Stochastic Learning Robust Beamforming for Millimeter-Wave Systems With Path Blockage

Hiroki Iimori, *Graduate Student Member, IEEE*, Giuseppe Thadeu Freitas de Abreu, *Senior Member, IEEE*, Omid Taghizadeh, *Member, IEEE*, Razvan-Andrei Stoica, *Member, IEEE*, Takanori Hara, *Graduate Student Member, IEEE*, and Koji Ishibashi, *Member, IEEE*

***Abstract*—We introduce a new robust, outage minimum, millimeter wave (mmWave) coordinated multipoint (CoMP) beamforming scheme to combat the random path blockages typical of mmWave systems. Unlike state-of-the-art methods, which are of limited applicability in practice due to their combinatorial nature which leads to prohibitive complexity, the proposed method is based on a stochastic-learning-approach, which learns crucial blockage patterns without resorting to the well-known worst-case optimization framework. Simulation results demonstrate the superior performance of the proposed method both in terms of outage probability and effective achievable rate.**

***Index Terms*—Millimeter wave communication, MIMO, blockages, coordinated beamforming, machine learning.**

I. Introduction

THE EXPONENTIALLY increasing demand for wireless data rates imply that future wireless systems such as the fifth generation and beyond (5G+) cellular networks must be designed to cope with severe spectrum shortages. Millimeter wave (mmWave) technology, which aims at the realization of multi giga bits per second (Gbps) data rates by operating at higher frequency bands between 24 GHz and 300 GHz, has been studied in recent years as a key solution to this problem [1], [2].

Indeed, mmWave communications have the great advantage of enabling systems to be equipped with more antenna elements, due to the shorter carrier wavelengths in comparison to traditional microwave systems. On the other hand, a major challenge of mmWave systems is their severer signal attenuation compared to microwave systems, which stems not only from the fact that received signal power decreases in proportional to the carrier frequency squared, as described by Friis formula, but also from less diffraction and higher levels of absorption by atmospheric gases [3]. In light of the above, many advanced channel estimation and hybrid beamforming methods have been proposed in the literature [4]–[9] in order to both exploit the system dimension advantage and compensate for the propagation limitation of mmWave systems.

A fundamental challenge that remains, however, is that mmWave signals are susceptible to random blockage resulting in unpredictable channel uncertainties [10]–[12]. To overcome this difficulty, new blockage-robust mmWave beamforming approaches have been proposed in recent literature. For instance, in [13], [14] a machine-learning based proactive received power prediction method was proposed, which leverages spatiotemporal image sensing side-information to enable the acquisition of blockage probability.

In turn, in [15], [16], a robust coordinated multipoint (CoMP) beamforming design based on the worst case optimization approach has been proposed to ensure a prescribed communication performance under such random blockages.

In this letter we contribute to this trending topic by proposing a stochastic framework for the design of outage-minimum robust mmWave CoMP systems, which has two main differences over the state-of-the-art proposed in [15], [16]. First, while the algorithm proposed in [15], [16] relies on the well-known worst-case optimization approach [17], which requires all conceivable quality of service (QoS) constraints to be considered thus leading to *high computational complexities* and *possible instabilities*,[1] our proposed method is based on an stochastic optimization framework that enables us to *avoid dealing with such combinatorial operations* and yields decent solutions even in ill-conditioned scenarios. Secondly, while a maximization problem with a sum-rate objective was considered in [15], [16], the fundamental goal of robust beamforming design for mmWave systems subjected to uncertain blockage is rather to minimize the required QoS violation (i.e., outage probability), indicating that the formulation in the latter is somewhat sub-optimal.

II. System Model and Problem Formulation

A. System Model

Consider the downlink of a single carrier CoMP narrowband mmWave system in which multiple synchronized base stations (BSs) cooperatively serves single-antenna downlink users subjected to unpredictable blockages. For such a scenario, let $b \in \mathscr{B} \triangleq \{1, 2, \ldots, B\}$ and $u \in$

[1] The worst-case method requires that the target QoS be satisfied even in ill-conditioned scenarios, which may not admit a feasible solution in some realistic circumstances.

Manuscript received May 5, 2020; accepted May 20, 2020. Date of publication May 26, 2020; date of current version September 9, 2020. This work was supported in part by the Ministry of Internal Affairs and Communications in Japan under Grant JPJ000254. The associate editor coordinating the review of this article and approving it for publication was C. Pan. *(Corresponding author: Hiroki Iimori.)*

Hiroki Iimori and Giuseppe Thadeu Freitas de Abreu are with the Focus Area Mobility, Department of Computer Science and Electrical Engineering, Jacobs University Bremen, 28759 Bremen, Germany (e-mail: h.iimori@ieee.org; g.abreu@jacobs-university.de).

Omid Taghizadeh is with the Network Information Theory Group, Technische Universität Berlin, 10623 Berlin, Germany (e-mail: taghizadehmotlagh@tu-berlin.de).

Razvan-Andrei Stoica is with the Department of Research and Development, WIOsense GmbH & Company KG, 28717 Bremen, Germany (e-mail: rstoica@ieee.org).

Takanori Hara and Koji Ishibashi are with the Advanced Wireless and Communication Research Center, University of Electro-Communications, Tokyo 182-8585, Japan (e-mail: hara@awcc.uec.ac.jp; koji@ieee.org).

Digital Object Identifier 10.1109/LWC.2020.2997366

$\mathscr{U} \triangleq \{1, 2, \ldots, U\}$ denote the BS and downlink user indices, respectively, with B and U denoting the total number of BSs and downlink users.

It has been shown [11], [15], [18], [19] that unpredictable blockages occur with 20%-60% probability, leading to significant loss in the achievable throughput. In order to abstract such non-ideal effects, we consider a probabilistic model [15], [16], in which each path undergoes random and independent blockage, whose frequency of occurrence can however be obtained as side-information for the system optimization, as shown in [13]. In other words, the blockage probability of each component channel path can be assumed to be known at the BSs and employed in robust beamforming designs.

Following [20], it is assumed that the channel contains a random number $K_{b,u}$ of clusters, which is modeled as [21] $K_{b,u} \sim \max(1, \mathrm{Possion}(\lambda))$ with the intensity parameter λ, where one of the number of clusters $K_{b,u}$ corresponds to line-of-sight (LoS) and the rest of them are regarded as non-line-of-sight (NLoS). It is furthermore assumed that channel state information (CSI) is acquired and tracked continuously exploiting the reciprocity between uplink and downlink of standard time division duplex (TDD) systems, such that channel estimates can be modeled as

$$\hat{\boldsymbol{h}}_{b,u} = \sqrt{\frac{1}{K_{b,u}}} \sum_{m=1}^{K_{b,u}} g_{b,u}^{m} \boldsymbol{a}_{T}(\phi_{b,u}^{m}), \tag{1}$$

where $\phi_{b,u}^{m}$ denotes the angles of departure (AoD) of the m-th cluster from the b-th BS towards the u-th downlink user and $\boldsymbol{a}_{T}(\phi_{b,u}^{m})$ is an array response vector at the transmitter, while $g_{b,u}^{m}$ is the associated channel gain modeled as $g_{b,u}^{m} \sim \mathcal{CN}(0, 10^{\frac{-\mathrm{PL}_{b,u}^{m}}{10}})$ with $\mathrm{PL}_{b,u}^{m} = \alpha + 10\beta \log_{10}(d_{b,u}) + \xi$ [dB] in which $d_{b,u}$ is the distance (in meters) between the b-th BS and the u-th user and the parameters α, β and ξ are determined according to [21, Table I].

In spite of the knowledge of $\hat{\boldsymbol{h}}_{b,u}$, during actual downlink the system might be subjected to partial outage if and when any or some of the LoS and non-line-of-sight (NLoS) clusters become temporarily blocked, such that the actual channel between the b-th BS and the u-th user can then be modeled as

$$\boldsymbol{h}_{b,u} = \sqrt{\frac{1}{K_{b,u}}} \sum_{m=1}^{K_{b,u}} \omega_{b,u}^{m} g_{b,u}^{m} \boldsymbol{a}_{T}(\phi_{b,u}^{m}), \tag{2}$$

where $\omega_{b,u}^{m} \in \{0, 1\}\, \forall m \in \{1, \ldots, K_{b,u}\}$ denotes a Bernoulli random variable with mean $p_{b,u}^{m}$ depicting the corresponding outage probability.[2]

Assuming that each BS is equipped with a uniform linear array (ULA) with N_t transmit antenna elements and half-wavelength spacing, let $\boldsymbol{f}_{b,u} \in \mathbb{C}^{N_t \times 1}$ denote the transmit beamforming vector from the b-th BS towards the u-th user subject to a maximum transmit power constraint $\sum_{u \in \mathscr{U}} \|\boldsymbol{f}_{b,u}\|_2^2 \leq P_{\max,b}$, such that the received signal y_u at the u-th downlink user can be written as

$$y_u = \sum_{b \in \mathscr{B}} \boldsymbol{h}_{b,u}^{\mathrm{H}} \boldsymbol{f}_{b,u} x_u + \sum_{u' \in \mathscr{U} \setminus u} \sum_{b \in \mathscr{B}} \boldsymbol{h}_{b,u}^{\mathrm{H}} \boldsymbol{f}_{b,u'} x_{u'} + n_u$$

$$= \boldsymbol{h}_{u}^{\mathrm{H}} \boldsymbol{f}_{u} x_u + \sum_{u' \in \mathscr{U} \setminus u} \boldsymbol{h}_{u}^{\mathrm{H}} \boldsymbol{f}_{u'} x_{u'} + n_u, \tag{3}$$

where x_u is the unit-power transmit signal intended to the u-th user, n_u denotes independent and identically distributed (i.i.d.) circularly symmetric additive white Gaussian noise (AWGN) at the u-th user, i.e., $n_u \sim \mathcal{CN}(0, \sigma_u^2)$, and the vectors $\boldsymbol{h}_u$ and $\boldsymbol{f}_u$ are respectively defined for all $u \in \mathscr{U}$ as $\boldsymbol{h}_u \triangleq [\boldsymbol{h}_{1,u}^{\mathrm{T}}, \ldots, \boldsymbol{h}_{B,u}^{\mathrm{T}}]^{\mathrm{T}}$ and $\boldsymbol{f}_u \triangleq [\boldsymbol{f}_{1,u}^{\mathrm{T}}, \ldots, \boldsymbol{f}_{B,u}^{\mathrm{T}}]^{\mathrm{T}}$.

B. Problem Formulation

Given the system model described above, in this subsection we offer an optimization problem formulation corresponding to the mmWave robust CoMP downlink beamforming design subject to unpredictable path blockage, which does not rely on a standard deterministic robust optimization framework such as the ones considered, e.g., in [15], [16].

In particular, our formulation consists of the following stochastic sum-outage-probability minimization problem[3]:

$$\underset{\boldsymbol{f}}{\text{minimize}} \quad \sum_{u \in \mathscr{U}} \Pr\{\Gamma_u(\boldsymbol{h}_u, \boldsymbol{f}) < \gamma_u\} \tag{4a}$$

$$\text{subject to} \quad \sum_{u \in \mathscr{U}} \|\boldsymbol{f}_{b,u}\|_2^2 \leq P_{\max,b}\ \forall\, b, \tag{4b}$$

where γ_u denotes the target signal-to-interference-plus-noise ratio (SINR) for the u-th user and the corresponding effective SINR Γ_u can be written as

$$\Gamma_u(\boldsymbol{h}_u, \boldsymbol{f}) = \frac{|\boldsymbol{h}_u^{\mathrm{H}} \boldsymbol{f}_u|^2}{\sum_{u' \in \mathscr{U} \setminus u} |\boldsymbol{h}_u^{\mathrm{H}} \boldsymbol{f}_{u'}|^2 + \sigma_u^2}, \tag{5}$$

where $\boldsymbol{f} \triangleq [\boldsymbol{f}_1^{\mathrm{T}}, \ldots, \boldsymbol{f}_U^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{BUN_t \times 1}$.

Notice that in contrast to deterministic formulations in [15], [16] – which are based on a worst-case sum throughput maximization framework and Taylor convexification with complexity that grows exponentially with B, U and $K_{b,u}$ – Equation (4) is stochastic and has only as many constraints as the number of BSs due to the power constraints. In addition, compared to state-of-the-art robust beamforming methods such as those in [15]–[17], which aim at maximizing sum throughput in detriment of outage, the objective of the mmWave optimization problem formulated in equation (4) is to preserve the system's continuity subject to QoS guarantees and under maximum transmit power constraints, which we argue is a more direct solution to the problem posed by random blockage.

III. Proposed Method

In this section, we propose a novel robust beamforming algorithm based on stochastic learning, aiming at solving the intended nondeterministic optimization problem given in equation (4). To that end, we first introduce an indicator function $\mathbb{1}_{\Gamma_u(\boldsymbol{h}_u, \boldsymbol{f}) < \gamma_u}$ defined as

$$\mathbb{1}_{\Gamma_u(\boldsymbol{h}_u, \boldsymbol{f}) < \gamma_u} = \begin{cases} 1 & \text{if } \Gamma_u(\boldsymbol{h}_u, \boldsymbol{f}) < \gamma_u \\ 0 & \text{otherwise} \end{cases}, \tag{6}$$

which yields

$$\underset{\boldsymbol{f}}{\text{minimize}} \quad \sum_{u \in \mathscr{U}} \mathbb{E}_{\omega_{b,u}^{m}} \left[\mathbb{1}_{\Gamma_u(\boldsymbol{h}_u, \boldsymbol{f}) < \gamma_u} \right] \tag{7a}$$

[2]Although one may consider the case that some paths are blocked at the channel estimation and unblocked for the data transmission, we assumed that such scenarios are negligible as the channel would not simply be considered for data transmission due to the few number of mmWave paths [21].

[3]In this letter, we consider the sum-outage probability as objective for the sake of tractability and effectiveness, which can be seen as a variate of the conventional sum-rate maximization or sum-mean square error (MSE) minimization problems.

$$\text{subject to} \sum_{u\in\mathscr{U}} \|\boldsymbol{f}_{b,u}\|_2^2 \leq P_{\max,b} \ \forall b. \tag{7b}$$

One may notice that equation (7) can be seen as an empirical risk minimization (ERM) problem [22], studied thoroughly in the machine learning and stochastic optimization literature and known to be solved efficiently via stochastic approximation approaches, which are widely adopted due to their complexity and memory requirements, as well as their easy-to-implement nature [23].

In light of the fact that the channel gains and corresponding AoDs are assumed to be known at BSs, as described in equations (1) and (2), we furthermore remark that possible combinations of blockage patterns due to $\omega_{b,u}^m$ can be randomly generated under the assumption that path blockage probabilities are somehow predictable [13] and can be utilized during stochastic optimization as part of the data set design.

Let $\tilde{\boldsymbol{h}}_u^t$ be the t-th data batch for $\boldsymbol{h}_u$ and $t \in \{1, 2, \ldots, N\}$ with N denoting the size of training data. Then, Equation (7) can be further rewritten in an ERM fashion as

$$\underset{\boldsymbol{f}}{\text{minimize}} \ \frac{1}{N}\sum_{t=1}^{N}\sum_{u\in\mathscr{U}}\left[\mathbb{1}_{\Gamma_u\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right)<\gamma_u}\right] \tag{8a}$$

$$\text{subject to} \sum_{u\in\mathscr{U}} \|\boldsymbol{f}_{b,u}\|_2^2 \leq P_{\max,b} \ \forall b. \tag{8b}$$

A universal technique to deal with variety of ERM problems such as that described by equation (8) is the gradient descent (GD) approach. This approach requires, however, the successive evaluation of sum-gradients which imposes high complexity especially in setups with large data sizes. In order to circumvent this issue, an efficient alternative to GD is stochastic gradient descent (SGD) [24], in which the solution at each t-th iteration is updated based on the recursion

$$\boldsymbol{f}^{(t)} = \boldsymbol{f}^{(t-1)} - \alpha_t \sum_{u\in\mathscr{U}} \nabla\mathbb{1}_{\Gamma_u\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right)<\gamma_u}, \tag{9}$$

where α_t is a suitable stepsize at the t-th iteration.

Notice that it is difficult to directly compute the gradient direction in Equation (9) due to the non-smoothness of the indicator function given by equation (6). Hence, we further introduce the generalized smooth hinge surrogate function [25], which is given by

$$\nu\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right) = \begin{cases} 0 & \text{if } 1 - \frac{\Gamma_u\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right)}{\gamma_u} < 0 \\ \frac{1}{2\varepsilon}\left(1 - \frac{\Gamma_u\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right)}{\gamma_u}\right)^2 & \text{otherwise} \\ 1 - \frac{\Gamma_u\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right)}{\gamma_u} - \frac{\varepsilon}{2} & \text{if } 1 - \frac{\Gamma_u\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right)}{\gamma_u} > \varepsilon, \end{cases} \tag{10}$$

where $0 < \varepsilon \ll 1$ and its gradient with respect to $\boldsymbol{f}$ is

$$\nabla\nu\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right) = \begin{cases} 0 & \text{if } 1 - \frac{\Gamma_u\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right)}{\gamma_u} < 0 \\ \frac{\frac{\Gamma_u\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right)}{\gamma_u} - 1}{\varepsilon}\frac{\nabla\Gamma_u\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right)}{\gamma_u} & \text{otherwise} \\ -\frac{\nabla\Gamma_u\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right)}{\gamma_u} & \text{if } 1 - \frac{\Gamma_u\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right)}{\gamma_u} > \varepsilon. \end{cases} \tag{11}$$

Algorithm 1 SGD-OutMin Beamforming Scheme

Inputs: Initial estimate: $\boldsymbol{f}^{(0)}$, Data set: $\mathcal{H}$
Outputs: Optimized beamforming vector: $\boldsymbol{f}$

Set $t = 1$, $\mu = \infty$ and $t_{\max} = 10^5$.
while $t \leq t_{\max}$ and $\mu \geq 10^{-6}$ **do**
  Sample data batch $\tilde{\boldsymbol{h}}_u^t$ $\forall u$ from $\mathcal{H}$.
  Compute $\sum_{u\in\mathscr{U}} \nabla\nu\left(\tilde{\boldsymbol{h}}_u^t, \boldsymbol{f}^{(t-1)}\right)$ from eq. (11).
  Update $\boldsymbol{f}^{(t)}$ according to eq. (13).
  Project $\boldsymbol{f}^{(t)}$ onto the feasible set (12b) [26], [27]
  $\mu = \|\boldsymbol{f}^{(t)} - \boldsymbol{f}^{(t-1)}\|$ and $t = t + 1$.
end while
$\boldsymbol{f} = \boldsymbol{f}^{(t)}$.

From the above, Equation (8) can be rewritten as

$$\underset{\boldsymbol{f}}{\text{minimize}} \ \frac{1}{N}\sum_{u\in\mathscr{U}}\sum_{t=1}^{N}\nu\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right) \tag{12a}$$

$$\text{subject to} \sum_{u\in\mathscr{U}} \|\boldsymbol{f}_{b,u}\|_2^2 \leq P_{\max,b} \ \forall b, \tag{12b}$$

which can be solved by the projected SGD algorithm with the following update

$$\boldsymbol{f}^{(t)} = \boldsymbol{f}^{(t-1)} - \alpha_t \sum_{u\in\mathscr{U}} \nabla\nu\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}^{(t-1)}\right), \tag{13}$$

where $\boldsymbol{f}^{(t-1)}$ indicates the solution at $(t-1)$-th iteration.

Note that in equation (13), we omit the normalizing factor $1/N$ as it is independent from the optimal condition. In order to compute the gradient $\nabla\Gamma_u$ with respect to the stacked beamforming vector $\boldsymbol{f}$, it is necessary to equivalently reformulate the SINR formula given in equation (5) as

$$\Gamma_u\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right) = \frac{\boldsymbol{f}^{\mathrm{H}}\tilde{\boldsymbol{H}}_u^t\boldsymbol{f}}{\boldsymbol{f}^{\mathrm{H}}\tilde{\boldsymbol{H}}_{\bar{u}}^t\boldsymbol{f} + \sigma_u^2}, \tag{14}$$

with

$$\tilde{\boldsymbol{H}}_u^t = \mathrm{diag}(\boldsymbol{e}_u) \otimes \tilde{\boldsymbol{h}}_u^t\tilde{\boldsymbol{h}}_u^{t\mathrm{H}}, \tag{15a}$$

$$\tilde{\boldsymbol{H}}_{\bar{u}}^t = \mathrm{diag}(\bar{\boldsymbol{e}}_u) \otimes \tilde{\boldsymbol{h}}_u^t\tilde{\boldsymbol{h}}_u^{t\mathrm{H}}, \tag{15b}$$

where $\boldsymbol{e}_u \in \{0,1\}^{U\times 1}$ is the standard basis of length U in which only its u-th element is 1, $\bar{\boldsymbol{e}}_u$ denotes the ones' complement of $\boldsymbol{e}_u$ (i.e., negation of $\boldsymbol{e}_u$) and $\otimes$ expresses the Kronecker product operator.

In light of all the above, the gradient $\nabla\Gamma_u(\tilde{\boldsymbol{h}}_u^i,\boldsymbol{f})$ with respect to $\boldsymbol{f}$ can be finally written in closed-form as

$$\nabla\Gamma_u\left(\tilde{\boldsymbol{h}}_u^t,\boldsymbol{f}\right) = \frac{\tilde{\boldsymbol{H}}_u^t\boldsymbol{f}}{\boldsymbol{f}^{\mathrm{H}}\tilde{\boldsymbol{H}}_{\bar{u}}^t\boldsymbol{f}+\sigma_u^2} - \frac{\boldsymbol{f}^{\mathrm{H}}\tilde{\boldsymbol{H}}_u^t\boldsymbol{f}}{\left(\boldsymbol{f}^{\mathrm{H}}\tilde{\boldsymbol{H}}_{\bar{u}}^t\boldsymbol{f}+\sigma_u^2\right)^2}\tilde{\boldsymbol{H}}_{\bar{u}}^t\boldsymbol{f}. \tag{16}$$

A summary of the here-proposed SGD-based outage-minimum robust beamforming design for mmWave CoMP systems with unpredictable blockages, dubbed the SGD-OutMin for short, is offered as a pseudo code in Algorithm 1.

A. Remarks on the Proposed Algorithm

Regarding the complexity of the proposed algorithm, one may readily notice that the most expensive operation at each iteration is calculation of the sum gradient direction in equation (13) and (16), which can be computed with the order $\mathcal{O}(N_t^2 B^2 U)$ by leveraging the sparsity and conjugate symmetry of the Hermitian matrix. Although further accelerations and extensions of the algorithm can be achieved by incorporating various emerging techniques such as stochastic variance reduced gradient (SVRG) [23] and momentum methods [28], they are left to be tackled in future works due to the fact that the scope of this letter is to provide a scalable and effective robust beamforming framework for mmWave CoMP systems with unpredictable blockages.

In addition, given [29], the learning rate α_t is adjusted to be shrunk as the number of iterations increases such that sufficient convergence conditions of the standard SGD for smooth non-convex optimization problems:

$$\sum_{t=1}^{\infty} \alpha_t \to \infty, \quad \sum_{t=1}^{\infty} \alpha_t^2 < \infty, \tag{17}$$

are satisfied.

IV. Simulation Results

In this section we evaluate the proposed method via software simulations both in terms of achieved outage probability and of the corresponding effective throughput defined as $R_{\text{eff},u} \triangleq W\mathbb{E}[\log_2(1+x)]$, where $x = \Gamma_u(\boldsymbol{h}_u, \boldsymbol{f})$ if $\Gamma_u(\boldsymbol{h}_u, \boldsymbol{f}) \geq \gamma_u$, $x = 0$ otherwise, and W is the sub-carrier bandwidth. Also we define $P_{\text{out},u}$ being the outage probability of the u-th user.

As already remarked earlier, state-of-the-art sum rate maximization (SRM) methods rely on a worst case optimization framework where all conceivable blockage patterns are taken into account as constraints so that the minimum throughput among such combinations can be maximized. As a consequence, probabilistic blockage was considered thereby only at LoS for a fixed number of clusters, because the complexity of the approach makes it prohibitive to apply to a more general and stochastic model such as the one considered here.

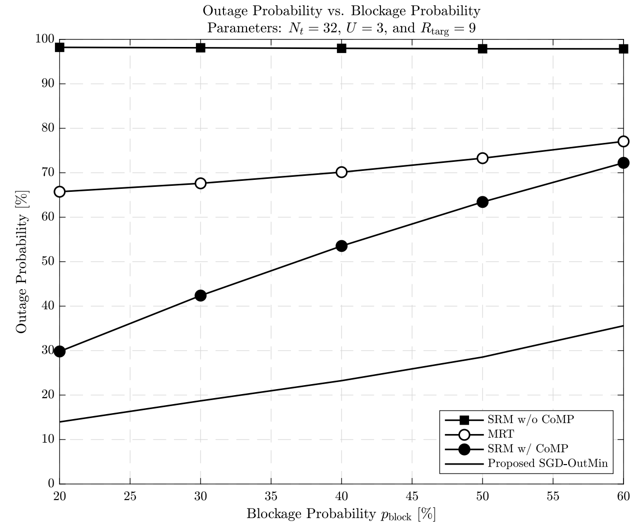
In light of the above, we compare the performance of our proposed beamforming design with the well-known maximum ratio transmission (MRT) beamforming (also known as conjugate beamforming), SRM beamforming without CoMP, and SRM beamforming with CoMP as a special case of the state-of-the-art proposed in [15], [16]. In order to maintain comparisons in line with the latter references, it is assumed that $B = 4$ BSs are respectively placed at each corner of a square cell with inter-cell spacing 100 [m], with each BS subject to a transmit power cap of $P_{\max,b} \leq 30$ [dbm], where the blockage probability can be accurately predictable.

Downlink users are assumed to be randomly placed within a square region and the noise floor σ_u^2 at each downlink user is assumed to be modeled as

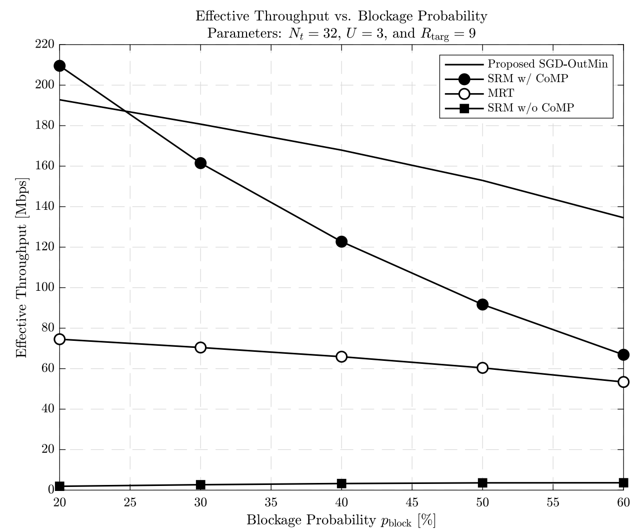
$$\sigma_u^2 = 10\log_{10}(1000\kappa T) + \text{NF} + 10\log_{10}(W)\,[\text{dBm}], \tag{18}$$

where κ is the Boltzmann's constant, $T = 293.15$ denotes the physical temperature at each user in kelvins, the noise figure at the downlink users NF is assumed to be 5 [dB] and the subcarrier bandwidth W is 20 [MHz].

For the comparisons, we consider a mmWave system operating at a carrier frequency of 28 [GHz], with the associated channel parameters according to [21, Table I] as described in Section II-A. For the sake of simplicity, it is assumed that blockage events occur with equal probability ($p_{b,u}^m = p_{\text{block}}$ $\forall u, b, m$), and that the target SINR is uniformly set to $\gamma_u = \gamma$. Finally, we define $R_{\text{targ}} \triangleq \log_2(1+\gamma)$ for the sake of brevity.



(a) Outage probability as a function of blockage probability p_{block}



(b) Effective throughput as a function of blockage probability p_{block}

Fig. 1. Outage probability and effective throughput comparisons with transmit antennas $N_t = 32$, number of users $U = 3$, the target throughput $R_{\text{targ}} = 9$, and the subcarrier bandwidth $W = 20$ [MHz].

In order to demonstrate the scalability and robustness of the proposed method, we consider two different simulation setups as follows. First, Figure 1 compares the outage probability and effective throughput performance of proposed and state-of-the-art (SotA) schemes as a function of blockage probability p_{block}, with $B = 4$ BSs each equipped with $N_t = 32$ antennas transmitting simultaneously to $U = 3$ single-antenna downlink users while targeting $R_{\text{targ}} = 9$, respectively.[4]

It is found non-surprisingly that the SRM scheme without CoMP transmission yields the worst performance, being easily beaten by the relatively simple MRT method due the CoMP advantage. In comparison, the SRM scheme implemented in a CoMP setting improves on the latter by demonstrating the ability to handle inter-user interference. Besides that, it is also found that the proposed beamforming algorithm outperforms all the aforementioned SotA methods in terms of both outage probability and effective throughput over a wide range

[4]Rather than arbitrary, this setup is motivated by the conditions of a real-life implementation of the proposed method, currently under pursuit in a collaborative research project in Japan.

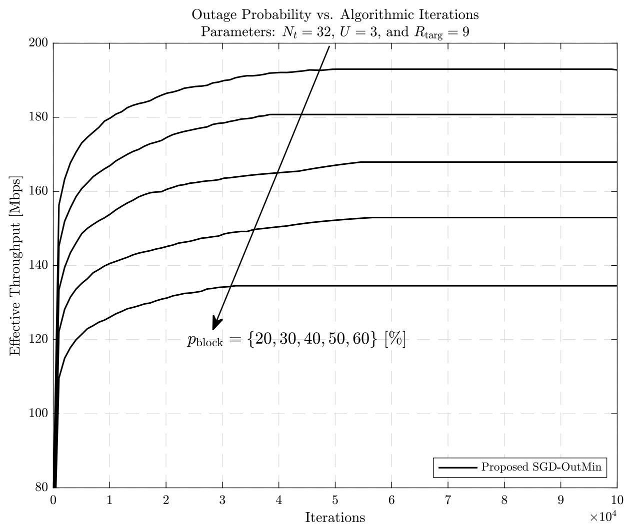


Fig. 2. Convergence behavior of the proposed algorithm with transmit antennas $N_t = 32$, number of users $U = 3$, the target throughput $R_{\rm targ} = 9$, and the subcarrier bandwidth $W = 20$ [MHz].

of blockage probability, demonstrating the ability to cope with path blockage over a wider range of block probabilities. Furthermore, we remark that the proposed algorithm can maintain a high data rate even under severe blockage conditions (i.e., $p_{\rm block} \in [50, 60]\%$).

Finally, in Figure 1(d), the convergence behavior of the proposed method for different blockage probabilities is shown as a function of algorithmic iterations. The curves show a clearly sub-linear rate regardless of the blockage probability, which can be attributed to the inherent variance in gradient direction as discussed in [23], implying that further acceleration may be possible by exploiting variance reduction techniques, which will be pursued in a future work.

V. Conclusion

We introduced a new stochastic-learning-based robust beamforming design framework for CoMP systems to combat unpredictable random blockages in mmWave channels. Unlike state-of-the-art methods which are based on the worst-case robust optimization framework, the complexity of our contributed method does not scale with the number of active clusters, such that the scheme is applicable to realistic mmWave systems. Simulated comparisons indicate that the proposed technique outperforms the state-of-the-arts both in terms of outage and achievable rates, which further demonstrate that the proposed scheme can retain decently high effective throughput even under a high blockage probability condition.

References

[1] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[2] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5G) wireless networks—With a focus on propagation models," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6213–6230, Dec. 2017.

[3] E. Björnson, L. V. der Perre, S. Buzzi, and E. G. Larsson, "Massive MIMO in sub-6 GHz and mmWave: Physical, practical, and use-case differences," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 100–108, Apr. 2019.

[4] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.

[5] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.

[6] R.-A. Stoica, H. Iimori, G. Abreu, and K. Ishibashi, "Frame theory and fractional programming for sparse recovery-based mmWave channel estimation," *IEEE Access*, vol. 7, pp. 150757–150774, 2019.

[7] P. Dong, H. Zhang, G. Y. Li, N. NaderiAlizadeh, and I. S. Gaspar, "Deep CNN for wideband mmWave massive MIMO channel estimation using frequency correlation," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Brighton, U.K., 2019, pp. 4529–4533.

[8] H. Iimori, R.-A. Stoica, K. Ishibashi, and G. T. F. de Abreu, "Robust sparse reconstruction of mmWave channel estimates via fractional programming," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Barcelona, Spain, 2020, pp. 318–323.

[9] G. Kwon and H. Park, "Joint user association and beamforming design for millimeter wave UDN with wireless backhaul," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 12, pp. 2653–2668, Dec. 2019.

[10] G. R. MacCartney, T. S. Rappaport, Jr., and S. Rangan, "Rapid fading due to human blockage in pedestrian crowds at 5G millimeter-wave frequencies," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–7.

[11] S. Ju, O. Kanhere, Y. Xing, and T. S. Rappaport, "A millimeter-wave channel simulator NYUSIM with spatial consistency and human blockage," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.

[12] I. K. Jain, R. Kumar, and S. S. Panwar, "The impact of mobile blockers on millimeter wave cellular systems," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 4, pp. 854–868, Apr. 2019.

[13] 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.

[14] T. Mikuma, T. Nishio, M. Morikura, K. Yamamoto, Y. Asai, and R. Miyatake, "Transfer learning-based received power prediction using RGB-D camera in mmWave networks," in *Proc. IEEE 89th Veh. Technol. Conf. (VTC2019-Spring)*, Kuala Lumpur, Malaysia, 2019, pp. 1–5.

[15] D. Kumar, J. Kaleva, and A. Tölli, "Multi-point connectivity for reliable mmWave," in *Proc. 27th Eur. Signal Process. Conf. (EUSIPCO)*, A Coruna, Spain, 2019, pp. 1–5.

[16] D. Kumar, J. Kaleva, and A. Tölli, "Rate and reliability trade-off for mmWave communication via multi-point connectivity," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, 2019, pp. 1–6.

[17] H. Iimori, G. Abreu, and K. Ishibashi, "Fractional programming for robust TX BF design in multi-user/single-carrier PD-NOMA," in *Proc. Int. Symp. Model. Optim. Mobile Ad Hoc Wireless Netw. (WiOpt)*, Avignon, France, 2019, pp. 1–7.

[18] G. R. MacCartney and T. S. Rappaport, "Millimeter-wave base station diversity for 5G coordinated multipoint (CoMP) applications," *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3395–3410, Jul. 2019.

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

[20] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[21] M. R. Akdeniz *et al.*, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[22] V. N. Vapnik, "Principles of risk minimization for learning theory," in *Proc. Int. Conf. Neural Inf. Process. (NIPS)*, Dec. 1991, pp. 831–838.

[23] R. Johnson and T. Zhang, "Accelerating stochastic gradient descent using predictive variance reduction," in *Proc. 26th Int. Conf. Neural Inf. Process. (NIPS)*, 2013, pp. 315–323.

[24] L. Bottou, *Online Learning and Stochastic Approximations*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[25] P. L. Bartlett and M. H. Wegkamp, "Classification with a reject option using a hinge loss," *J. Mach. Learn. Res.*, vol. 9, pp. 1823–1840, Jun. 2008.

[26] O. Taghizadeh, A. C. Cirik, and R. Mathar, "Hardware impairments aware transceiver design for full-duplex amplify-and-forward MIMO relaying," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1644–1659, Mar. 2018.

[27] H. Iimori, G. T. F. de Abreu, and G. C. Alexandropoulos, "MIMO beamforming schemes for hybrid SIC FD radios with imperfect hardware and CSI," *IEEE Trans. Wireless Commun.*, vol. 18, no. 10, pp. 4816–4830, Oct. 2019.

[28] D. P. Kingma and J. L. Ba, "ADAM: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, Atlanta, GA, USA, May 2015, pp. 1–15.

[29] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Stat.*, vol. 22, no. 3, pp. 400–407, 1951.