angle of a complex number, and $\exp\{\mathrm{j}\angle(\cdot)\}$ is an element-wise operation.

C. Algorithm Development

Under the SMM framework, closed-form solutions of $\mathbf{f}$ in (20) and of $\mathbf{e}$ in (24) at each iteration are obtained. Such simple closed-form solutions can greatly reduce the computational complexity. Algorithm 1 summarizes the proposed robust beamforming design based on the SMM-based outage probability minimum problem for RIS-aided single-user mmWave systems in which the BS-user links experience random blockages. The proposed algorithm is referred to as SMM-OutMin. Note that the convergence speed of the SMM-based algorithm might be affected by the tightness of the upper bounds in Lemma 1 and Lemma 2, thus SQUAREM [37] is adopted to accelerate the SMM-based algorithm. Denote by $F(\mathbf{f}^n)$ and $F(\mathbf{e}^n)$ the objective function values of Problem (16) and Problem (23) in the n-th iteration, respectively. $\mathcal{P}_{\mathcal{S}}(\cdot)$ is an operation mapped to nonlinear constraint set. For the power constraint set $\mathcal{S}_f$, $\mathcal{P}_{\mathcal{S}}(\cdot)$ can be $\mathcal{P}_{\mathcal{S}}(\mathbf{x}) = \frac{(\mathbf{x})}{||\mathbf{x}||_2}||\mathbf{f}_2||_2$. For the unit-modulus constraint set $\mathcal{S}_e$, $\mathcal{P}_{\mathcal{S}}(\cdot)$ can be element-wise operation, i.e., $\mathcal{P}_{\mathcal{S}}(\mathbf{x}) = j\angle(\mathbf{x})$. Steps 10 to 13 and steps 21 to 24 are used to maintain the convergence property of the objective function values.

Algorithm 1 SMM-OutMin Algorithm

Initialize: Initialize $\mathbf{f}^0 \in \mathcal{S}_f$ and $\mathbf{e}^0 \in \mathcal{S}_e$. Set $n = 1$.

repeat
  Obtain the sample channel $\mathbf{G}^n$.
  Set $\mathbf{e} = \mathbf{e}^{n-1}$.
  Obtain $\mathbf{f}_1$ according to (20) based on $\mathbf{f}^{n-1}$.
  Obtain $\mathbf{f}_2$ according to (20) based on $\mathbf{f}_1$.
  $\mathbf{j}_1 = \mathbf{f}_1 - \mathbf{f}^{n-1}$.
  $\mathbf{j}_2 = \mathbf{f}_2 - \mathbf{f}_1 - \mathbf{j}_1$.
  $\omega = -\frac{||\mathbf{j}_1||_2}{||\mathbf{j}_2||_2}$.
  $\mathbf{f}^n = -\mathcal{P}_{\mathcal{F}}(\mathbf{f}^{n-1} - 2\omega\mathbf{j}_1 + \omega^2\mathbf{j}_2)$.
  while $|F(\mathbf{f}^n) - F(\mathbf{f}^{n-1})| \leq |F(\mathbf{f}^{n-1}) - F(\mathbf{f}^{n-2})|$ **do**
    $\omega = (\omega - 1)/2$.
    $\mathbf{f}^n = -\mathcal{P}_{\mathcal{S}}(\mathbf{f}^{n-1} - 2\omega\mathbf{j}_1 + \omega^2\mathbf{j}_2)$.
  end while
  Set $\mathbf{f} = \mathbf{f}^n$.
  Obtain $\mathbf{e}_1$ according to (24) based on $\mathbf{e}^{n-1}$.
  Obtain $\mathbf{e}_2$ according to (24) based on $\mathbf{e}_1$.
  $\mathbf{j}_1 = \mathbf{e}_1 - \mathbf{e}^{n-1}$.
  $\mathbf{j}_2 = \mathbf{e}_2 - \mathbf{e}_1 - \mathbf{j}_1$.
  $\omega = -\frac{||\mathbf{j}_1||_2}{||\mathbf{j}_2||_2}$.
  $\mathbf{e}^n = -\mathcal{P}_{\mathcal{S}}(\mathbf{e}^{n-1} - 2\omega\mathbf{j}_1 + \omega^2\mathbf{j}_2)$.
  while $|F(\mathbf{e}^n) - F(\mathbf{e}^{n-1})| \leq |F(\mathbf{e}^{n-1}) - F(\mathbf{e}^{n-2})|$ **do**
    $\omega = (\omega - 1)/2$.
    $\mathbf{e}^n = -\mathcal{P}_{\mathcal{S}}(\mathbf{e}^{n-1} - 2\omega\mathbf{j}_1 + \omega^2\mathbf{j}_2)$.
  end while
  $n = n + 1$.
until $||F(\mathbf{f}^n) - F(\mathbf{f}^{n-1})||_2 \to 0$ and $||F(\mathbf{e}^n) - F(\mathbf{e}^{n-1})||_2 \to 0$.

1) Convergence analysis: The convergence of Algorithm 1 is given in the following theorem.

Theorem 1 *Suppose Assumptions A and B are satisfied. Then the sequence of the solutions obtained in each iteration of Algorithm 1 converge to the set of stationary points of Problem (11) almost surely.*

Proof: See Appendix B. ■

2) Complexity analysis: The computational complexity for updating $\mathbf{f}^n$ and $\mathbf{e}^n$ at each iteration mainly depends on the computation of (20) and (24), respectively. In particular, due to the update rule in $\{\sum_{i=1}^n \alpha_f^i, \sum_{i=1}^n \mathbf{d}_f^i, \sum_{i=1}^n \mathbf{d}_e^i\}$, only $\{\alpha_e^n, \mathbf{d}_f^n, \mathbf{d}_e^n\}$ needs to be calculated at the n-th iteration. Therefore, the approximate complexity of each iteration is given by $\mathcal{O}(4UMN+12N)$.

IV. Multiuser System

In this section, we consider the general multiuser setup and solve Problem (8). Problem (8) is more challenging than problem (9) due to the non-differentiable objective function. Furthermore, the complex objective function complicates the use of the SMM algorithm, thus we extend the SMM method to a general algorithm to solve Problem (8).

A. Problem Reformulation

We first approximate the probability function in the original Problem (8) by using the expectation of the smooth function in (10). By defining $f_k\left(\mathbf{F}, \mathbf{e}|\mathbf{G}\right) = u\left(\mathbf{e}^{\mathrm{H}}\mathbf{G}_k\mathbf{F}\boldsymbol{\Upsilon}_k\mathbf{F}^{\mathrm{H}}\mathbf{G}_k^{\mathrm{H}}\mathbf{e} + \gamma_k\sigma_k^2\right)$, in which $\boldsymbol{\Upsilon}_k$ is a diagonal matrix whose diagonal entry is γ_k except in the k-th diagonal element that is -1, the approximate objective function is $\max_{k\in\mathcal{K}} \mathbb{E}\left[f_k\left(\mathbf{F}, \mathbf{e}|\mathbf{G}\right)\right]$. However, the obtained function is still intractable since the maximization operation couples $f_k, \forall k$ and the different channel states due to the expectation operation. This issue motivates us to use the following Jensen inequality

$$\max_{k\in\mathcal{K}} \mathbb{E}\left[f_k\left(\mathbf{F}, \mathbf{e}|\mathbf{G}\right)\right] \leq \mathbb{E}\left[\max_{k\in\mathcal{K}} f_k\left(\mathbf{F}, \mathbf{e}|\mathbf{G}\right)\right], \tag{25}$$

due to the fact that the max function $\max_{k\in\mathcal{K}}\{x_1, \ldots, x_K\}$ is convex [38].

Furthermore, the non-differentiable max function, $\max_{k\in\mathcal{K}} f_k\left(\mathbf{F}, \mathbf{e}|\mathbf{G}\right)$, is approximated by adopting a smooth log–sum–exp upper-bound [39]

$$\max_{k\in\mathcal{K}} f_k\left(\mathbf{F}, \mathbf{e}|\mathbf{G}\right) \approx F\left(\mathbf{F}, \mathbf{e}|\mathbf{G}\right)$$

$$= \mu \ln\Big(\sum_{k\in\mathcal{K}} \exp\left\{\frac{1}{\mu} f_k\left(\mathbf{F},\mathbf{e}|\mathbf{G}\right)\right\}\Big), \tag{26}$$

where $\mu > 0$ is a smoothing parameter satisfying

$$\max_{k\in\mathcal{K}} f_k\left(\mathbf{F},\mathbf{e}|\mathbf{G}\right) \le F\left(\mathbf{F},\mathbf{e}|\mathbf{G}\right) \le \max_{k\in\mathcal{K}} f_k\left(\mathbf{F},\mathbf{e}|\mathbf{G}\right) + \frac{1}{\mu}\log\left(|\mathcal{K}|\right). \tag{27}$$

When μ is chosen appropriately, the smooth approximation of Problem (8) is approximately reformulated as

$$\min_{\mathbf{F}\in\mathcal{S}_f,\mathbf{e}\in\mathcal{S}_e} \; G\left(\mathbf{F},\mathbf{e}|\mathbf{G}\right) = \mathbb{E}\left[F\left(\mathbf{F},\mathbf{e}|\mathbf{G}\right)\right]. \tag{28}$$

B. Stochastic Successive Convex Approximation Method

Similar to Problem (11), Problem (28) can still be solved by adopting the SMM method. However, the function $F\left(\mathbf{F},\mathbf{e}|\mathbf{G}\right)$ in (26) is much more complex and its second-order derivative, which is necessary to construct the upper bound surrogate function of $F\left(\mathbf{F},\mathbf{e}|\mathbf{G}\right)$ as shown in Appendix A, is not easy to be calculated. Furthermore, the coefficient of the second-order term in the final upper-bound surrogate function of $F\left(\mathbf{F},\mathbf{e}|\mathbf{G}\right)$ (α_f^i in (14)) can be quite loose, resulting in a very slow convergence rate of the SMM algorithm.

Therefore, in this section, we adopt a flexible SSCA method to address the above issues. The surrogate functions employed by the SSCA method do not need to be an upper bound of the original function but they need only to preserve the first-order property of the original function. Accordingly, the surrogate functions need to satisfy Assumption B and the following assumptions [40].

Assumption C

$$(\mathrm{C1}): \hat{F}\left(\mathbf{x},\mathbf{x}^{i-1}|\mathbf{G}\right) \text{ is strongly convex in } \mathbf{x} \text{ for } \forall \mathbf{x}^{i-1}\in\mathcal{S}_x.$$

$$(\mathrm{C2}): \hat{F}(\mathbf{x}^{i-1},\mathbf{x}^{i-1}|\mathbf{G}) = F(\mathbf{x}^{i-1}|\mathbf{G}), \forall \mathbf{x}^{i-1}\in\mathcal{S}_x.$$

$$(\mathrm{C3}): \nabla_{\mathbf{x}}\hat{F}(\mathbf{x}^{i-1},\mathbf{x}^{i-1}|\mathbf{G}) = \nabla_{\mathbf{x}}F(\mathbf{x}^{i-1}|\mathbf{G}), \forall \mathbf{x},\mathbf{x}^{i-1}\in\mathcal{S}_x.$$

Assumption C cannot ensure that the sequences of the approximate objective values are monotonically decreasing at each iteration. Nevertheless, to guarantee convergence, the variables can be updated by choosing an appropriate step size at each iteration that yields a sufficient decrease of the objective value. Based on the above assumptions, we choose the proximal gradient-like approximation to construct the surrogate function, which is

$$\hat{F}\left(\mathbf{x},\mathbf{x}^{i-1}|\mathbf{G}\right) = F\left(\mathbf{x}^{i-1}|\mathbf{G}\right) + \nabla_{\mathbf{x}}F(\mathbf{x}^{i-1}|\mathbf{G})^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^{i-1}) + \frac{\tau^i}{2}||\mathbf{x}-\mathbf{x}^{i-1}||^2, \tag{29}$$

where τ^i can be any positive number.

1) Optimizing $\mathbf{f}$*:* By using (29), we construct a surrogate function as

$$\begin{aligned}
&\hat{F}\left(\mathbf{F},\mathbf{F}^{i-1}|\mathbf{G}\right)\\
&=F\left(\mathbf{F}^{i-1}|\mathbf{G}\right)+\mathrm{Tr}\left(\nabla_{\mathbf{F}}F\left(\mathbf{F}^{i-1}|\mathbf{G}\right)^{\mathrm{T}}\left(\mathbf{F}-\mathbf{F}^{i-1}\right)\right)+\mathrm{Tr}\left(\nabla_{\mathbf{F}^*}F\left(\mathbf{F}^{i-1}|\mathbf{G}\right)^{\mathrm{T}}\left(\mathbf{F}^*-\mathbf{F}^{i-1,*}\right)\right)\\
&\quad+\frac{\tau^i}{2}||\mathbf{F}-\mathbf{F}^{i-1}||_F^2\\
&=F\left(\mathbf{F}^{i-1}|\mathbf{G}\right)+2\sum_{k\in\mathcal{K}}l_k^i\mathrm{Re}\left\{\mathrm{Tr}\left(\boldsymbol{\Upsilon}_k\mathbf{F}^{i-1,\mathrm{H}}\mathbf{G}_k^{\mathrm{H}}\mathbf{e}^{i-1}\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}_k(\mathbf{F}-\mathbf{F}^{i-1})\right)\right\}+\frac{\tau^i}{2}||\mathbf{F}-\mathbf{F}^{i-1}||_F^2\\
&=2\mathrm{Re}\left\{\mathrm{Tr}\left(\mathbf{P}_f^{i,\mathrm{H}}\mathbf{F}\right)\right\}+\frac{\tau^i}{2}||\mathbf{F}||_F^2+\mathrm{cons1}^i,
\end{aligned}\tag{30}$$

for $\mathbf{F}$ around the fixed $\mathbf{F}^{i-1}$ when $\mathbf{e}$ is given.

The parameters in (30) are as follows

$$\mathbf{P}_f^i=\mathbf{W}_f^i-\frac{\tau^i}{2}\mathbf{F}^{i-1},\tag{31a}$$

$$\mathbf{W}_f^i=\sum_{k\in\mathcal{K}}l_k^i\mathbf{G}_k^{i,\mathrm{H}}\mathbf{e}^{i-1}\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}_k^i\mathbf{F}^{i-1}\boldsymbol{\Upsilon}_k,\tag{31b}$$

$$l_k^i=\frac{\exp\left\{\frac{1}{\mu}f_k\left(\mathbf{F}^{i-1},\mathbf{e}^{i-1}|\mathbf{G}^i\right)\right\}}{\sum_{k\in\mathcal{K}}\exp\left\{\frac{1}{\mu}f_k\left(\mathbf{F}^{i-1},\mathbf{e}^{i-1}|\mathbf{G}^i\right)\right\}}\frac{\theta e^{-\theta x_k^i}}{\left(1+e^{-\theta x_k^i}\right)^2},\tag{31c}$$

$$x_k^i=\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}_k^i\mathbf{F}^{i-1}\boldsymbol{\Upsilon}_k\mathbf{F}^{i-1,\mathrm{H}}\mathbf{G}_k^{i,\mathrm{H}}\mathbf{e}^{i-1}+\gamma_k\sigma_k^2,\tag{31d}$$

$$\mathrm{cons1}^i=F\left(\mathbf{F}^{i-1}|\mathbf{G}^i\right)+\frac{\tau^i}{2}||\mathbf{F}^{i-1}||_F^2-2\mathrm{Re}\left\{\mathrm{Tr}\left(\mathbf{W}_f^{i,\mathrm{H}}\mathbf{F}^{i-1}\right)\right\}.\tag{31e}$$

By using (30), the subproblem of Problem (28) corresponding to $\mathbf{F}$ at the n-th iteration is formulated as

$$\min_{\mathbf{F}\in\mathcal{S}_F}\ \frac{1}{n}\sum_{i=1}^n\hat{F}\left(\mathbf{F},\mathbf{F}^{i-1}|\mathbf{G}^i\right).\tag{32}$$

The method to solve Problem (32) is the same as that of solving Problem (16), thus we directly provide the global minimizer of Problem (32) as

$$\widehat{\mathbf{F}}^n=\begin{cases}\frac{-2}{\sum_{i=1}^n\tau^i}\sum_{i=1}^n\mathbf{P}_f^i, & \text{if } \frac{4||\sum_{i=1}^n\mathbf{P}_f^i||_F^2}{(\sum_{i=1}^n\tau^i)^2}\le P_{max},\\ -\sqrt{\frac{P_{max}}{||\sum_{i=1}^n\mathbf{P}_f^i||_F^2}}\sum_{i=1}^n\mathbf{P}_f^i, & \text{otherwise.}\end{cases}\tag{33}$$

2) Optimizing $\mathbf{e}$*:* Furthermore, with fixed $\mathbf{F}$, the subproblem of Problem (28) corresponding to $\mathbf{e}$ at the n-th iteration is

$$\min_{\mathbf{e}\in\mathcal{S}_e}\ \frac{1}{n}\sum_{i=1}^n\hat{F}\left(\mathbf{e},\mathbf{e}^{i-1}|\mathbf{G}^i\right),\tag{34}$$

where $\hat{F}\left(\mathbf{e}, \mathbf{e}^{i-1}|\mathbf{G}^i\right) = 2\mathrm{Re}\left\{\mathbf{p}_e^{i,\mathrm{H}}\mathbf{e}\right\} + \mathrm{cons2}^i$, and

$$\mathbf{p}_e^i = \mathbf{w}_e^i - \frac{\tau^i}{2}\mathbf{e}^{i-1}, \tag{35a}$$

$$\mathbf{w}_e^i = \sum_{k\in\mathcal{K}} l_k \mathbf{G}_k \mathbf{F}^{i-1} \boldsymbol{\Upsilon}_k \mathbf{F}^{i-1,\mathrm{H}} \mathbf{G}^{\mathrm{H}} \mathbf{e}^{i-1}, \tag{35b}$$

$$\mathrm{cons2}^i = F\left(\mathbf{e}^{i-1}|\mathbf{G}^i\right) + \tau^i(UM+1) - 2\mathrm{Re}\left\{\mathbf{w}_e^{i,\mathrm{H}}\mathbf{e}^{i-1}\right\}. \tag{35c}$$

Finally, the minimizer of Problem (34) is

$$\widehat{\mathbf{e}}^n = \exp\left\{\mathrm{j}\angle\left(\left(\sum_{i=1}^{n}\mathbf{p}_e^i\right) / \left[\sum_{i=1}^{n}\mathbf{p}_e^i\right]_{UM+1}\right)\right\}. \tag{36}$$

C. Algorithm development

The simple closed-form solutions of $\mathbf{f}$ in (33) and of $\mathbf{e}$ in (36) can greatly reduce the computational complexity. Algorithm 2 summarizes the proposed SSCA-based robust beamforming design for RIS-aided multiuser mmWave systems in which the BS-user links experience random blockages. The proposed algorithm is referred to as SSCA-OutMin.

Algorithm 2 SSCA-OutMin Algorithm

Initialize: Initialize $\mathbf{F}^0 \in \mathcal{S}_f$ and $\mathbf{e}^0 \in \mathcal{S}_e$. Set $n = 0$.

1: **repeat**
2: $n = n + 1$.
3: Obtain the sample channel $\mathbf{G}^n$.
4: Calculate $\widehat{\mathbf{F}}^n$ according to (33).
5: Update $\mathbf{F}^n = \mathbf{F}^{n-1} + \xi_f^n\left(\widehat{\mathbf{F}}^n - \mathbf{F}^{n-1}\right)$.
6: Calculate $\widehat{\mathbf{e}}^n$ according to (36).
7: Update $\mathbf{e}^n = \mathbf{e}^{n-1} + \xi_e^n\left(\widehat{\mathbf{e}}^n - \mathbf{e}^{n-1}\right)$.
8: **until** $||\mathbf{F}^n - \mathbf{F}^{n-1}||_F^2 \to 0$ and $||\mathbf{e}^n - \mathbf{e}^{n-1}||_2 \to 0$.

1) Step-size selection: It is worth noting that the approximation in (30) has the same form as that in (14). However, τ^i in the SSCA method can be any positive number, and $\hat{F}\left(\mathbf{F}, \mathbf{F}^{i-1}|\mathbf{G}\right)$ might no longer be a global upper bound of $F\left(\mathbf{F}|\mathbf{G}\right)$. In this case, the step sizes ξ_f^n and ξ_e^n need to be carefully chosen to ensure convergence.

As an example, we take ξ_f^n to illustrate the update rule, which is a line-search (also called Armijo step-size) rule: Consider $\xi_f^0 > 0$ and $c_{1,f}, c_{2,f} \in (0,1)$. Let ξ_f^n be the largest element in $\{\xi_f^0 c_{2,f}^t\}_{t=0,1,...}$ such that

$$F\left(\mathbf{F}^{n-1}+\xi_f^n\left(\widehat{\mathbf{F}}^n-\mathbf{F}^{n-1}\right)\right) \leq F(\mathbf{F}^{n-1}) + c_{1,f}\xi_f^n \mathrm{Tr}\left(\nabla_{\mathbf{F}} F(\mathbf{F}^{n-1})^{\mathrm{T}}\left(\widehat{\mathbf{F}}^n-\mathbf{F}^{n-1}\right)\right). \quad (37)$$

Theorem 2 *If $\{\xi_f^n\}_{n=1,2,...}$ is chosen according to the line-search rule, then*

$$\lim_{n\to\infty} ||\widehat{\mathbf{F}}^n - \mathbf{F}^{n-1}|| = 0.$$

Proof: See Theorem 7 in [41]. ■

2) Convergence analysis: The convergence of Algorithm 2 is given in the following theorem.

Theorem 3 *Suppose Assumptions B and C are satisfied. Then every limit point of the iterations generated by Algorithm 2 is a stationary point of Problem (28) almost surely.*

Proof: See Appendix C. ■

3) Complexity analysis: The computational complexity for updating $\mathbf{f}^n$ and $\mathbf{e}^n$ at each iteration mainly depend on the computation of (33) and (36), respectively. In particular, only $\{\mathbf{P}_f^n, \mathbf{p}_e^n\}$ needs to be calculated at the n-th iteration. Therefore, the approximate complexity of each iteration is given by $\mathcal{O}((K+2)2UMN + UMK + NK + (N+2)K^2 + 2N)$.

4) Initial point: Problem (28) has, in general, multiple local minima points due to the non-convex unit-modulus constraint and $\mathbf{e} \in \mathcal{S}_e$. The accurate selection of the initial points in Algorithm 2 plays an important role for the convergence speed and the quality of the obtained local solution. To that end, we first initialize $\mathbf{e}$ to maximize the minimum equivalent total channel gain, resulting in the following optimization problem

$$\mathbf{e}^0 = \arg\max_{\mathbf{e}\in\mathcal{S}_e} \min_{k\in\mathcal{K}} ||\mathbf{e}^{\mathrm{H}}\mathbf{G}_k^0||_2^2. \quad (38)$$

Problem (38) can be efficiently solved by using the SDR method as follows

$$\max_{\mathbf{E}} \ t \quad (39a)$$

$$\text{s.t. } \mathrm{Tr}\{\mathbf{G}_k^0\mathbf{G}_k^{0,\mathrm{H}}\mathbf{E}\} \geq t, \forall k \in \mathcal{K} \quad (39b)$$

$$\mathbf{E} \succeq 0, \mathrm{rank}(\mathbf{E}) = 1, [\mathbf{E}]_{m,m} = 1, \forall m, \quad (39c)$$

where $\mathbf{E} = \mathbf{e}\mathbf{e}^{\mathrm{H}}$ and t is an auxiliary variable.

Furthermore, $\mathbf{F}$ is initialized by using the maximum-ratio transmission (MRT) method as

$$\mathbf{F}^0 = P_{max}\frac{\mathbf{G}^0\mathbf{e}^0}{||\mathbf{G}^0\mathbf{e}^0||}. \quad (40)$$

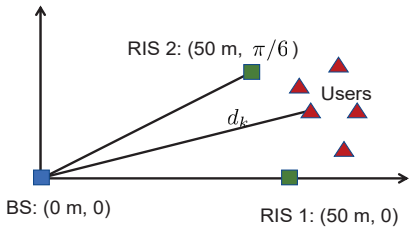


Fig. 2: The simulated system setup.

V. Numerical results and discussion

A. Simulation setup

In this section, we numerically evaluate the performance of the proposed algorithms. All experiments are performed on a PC with a 1.99 GHz i7-8550U CPU and 16 GB RAM. We adopt a polar coordinate to describe the simulated system setup as shown in Fig. 2, where the BS is located at (0 m, 0), and the two RISs are deployed in the locations (50 m, 0) and (50 m, $\pi/6$) which are close to the users. The users are randomly placed in a range with polar diameter $d_k \in$[50 m, 80 m] and polar angle $\vartheta \in [0, \pi/6]$, where d_k is used to calculate the distance-dependent blockage probability. The large-scale fading, reported in urban micro (UMi)-street canyon scenario [42], is modeled as $\mathrm{PL} = 32.4 + 20\log_{10}(f_c) + 10\alpha\log_{10}(D) + \xi$ in dB with link distance D (in meters), path loss exponent α, and log-normal shadowing $\xi \sim \mathcal{CN}(0, \sigma_{\xi}^2)$ where σ_{ξ}^2 denotes the log-normal shadowing variance. The mmWave system operates on carrier frequency $f_c = 28$ GHz and bandwidth 20 MHz. Since the macro-scattering environment between the BS and the users are complex, only NLoS clusters are assumed to exist in the BS-user links, i.e., the Rician factor is $\kappa = 0$. The parameter settings in PL of NLoS are $\alpha = 3.5$ and $\sigma_{\xi} = 8.2$ dB [42]. In practice, the RISs can be installed such that the BS-RIS links and the RIS-user links are blockage-free. Thus, the channels in (5) and (6) contain only LoS cluster with a Rician factor $\kappa \to \infty$. The parameters in PL of LoS are $\alpha = 2$ and $\sigma_{\xi} = 4$ dB according to [42]. Unless stated otherwise, we assume $L_{\mathrm{b},k} = L_{u,k} = L_{\mathrm{b},u} = 5$ and $I = 20$. The transmit power limit of the BS is $P_{max} = 30$ dBm and the noise power at each user is $\sigma_1^2 = \ldots = \sigma_K^2 = -94$ dBm. For simplicity, we consider an equal blockage probability, $p_{k,l} = p_{\mathrm{block}}, \forall k, l$, and an equal target SINR, $\gamma = \gamma_1 = \ldots = \gamma_K$, which yields the target rate $R_{\mathrm{targ}} = \log_2(1+\gamma)$. The smooth parameters are chosen to be $\theta = \frac{1}{\max_{\forall k \in \mathcal{K}} |x_k^0|}$ and $\mu = \frac{1}{100K}$.

To evaluate the performance of the proposed stochastic optimization algorithms, we consider

the following benchmark schemes. 1) SMRT: In this scheme, the active precoding is updated by using stochastic maximum-ratio transmission (SMRT), that is $\mathbf{F}^{n+1} = \frac{\frac{1}{n}\sum_{i=1}^{n}\mathbf{G}^{i,\mathrm{H}}\mathbf{e}^{i}}{||\frac{1}{n}\sum_{i=1}^{n}\mathbf{G}^{i,\mathrm{H}}\mathbf{e}^{i}||_2}\sqrt{P_{max}}$. The passive beamforming is still updated by using the SMM or SSCA methods. 2) NoRIS: In this case, no RIS is employed and the optimal active precoding is obtained by using the SMM or SSCA methods. 3) No-robust: In this scheme, the beamforming is designed by using the SMM or SSCA methods by taking into account the random small-scale parameters while assuming $p_{\mathrm{block}} = 0$. 4) Imperfect CSI: In this scheme, the beamforming is designed by using the SMM or SSCA methods based on the imperfect CSI of cluster central angles. In specific, we assume that the modulus of the cluster central angle estimation error is 0.01. 5) SAA: In this scheme, we generate 300 independent channel realizations in advance, the solutions at each iteration is the average over these 300 channel samples, and the surrogate function used at each iteration is obtained by adopting the MM or SCA methods. We take the beamforming design in the single-user case as an example. By modifying Problems (12) and (13), the beamforming designed by using the SAA-MM method is updated as follows

$$\mathbf{f}^{n} = \arg\min_{\mathbf{f}\in\mathcal{S}_f} \ \frac{1}{300}\sum_{i=1}^{300}\hat{f}\left(\mathbf{f}, \mathbf{f}^{n-1}|\mathbf{G}^{i}\right), \tag{41}$$

and

$$\mathbf{e}^{n} = \arg\min_{\mathbf{e}\in\mathcal{S}_e} \ \frac{1}{300}\sum_{i=1}^{300}\hat{f}\left(\mathbf{e}, \mathbf{e}^{n-1}|\mathbf{G}^{i}\right). \tag{42}$$

In order to demonstrate the robustness of the proposed algorithms, we consider two performance metrics: the outage probability and the effective rate. In particular, the outage probability of each user is calculated by averaging over 1000 independent channel realizations. The corresponding effective rate of the k-th user is defined as $R_{\mathrm{eff},k} \triangleq \mathbb{E}[\log_2(1+\Gamma_k(\mathbf{F},\mathbf{e}))]$ if $\Gamma_k(\mathbf{F},\mathbf{e}) \geq \gamma$ and $R_{\mathrm{eff}} \triangleq 0$ otherwise.

B. Convergence

Fig. 3 investigates the convergence behavior of the considered stochastic optimization algorithms. For comparison, we consider a single-user case containing RIS 1 in Fig. 2, and the other parameters are given in Fig. 3. In Fig. 3, the coordinate value on the y-axis is the objective value of Problem (16) or (32), and not the actual outage probability of the original problem. It is observed from Fig. 3 that the SMM and SSCA algorithms are characterized by an oscillatory convergence which depends on the random channel generations at each iteration. On the other

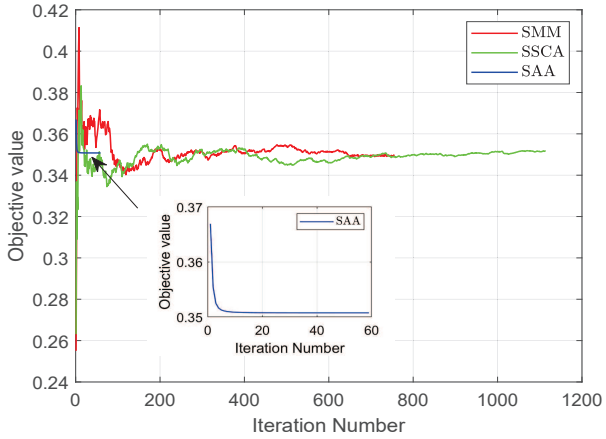


Fig. 3: Convergence behavior of different algorithms, when $N = 8$, $M = 128$, $K = 1$, $U = 1$, and $R_{\text{targ}} = 0.5$ bps/Hz.

TABLE I: Comparison of the CPU time

Algorithms	The CPU time (sec) per iteration	The CPU time (sec)
SMM	0.0025	1.8750
SSCA	0.0042	4.6719
SAA	0.3557	20.9844

hand, using 300 channel realizations for each iteration leads to the monotonic convergence of the SAA algorithm when adopting a monotonically decreasing surrogate function for each channel realization. Although the SAA algorithm requires the least number of iterations to converge, it is much more computationally demanding than the other two algorithms. This fact can be observed in Table I which compares the CPU time consumption of each iteration and total CPU time consumption of iteration convergence for the three considered algorithms. Theoretically, the computational complexity of each iteration of the SAA algorithm is 300 times higher than that of the SMM or SAA algorithms, because at each iteration of the SAA algorithm, each parameter needs to be calculated 300 times for all channel realizations.

C. Single-user system

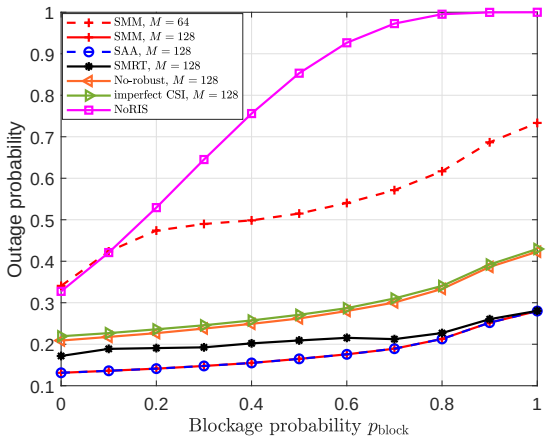
We consider a single-user system containing RIS 1 in Fig. 2 with the target rate of $R_{\text{targ}} = 0.5$ bps/Hz. Fig. 4 illustrates the performance of different algorithms as a function of the blockage probability. First, it can be seen that SMM-based beamforming in the RIS-aided mmWave system

of $M = 64$ outperforms the NoRIS scheme when $p_{\text{block}} \geq 0.1$. The main reason is that the direct BS-user channel may be much stronger than the cascaded BS-RIS-user channel as the latter experiences the double path loss effect, which is significant if the path-loss is large. Therefore, when the blockage probability is small ($p_{\text{block}} \leq 0.1$), the BS tends to allocate the transmit power to the stronger direct path, thus reducing the contribution of the RIS to the system performance. However, by increasing the number of reflecting elements at the RIS to $M = 128$, our proposed SMM algorithm outperforms the NoRIS system in the whole blockage probability region (i.e., $0 \leq p_{\text{block}} \leq 1$). The reason is that the RIS-aided reflecting channels starts to compensate for the performance loss caused by the blockages when $p_{\text{block}} = 0$. In addition, it can be seen that the performance of the SMM-based beamforming is the same as that of the SAA-based beamforming. Last but not least, compared with the No-robust RIS-aided scenario and the scheme with imperfect CSI, the robust RIS beamforming can significantly compensate for the performance loss caused by the presence of random blockages.

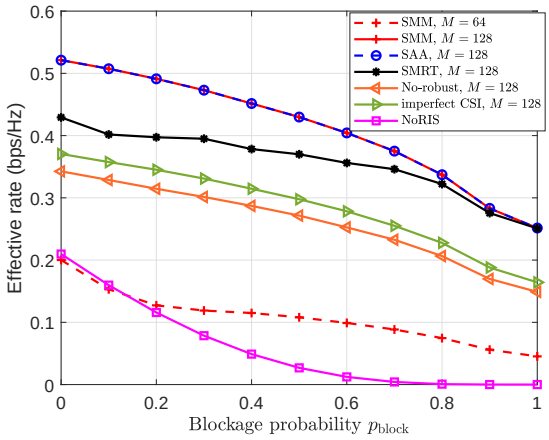
Fig. 5 shows the impact of the size of the RIS and the size of the antenna array at the BS on the outage probability by using the SMM algorithm. It can be observed from Fig. 5(a) that when BS is equipped with $N = 8$ antennas, the RIS plays a significant role in guaranteeing the desired user's QoS and improving the system robustness with increased number of reflection elements ($M : 64 \rightarrow 256$). In other words, a large-size RIS with $M \geq 224$ can still reduce the outage probability to about 0.1, even if the direct channel is completely blocked. Finally in Fig. 5(b) we observe that increasing the number of transmit antennas ($N : 8 \rightarrow 64$) has the same ability to reduce the outage probability as increasing the number of reflection elements ($M : 64 \rightarrow 256$), except when the direct channel is completely blocked.

D. Multiuser system

In this section, we consider a multiuser system with $K = 3$ users, in which the target rate is $R_{\text{targ}} = 0.1$ bps/Hz. As we can see from Fig. 6(a), compared with the No-robust RIS-aided scenario, the robust beamforming in the RIS-aided scheme in a multiuser system can efficiently improve the maximum outage probability. In addition, different from the single-user system of $M = 128$, the RIS-aided setup outperforms the NoRIS scheme only when $p_{\text{block}} \geq 0.2$. This fact is consistent with the observation of $M = 64$ in Fig. 4 (a), that is, the performance gain of an RIS starts to manifest when $p_{\text{block}} = 0.2$. Further increasing the number of RISs and the size of each RIS can reduce the maximum outage probability in the whole blockage probability



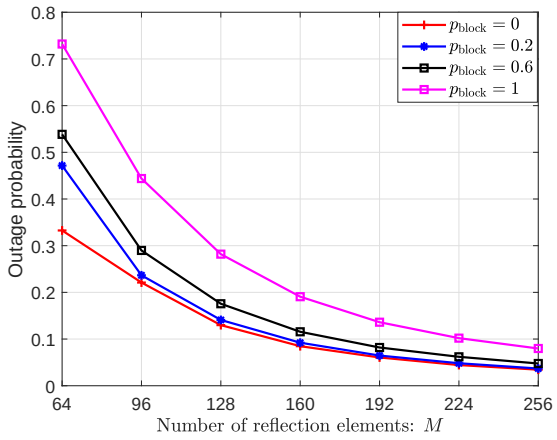
(a) Outage probability



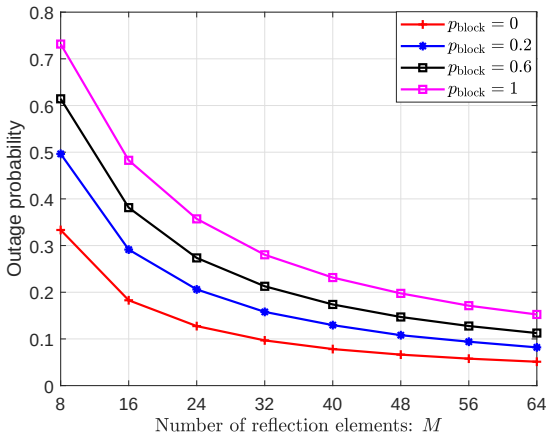
(b) Effective rate

Fig. 4: Comparison of outage probability and effective rate as a function of the blockage probability p_{block} for $N = 8$, $K = 1$, $U = 1$, and $R_{\text{targ}} = 0.5$ bps/Hz.

region (i.e., $0 \leq p_{\text{block}} \leq 1$). Although the improvement in the performance in terms of the maximum outage probability of the large-size RIS beamforming in Fig. 6(a) is slightly weak, the enhancement of the corresponding minimum effective rate shown in Fig. 6(b) is apparent. Specifically, the minimum effective rate of the robust designs is better than that of the non-robust schemes. Moreover, the contribution of the RIS in improving the minimum effective rate is also apparent for the almost blockage probability region (i.e., $p_{\text{block}} \leq 0.8$). Finally, we observe that multiple RISs with a relatively small size can improve the worst-case user's performance, and outperform the performance offered by a single RIS with large size, especially in the multi-user



(a) Number of reflection elements

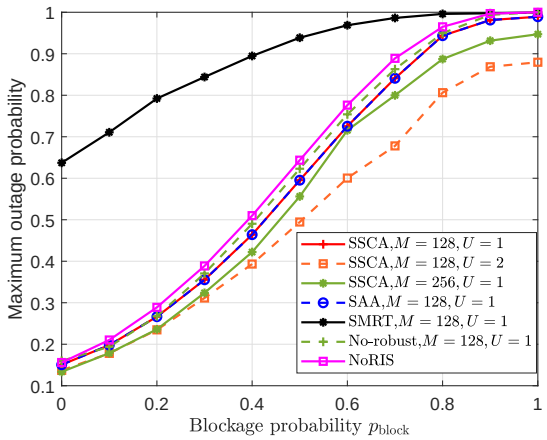


(b) Number of transmit antennas

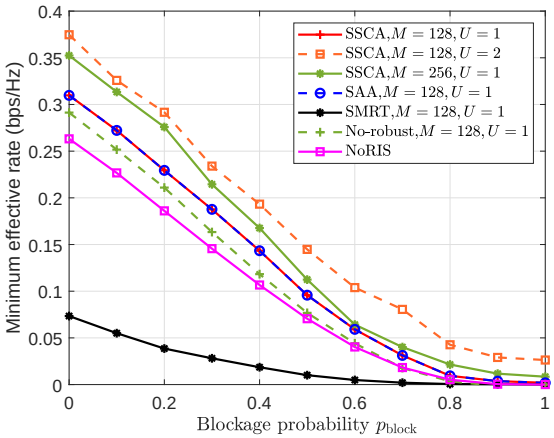
Fig. 5: Comparison of outage probability as a function of M with fixed $N = 8$ and of N with fixed $M = 128$, when $K = 1$, $U = 1$, and $R_{\text{targ}} = 0.5$ bps/Hz.

scenario. This is because the optimal deployment of multiple RISs can provide more spatial diversity for users who are far away from the RIS, thus improving the system performance.

Finally, Fig. 7 investigates the performance as a function of the number of users. For a fair comparison, we consider the setup $N = 32, M = 128$ and $U = 2$. Compared with the NoRIS scheme, the contribution of multiple RISs in reducing the maximum outage probability shown in Fig. 7 decreases as the increase of K. When $K \geq 5$, the RIS scheme cannot guarantee the QoS performance of the worst-case user due to the high outage probability.



(a) Outage probability



(b) Effective rate

Fig. 6: Comparison of maximum outage probability and minimum effective rate as a function of the blockage probability p_{block} for $N = 16$, $K = 3$, and $R_{\text{targ}} = 0.1$ bps/Hz.

VI. Conclusions

In this work, we have improved the reliability of a mmWave system in the presence of random blockages by employing multiple RISs and designing the corresponding robust beamforming. In order to reduce the system outage, we have formulated and solved a maximum outage probability minimization problem which belongs to the family of stochastic optimization problems . More precisely, we have taken into account the statistical CSI and the blockage probability, and have solved the formulated optimization problem by adopting a stochastic optimization framework. Closed-form solutions have been derived at each iteration by adopting the SMM and SSCA

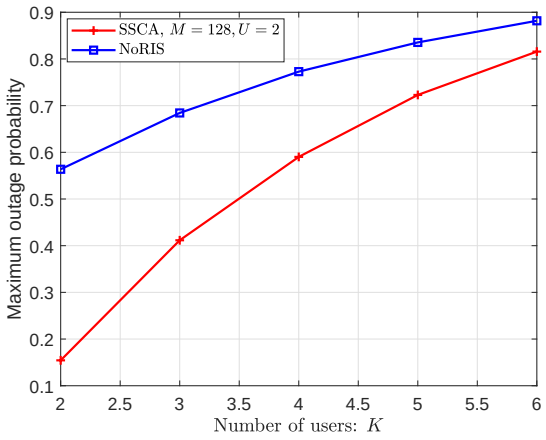


Fig. 7: Comparison of maximum outage probability as a function of the number of users K for $N = 32$, $p_{\text{block}} = 0.6$, and $R_{\text{targ}} = 0.1$ bps/Hz.

methods. The two proposed stochastic methods are guaranteed to converge to the set of stationary points of the original stochastic problems. Selected numerical results have demonstrated the performance gain in terms of outage probability and effective rate of the RIS-aided mmWave systems in the presence of random blockages.

APPENDIX A
THE PROOF OF LEMMA 1

In this subsection, $\mathbf{G}^i$ is dropped for simplicity, i.e., $f\left(\mathbf{f}|\mathbf{G}^i\right)$ is replaced by $f\left(\mathbf{f}\right)$. Since $f\left(\mathbf{f}\right)$ is twice differentiable, we propose a second-order approximation to upper bound $f\left(\mathbf{f}\right)$ at any fixed point $\mathbf{f}^{i-1}$:

$$\begin{aligned} f\left(\mathbf{f}\right) \leq & \hat{f}\left(\mathbf{f}, \mathbf{f}^{i-1}\right) \\ = & f\left(\mathbf{f}^{i-1}\right) + 2\mathrm{Re}\left\{\mathbf{m}_f^{i,\mathrm{H}}(\mathbf{f}-\mathbf{f}^{i-1})\right\} + (\mathbf{f}-\mathbf{f}^{i-1})^{\mathrm{H}}\mathbf{M}_f^i(\mathbf{f}-\mathbf{f}^{i-1}), \end{aligned} \tag{43}$$

where $\mathbf{m}_f^i$ and $\mathbf{M}_f^i$ are to be designed to satisfy Assumption A.

Assumptions (A1) and (A2) are readily satisfied. Assumption (A4) is a derivative consistency condition. Denote $\tilde{\mathbf{f}} \in \mathcal{S}_f$, then the directional derivative of $\hat{f}\left(\mathbf{f}, \mathbf{f}^{i-1}\right)$ at $\mathbf{f}^{i-1}$ with direction $\tilde{\mathbf{f}} - \mathbf{f}^{i-1}$ is

$$2\mathrm{Re}\left\{\mathbf{m}_f^{i,\mathrm{H}}(\tilde{\mathbf{f}}-\mathbf{f}^{i-1})\right\}. \tag{44}$$

The corresponding directional derivative of $f(\mathbf{f})$ is

$$\frac{-\theta e^{-\theta x^i}}{\left(1+e^{-\theta x^i}\right)^2} 2\mathrm{Re}\left\{\mathbf{f}^{i-1,\mathrm{H}}\mathbf{G}^{i,\mathrm{H}}\mathbf{e}^{i-1}\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}\left(\tilde{\mathbf{f}}-\mathbf{f}^{i-1}\right)\right\}, \tag{45}$$

where x^i is given in (15c).

Assumption (A4) is satisfied only when (44) and (45) are equal, yielding

$$\mathbf{m}_f^i = -\frac{\theta e^{-\theta x^i}}{\left(1+e^{-\theta x^i}\right)^2}\mathbf{G}^{i,\mathrm{H}}\mathbf{e}^{i-1}\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}^i\mathbf{f}^{i-1}. \tag{46}$$

In order for Assumption (A3) to hold, it is sufficient to show that $\hat{f}(\mathbf{f},\mathbf{f}^{i-1})$ is an upper bound for each linear cut in any direction. In particular, let $\mathbf{f}=\mathbf{f}^{i-1}+\xi(\tilde{\mathbf{f}}-\mathbf{f}^{i-1}), \forall \xi\in[0,1]$, we need to show

$$f\left(\mathbf{f}^{i-1}+\xi(\tilde{\mathbf{f}}-\mathbf{f}^{i-1})\right) \le f\left(\mathbf{f}^{i-1}\right)+2\xi\mathrm{Re}\left\{\mathbf{m}_f^{i,\mathrm{H}}(\tilde{\mathbf{f}}-\mathbf{f}^{i-1})\right\}+\xi^2(\tilde{\mathbf{f}}-\mathbf{f}^{i-1})^{\mathrm{H}}\mathbf{M}_f^i(\tilde{\mathbf{f}}-\mathbf{f}^{i-1}). \tag{47}$$

Define $L(\xi)=f(\mathbf{f}^{i-1}+\xi(\tilde{\mathbf{f}}-\mathbf{f}^{i-1}))$ and $l(\xi)=\gamma\sigma^2-|\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}^i(\mathbf{f}^{i-1}+\xi(\widetilde{\mathbf{f}}-\mathbf{f}^{i-1}))|^2$. (47) holds if the second-order derivative of $L(\xi)$ is no more than that of the right hand side of (47). The corresponding sufficient condition can be formulated as

$$\frac{\partial^2 L(\xi)}{\partial\xi^2} \le 2\mathrm{Tr}\left[(\tilde{\mathbf{f}}-\mathbf{f}^{i-1})^{\mathrm{H}}\mathbf{M}_f^i(\tilde{\mathbf{f}}-\mathbf{f}^{i-1})\right]. \tag{48}$$

Before deriving the expression of $\partial^2 L(\xi)/\partial\xi^2$, we calculate the first-order derivative of $L(\xi)$, as follows

$$\frac{\partial L(\xi)}{\partial\xi} = g(\xi)\nabla_\xi l(\xi), \tag{49}$$

where $g(\xi)=\frac{\theta e^{-\theta l(\xi)}}{(1+e^{-\theta l(\xi)})^2}$, $\nabla_\xi l(\xi)=-2\mathrm{Re}\{\mathbf{q}^{\mathrm{H}}(\widetilde{\mathbf{f}}-\mathbf{f}^{i-1})\}$, and $\mathbf{q}=\mathbf{G}^{i,\mathrm{H}}\mathbf{e}^{i-1}\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}^i(\mathbf{f}^{i-1}+\xi(\widetilde{\mathbf{f}}-\mathbf{f}^{i-1}))$.

Then, the second-order derivative is derived as

$$\frac{\partial^2 L(\xi)}{\partial\xi^2} = g(\xi)\nabla_\xi^2 l(\xi)-\theta g(\xi)\nabla_\xi l(\xi)\left(\nabla_\xi l(\xi)\right)^{\mathrm{T}}+2\left(1+e^{-\theta l(\xi)}\right)g(\xi)\nabla_\xi l(\xi)\left(g(\xi)\nabla_\xi l(\xi)\right)^{\mathrm{T}}, \tag{50}$$

where $\nabla_\xi^2 l(\xi)=-2\mathrm{Re}\{(\widetilde{\mathbf{f}}-\mathbf{f}^{i-1})^{\mathrm{H}}\mathbf{\Theta}(\widetilde{\mathbf{f}}-\mathbf{f}^{i-1})\}$ and $\mathbf{\Theta}=\xi\mathbf{G}^{i,\mathrm{H}}\mathbf{e}^{i-1}\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}^i$.

(50) is rewritten as a quadratic form of $\mathbf{t}=\widetilde{\mathbf{f}}-\mathbf{f}^{i-1}$, as follows

$$\frac{\partial^2 L(\xi)}{\partial\xi^2} = \begin{bmatrix}\mathbf{t}\\ \mathbf{t}^*\end{bmatrix}^{\mathrm{H}}\mathbf{\Phi}\begin{bmatrix}\mathbf{t}\\ \mathbf{t}^*\end{bmatrix}, \tag{51}$$

where

$$\mathbf{\Phi} = g(\xi)\left(2\left(1+e^{-\theta l(\xi)}\right)g(\xi)-\theta\right)\begin{bmatrix}\mathbf{q}\\ \mathbf{q}^*\end{bmatrix}\begin{bmatrix}\mathbf{q}\\ \mathbf{q}^*\end{bmatrix}^{\mathrm{H}} - g(\xi)\mathbf{I}_2\otimes\mathbf{\Theta}.$$

Furthermore, we also manipulate the right hand side of (48) into the same form as in (51), as follows

$$\begin{bmatrix}\mathbf{t}\\ \mathbf{t}^*\end{bmatrix}^{\mathrm{H}}\begin{bmatrix}\mathbf{I}\otimes\mathbf{M}_f^i & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\otimes\mathbf{M}_f^{i,\mathrm{T}}\end{bmatrix}\begin{bmatrix}\mathbf{t}\\ \mathbf{t}^*\end{bmatrix}. \tag{52}$$

Combining (51) and (52), the sufficient condition in (48) is equivalent to

$$\begin{bmatrix}\mathbf{t}\\ \mathbf{t}^*\end{bmatrix}^{\mathrm{H}}\mathbf{\Phi}\begin{bmatrix}\mathbf{t}\\ \mathbf{t}^*\end{bmatrix} \le \begin{bmatrix}\mathbf{t}\\ \mathbf{t}^*\end{bmatrix}^{\mathrm{H}}\begin{bmatrix}\mathbf{I}\otimes\mathbf{M}_f^i & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\otimes\mathbf{M}_f^{i,\mathrm{T}}\end{bmatrix}\begin{bmatrix}\mathbf{t}\\ \mathbf{t}^*\end{bmatrix},$$

which can be satisfied when $\mathbf{M}_f^i$ satisfies

$$\mathbf{\Phi} \preceq \begin{bmatrix}\mathbf{I}\otimes\mathbf{M}_f^i & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\otimes\mathbf{M}_f^{i,\mathrm{T}}\end{bmatrix}.$$

For convenience, we select $\mathbf{M}_f^i = \alpha_f^i\mathbf{I} = \lambda_{\max}(\mathbf{\Phi})\mathbf{I}$. Then, $\hat{f}(\mathbf{f},\mathbf{f}^{i-1})$ in (43) is designed to be

$$\begin{aligned}&\hat{f}\left(\mathbf{f},\mathbf{f}^{i-1}\right)\\ &= f\left(\mathbf{f}^{i-1}\right)+2\mathrm{Re}\left\{\mathbf{m}_f^{i,\mathrm{H}}(\mathbf{f}-\mathbf{f}^{i-1})\right\}+\alpha_f^i||\mathbf{f}-\mathbf{f}^{i-1}||_2^2\\ &= 2\mathrm{Re}\left\{\mathbf{d}_f^{i,\mathrm{H}}\mathbf{f}\right\}+\alpha_f^i||\mathbf{f}||_2^2+\mathrm{const}_f^i,\end{aligned}$$

where $\mathbf{d}_f^{i,\mathrm{H}}$, α_f^i and const_f^i are defined in Lemma 1. The deterministic expression of $\lambda_{\max}(\mathbf{\Phi})$ is difficult to obtain, therefore we derive the upper bound shown as follows

$$\begin{aligned}&\lambda_{\max}(\mathbf{\Phi})\\ &\overset{(p1)}{\le} 2\left(1+e^{-\theta l(\xi)}\right)g^2(\xi)\lambda_{\max}\left(\begin{bmatrix}\mathbf{q}\\ \mathbf{q}^*\end{bmatrix}\begin{bmatrix}\mathbf{q}\\ \mathbf{q}^*\end{bmatrix}^{\mathrm{H}}\right)\\ &\quad - g(\xi)\lambda_{\min}(\mathbf{I}_2\otimes\mathbf{\Theta}) - \theta g(\xi)\lambda_{\min}\left(\begin{bmatrix}\mathbf{q}\\ \mathbf{q}^*\end{bmatrix}\begin{bmatrix}\mathbf{q}\\ \mathbf{q}^*\end{bmatrix}^{\mathrm{H}}\right)\\ &\overset{(p2)}{=} 4\left(1+e^{-\theta l(\xi)}\right)g^2(\xi)||\mathbf{q}||_2^2\\ &\overset{(p3)}{<} \frac{\theta^2}{2}||\mathbf{q}||_2^2\end{aligned}$$

$$\overset{\text{(p4)}}{\leq} \frac{\theta^2}{2}\lambda_{\max}(\mathbf{G}^{i,\mathrm{H}}\mathbf{e}^{i-1}\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}^{i}\mathbf{G}^{i,\mathrm{H}}\mathbf{e}^{i-1}\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}^{i}) \cdot ||\mathbf{f}^{i-1}+\gamma(\widetilde{\mathbf{f}}-\mathbf{f}^{i-1})||_2^2$$

$$\overset{\text{(p5)}}{\leq} \frac{\theta^2}{2}P_{max}|\mathbf{e}^{i-1,\mathrm{H}}\mathbf{G}^{i}\mathbf{G}^{i,\mathrm{H}}\mathbf{e}^{i-1}|^2.$$

The above inequalities are due to the following mathematical properties:

(p1): $\mathbf{A}$ and $\mathbf{B}$ are Hermitian matrices: $\lambda_{\max}(\mathbf{A})+\lambda_{\max}(\mathbf{B}) \geq \lambda_{\max}(\mathbf{A}+\mathbf{B})$ [43].

(p2): $\mathbf{A}$ is rank one: $\lambda_{\max}(\mathbf{A}) = \mathrm{Tr}\left[\mathbf{A}\right], \lambda_{\min}(\mathbf{A}) = 0$ [43].

(p3): $\left(1+e^{-\theta l(\xi)}\right) g^2(\xi) \leq \theta^2/8$, where the equality holds when $l(\xi)=0$.

(p4): $\mathbf{A}$ is positive semidefinite with maximum eigenvalue $\lambda_{\max}(\mathbf{A})$ and $\mathbf{B}$ is positive semidefinite: $\mathrm{Tr}\left[\mathbf{AB}\right] \leq \lambda_{\max}(\mathbf{A})\mathrm{Tr}\left[\mathbf{B}\right]$ [43].

(p5): Power constraint: $||\mathbf{f}^{i-1}+\gamma(\widetilde{\mathbf{f}}-\mathbf{f}^{i-1})||_2^2 \leq P_{max}$.

Hence, the proof is completed.

APPENDIX B
THE PROOF OF THEOREM 1

Define the random functions

$$g^n(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} f\left(\mathbf{x}|\mathbf{G}^i\right), \tag{53}$$

$$\hat{g}^n(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} \hat{f}\left(\mathbf{x},\mathbf{x}^{i-1}|\mathbf{G}^i\right). \tag{54}$$

To state the convergence result, we need the following lemmas.

Lemma 3 *Suppose Assumptions A and B are satisfied and define a limit point $\bar{\mathbf{x}}$ of a subsequence $\{\mathbf{x}^{n_j}\}_{j=1}^{\infty}$, then there exists uniformly continuous functions $g(\mathbf{x})$ and $\hat{g}(\mathbf{x})$ such that*

$$g(\mathbf{x}) = \lim_{n\to\infty} g^n(\mathbf{x}) = \mathbb{E}\left[f\left(\mathbf{x}|\mathbf{G}\right)\right], \forall \mathbf{x} \in \mathcal{S}_x, \tag{55}$$

$$g(\bar{\mathbf{x}}) = \lim_{j\to\infty} g^{n_j}(\mathbf{x}^{n_j}), \tag{56}$$

$$\hat{g}(\mathbf{x}) = \lim_{j\to\infty} \hat{g}^{n_j}(\mathbf{x}), \forall \mathbf{x} \in \mathcal{S}_x, \tag{57}$$

$$\hat{g}(\bar{\mathbf{x}}) = \lim_{j\to\infty} \hat{g}^{n_j}(\mathbf{x}^{n_j}). \tag{58}$$

Proof: First, it follows that $f\left(\mathbf{x},\mathbf{G}\right)$ is bounded for $\forall \mathbf{x} \in \mathcal{S}_x$ and all channel realizations due to the assumption (B2), and therefore, (55) holds by using the strong law of large numbers [44]. Also, the families of functions $\{g^{n_j}(\mathbf{x})\}$ are equicontinuous and bounded over a

compact set $\mathcal{S}_x$ due to the assumption (B2) and the use of mean value theorem. Thus, by restricting to a subsequence, we have (56). Furthermore, the families of functions $\{\hat{g}^n(\mathbf{x})\}$ are also equicontinuous and bounded over a compact set $\mathcal{S}_x$ due to the assumption (B2) that $||\nabla_{\mathbf{x}}\hat{f}(\mathbf{x}, \mathbf{x}^{i-1}, \mathbf{G})||$ is bounded. Hence the ArzelÃ -Ascoli theorem [45] implies that, by restricting to a subsequence, there exists a uniformly continuous function $\hat{g}(\mathbf{x})$ such that (57) and (58) hold. ■

On the other hand, the update rule of Algorithm 1 leads the following lemma.

Lemma 4 $\lim_{n\to\infty}|\hat{g}^n(\mathbf{x}^n) - g^n(\mathbf{x}^n)|= 0$*, almost surely.*

Proof: The proof of Lemma 4 is the same with that of ([36], Lemma 1) and is omitted for conciseness. ■

Assumption (A3) implies that $\hat{g}^{n_j}(\mathbf{x}) \geq g^{n_j}(\mathbf{x}), \forall \mathbf{x} \in \mathcal{S}_x$. combining it with (55) and (57), we obtain

$$\hat{g}(\mathbf{x}) \geq g(\mathbf{x}), \forall \mathbf{x} \in \mathcal{S}_x. \tag{59}$$

Moreover, combining Lemma 4 with (56) and (58), it yields

$$\hat{g}(\bar{\mathbf{x}}) = g(\bar{\mathbf{x}}). \tag{60}$$

Then, (59) and (60) imply that $\bar{\mathbf{x}}$ is a minimizer of function $\hat{g}(\mathbf{x}) - g(\mathbf{x})$, hence the first-order optimality condition satisfies

$$\nabla\hat{g}(\bar{\mathbf{x}}) - \nabla g(\bar{\mathbf{x}}) = 0. \tag{61}$$

Due to the fact that $\bar{\mathbf{x}}$ is the limit point of Problem (12) or Problem (13), we have $\hat{g}(\bar{\mathbf{x}}) \leq \hat{g}(\mathbf{x}), \forall \mathbf{x} \in \mathcal{S}_x$, which implies that

$$\langle \nabla\hat{g}(\bar{\mathbf{x}}), \mathbf{x} - \bar{\mathbf{x}}\rangle \geq 0, , \forall \mathbf{x} \in \mathcal{S}_x. \tag{62}$$

Combining this with (61), we obtain

$$\langle \nabla g(\bar{\mathbf{x}}), \mathbf{x} - \bar{\mathbf{x}}\rangle \geq 0, , \forall \mathbf{x} \in \mathcal{S}_x, \tag{63}$$

which means that the directional derivative of the objective function $g(\mathbf{x})$ is non-negative for every feasible direction at $\bar{\mathbf{x}}$. Recalling that $\mathbf{x} \in \{\mathbf{f}, \mathbf{e}\}$ and defining the limit points $\{\bar{\mathbf{f}}, \bar{\mathbf{e}}\}$, (63) is equivalent to

$$\begin{cases} \left\langle \nabla g(\bar{\mathbf{f}}), \mathbf{f} - \bar{\mathbf{f}} \right\rangle \geq 0, \forall \mathbf{f} \in \mathcal{S}_f, \\ \langle \nabla g(\bar{\mathbf{e}}), \mathbf{e} - \bar{\mathbf{e}}\rangle \geq 0, \forall \mathbf{e} \in \mathcal{S}_e. \end{cases}$$

Therefore, according to [41], $\{\bar{\mathbf{f}}, \bar{\mathbf{e}}\}$ is a stationary point of Problem (11) due to the regularity of $g(\cdot)$.

APPENDIX C
THE PROOF OF THEOREM 1

Define the random functions

$$G^n(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} F\left(\mathbf{x}|\mathbf{G}^i\right), \tag{64}$$

$$\hat{G}^n(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} \hat{F}\left(\mathbf{x}, \mathbf{x}^{i-1}|\mathbf{G}^i\right). \tag{65}$$

To state the convergence result, we need the following lemmas.

Lemma 5 *Suppose Assumptions B and C are satisfied and define a limit point $\bar{\mathbf{x}}$ of a subsequence $\{\mathbf{x}^{n_j}\}_{j=1}^{\infty}$, then there exists uniformly continuous functions $G(\mathbf{x})$ and $\hat{G}(\mathbf{x})$ such that*

$$G(\mathbf{x}) = \lim_{n\to\infty} G^n(\mathbf{x}) = \mathbb{E}\left[F\left(\mathbf{x}|\mathbf{G}\right)\right], \forall \mathbf{x} \in \mathcal{S}_x, \tag{66}$$

$$G(\bar{\mathbf{x}}) = \lim_{j\to\infty} G^{n_j}(\mathbf{x}^{n_j}), \tag{67}$$

$$\hat{G}(\mathbf{x}) = \lim_{j\to\infty} \hat{G}^{n_j}(\mathbf{x}), \forall \mathbf{x} \in \mathcal{S}_x, \tag{68}$$

$$\hat{G}(\bar{\mathbf{x}}) = \lim_{j\to\infty} \hat{G}^{n_j}(\mathbf{x}^{n_j}). \tag{69}$$

Proof: The proof of Lemma 5 is the same with that of Lemma 3 and is omitted for brevity. ■

On the other hand, $\mathbf{x}^{n_j}$ is the minimizer of $\hat{G}^{n_j}(\mathbf{x})$, thus

$$\hat{G}^{n_j}(\mathbf{x}^{n_j}) \leq \hat{G}^{n_j}(\mathbf{x}), \forall \mathbf{x} \in \mathcal{S}_x. \tag{70}$$

Assuming $j \to \infty$, and combining (68) and (69), we obtain $\hat{G}(\bar{\mathbf{x}}) \leq \hat{G}(\mathbf{x}), \forall \mathbf{x} \in \mathcal{S}_x$, which implies that its first-order optimality condition satisfies

$$\left\langle \nabla \hat{G}(\bar{\mathbf{x}}), \mathbf{x} - \bar{\mathbf{x}} \right\rangle \geq 0, , \forall \mathbf{x} \in \mathcal{S}_x. \tag{71}$$

By combining (71) and Assumption (C3), we finally obtain

$$\left\langle \nabla G(\bar{\mathbf{x}}), \mathbf{x} - \bar{\mathbf{x}} \right\rangle \geq 0, , \forall \mathbf{x} \in \mathcal{S}_x. \tag{72}$$

Since that $\mathbf{x} \in \{\mathbf{F}, \mathbf{e}\}$, we define the limit points $\{\bar{\mathbf{F}}, \bar{\mathbf{e}}\}$, (72) is then equivalent to

$$\begin{cases} \left\langle \nabla G(\bar{\mathbf{F}}), \mathbf{F} - \bar{\mathbf{F}} \right\rangle \geq 0, \forall \mathbf{F} \in \mathcal{S}_f, \\ \langle \nabla G(\bar{\mathbf{e}}), \mathbf{e} - \bar{\mathbf{e}} \rangle \geq 0, \forall \mathbf{e} \in \mathcal{S}_e. \end{cases}$$

Therefore, according to [41], $\{\bar{\mathbf{F}}, \bar{\mathbf{e}}\}$ is a stationary point of Problem (28) due to the regularity of $G(\cdot)$.

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] V. Raghavan *et al.*, "Spatio-temporal impact of hand and body blockage for millimeter-wave user equipment design at 28 GHz," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 46–52, Dec. 2018.

[3] V. Raghavan, M.-L. Chi, M. A. Tassoudji, O. H. Koymen, and J. Li, "Antenna placement and performance tradeoffs with hand blockage in millimeter wave systems," *IEEE Trans. Commun.*, vol. 67, no. 4, pp. 3082–3096, Apr. 2019.

[4] 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.

[5] T. Bai, R. Vaze, and R. W. Heath, "Analysis of blockage effects on urban cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 5070–5083, Sept. 2014.

[6] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[7] 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.

[8] M. Alrabeiah and A. Alkhateeb, "Deep learning for mmWave beam and blockage prediction using sub-6 GHz channels," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5504–5518, Sept. 2020.

[9] 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.

[10] H. Iimori *et al.*, "Stochastic learning robust beamforming for millimeter-wave systems with path blockage," *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1557–1561, Sept. 2020.

[11] G. Zhou, C. Pan, H. Ren, K. Wang, M. Elkashlan, and M. Di Renzo, "Stochastic learning-based robust beamforming design for RIS-aided millimeter-wave systems in the presence of random blockages," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 1057–1061, Jan. 2021.

[12] 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).

[13] ——, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

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

[15] X. Yu, D. Xu, D. W. K. Ng, and R. Schober, "IRS-assisted green communication systems: Provable convergence and robust optimization," *IEEE Trans. Commun., Early Access*, 2021.

[16] C. Pan, H. Ren, K. Wang *et al.*, "Reconfigurable intelligent surface for 6G and beyond: Motivations, principles, applications, and research directions," *IEEE Commun. Mag.*, 2021 (to appear).

[17] 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, Aug. 2020.

[18] T. Bai, C. Pan, Y. Deng *et al.*, "Latency minimization for intelligent reflecting surface aided mobile edge computing," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2666–2682, Nov. 2020.

[19] A. Abrardo, D. Dardari, and M. Di Renzo, "Intelligent reflecting surfaces: Sum-rate optimization based on statistical CSI," *IEEE Trans. Commun., Early Access*, pp. 1–1, 2021.

[20] G. Zhou, C. Pan, H. Ren *et al.*, "Robust beamforming design for intelligent reflecting surface aided MISO communication systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1658–1662, Oct. 2020.

[21] K. Zhi, C. Pan, H. Ren, and K. Wang, "Power scaling law analysis and phase shift optimization of RIS-aided massive MIMO systems with statistical CSI," Oct. 2020. [Online]. Available: https://arxiv.org/abs/2010.13525

[22] C. Pan, H. Ren, K. Wang *et al.*, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.

[23] Q. Wu and R. Zhang, "Joint active and passive beamforming optimization for intelligent reflecting surface assisted SWIPT under QoS constraints," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1735–1748, Aug. 2020.

[24] S. Hong, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Artificial-noise-aided secure mimo wireless communications via intelligent reflecting surface," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7851–7866, Dec. 2020.

[25] P. Mursia, V. Sciancalepore, A. Garcia-Saavedra, L. Cottatellucci, X. Costa-Pérez, and D. Gesbert, "RISMA: Reconfigurable intelligent surfaces enabling beamforming for IoT massive access," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 1072–1085, Apr. 2021.

[26] B. Di, H. Zhang, L. Li, L. Song, Y. Li, and Z. Han, "Practical hybrid beamforming with finite-resolution phase shifters for reconfigurable intelligent surface based multi-user communications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4565–4570, Apr. 2020.

[27] P. Wang, J. Fang, X. Yuan, Z. Chen, and H. Li, "Intelligent reflecting surface-assisted millimeter wave communications: Joint active and passive precoding design," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 14 960–14 973, Dec. 2020.

[28] 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.

[29] Y. Pan, K. Wang, C. Pan, H. Zhu, and J. Wang, "Sum rate maximization for intelligent reflecting surface assisted terahertz communications," 2020. [Online]. Available: https://arxiv.org/abs/2008.12246

[30] B. Ning, Z. Chen, W. Chen, and Y. Du, "Channel estimation and transmission for intelligent reflecting surface assisted THz communications," in *IEEE ICC 2020*, pp. 1–7.

[31] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "A framework of robust transmission design for IRS-aided MISO communications with imperfect cascaded channels," *IEEE Trans. Signal Process.*, vol. 68, pp. 5092–5106, Aug. 2020.

[32] S. Shen, B. Clerckx, and R. Murch, "Modeling and architecture design of intelligent reflecting surfaces using scattering parameter network analysis," *IEEE Trans. Wireless Commun., early access*, 2021.

[33] A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Sel. Areas Commun.*, vol. 5, no. 2, pp. 128–137, Feb. 1987.

[34] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Wireless Communications. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[35] J. Mairal, "Stochastic majorization-minimization algorithms for largescale optimization," *Advances in Neural Information Processing Systems*, pp. 2283–2291, 2013.

[36] M. Razaviyayn, M. Sanjabi, and Z.-Q. Luo, "A stochastic successive minimization method for nonsmooth nonconvex optimization with applications to transceiver design in wireless communication networks," *Springer Verlag New York*, pp. 515–545, 2016.

[37] R. Varadhan and C. Roland, "Simple and globally convergent methods for accelerating the convergence of any EM algorithm," *Scand. J. Statist.*, vol. 35, no. 2, pp. 335–353, 2008.

[38] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge Univ. Press, 2004.

[39] S. Xu, "Smoothing method for minimax problems," *Comput. Optim. Appl.*, vol. 20, no. 3, pp. 267–279, 2001.

[40] M. Razaviyayn, *Successive convex approximation: Analysis and applications*. Ph.D. dissertation, Univ. Minnesota, Minneapolis, MN, USA, 2014.

[41] N. Dunford and J. Schwartz, *Successive convex approximation: Analysis and applications*. Ph.D. dissertation, Univ. Minnesota, Minneapolis, MN, USA, 2014.

[42] *Technical Specification Group Radio Access Network; Study on Channel Model for Frequencies from 0.5 to 100 GHz (Release 16)*. document 3GPP TR 38.901 V16.1.0, Dec. 2019.

[43] P. Maher and H. Luthepohl, "Handbook of matrices," *The Math. Gaz.*, vol. 83, no. 498, p. 557, 1999.

[44] B. Fristedt and L. Gray, *A Modern Approach to Probability Theory*. Birkhuser, Boston, 1996.

[45] N. Dunford and J. Schwartz, *Linear operators. Part 1: General theory*. Interscience Publications, New York, 1958.