CHALMERS
UNIVERSITY OF TECHNOLOGY

Initial Access in mm-wave 5G Mobile Communications

Master's thesis in Master Programme of Communication Engineering

Hao Guo

Department of Electrical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2017

Master's thesis: EX013/2017

Initial Access in mm-wave 5G Mobile Communications

Hao Guo

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
Communication and Antenna systems Group
Master Programme of Communication Engineering
Chalmers University of Technology
Gothenburg, Sweden 2017

Initial Access in mm-wave 5G Mobile Communications
Hao Guo

Supervisor: Behrooz Makki
Examiner: Tommy Svensson

Master's Thesis: EX013/2017
Department of Electrical Engineering
Communication and Antenna systems Group
Master Programme of Communication Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2017

Abstract

Initial access beamforming is identified as a key challenge for the upcoming 5G mobile communication systems operating at high carrier frequencies, and several techniques are currently being proposed. In this thesis, we first give a brief introduction to 5G millimeter wave systems, initial access procedures as well as the codebook-based beamforming. Then, we study the performance of initial access beamforming schemes in the cases with large-but-finite number of transmit antennas and single-antenna users. Particularly, we develop an efficient beamforming scheme using genetic algorithms with the algorithm running delay taken into account. As shown, our proposed algorithm is generic in the sense that it can be effectively applied with different channel models, metrics and beamforming methods. Also, our results indicate that the proposed scheme can reach (almost) the same end-to-end throughput as the exhaustive search-based optimal approach with considerably less implementation complexity.

As the second part of the work, we extend our proposed genetic algorithm-based beamforming scheme to include beamforming at both the transmitter and the receivers. Also, we compare the system performance with different state-of-the-art beamforming approaches in the millimeter wave multi-user multiple-input- multiple-output networks. Taking the millimeter wave communication characteristics and various metrics into account, we investigate the effect of different parameters such as the number of transmit antennas/users/per-user receive antennas, beamforming resolutions as well as hardware impairments on the system performance employing different beamforming algorithms. As shown, our proposed genetic algorithm-based approach performs well in delay-constrained networks with multi-antenna users. Compared to the considered state-of-the-art schemes, our method reaches the highest service outage-constrained end-to-end throughput with considerably less implementation complexity.

Keywords: Initial access beamforming; Millimeter wave communications; Multi-user multiple-input-multiple-output networks; Throughput

Acknowledgements

I would like to take this opportunity to thank Prof. Tommy Svensson and Dr. Behrooz Makki for giving me this opportunity to join the Communication group as a one-year master thesis student. It has been very enjoyable learning and working here.

I would like to express my thanks to Prof. Tommy Svensson for being a great advisor since we first met in Nankai University, Tianjin. Your enthusiasm and professional presentation gave me a good first impression of Chalmers and Sweden. During my master study, you helped me to get used to the new environment and gave me useful advices on choosing courses. In the second year, you offered me a good chance to start my research work with a one-year master thesis training. I am always impressed by all your awesome research ideas and your advices on presentations and writings help me a lot.

My deep gratitude goes to my supervisor Dr. Behrooz Makki. We started our at-any-moment meetings even before my thesis topic were settled, and it has been a long and enjoyable journey since then. You have helped me through so many tough moments with full supports and great patience. Nothing would have been possible without you.

I am thankful for all the valuable suggestions and help from professors and PhD students in the S2 department. Especially thanks to Chao Fang, you have helped me so much with thesis writing and attending the academic meeting. I would also like to thank JinLin Liu as well. You gave me great advices both professionally and personally. My great thanks to Pavel Gueorguiev for sharing your impressive knowledge on programming skills. My spoken English improved a lot with you. I also thank people in the thesis room for being nice and friendly.

Finally, I sincerely appreciate my parents for their unconditional love.

The research leading to these results received funding from the European Commission H2020 programme under grant agreement $n°671650$ (5G PPP mmMAGIC project), and from the Swedish Governmental Agency for Innovation Systems (VINNOVA) within the VINN Excellence Center Chase, which are gratefully acknowledged.

Hao Guo, Gothenburg, April 2017

List of acronyms

ADC	Analog-to-Digital Converter
BF	Beamforming
bpcu	bit-per-channel-use
BS	Base Station
CSI	Channel State Information
DFT	Discrete Fourier Transform
GA	Genetic Algorithm
IA	Initial Access
IID	Independent and Identically Distributed
LOS	Line-of-Sight
LTE	Long Term Evolution
mm-wave	Millimeter Wave
MU-MIMO	Multi-User Multiple-Input-Multiple-Output
NLOS	Non-Line-of-Sight
PA	Power Amplifier
PSS	Primary Synchronization Signal
RA	Random Access
SINR	Signal-to-Interference-plus-Noise Ratio
UE	User Equipment
WLANs	Wireless Local Area Networks
5G	5th Generation Wireless Systems

Contents

Part I

Introductory Chapters

1
Background

The 5th generation wireless systems (5G) is referred to as the proposed next telecommunications standards that should be rolled out by 2020 to meet business and consumer demands [1]. As one of the key features, 5G will support operations in the new range of 6-100 GHz, also known as the millimeter wave (mm-wave) spectrum.

Due to peak power limitation and severe path loss, mm-wave systems generally require high base station density to reach acceptable coverage. Extremely dense mm-wave networks will have the interference management issue. As one of the key techniques in wireless communications, beamforming (BF) can be a good technique to mitigate the poor coverage range and improve the system throughput. Fortunately, the physical size of antennas at mm-wave frequencies is normally small, which gives a strong hardware support for employing large-scale BF. However, we need to have BF also in the control and pilot channels, implying that it is much more difficult to obtain the channel state information (CSI) since we have no sector-wide broadcast channel in the mm-wave system. One of the solutions for these problems is the codebook-based BF during initial access (IA) procedure which can improve the system performance without the requirements of the perfect CSI.

In this perspective, this thesis studies, designs and compares different schemes for IA BF in mm-wave communications. The main contents of this MSc thesis are:

* Introduce the background knowledge of 5G, IA procedures as well as the codebook-based BF,
* Perform a literature review on related research works,
* Propose a genetic algorithm (GA)-based BF approach for mm-wave multi-user multiple-input-multiple-output (MU-MIMO) networks,
* Evaluate the system performance with different parameter settings,
* Compare an extended GA-based scheme using multi-antennas also on the user side with the state-of-the-art approaches in terms of both service outage-constrained end-to-end throughput and implementation complexity.

1.1 5G Millimeter Wave (mm-wave) Networks

In this section, we start with the recent developments of 5G, explaining two key requirements and pointing out one practical solution. These background knowledge could give us a clear picture of the exact purposes of our work.

1.1.1 Starting Point: New Frequency Bands, More Users

New standards in wireless communications are continuously proposed. The reasons are two-fold: First, the number of users are increasing rapidly, and they always ask for better quality-of-service. For example, automobiles/pedestrians would like to get access quickly even with mobility and increasing interferences (see Fig. 1.1). On the other hand, all sub-3 GHz frequencies are already utilized and there is not much left for microwave communications (see Fig. 1.2). Thus, there is a need to exploit new and higher frequency bands where more users can be served with higher data rates. The 5G new air interface starts right here: New frequency bands for more users.

Figure 1.1: Mobility with more connected users.

5G operating bands consist of three parts: evolution for existing bands in 4G (below 6 GHz), potentially a few new bands above 3 GHz and new frequency band (6-100 GHz). Based on the relationship between frequency f and wavelength $\lambda = \frac{c}{f}$ with c being the speed of the light, the new supported bands (above 6 GHz) are also named as mm-wave spectrum. There are lots of special characteristics in mm-wave communications: high path loss, strict peak power limitation, blockage-sensitive signals. These constraints increase the difficulty of the system design remarkably.

Moreover, connecting everything is one of the main goals of 5G. Smart home is a typical application. However, more connected users means more interference and lower data speed, which also calls for new techniques to improve it.

1.1.2 Key Requirements: High Data Rate and Low Latency

Depending on the considered scenarios and use-cases, there are some basic requirements for 5G provided by the Next Generation Mobile Networks Alliance [2]:

- very high data rates: 1 to 10 Gb/s (e.g., virtual reality office),

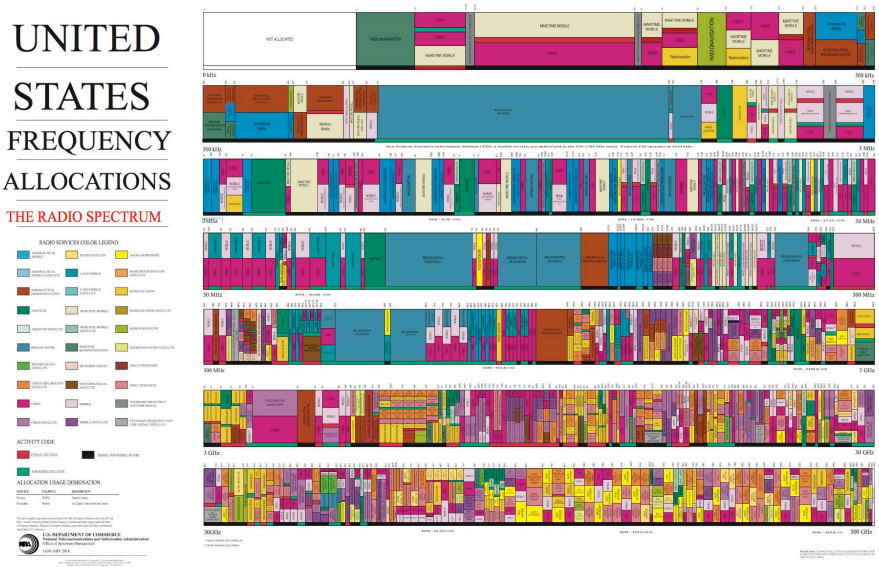


Figure 1.2: U.S. Frequency Allocation Chart as of January 2016. (Available at: https://www.ntia.doc.gov/page/2011/united-states-frequency-allocation-chart)

- very low end-to-end latency: less than 5 ms (e.g., traffic safety),
- large number of connected devices: up to 300,000 devices per access point (e.g., massive deployment of sensors),
- high reliability: 99.999% (e.g., tele-protection in smart grid networks),

in which the most important requirements are high data rates and low latency. This means that we need to find low-complexity techniques which can increase the data rate and reduce the latency remarkably.

1.1.3 One Solution: Effective Initial Access Beamforming

BF, also known as spatial filtering, is an important technique in wireless communications. It aims at forming high directional transmit/receive beam patterns by antenna arrays and phase shifters [3]. Conventional BF is performed by precoding/-combing with CSI after stable links have been established. However, it is difficult to acquire CSI at mm-wave bands. Moreover, due to the high path loss, the coverage range for mm-wave frequencies is still small even with the extension by BF. To mitigate these constraints in the mm-wave spectrum, employing the codebook-based BF with large antenna arrays before links have been established, i.e., the IA BF, would be one of the feasible solutions to satisfy the throughput requirement with imperfect CSI (see Fig. 1.3).

Moreover, the word access means that there is a threshold to decide whether or not the user equipments (UEs) have been connected to the base stations (BSs). One reasonable situation is that the users' received signal-to-interference-plus-noise ratio (SINR) should be above predefined thresholds, as it includes both the received power limitation and the interferences from other users.

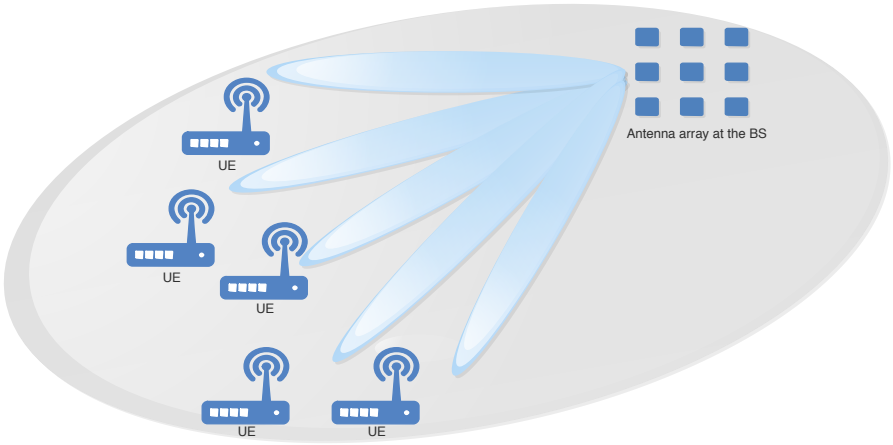


Figure 1.3: A sketch map for a multi-user system employing BF to realise high directional communications.

As mentioned before, low end-to-end latency is one of the key requirements for mm-wave mobile communications. For this reason, the time cost for the BF scheme should be carefully considered during the design procedure. To present the trade-off that more BF time can get better beam resolution but reduces data transmission time (see Fig. 1.4), in our system, we take the cost of running the algorithm into account so that the maximum throughput is obtained by finding a suboptimal BF matrix while the rest of the time slot is used for data transmissions. A convergence example in section 5.1.1 also describes this trade-off in detail.

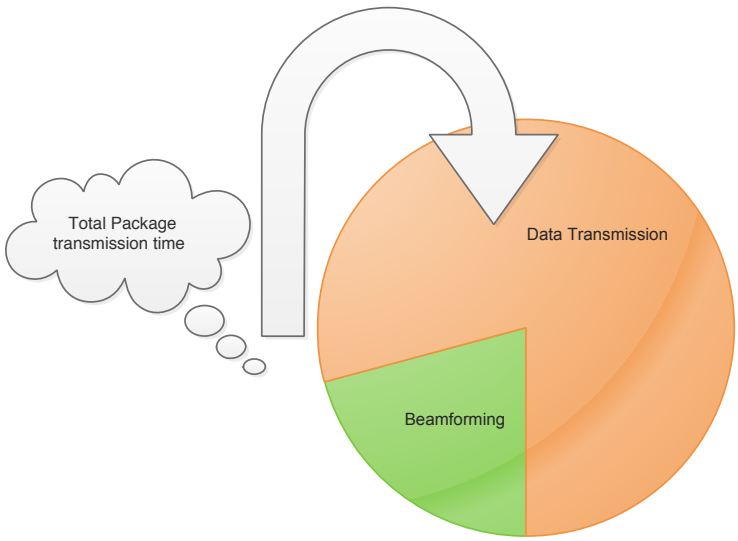


Figure 1.4: Time structure for the delay-constrained system.

1.2 Initial Access Procedure

Basically, the steps for the mm-wave IA is similar to those in Long Term Evolution (LTE), which are proposed in some standard texts such as [4]. However, the procedures need to be modified for mm-wave because we want to employ BF before the UE and the BS detect each other. These steps are, namely, Cell Search, Random Access, and Scheduled Communications, which are also presented in Fig. 1.5. Note that the most challenging part of mm-wave IA is to decide the beam directions of the BS and the UE and, for this reason, we mainly focus on the first and the third step.

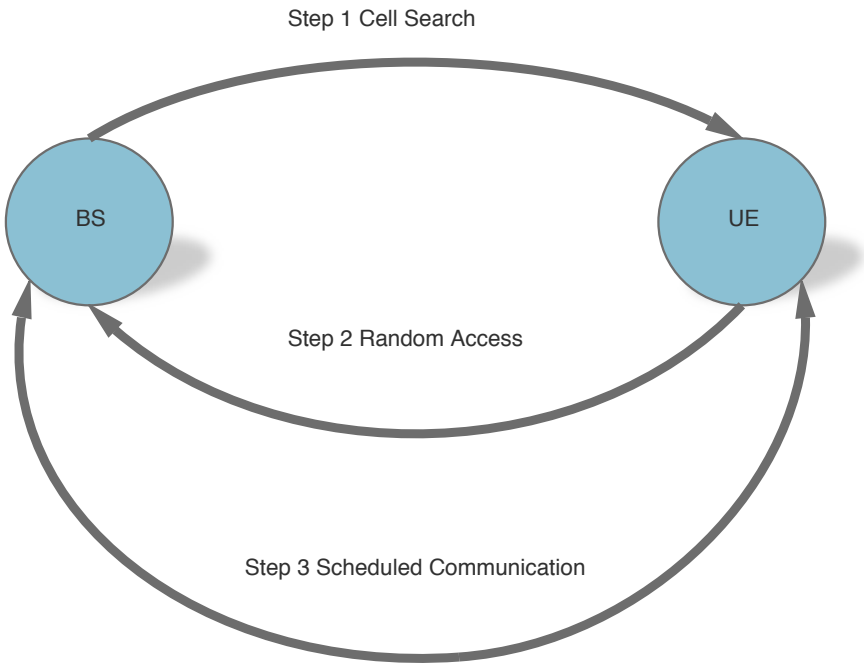


Figure 1.5: Steps for IA procedure.

1) **Cell Search**: The BS periodically broadcasts the synchronization signal with different beam direction to the UE, which is similar to the Primary Synchronization Signal (PSS) in LTE. The only difference is the synchronization signal is also used to perform BF. At the receive side, the UE will decide their beam direction if the synchronization signal is detected. There would be some certain thresholds for the UE to determine whether or not it has been connected to the link. After trying a certain number of synchronization signals, the UE can make a decision of the optimal beam direction.
2) **Random Access (RA)**: In this step, both the BS and the UE know the preliminary directions through which they should steer their beams. They will exchange the random access messages in order to establish the link, including
   - *RA preamble transmission*: The UE randomly selects a small number of waveforms (named as preamble) and transmits to the BS.
   - *RA response*: The BS sends a response message with an initial timing

and power correction as well as some access control identifier to precisely identify the UE.

– *Connection request*: The UE responses back the connection request message with some authentication and identification information.

3) **Scheduled Communications**: Beam training/refinement can be applied in this part to further optimize beam searching results in the previous steps by steering beams, where the user mobility can also be handled. After this step, messages can be transmitted/received with full BF gain using scheduled channels.

1.3 Codebook-based Beamforming

A codebook is a matrix, where each column represents a weight vector of antennas and can form a certain beam pattern (See Fig. 1.6). As we mentioned before, due to the large feedback/estimation overhead and high path loss, it is difficult to acquire CSI at mm-wave, which means conventional CSI-based precoding/combing schemes may be infeasible. In this case, the pre-defined codebook can be used to generate directional transmitted/received beams in Step 1) and Step 3). Our goal is to propose an efficient IA BF algorithm. From the codebook perspective, this means we want to design a low-delay algorithm to find the (sub)optimal sets of codebook vectors. We use the GA principle, running iterations continuously based on previous-round solutions, and finally reaching an appropriate (sub)optimal solution. Details of the beam searching algorithm can be seen in Chapter 4.

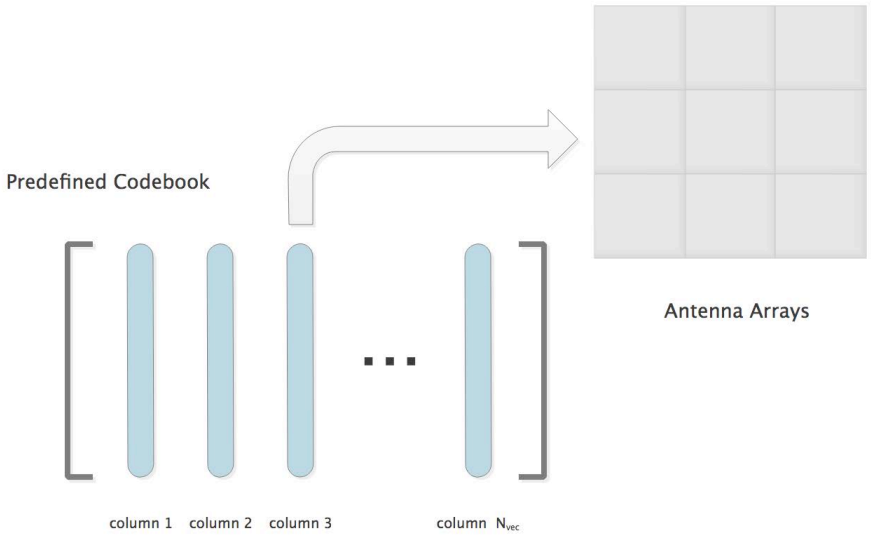


Figure 1.6: Codebook-based BF.

1.4 Analog/Digital/Hybrid Beamforming

This section can be part of the thesis scope. It is important to know the BF type of our system, in terms of analog/digital/hybrid BF, and more specifically, amplitude/phase-based BF.

The definition of the BF type is based on the signal type that the BF is applied to [5]. For example, analog BF means amplitude/phase variation is applied to the analog signal before the analog-to-digital converter (ADC) while digital BF means amplitude/phase variation is applied to the digital signal resulted from the ADC.

For the analog BF, the receiver can "look" in only one or a small number of directions at a time, while in the digital BF the receiver can look in all directions at once. However, digital BF requires more ADCs compared to analog BF. This increases the power consumption to unacceptable levels, especially at the UE side. One possible solution can be to apply low-complexity digital BF [6] which means using few bits ADCs to get a good trade-off. Indeed, the advantage of analog and digital BF can be combined by employing hybrid BF, in which the BF efficiency is favorable with acceptable implementation complexity.

In this thesis work, we consider the analog BF with phase variation only. In this case, there are a limited number of columns in the codebook and it's possible to search for the optimal BF within a short period. Our proposed algorithm can be applied to digital/hybrid BF, and it will be part of our future work.

2
Related Works

In this chapter, we want to collect the recent research works on the IA BF in three aspects: 1) IA in 5G systems, 2) codebook-based BF, 3) different beam searching algorithms.

Different standards have developed mm-wave BF technique such as IEEE 802.15.3c (TG3c) [7], IEEE 802.11ad (TGad)[8] and ECMA-387[9].

- **Cell discovery**: In [10], the technical details of BF at mm-wave frequency are described and a hierarchical search strategy is designed. Then, [11] proposes a directional downlink cell search method in a mm-wave cell. The results of [11] are extended in [6] where it looks at the whole IA procedure with different design options and evaluates the overall delay to determine the control plane latency. Also, a power-delay-profile-based IA approach for mm-wave small cells is presented in [12].
- **RA**: [13] presents four simple but representative IA protocols during the cell search and the RA phase, where the best trade-off between IA delay and user-perceived downlink throughput is achieved under a fast cell search protocol. A multi-user hybrid precoding algorithm is proposed in [14] and higher sum-rates can be achieved using hybrid precoding compared to analog-only BF solutions. Also, in [15] an integrated view on medium access control layer issues for mm-wave cellular networks is provided, one of them is the RA.
- **Beam training/refinement**: In [16] a two-stage-scanning-based iterative beam search approach is designed and compared with a brute-force sequential search method. Also, as the extension work, [17] makes a comparison of three different IA BF approaches, namely, exhaustive search, two-step search and context information-based search in terms of miss-detection probability and discovery time. Moreover, [18] compares different IA protocols with access delay and user throughput. In our paper [19], we propose a GA based initial BF approach and evaluate the system performance with different parameter settings. A GA-based selection approach is also used in [20] and [21], where GA is used for resource allocation in the return link of multi-beam satellite systems and antenna selection in multi-user networks, respectively.

For the codebook-based BF, [22] designs limited feedback-type directional codebooks for mm-wave systems. In [23], a discrete Fourier transform (DFT)-based codebook is designed and a low-complexity antenna vector training algorithm is presented. Also, [24] presents a tree-structured multi-level BF codebook used for mm-wave backhaul

systems. Moreover, an efficient multi-stage codebook is designed to support the IEEE 802.15.3c three-stage protocol in [25].

Considering different beam searching algorithms, [26] proposes a turbo-like BF scheme in order to reduce both the searching complexity and the system overhead. Moreover, [27] designs a link-by-link concurrent BF protocol using the sum throughput as the performance metric. A low-complexity mm-wave BF algorithm premised on the sparse multipath structure is presented in [28]. Also, a one-sided codebook-based beam alignment strategy in wireless local area networks (WLANs) is proposed in [29]. These works are based on machine learning-based iterative algorithms where the proposed algorithms are applicable for different channel models, precoding schemes, and performance metrics.

3

System Model

3.1 Channel Model

We consider a MU-MIMO setup with M transmit antennas at a BS and τ multi-antenna users, each with β antennas. As a result, there are $N = \tau \times \beta$ total antennas at the receiver side (see Fig. 3.1). Note that in our paper [19] we set $\beta = 1$, which means we do not employ receive combining there. In our second paper, however, we study the system performance in the cases with receive combining and different number of antennas at the users. We also set $M > N$. At each time slot t, the aggregated received signal vector $\mathbf{Y}(t)$ over the users after receive BF can be described as

$$\mathbf{Y}(t) = \sqrt{\frac{P}{M}}\mathbf{U}(t)^H\mathbf{H}(t)\mathbf{V}(t)\mathbf{X}(t) + \mathbf{Z}(t), \tag{3.1}$$

where P is the total power budget, $\mathbf{H}(t) \in \mathcal{C}^{N\times M}$ is the channel matrix, $\mathbf{X}(t) \in \mathcal{C}^{M\times 1}$ is the intended message signal, $\mathbf{V}(t) \in \mathcal{C}^{M\times M}$ is the precoding matrix at the BS, $\mathbf{U}(t) \in \mathcal{C}^{N\times N}$ is the aggregated combining matrix at the users' side, and $\mathbf{Z}(t) \in \mathcal{C}^{N\times 1}$ denotes the independent and identically distributed (IID) Gaussian noise matrix. For simplicity, we drop the time index t in the following.[1]

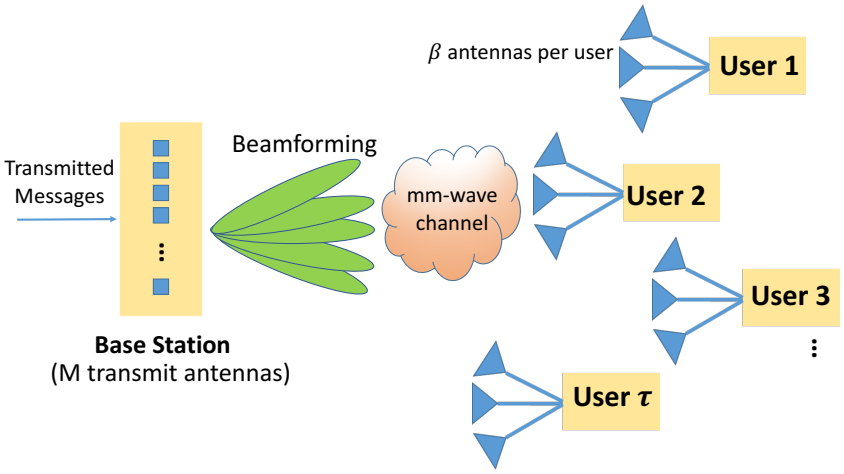


Figure 3.1: Mm-wave multiuser MIMO system model.

[1] We study slow-fading models where the channels remain constant for long time.

Our proposed IA schemes are applicable for different channel models. In the simulations, however, we consider the channel model

$$\mathbf{H} = \sqrt{\frac{k}{k+1}}\mathbf{H}_{\text{LOS}} + \sqrt{\frac{1}{k+1}}\mathbf{H}_{\text{NLOS}}, \tag{3.2}$$

where $\mathbf{H}_{\text{LOS}}$ and $\mathbf{H}_{\text{NLOS}}$ denote the line-of-sight (LOS) and the non-line-of-sight (NLOS) components of the channel, respectively, and the NLOS component is assumed to follow a complex Gaussian distribution. Also, k controls the relative strength of the LOS and the NLOS components. In (3.2), setting $k = 0$ represents an NLOS condition while $k \to \infty$ gives an LOS channel.

3.2 Pre-defined Codebook

Unlike the conventional BF procedure acquiring CSI, in mm-wave systems we suggest to perform codebook-based BF, which means selecting a precoding matrix $\mathbf{V}$ out of a predefined codebook $\mathbf{W}_{\text{T}}$ at the BS while selecting a combining matrix $\mathbf{U}$ out of a predefined codebook $\mathbf{W}_{\text{R}}$ at the receiver side, sending test signal and finally making decisions on transmit/receive beam patterns based on the users' feedback about their performance metrics. The IA will be finished as soon as a stable control link is established. The time structure for a packet transmission can be seen in Fig. 1.4, where part of the packet period is dedicated to design appropriate beams and the rest is used for data transmission. Thus, we need to find a balance between the beam design delay and the data transmission period by choosing an efficient approach.
Here, we use DFT-based codebooks [30] at both sides which are defined as

$$\begin{aligned} &\mathbf{W}_{\text{T}} = \{w(m,u)\} = \{e^{-j2\pi(m-1)(u-1)/N_{\text{vec}}}\}, \\ &m = 1, 2, \ldots, M, u = 1, 2, \ldots, N_{\text{vec}}, \end{aligned} \tag{3.3}$$

for the BS, while

$$\begin{aligned} &\mathbf{W}_{\text{R}} = \{w(n,u)\} = \{e^{-j2\pi(n-1)(u-1)/N_{\text{vec}}}\}, \\ &n = 1, 2, \ldots, N, u = 1, 2, \ldots, N_{\text{vec}}, \end{aligned} \tag{3.4}$$

for the users, where $N_{\text{vec}} \geq \max(M, N)$ is the number of codebook vectors.

3.3 Performance Metrics

The machine learning based schemes of [24]-[27] and two included papers in this thesis are generic, in the sense that they can be implemented for different metrics. For the simulations, however, we consider the service outage-constrained end-to-end throughput and the average number of required iterations as the system performance metrics. In some scenarios, it may be required to serve the users with some minimum required rates, otherwise *service outage* occurs. In the K-th iteration round of the

algorithm, the service outage-constrained end-to-end throughput in bit-per-channel-use (bpcu) is defined as

$$\begin{aligned} R(K) &= (1-\alpha K)\sum_{i=1}^{\tau} r_i^K U(r_i^K, \log_2(1+\theta)), \\ r_i^K &= \log_2\left(1+\mathbf{SINR}_i^K\right), \\ U(r_i^K, \log_2(1+\theta)) &= \begin{cases} 1 & r_i^K \geq \log_2(1+\theta) \\ 0 & r_i^K < \log_2(1+\theta). \end{cases} \end{aligned} \tag{3.5}$$

Here, r_i^k denotes the achievable rate of the user i at the end of the K-th iteration. Also, parameter α is the relative delay cost for running each iteration of the algorithm which fulfills $\alpha N_{\text{it}} < 1$ with N_{it} being the maximum possible number of iterations. Then, $\log_2(1+\theta)$ is the minimum per-user rate while θ represents the minimum required SINR of each user. Also,

$$\mathbf{SINR}_i^K = \frac{\frac{P}{M} g_{i,i}^K}{BN_0 + \frac{P}{M}\sum_{i\neq j}^{N} g_{i,j}^K} \tag{3.6}$$

is the received SINR at the user i in the iteration round K.[2] Here, $g_{i,j}$ is the (i,j)-th element of the matrix $\mathbf{G}_K = |\mathbf{U}_K^H \mathbf{H} \mathbf{V}_K|^2$ which is referred to as the channel gain throughout the paper. Moreover, B is the system bandwidth and N_0 is the power spectral density of the noise. We set $BN_0 = 1$ to simplify the system so that the power P (in dB, $10\log_{10} P$) denotes the signal-to-noise ratio (SNR) as well.
As opposed to, e.g., [6, Eq. 1], [27, Eq. 3], [31, Eq. 43], [32, Eq. 3], [33, Eq. 5] and [34, Eq. 5], we consider the algorithm running delay in the performance analysis. As seen in the following, there is a trade-off between optimizing BF matrices and reducing the data transmission period. In this case, the optimal solution may be achieved by running the algorithms for a limited number of iterations (see (3.5)).

3.4 Power Amplifier Efficiency

The efficiency of the radio-frequency high Power Amplifier (PA) should be taken into consideration in the multi-antenna systems. Here, we consider the state-of-the-art PA efficiency model [35, Eq. 13], [36, Eq. 3] described as:

$$\rho_{\text{cons}} = \frac{\rho_{\max}^{\mu}}{\epsilon \times \rho_{\text{out}}^{\mu-1}}, \tag{3.7}$$

where ρ_{cons}, ρ_{out}, $\rho_{\max}$ denote the consumed power, the output power and the maximum output power of the PA, respectively. Also, $\epsilon \in [0,1]$ represents the power efficiency and $\mu \in [0,1]$ is a parameter depending on the PA class. Setting $\epsilon = 1$, $P_{\max} = \infty$ and $\mu = 0$ in (3.7) represents the special case (with an ideal PA).

[2] This is based on the assumption that we know the channels at the receivers. But we do not know the channels at the transmitters (because $\mathbf{H}$ is a big matrix and it is costly to feedback such big matrix information to the transmitter). Even with perfect CSI at receiver we try to find the best BF among a limited number of possible beams because in practice we always have a few discrete options for BF.

4

Genetic Algorithm (GA)-based Beamforming Approach

4.1 Beamforming Approach in Single-Antenna User Networks

We use a GA-based approach for beam selection during IA BF and details are explained in Algorithm 1. The algorithm starts by getting L possible beam selection sets randomly and each of them means a certain beam formed by transmit antennas, i.e., a submatrix of the codebook $\mathbf{W}_{\mathrm{T}}$ in (3.3). During each iteration, we determine the best selection result, named as the *Queen*, based on our objective metric. For instance, we choose the BF matrix with the highest end-to-end throughput if (3.5) is considered as the objective function. Next, we keep the Queen and regenerate $S < L$ matrices around the Queen. This can be done by making small changes to the Queen such as changing a number of columns in the Queen matrix. In the simulations, we replace 10% of the columns of the Queen by other random columns from the codebook. Finally, during each iteration $L-S-1$ BF matrices are selected randomly. After N_{it} iterations, considered by the algorithm designer, the Queen is returned as the beam selection rule in the considered time slot.
As demonstrated, the algorithm is generic in the sense that it is independent of the channel model, objective function or precoding matrix, thus this can be used in different scenarios and as a benchmark for comparison of different IA schemes. Also, our proposed algorithm converges to the (sub)optimal value of the considered metrics with few iterations. Therefore, the proposed algorithm can be useful in IA BF where the delay is one of the most important factors.

4.2 Extension of the GA-based Scheme and Considered State-of-the-art Methods

Here, we compare the performance of different IA BF methods as follows.
Extended GA-based Search [19]: The algorithm starts by making L possible beam selection sets at both transmitter and receiver, i.e., submatrices of each codebook. We follow the same procedure as in Algorithm 1 but now each of the L possible solutions has one transmit and one receive BF matrix. In this way, this is an extended version of our GA-based approach in [19] with BF at both the transmitter and the receiver.

Algorithm 1 GA-based Beam Selection Algorithm

In each time slot with instantaneous channel realization $\mathbf{H} \in \mathcal{C}^{N \times M}$, do the followings:

I. Consider L, e.g., $L = 10$, sets of BF matrices $\mathbf{V}_l$, $l = 1,..., L$, randomly selected from the pre-defined codebook $\mathbf{W}$.

II. For each $\mathbf{V}_l$, evaluate the instantaneous value of the objective metric R_l, $l = 1,..., L$, for example end-to-end throughput (3.5).

III. Selection: Find the best BF matrix which results in the best value of the considered metric, named as the *Queen*, e.g., $\mathbf{V}_q$ satisfies $R(\mathbf{V}_l) \leq R(\mathbf{V}_q)$, $\forall$ $l = 1,..., L$ if the end-to-end throughput is the objective function.

IV. $\mathbf{V}_1 \leftarrow \mathbf{V}_q$

V. Create $S \ll L$, e.g., $S = 5$, BF matrices $\mathbf{V}_s^{\text{new}}$, $s = 1,..., S$, around the Queen $\mathbf{V}_1$. These sets are generated by making small changes in the Queen.

VI. $\mathbf{V}_{s+1} \leftarrow \mathbf{V}_s^{\text{new}}$, $s = 1,..., S$.

VII. Go back to Step II and run for N_{it} iterations, N_{it} is a fixed number decided by designer.

Return the final Queen as the beam selection rule for the current time slot.

Tabu Search [26]: The Tabu-search approach follows the basic idea as in the GA-based scheme [19] where we choose and update the Queen in successive iterations. The only difference is the evolution method of the Queen in successive iterations. With Tabu, we use the definition of *neighborhood* in [26]: One matrix $\mathbf{A}$ is defined as another matrix $\mathbf{B}$'s neighborhood if 1) $\mathbf{A}$ has only one different column compared with $\mathbf{B}$ or 2) the index difference between the two corresponding columns in $\mathbf{A}$ and $\mathbf{B}$ is equal to one. To make S beam selection sets, we change the Queen from previous round to its neighbors.

Link-by-link Search [27]: In this strategy, the beam design of τ users is not optimized simultaneously. Instead, with a greedy approach, the BF solution is settled user-by-user by considering the interference from the other $\tau - 1$ links. The system performance improves in successive iterations until it converges to some (sub)optimal BF rules.

Two-level Search [24]-[25]: Being inspired by multi-stage BF techniques, e.g., [24] and [25], we design a two-level-codebook search scheme for our system. In the first level, the BS transmits messages over wider sectors using the codebook with $N_{\text{vec}}/2$ columns, while in the second level it searches the optimal solution within the best such sector by steering narrower beams with an N_{vec}-column codebook.

5

Simulation Results

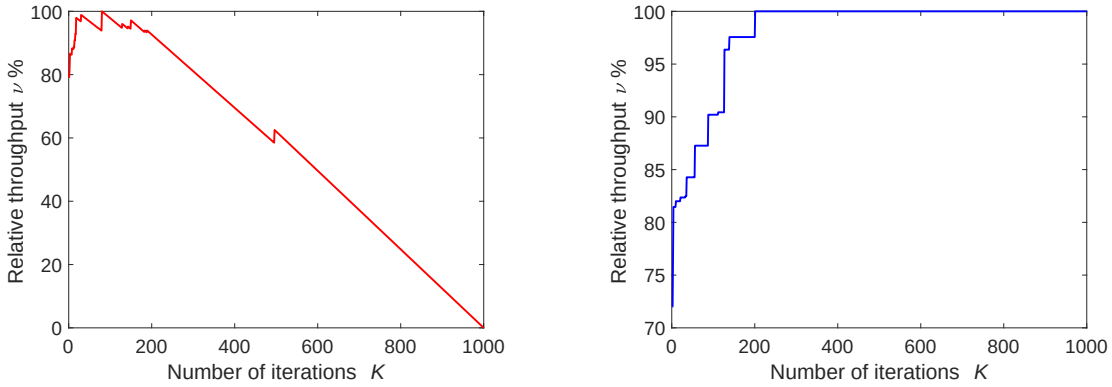
In this chapter, we present some interesting results which are also included in the papers. We present the figures in two parts: Section 5.1 presents results about the advantages of the GA-based IA BF approach in the single-antenna-user MU-MIMO networks. A convergence example with two different cases, i.e., running the GA with the running delay ($\alpha = 0.001$) and without the delay ($\alpha = 0$), shows what will happen if we introduce a delay factor in the performance metric and how many iterations are required to reach the (sub)optimal results. Then, couple of results present that our algorithm works well with different parameter settings. In section 5.2, we extend our algorithm to include BF at both the transmitter and the receiver, and we compare the performance with alternative BF approaches in the multi-antenna-user MU-MIMO networks. As shown, our proposed GA-based approach performs well in delay-constrained networks with multi-antenna users. Compared to the considered state-of-the-art schemes, our method reaches the highest service outage-constrained end-to-end throughput with considerably less implementation complexity.

5.1 On the Advantages of the GA-based Approach

For the simulation results, we consider the channel model described by (3.2), with $k = 0, 1, 2, 3, 10$. Except for Fig. 5.1 which shows an example of the algorithm procedure by plotting the relative throughput $\nu = \frac{R(K)}{\max\limits_{\forall K} R(K)}\%$, for each point in the curves we run the simulation for 10^4 different channel realizations of the NLOS component in (3.2). Then, the LOS component of the channel in (3.2) is set to $\mathbf{H}_{\text{LOS}}(i,i) = \beta(1+j)$, $\forall i = 1, \ldots, \min(M,N)$, $\mathbf{H}_{\text{LOS}}(i,k) = \sqrt{\frac{MN - \min(M,N)\beta^2}{\min(M,N)}}(1+j)$, $\forall i \neq k$ or $i = k > \min(M,N)$, $j = \sqrt{-1}$, which leads to $|\mathbf{H}_{\text{LOS}}|^2 = 1$, $\forall M, N, \beta$. Here, β is a constant which we set $\beta = 0.2$ in the simulations. In all figures, the GA approach is run for $N_{\text{it}} = 1000$ iterations, which is a sufficiently large number with no performance improvement observed. Then, Tables 5.1-5.2 show the average number of required iterations to reach the (sub)optimal solution. Also, we set $L = 10, S = 5$ in Algorithm 1. Finally, in all figures, except for Fig. 5.5, we consider an ideal PA, i.e., set $P_{\max} = \infty, \mu = 0, \epsilon = 1$ in (3.7). The effect of imperfect PAs is studied in Fig. 5.5. In Figs. 5.1-5.2 and Tables 5.1-5.2, we use the relative end-to-end throughput $\nu = \frac{R(K)}{\max\limits_{\forall K} R(K)}\%$ as the performance metric to make the data analysis easier. Throughput (3.5) optimization with a service outage constraint is studied in

Fig. 5.4.

5.1.1 A Convergence Example



(a) (b)

Figure 5.1: Examples of the convergence process of the GA-based BF for systems with (subplot a) and without (subplot b) delay cost of the algorithm. $M = 32$, $N = 8$, $N_{\text{vec}} = 128$, $P = 10$ dB, $k = 0$, $\theta = -100$ dB.

Figures 5.1 (a) and 5.1 (b) show examples of the GA performance in different iterations in the cases with ($\alpha = 0.001$) and without costs of running the algorithm ($\alpha = 0$), respectively (see (3.5)). Here, we set $M = 32$, $N = 8$, $N_{\text{vec}} = 128$, $P = 10$ dB, $k = 0$, $\theta = -100$ dB. From Fig. 5.1 (a) we observe that very few iterations are required to reach the maximum relative throughput, if the running delay of the algorithm is taken into account. That is, considering the cost of running the algorithm, the maximum relative throughput is obtained by finding a suboptimal BF matrix and leaving the rest of the time slot for data transmission (see Fig. 1.4). On the other hand, as the number of iterations increases, the cost of running the algorithm reduces the relative throughput and it converge to zero at $K = \frac{1}{\alpha}$ (see (3.5)). With no cost for running the algorithm, on the other hand, the system performance improves with the number of iterations monotonically (Fig. 5.1 (b)). From Fig. 5.1 (b) we can see that $K \to \infty$, enabled by $N_{\text{it}} \to \infty$, represents exhaustive search. In this case, the developed algorithm leads to (almost) the same performance as the exhaustive search-based scheme with very limited number of iterations. Note that with the parameter settings of Fig. 5.1, exhaustive search implies testing in the order of 10^{30} possible BF matrices. However, with the parameter settings of Fig. 5.1 (b), our algorithm reaches more than 95% of the maximum achievable relative throughput with less than 200 iterations.

Table 5.1: Average number of required iterations ($\alpha = 0.001$).

M/N	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 10$
32/8	58	52	47	44	35
64/8	60	52	46	43	29

Table 5.2: Average number of required iterations ($\alpha = 0$).

M/N	$k=0$	$k=1$	$k=2$	$k=3$	$k=10$
32/8	495	488	484	483	474
64/8	497	496	491	490	488

Furthermore, Tables 5.1 and 5.2 show the average number of iterations that is required in delay-constrained and delay-unconstrained systems to achieve the (sub)optimal system throughput, in the case with $k = 0, 1, 2, 3, 10$ and different number of transmit antennas/users. As demonstrated, for both delay-constrained and delay-unconstrained cases the maximum end-to-end throughput is achieved with few iterations of the algorithm. Also, the required number of iterations is almost independent of the channel model and decreases if the cost of running the algorithm is taken into account (Tables 5.1 and 5.2).

5.1.2 Performance Analysis for Different Parameters

In this part, we evaluate the system performance with different parameter settings.

5.1.2.1 On the Effect of the Channel Condition

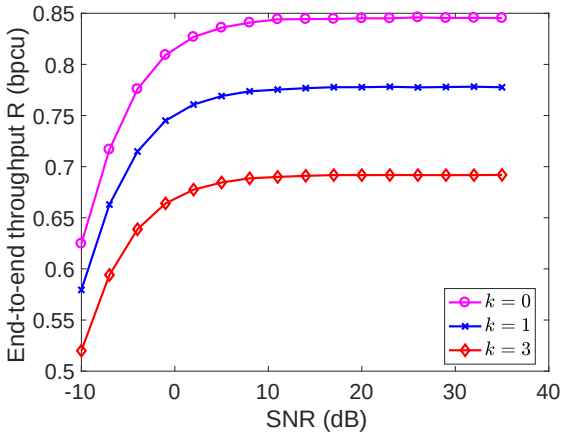


Figure 5.2: The effect of different channel conditions on the end-to-end throughput (3.5), $M = 32$, $N = 8$, $\alpha = 0.001$, $\theta = -100$ dB.

Figure 5.2 shows the effect of the channel condition on the end-to-end throughput. Here, we set $M = 32$, $N = 8$, $k = 0, 1, 3$, and plot the end-to-end throughput versus the SNR. As seen in the figure, the throughput decreases with k, because with the parameter settings of the figure the interference power increases more than the useful signal power as the power of the LOS signal components increase.

5.1.2.2 On the Effect of the Codebook Size

In Fig. 5.3, we analyze the effect of the codebook size N_{vec} in (3.3) on the end-to-end throughput. Here, we set $M = 32$, $N = 8$, $k = 1$, $P = 8, 10$ dB and consider

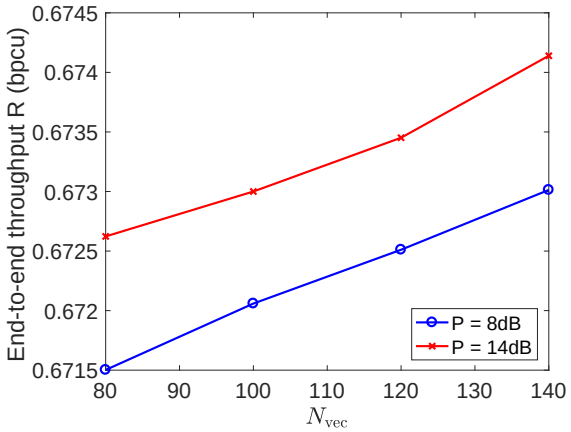


Figure 5.3: The effect of the codebook size on the end-to-end throughput (3.5), $\alpha = 0.001$, $M = 32$, $N = 8$, $k = 1$, $\theta = -100$ dB.

a delay-constrained system with $\alpha = 0.001$. As demonstrated in the figure, with the considered values of N_{vec} the end-to-end throughput is almost insensitive to N_{vec}. However, for small values of N_{vec}, the codebook size is expected to affect the throughput remarkably.

5.1.2.3 On th Effect of Service Outage Constraint

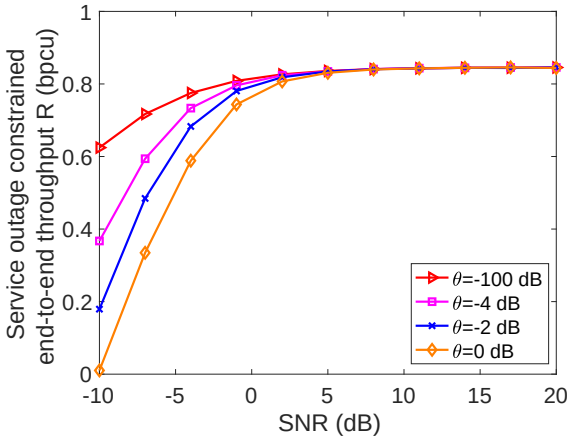


Figure 5.4: Service outage constrained end-to-end throughput with $M = 32$, $N = 8$, $k = 0$, $N_{\text{vec}} = 128$, $\theta = -100, -4, -2, 0$ dB.

Figure 5.4 demonstrates the service outage constrained end-to-end throughput (3.5) for different values of the required received SNR threshold θ in (3.5). It can be seen that the service outage constraint affects the end-to-end and the per-user throughput significantly at severe service outage constraints. However, the effect of service outage constraint on the throughput decreases for small values of θ.

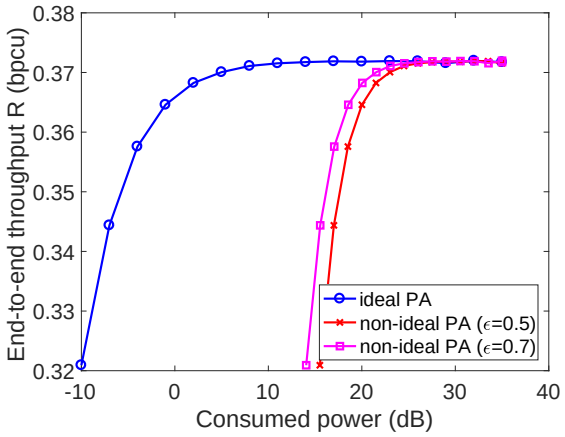


Figure 5.5: The effect of power budget and PAs efficiency on the end-to-end throughput (3.5). $M = 64$, $N = 8$, $k = 3$, $N_{\text{vec}} = 128$, $\alpha = 0.001$, $\theta = -100$ dB.

5.1.3 On the Effect of the Imperfect Power Amplifier

Figure 5.5 evaluates the effect of power budget on the end-to-end throughput. Considering $M = 64$, $N = 8$, $N_{\text{vec}} = 128$, $k = 3$, we plot the end-to-end throughput versus the consumed power (see (3.7)). The effect of non-ideal PAs is also considered where we set $\rho_{\max} = 35$ dB, $\mu = 0.5$, $\epsilon = 0.5, 0.7$. As demonstrated in the figure, the inefficiency of the PA affects the end-to-end throughput remarkably. However, the effect of the PA inefficiency decreases with the consumed power. This is intuitively because the effective efficiency of the PAs $\epsilon^{\text{effective}} = \epsilon(\frac{\rho_{\text{out}}}{\rho_{\max}})^{\mu}$ increases with the power.

5.2 Different State-of-the-art Approaches

In this section, we compare different machine learning based IA BF approaches which have been introduced in chapter 4, including an extended version of our proposed GA-based BF in [19], Tabu search BF [26], link-by-link BF [27] and two-level codebook BF [24]-[25] in large-but-finite MU-MIMO mm-wave communication systems. Also, we extend our algorithm in [4] to the cases with BF at both the transmitter and the receiver. We study the system performance in terms of the end-to-end throughput with service outage constraints (3.5) as well as the implementation complexity.

5.2.1 Comparison of the Service Outage-Constrained End-to-end Throughput

In Fig. 5.6, we compare the throughput (3.5) reached by different considered algorithms. Here, the results are obtained for $M = 32$, $\tau = 4$, $N = 4$, $k = 0$, $\alpha = 0$, $\theta = -4$ dB. It can be seen from the figure that for a broad range of SNRs our GA-based BF [19] leads to the best system throughput, followed by the link-by-link search [27], two-level search [24], [25] and Tabu search [26].

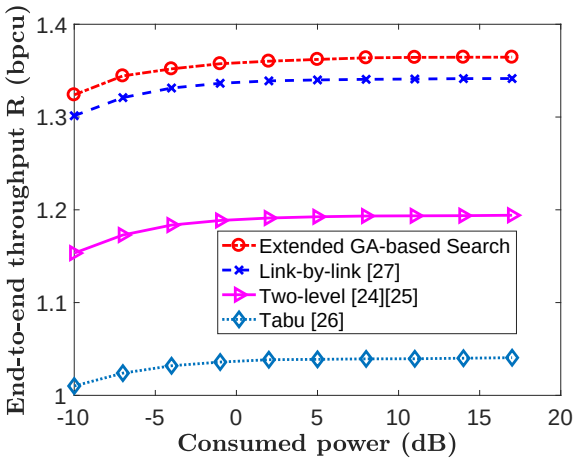


Figure 5.6: Service outage-constrained end-to-end throughput of different methods. $M = 32$, $\tau = 4$, $N = 4$, $k = 0$, $\alpha = 0$.

Figure 5.7 shows the effect of number of receive antennas per user β on the throughput (3.5). As seen in the figure, the end-to-end throughput increases with the number of per-user antennas as expected, since multi-antenna techniques can improve the data rate remarkably. Moreover, the relative performance gain of the GA-based scheme, compared to other considered schemes, increases with the number of receive antennas, which is an interesting point when designing large-scale networks.

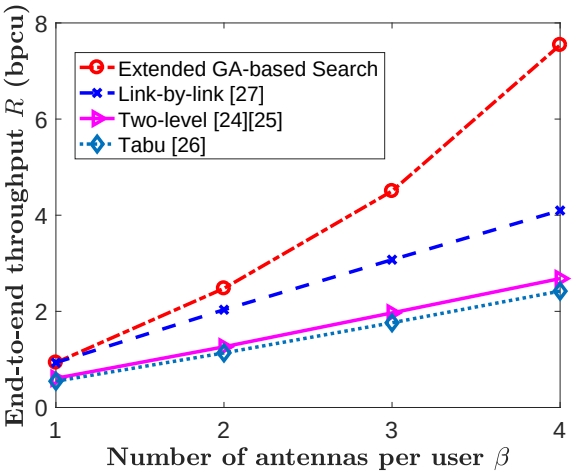


Figure 5.7: Throughput (3.5) with different numbers of receive antennas at the user side β. $M = 32$, $\tau = 4$, $\beta = 1, 2, 3, 4$, $k = 3$, $\alpha = 0$, $\theta = -4$ dB, $P = 2$ dB.

5.2.2 Comparison of the Implementation Complexity

To compare different methods, it is necessary that we consider the implementation complexity of each algorithm. For this reason, we derive the per-iteration complexity of different algorithms based on the fact that the product of matrices of size $N \times M$ and $M \times M$ has the complexity $\mathcal{O}(NM^2)$ in MATLAB. In this way, the per-iteration complexity for the GA-based approach is given by

$$C_{\text{GA}} = L(2\mathcal{O}(N^2M) + \mathcal{O}(NM^2) + \mathcal{O}(NM)), \tag{5.1}$$

and $C_{\text{Tabu}} = C_{\text{GA}}$, $C_{\text{link-by-link}} = \tau \times C_{\text{GA}}$, $C_{\text{two-level}} = 2 \times C_{\text{GA}}$. Here, L is the number of beam selection sets within each iteration.

6

Summary of Included Papers

6.1 A Genetic Algorithm-based Beamforming Approach for Delay-constrained Networks

Accepted in 15th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (Proc. IEEE WiOpt' 2017), Telecom ParisTech, Paris, France, 15th-19th May, 2017.

In this paper, we study the performance of IA beamforming schemes in the cases with large-but-finite number of transmit antennas and users. Particularly, we develop an efficient beamforming scheme using genetic algorithms. With the proposed algorithm, the appropriate beamforming matrix is selected from a set of predefined matrices such that the network performance is optimized. Moreover, taking the mm-wave communications characteristics and different metrics into account, we investigate the effect of various parameters such as the number of antennas/receivers, beamforming resolution as well as hardware impairments on the system performance. As shown, our proposed algorithm is generic in the sense that it can be effectively applied with different channel models, metrics and beamforming methods. Also, our results indicate that the proposed scheme can reach (almost) the same end-to-end throughput as the exhaustive search-based optimal approach with considerably less implementation complexity.

For the simulation results, we consider the end-to-end throughput, the end-to-end service outage-constrained throughput as well as the service outage probability. In this way, we take the algorithm running time into account and, as opposed to conventional iterative schemes, the system performance is not necessarily improved in successive iterations. Instead, as shown via simulations, the maximum throughput is achieved with few iterations, i.e., by picking up a suboptimal beamforming approach, and using the remaining time for information transmission.

The simulation results show that 1) the proposed scheme can reach (almost) the same performance as in the exhaustive search-based scheme with considerably less implementation complexity. Moreover, 2) non-ideal PAs affect the system performance significantly and should be carefully considered/compensated in the network design, and 3) the network throughput increases almost linearly with the number of codebook vectors. 4) The proposed algorithm is effectively applicable for various convex and non-convex performance metrics. 5) In general, the users service out-

age constraints affect the end-to-end throughput remarkably, while the effect of the constraint decreases at high signal-to-noise ratios (SNRs).

6.2 A Comparison of IA Beamforming Algorithms for Millimeter Wave Networks

Submitted to 2017 IEEE 86th Vehicular Technology Conference: VTC2017-Fall, 24–27 September 2017, Toronto, Canada.[1]

In this paper, we extend our previously proposed GA-based beamforming scheme to include beamforming at both the transmitter and the receiver, and we compare the system performance with alternative beamforming approaches in the mm-wave MU-MIMO networks. Taking the mm-wave communications characteristics and various metrics into account, we investigate the effect of different parameters such as the number of transmit antennas/users/per-user receive antennas, beamforming resolution as well as hardware impairments on the system performance employing different beamforming algorithms. As shown, our proposed GA-based approach performs well in delay-constrained networks with multi-antenna users. Compared to the considered state-of-the- art schemes, our method reaches the highest service outage- constrained end-to-end throughput with considerably less implementation complexity.

[1] To be attached in this report after the VTC review.

7
Future Work

The main contribution of this thesis is to study, understand and design efficient IA BF schemes for large-but-finite MU-MIMO networks. However, as a relatively broad topic, there are many other interesting directions for the future work:

- Present a comprehensive picture of how the overall network access procedure should be,
- Extend our beam selection method to the cases with digital/hybrid BF schemes,
- Test our algorithm with realistic channel models and more specific access threshold in the mm-wave bands,
- Invent novel transmission and reception techniques using large antenna arrays for multi-node mm-wave systems employing base station cooperation and relaying for mobile users.

Bibliography

[1] N. Alliance, "NGMN 5G white paper," *Next Generation Mobile Networks, White paper*, Feb. 2015. [Online]. Available: https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0.pdf

[2] A. Osseiran *et al.*, "Scenarios for 5G mobile and wireless communications: the vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, May. 2014.

[3] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, Aug. 1988.

[4] E. Dahlman, S. Parkvall, J. Skold, and P. Beming, *3G evolution: HSPA and LTE for mobile broadband*. Academic press, 2010.

[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, Jan. 2016.

[6] C. N. Barati, S. A. Hosseini, M. Mezzavilla, T. Korakis, S. S. Panwar, S. Rangan, and M. Zorzi, "Initial Access in Millimeter Wave Cellular Systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 7926–7940, Dec. 2016.

[7] J. P. Gilb, "IEEE Standards 802.15. 3c?–Part 15.3: wireless medium access control (MAC) and physical layer (PHY) specifications for high rate wireless personal area networks (WPANs) Amendment 2: millimeter-wave-based alternative physical layer extension [S]," *IEEE Computer Society, New York*, 2009.

[8] C. Cordeiro *et al.*, "IEEE P802. 11 Wireless LANs, PHY/MAC Complete Proposal Specification (IEEE 802.11-10/0433r2)," 2010.

[9] H. Rate, "GHz PHY, MAC and PALs, Standard ECMA-387," Dec. 2010.

[10] V. Desai, L. Krzymien, P. Sartori, W. Xiao, A. Soong, and A. Alkhateeb, "Initial beamforming for mmwave communications," in *Proc. IEEE Asilomar'2014, CA, USA*, Nov. 2014, pp. 1926–1930.

[11] C. N. Barati, S. A. Hosseini, S. Rangan, P. Liu, T. Korakis, S. S. Panwar, and T. S. Rappaport, "Directional cell discovery in millimeter wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6664–6678, Jul. 2015.

[12] Y. Qi and M. Nekovee, "Coordinated Initial Access in Millimetre Wave Standalone Networks," *Millimeter-wave Networking Workshop (mmNet 2016)*, May. 2016.

[13] Y. Li, J. G. Andrews, F. Baccelli, T. D. Novlan, and C. Zhang, "Design and Analysis of Initial Access in Millimeter Wave Cellular Networks," *Submitted to IEEE Trans. Wireless Commun.*, Mar. 2017. [Online]. Available: arXiv:1609.05582

[14] A. Alkhateeb, R. W. Heath, and G. Leus, "Achievable rates of multi-user millimeter wave systems with hybrid precoding," in *Proc. IEEE ICCW'2015, London, UK*, Jun. 2015, pp. 1232–1237.

[15] H. Shokri-Ghadikolaei, C. Fischione, G. Fodor, P. Popovski, and M. Zorzi, "Millimeter wave cellular networks: A MAC layer perspective," *IEEE Trans. Commun.*, vol. 63, no. 10, pp. 3437–3458, Oct. 2015.

[16] M. Giordani, M. Mezzavilla, C. N. Barati, S. Rangan, and M. Zorzi, "Comparative analysis of initial access techniques in 5G mmwave cellular networks," in *Proc. IEEE CISS'2016, Princeton University, Princeton, NJ 08544*, Mar. 2016, pp. 268–273.

[17] M. Giordani, M. Mezzavilla, and M. Zorzi, "Initial access in 5G mmWave cellular networks," *IEEE Commun. Mag.*, vol. 54, no. 11, pp. 40–47, Nov. 2016.

[18] Y. Li, J. G. Andrews, F. Baccelli, T. D. Novlan, and J. Zhang, "On the initial access design in millimeter wave cellular networks," in *Proc. IEEE GLOBECOM'2016, Washington, USA*, Dec. 2016, pp. 1–6.

[19] H. Guo, B. Makki, and T. Svensson, "A Genetic Algorithm-based Beamforming Approach for Delay-constrained Networks," in *Proc. IEEE WiOpt'2017, Paris, France*, May. 2017. [Online]. Available: arXiv:1703.03792

[20] B. Makki, T. Svensson, G. Cocco, T. de Cola, and S. Erl, "On the throughput of the return-link multi-beam satellite systems using genetic algorithm-based schedulers," in *Proc. IEEE ICC'2015, London, UK*, Jun. 2015, pp. 838–843.

[21] B. Makki, A. Ide, T. Svensson, T. Eriksson, and M.-S. Alouini, "A Genetic Algorithm-based Antenna Selection Approach for Large-but-Finite MIMO Networks," *IEEE Trans. Veh. Technol.*, Dec. 2016.

[22] V. Raghavan, J. Cezanne, S. Subramanian, A. Sampath, and O. Koymen, "Beamforming tradeoffs for initial UE discovery in millimeter-wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 543–559, Jan. 2016.

[23] L. Zhou and Y. Ohashi, "Efficient codebook-based MIMO beamforming for millimeter-wave WLANs," in *Proc. IEEE PIMRC'2012, Sydney, Australia*, Sept. 2012, pp. 1885–1889.

[24] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Multilevel millimeter wave beamforming for wireless backhaul," in *Proc. IEEE GC Wkshps'2011, Houston, Texas, USA*, Dec. 2011, pp. 253–257.

[25] L. Chen, Y. Yang, X. Chen, and W. Wang, "Multi-stage beamforming codebook for 60GHz WPAN," in *Proc. IEEE ICST'2011, Harbin, China*, Aug. 2011, pp. 361–365.

[26] X. Gao, L. Dai, C. Yuen, and Z. Wang, "Turbo-like beamforming based on Tabu search algorithm for millimeter-wave massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5731–5737, Jul. 2016.

[27] J. Qiao, X. Shen, J. W. Mark, and Y. He, "MAC-layer concurrent beamforming protocol for indoor millimeter-wave networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 1, pp. 327–338, Jan. 2015.

[28] J. Singh and S. Ramakrishna, "On the feasibility of codebook-based beamforming in millimeter wave systems with multiple antenna arrays," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2670–2683, Jan. 2015.

[29] C. Cordeiro, D. Akhmetov, and M. Park, "IEEE 802.11 ad: Introduction and performance evaluation of the first multi-Gbps WiFi technology," in *Proceedings of the 2010 ACM international workshop on mmWave communications: from circuits to networks, Chicago, IL, USA*, 2010, pp. 3–8.

[30] L. Wan, X. Zhong, Y. Zheng, and S. Mei, "Adaptive codebook for limited feedback MIMO system," in *Proc. IEEE IFIP'2009, Cairo, Egypt*, Apr. 2009, pp. 1–5.

[31] J. Choi, "Beam selection in mm-Wave multiuser MIMO systems using compressive sensing," *IEEE Trans. Commun.*, vol. 63, no. 8, pp. 2936–2947, Jun. 2015.

[32] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[33] B. Li, Z. Zhou, W. Zou, X. Sun, and G. Du, "On the efficient beamforming training for 60GHz wireless personal area networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 504–515, Feb. 2013.

[34] H.-H. Lee and Y.-C. Ko, "Low complexity codebook-based beamforming for MIMO-OFDM systems in millimeter-wave WPAN," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3607–3612, Nov. 2011.

[35] B. Makki, T. Svensson, T. Eriksson, and M.-S. Alouini, "On the Required Number of Antennas in a Point-to-Point Large-but-Finite MIMO System: Outage-Limited Scenario," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1968–1983, May. 2016.

[36] D. Persson, T. Eriksson, and E. G. Larsson, "Amplifier-aware multiple-input single-output capacity," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 913–919, Jan. 2014.

Part II

Included Papers

A Genetic Algorithm-based Beamforming Approach for Delay-constrained Networks

Hao Guo, Behrooz Makki, Tommy Svensson
Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden
ghao@student.chalmers.se, {behrooz.makki, tommy.svensson}@chalmers.se

***Abstract*—In this paper, we study the performance of initial access beamforming schemes in the cases with large but finite number of transmit antennas and users. Particularly, we develop an efficient beamforming scheme using genetic algorithms. Moreover, taking the millimeter wave communication characteristics and different metrics into account, we investigate the effect of various parameters such as number of antennas/receivers, beamforming resolution as well as hardware impairments on the system performance. As shown, our proposed algorithm is generic in the sense that it can be effectively applied with different channel models, metrics and beamforming methods. Also, our results indicate that the proposed scheme can reach (almost) the same end-to-end throughput as the exhaustive search-based optimal approach with considerably less implementation complexity.**

I. Introduction

Developing key technical components and concepts for millimeter wave (MMW) communications in the range of 6-100 GHz is of interest for 5G. Different works have estimated/measured the channel characteristics in such MMW frequency bands [1]–[4]. The use of such high frequencies for mobile communications is challenging but necessary for supporting 5G which targets for peak data rates in the order of 10-100 Gbps with low end-to-end latencies (down to 1 ms) [5].

Due to peak power limitation and high path loss in MMW communications, there is a need for directional transmissions. Fortunately, the physical size of antennas at MMW frequency bands is small so that it is possible to use large antenna arrays and perform beamforming [3] [4]. For typical wireless systems, beamforming is performed by employing precoding with channel state information (CSI) feedback or estimation after the control link is established. However, even with the extended coverage from beamforming, the coverage range for MMW frequencies is typically small due to high pathloss. As a result, we need to employ beamforming also on initial access (IA) channels. This calls for the need to design novel IA procedures.

During the MMW initial access procedure, beamforming is different from the conventional one because it is hard to acquire CSI. Different works have been recently presented on both physical architecture and procedural algorithm to solve the problem (see Section II for literature review). However, in these works either the initial access algorithms are designed for specific metrics, channel models, and precoding schemes or their implementation complexity grows significantly with the number of antennas/users. Moreover, the running delay of the algorithm is an important issue which has been rarely considered in the performance evaluations.

In this paper, we study the performance of large-but-finite multiple-input-multiple-output (MIMO) MMW networks using codebook-based beamforming. The contributions of the paper are two-fold. First, we propose an efficient genetic algorithm (GA)-based approach for initial access beamforming. With the proposed algorithm, the appropriate beamforming matrix is selected from a set of predefined matrices such that the network performance is optimized. As we show, our proposed scheme is generic in the sense that it can be implemented in the cases with different channel models, beamforming methods as well as optimization metrics. Second, we evaluate the performance of the beamforning-based MIMO networks for different parameters such as hardware impairments, different channel models, beamforming resolution and number of antennas/receivers.

For the simulation results, we consider the end-to-end throughput, the end-to-end service outage-constrained throughput as well as the service outage probability. In this way, we take the algorithm running time into account and, as opposed to conventional iterative schemes, the system performance is not necessarily improved in successive iterations. Instead, as shown via simulations, the maximum throughput is achieved with few iterations, i.e., by picking up a suboptimal beamforming approach, and using the remaining time for information transmission.

The simulation results show that 1) the proposed scheme can reach (almost) the same performance as in the exhaustive search-based scheme with considerably less implementation complexity. Moreover, 2) non-ideal power amplifiers (PAs) affect the system performance significantly and should be carefully considered/compensated in the network design, and 3) the network throughput increases almost linearly with the number of codebook vectors. 4) The proposed algorithm is effectively applicable for various convex and non-convex performance metrics. 5) In general, the users service outage constraints affect the end-to-end throughput remarkably, while the effect of the constraint decreases at high signal-to-noise ratios (SNRs). 6) In practice, taking the algorithm running delay into account, the maximum end-to-end throughput is reached by dedicating a small fraction of the packet period to finding the suboptimal beamforming solution and using the rest of the packet for data transmission.

II. Literature Review

In this section, we review the recent results on initial access. The readers mainly interested in the technical discussions can skip this part and go to Sections III-V where we present the system model, our proposed algorithm and the simulation results, respectively. Beamforming techniques at MMW frequencies have been widely investigated and led to standard developments such as IEEE 802.15.3c (TG3c) [6], IEEE 802.11ad (TGad) [7] and ECMA-387 [8]. In wireless local area networks (WLANs), a one-sided beam search strategy using a beamforming codebook has been employed to establish initial alignment between large array antennas [9].

For mobile communication systems, on the other hand, there are few works on initial access beamforming. In general, most of the presented works are based on multi-level/greedy search algorithms and utilize the sparse nature of the MMW channel. Several issues for initial access beamforming in MMW frequencies are presented in [10] and a fast-discovery hierarchical search method is proposed. Moreover, [11] designs a novel greedy-geometric algorithm to synthesize antenna patterns featuring desired beamwidth. In [12], a survey of several recently proposed IA techniques is provided. Then, [13] performs IA for clustered MMW small cells with a power-delay-profile-based approach to reduce the IA set up time. Also, [14] shows the significant benefits of using low-resolution fully digital architectures during IA, in comparison to single stream analog beamforming.

In MMW multiuser multiple-input-single-output (MISO) downlink systems, an opportunistic random beamforming technique is provided in [15]. Also, [16] develops low-complexity algorithms for optimizing the choice of beamforming directions, premised on the sparse multipath structure of the MMW channel. Then, [17] studies a low-complexity beam selection method by designing low-cost analog beamformers. This beam selection method can be carried out without explicit channel estimation. A directional cell discovery method is proposed in [18] where the BS periodically transmits synchronization signals to scan the whole angular space in time-varying random directions. In [19], MMW precoder design is formulated as a sparsity-constrained signal recovery problem, and an algorithmic solution with orthogonal matching pursuit is proposed. Finally, [20] proposes a hybrid precoding algorithm based on a low training overhead channel estimation method to overcome the hardware constraints in MMW analog-only beamforming systems.

Considering codebook-based beamforming, a broadcast-based solution for MMW systems is proposed in [21], where limited feedback-type directional codebooks are used for the beamforming procedure. Moreover, [22] studies concurrent beamforming issues for achieving high capacity in indoor MMW networks. In [23], an efficient beam alignment technique is designed which uses adaptive subspace sampling and hierarchical beam codebooks. With pre-specified beam codebooks, [24] proposes a Rosenbrock numerical algorithm to accelerate the beam-switch process which is modeled as a 2-D plane optimization problem. Also, [25] adopts the discrete Fourier transform (DFT)-based codebooks and proposes an efficient iterative antenna vector training algorithm. Finally, [26] provides a codebook-based beamforming scheme with multi-level training and level-adaptive antenna selection, which can be used for MIMO orthogonal frequency division multiplexing (OFDM) systems in MMW wireless personal area networks (WPANs).

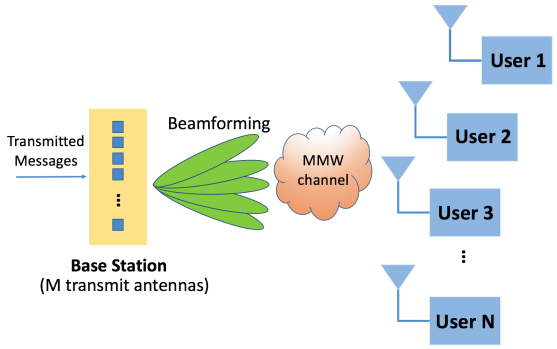


Fig. 1. MMW multiuser MIMO system model.

There are previous works using the GA-based selection approach. For instance, [27] elaborates on the performance of scheduling in the return-link of a multi-beam satellite system. Moreover, [28] uses a genetic algorithm to achieve a near-optimal array gain in all directions during codebook-based beamforming.

III. System Model

We use a multiuser MIMO setup with M transmit antennas in a BS and N single-antenna users (see Fig. 1). At each time slot t, the received signal can be described as

$$\mathbf{Y}(t) = \sqrt{\frac{P}{M}}\mathbf{H}(t)\mathbf{V}(t)\mathbf{X}(t) + \mathbf{Z}(t), \tag{1}$$

where P is the total power budget, $\mathbf{H}(t) \in \mathcal{C}^{N\times M}$ is the channel matrix, $\mathbf{X}(t) \in \mathcal{C}^{M\times 1}$ is the intended message signal, $\mathbf{V}(t) \in \mathcal{C}^{M\times M}$ is the precoding matrix, and $\mathbf{Z}(t) \in \mathcal{C}^{N\times 1}$ denotes the independent and identically distributed (IID) Gaussian noise matrix. For simplicity, we drop time index t in the following.

The channel $\mathbf{H}$ is modeled by

$$\mathbf{H} = \sqrt{\frac{k}{k+1}}\mathbf{H}_{\text{LOS}} + \sqrt{\frac{1}{k+1}}\mathbf{H}_{\text{NLOS}}, \tag{2}$$

where $\mathbf{H}_{\text{LOS}}$ and $\mathbf{H}_{\text{NLOS}}$ denote the line-of-sight (LOS) and non-line-of-sight (NLOS) components of the channel. Also, k controls the relative strength of the LOS and NLOS components. It can be seen that $k = 0$ represents an NLOS channel, while $k \to \infty$ gives a LOS condition. Also, the NLOS component is assumed to follow complex Gaussian distribution.

Fig. 2. Schematic of a packet transmission period.

A. Initial Beamforming Procedure

Conventional beamforming procedure schemes utilize CSI to generate the precoding matrix. However, it is almost impossible to acquire CSI in large scale MMW systems. Instead, we perform codebook-based beamforming, which means selecting a precoding matrix $\mathbf{V}$ out of a predefined codebook $\mathbf{W}$ at the BS, sending test signal and finally making decisions on transmit beam patterns based on users' feedback about their received metrics. The IA will be finished as soon as a stable control link is established. The time structure for the packet transmission can be seen in Fig. 2, where part of the packet period is dedicated to design the appropriate beams and the rest is used for data transmission. Thus, we need to find a balance between beamforming time and data transmission period by picking up a suboptimal beamforming approach and using the remaining time for information transmission. Here, we use a DFT-based codebook [29] defined as

$$\begin{aligned}&\mathbf{W} = |w(m,u)| = |e^{-j2\pi(m-1)(u-1)/N_{\text{vec}}}|,\\ &m = 1,2,...,M, u = 1,2,...,N_{\text{vec}},\end{aligned} \tag{3}$$

where $N_{\text{vec}} \geq M$ is the number of codebook vectors. This codebook can achieve uniform antenna gain in all directions, however, as seen in the following, the proposed algorithm can be implemented for different codebook definitions.

B. Performance Metrics

As seen in the following, the proposed GA-based algorithm is generic, in the sense that it can be effectively applied for various performance metrics. For the simulations, however, we consider the end-to-end throughput, the service outage-constrained throughput and the service outage probability defined as follows.

Set N_{it} to be the maximum possible number of iterations which is decided by designer. Considering the K-th iteration round of the algorithm, K=1, 2,..., N_{it}, the end-to-end throughput in bit-per-channel-use (bpcu) is defined as

$$\begin{aligned}&R(K) = (1-\alpha K)\sum_{i=1}^{N} r_i^K,\\ &r_i^K = \log_2\left(1+\mathbf{SINR}_i^K\right).\end{aligned} \tag{4}$$

Here, α is the relative delay cost for running each iteration of the algorithm which fulfills $\alpha N_{\text{it}} < 1$. Also,

$$\mathbf{SINR}_i^K = \frac{\frac{P}{M} g_{i,i}}{BN_0 + \frac{P}{M}\sum_{i\neq j}^{N} g_{i,j}} \tag{5}$$

is the signal-to-interference-plus-noise ratio (SINR) of user i in iteration K, in which $g_{i,j}$ is the (i,j)-th element of the matrix $\mathbf{G}^K = |\mathbf{H}\mathbf{V}^K|^2$ (and $\mathbf{G}$ is referred to as the channel gain throughout the paper), B is the system bandwidth and N_0 is the power spectral density of the noise. Thus, r_i^K denotes the achievable rate of user i at the end of the K-th iteration of the algorithm. In this way, as opposed to, e.g., [14, Eq. 1] [17, Eq. 43] [19, Eq. 3], [22, Eq. 3], [24, Eq. 5] [26, Eq. 5], we take the algorithm running delay into account. Thus, there is a trade-off between finding the optimal beamforming matrices and reducing the data transmission time slot, and the highest throughput may be achieved by few iterations, i.e., a rough estimation of the optimal beamformer. To simplify the presentations, we set $BN_0 = 1$. As a result, the power P (in dB, $10\log_{10} P$) denotes the SNR as well.

In different applications, it may be required to serve the users with some minimum required rates, otherwise *service outage* occurs. For this reason, we analyze the end-to-end service outage-constrained throughput defined as

$$\begin{aligned}&\tilde{R}(K) = (1-\alpha K)\sum_{i=1}^{N} r_i U(r_i, \log_2(1+\theta)),\\ &U(r_i,\log_2(1+\theta)) = \begin{cases}1 & r_i \geq \log_2(1+\theta)\\ 0 & r_i < \log_2(1+\theta),\end{cases}\end{aligned} \tag{6}$$

where $\log_2(1+\theta)$ is the minimum rate required by the users and θ represents the minimum required SINR of the users. This is interesting for applications where each user is required to have a minimum rate $\log_2(1+\theta)$. Among our motivations for the service outage-constrained throughput analysis is to highlight the effectiveness of the proposed algorithm in optimizing the non-convex criteria.

Finally, as another performance metric, we study the service outage probability which is defined as

$$\phi = \Pr(r_i < \log_2(1+\theta), \forall i). \tag{7}$$

C. On the Effect of Power Amplifier

In multi-antenna systems, when the number of transmit antennas increases, the efficiency of radio-frequency PAs should be taken into account. Here, we consider the state-of-the-art PA efficiency model [30, Eq. 13], [31, Eq. 3]

$$\rho_{\text{cons}} = \frac{\rho_{\text{max}}^{\mu}}{\epsilon \times \rho_{\text{out}}^{\mu-1}} \tag{8}$$

where ρ_{cons}, ρ_{out}, ρ_{max} refer to the consumed power, output power and maximum output power of the PA, respectively. Also, $\epsilon \in [0,1]$ is the power efficiency and $\mu \in [0,1]$ is a parameter which depends on the PA classes. Note that setting $\epsilon = 1$, $P_{\text{max}} = \infty$ and $\mu = 0$ represents the special case with an ideal PA.

IV. Algorithm Description

We use a GA-based approach for beam selection during IA beamforming and details are explained in Algorithm 1. The algorithm starts by getting L possible beam selection sets randomly and each of them means a certain beam formed by transmit antennas, i.e., a submatrix of the codebook. During

each iteration, we determine the best selection result, named as the *Queen*, based on our objective metrics. For instance, we choose the beamforming matrix with the highest end-to-end throughput if (4) is considered as the objective function. Next, we keep the Queen and regenerate $S < L$ matrices around the Queen. This can be done by making small changes to the Queen such as changing a number of columns in the Queen matrix. In the simulations, we replace 10% of the columns of the Queen by other random columns from the codebook. Finally, during each iteration $L-S-1$ beamforming matrices are selected randomly. After N_{it} iterations, considered by the algorithm designer, the Queen is returned as the beam selection rule in the considered time slot.

Algorithm 1 GA-based Beam Selection Algorithm

In each time slot with instantaneous channel realization $\mathbf{H} \in \mathcal{C}^{N\times M}$, do the followings:

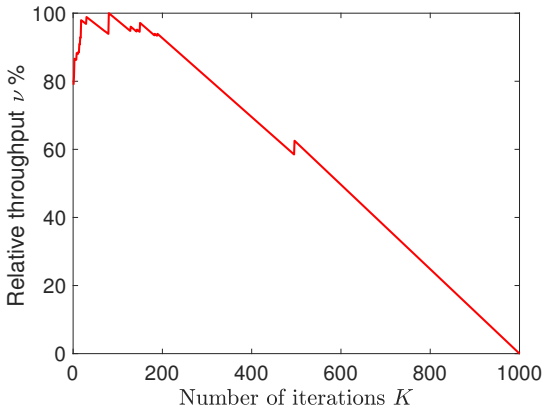
I. Consider L, e.g., $L = 10$, sets of beamforming matrices $\mathbf{V}_l$, $l = 1,..., L$, randomly selected from the pre-defined codebook $\mathbf{W}$.
II. For each $\mathbf{V}_l$, evaluate the instantaneous value of the objective metric R_l, $l = 1,..., L$, for example end-to-end throughput (4).
III. Selection: Find the best beamforming matrix which results in the best value of the considered metric, named as the *Queen*, e.g., $\mathbf{V}_q$ satisfies $R(\mathbf{V}_l) \leq R(\mathbf{V}_q)$, $\forall\ l = 1,..., L$ if the end-to-end throughput is the objective function.
IV. $\mathbf{V}_1 \leftarrow \mathbf{V}_q$
V. Create $S \ll L$, e.g., $S = 5$, beamforming matrices $\mathbf{V}_s^{\text{new}}$, $s = 1,..., S$, around the Queen $\mathbf{V}_1$. These sets are generated by making small changes in the Queen.
VI. $\mathbf{V}_{s+1} \leftarrow \mathbf{V}_s^{\text{new}}$, $s = 1,..., S$.
VII. Go back to Step II and run for N_{it} iterations, N_{it} is a fixed number decided by designer.

Return the final Queen as the beam selection rule for the current time slot.

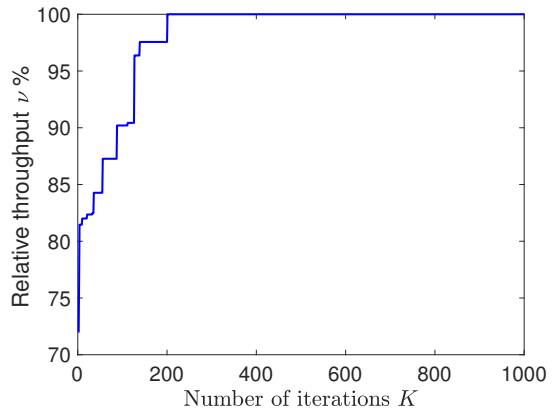
As demonstrated, the algorithm is generic in the sense that it is independent of the channel model, objective function or precoding matrix, thus this can be used in different scenarios and as a benchmark for comparison of different IA schemes. Also, our proposed algorithm converges to the (sub)optimal value of the considered metrics using few iterations. Therefore, the proposed algorithm can be useful in IA beamforming where the delay is one of the most important factors.

V. Simulation Results

For the simulation results, we consider the channel model described by (2), with $k = 0, 1, 2, 3, 10$. We have checked the results for a broad range of parameter settings. However, due to space limits and because they follow the same qualitative behaviors as in Fig. 3-9, they are not reported. The initial access beamforming process is performed based on the codebook defined in (3). However, one can use the algorithm for different beamforming schemes at the transmitter and receivers. Except for Fig. 3 which shows an example of the algorithm procedure by plotting the relative throughput $\nu = \frac{R(K)}{\max\limits_{\forall K} R(K)}\%$, for each point in the curves we run the simulation for 10^4 different channel realizations of the NLOS component in (2). Then, the LOS component of the channel in (2) is set to $\mathbf{H}_{\text{LOS}}(i,i) = \beta(1+j)$, $\forall i = 1, \ldots, \min(M,N)$, $\mathbf{H}_{\text{LOS}}(i,k) = \sqrt{\frac{MN-\min(M,N)\beta^2}{\min(M,N)}}(1+j)$, $\forall i \neq k$ or $i = k > \min(M,N)$, $j = \sqrt{-1}$, which leads to $|\mathbf{H}_{\text{LOS}}|^2 = 1$, $\forall M, N, \beta$. In the simulations we set $\beta = 0.2$. In all figures, the GA algorithm is run for sufficiently large number of iterations until no performance improvement is observed. Then, Tables I-II show the average number of required iterations to reach the (sub)optimal solution. Also, we set $L = 10, S = 5$ in Algorithm 1. Finally, in all figures, except for Fig. 4, we consider an ideal PA, i.e., set $P_{\max} = \infty, \mu = 0, \epsilon = 1$ in (8). The effect of imperfect PAs is studied in Fig. 4. In Figs. 3-6 and Tables I-II, we consider the end-to-end throughput (4) as the performance metric. Throughput optimization with a service outage constraint, i.e., (6), is studied in Figs. 7-9. Finally, Table III studies the system performance in the case



(a) System performance with delay ($\alpha = 0.001$)



(b) System performance without delay ($\alpha = 0$)

Fig. 3. Examples of the convergence process of the GA-based beamforming for systems with (subplot a) and without (subplot b) delay cost of the algorithm. $M = 32$, $N = 8$, $N_{\text{vec}} = 128$, $P = 10$ dB, $k = 0$.

TABLE I
AVERAGE NUMBER OF REQUIRED ITERATIONS ($\alpha = 0.001$)

M/N	$k=0$	$k=1$	$k=2$	$k=3$	$k=10$
32/8	58	52	47	44	35
64/8	60	52	46	43	29

TABLE II
AVERAGE NUMBER OF REQUIRED ITERATIONS ($\alpha = 0$)

M/N	$k=0$	$k=1$	$k=2$	$k=3$	$k=10$
32/8	495	488	484	483	474
64/8	497	496	491	490	488

minimizing the service outage probability (7).

Figures 3a and 3b show examples of the GA performance in different iterations in the cases with ($\alpha = 0.001$) and without costs of running the algorithm ($\alpha = 0$), respectively (see (4)). Here, we set $M = 32$, $N = 8$, $N_{\text{vec}} = 128$, $P = 10$ dB, $k = 0$. From Fig. 3a we observe that very few iterations are required to reach the maximum throughput, if the running delay of the algorithm is taken into account. That is, considering the cost of running the algorithm, the maximum throughput is obtained by finding a suboptimal beamforming matrix and leaving the rest of the time slot for data transmission (see Fig. 2). On the other hand, as the number of iterations increases, the cost of running the algorithm reduces the end-to-end throughput converging to zero at $K = \frac{1}{\alpha}$ (see (4)). With no cost for running the algorithm, on the other hand, the system performance improves with the number of iterations monotonically (Fig. 3b). However, the developed algorithm leads to (almost) the same performance as the exhaustive search-based scheme (from Fig. 3b we can see that $K \to \infty$, enabled by $N_{\text{it}} \to \infty$, represents exhaustive search) with very limited number of iterations (note that with the parameter settings of Fig. 3, exhaustive search implies testing in the order of 10^{30} possible beamforming matrices). For example, with the parameter settings of Fig. 3b, our algorithm reaches more than 95% of the maximum achievable throughput with less than 200 iterations. Furthermore, Tables I and II show the average number of iterations that is required in delay-constrained and delay-unconstrained systems to achieve the (sub)optimal system throughput, in the case with $k = 0, 1, 2, 3, 10$ and different number of transmit antennas/users. As demonstrated, for both delay-constrained and delay-unconstrained cases the maximum end-to-end throughput is achieved with few iterations of the algorithm. Also, the required number of iterations is almost independent of the channel model and decreases if the cost of running the algorithm is taken into account (Tables I and II).

Figure 4 evaluates the effect of power budget on the end-to-end throughput. Considering $M = 64$, $N = 8$, $N_{\text{vec}} = 128$, $k = 3$, we plot the end-to-end throughput versus the consumed power (see (8)). The effect of non-ideal PAs is also considered where we set $\rho_{\max} = 35$ dB, $\mu = 0.5$, $\epsilon = 0.5, 0.7$.

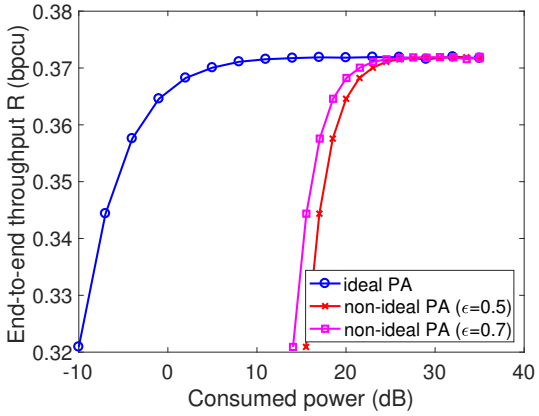


Fig. 4. The effect of power budget and PAs efficiency on the end-to-end throughput (4). $M = 64$, $N = 8$, $k = 3$, $N_{\text{vec}} = 128$. $\alpha = 0.001$.

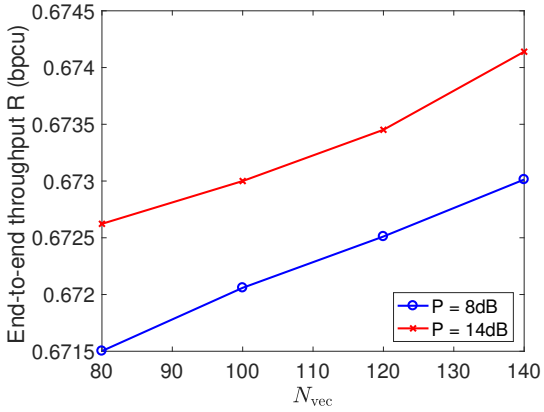


Fig. 5. The effect of the codebook size on the end-to-end throughput (4), $\alpha = 0.001$, $M = 32$, $N = 8$, $k = 1$.

As demonstrated in the figure, the inefficiency of the PA affects the end-to-end throughput remarkably. However, the effect of the PA inefficiency decreases with the SNR. This is intuitively because the effective efficiency of the PAs $\epsilon^{\text{effective}} = \epsilon(\frac{p_{\text{out}}}{p_{\max}})^{\mu}$ increases with SNR.

In Fig. 5, we analyze the effect of the codebook size N_{vec} on the end-to-end throughput. Here, we set $M = 32$, $N = 8$, $k = 1$, $P = 8$ dB, 10 dB. Figure 5 shows the end-to-end throughput in a delay-constrained system ($\alpha = 0.001$) with different numbers of N_{vec} in (3). As expected, the end-to-end throughput increases (almost) linearly with N_{vec}, because there are more options to select the appropriate beamformer as N_{vec} increases.

Figure 6 shows the effect of the channel condition on the end-to-end throughput. Here, we set $M = 32$, $N = 8$, $k = 0, 1, 3$, and plot the end-to-end throughput versus the SNR. As seen in the figure, the throughput decreases with k, because with the parameter settings of the figure the interference power increases more than the useful signal power as the power of the LOS signal components increase.

Figure 7 demonstrates the service outage constrained end-

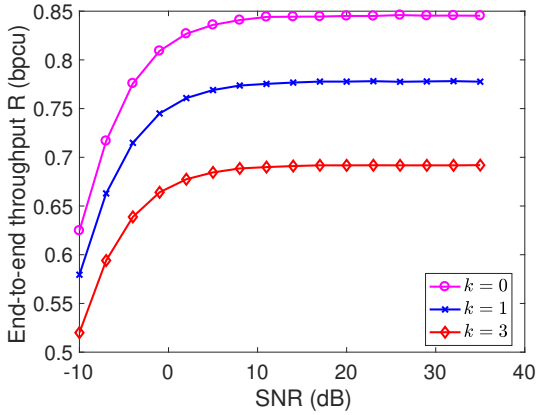


Fig. 6. The effect of different channel conditions on the end-to-end throughput (4), $M = 32$, $N = 8$, $\alpha = 0.001$.

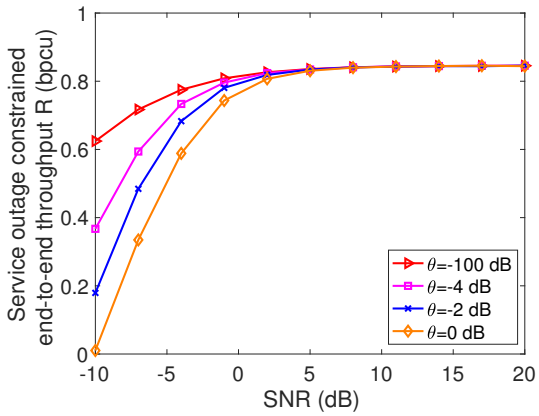


Fig. 7. Service outage constrained end-to-end throughput with $M = 32$, $N = 8$, $k = 0$, $N_{\text{vec}} = 128$, $\theta = -100, -4, -2, 0$ dB.

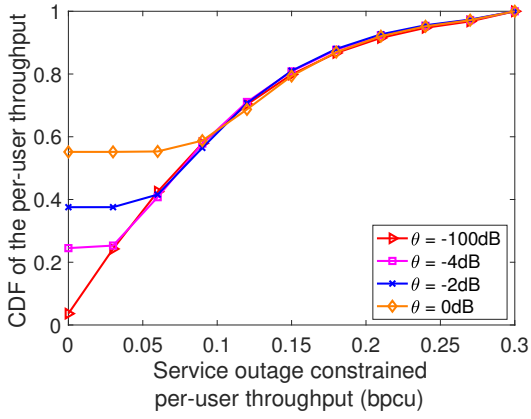


Fig. 8. CDF of service outage constrained per-user throughput in the cases optimizing (6), $M = 32$, $N = 8$, $k = 0$, $N_{\text{vec}} = 128$, $\theta = -100, -4, -2, 0$ dB.

to-end throughput (6) for different values of required received SNR thresholds θ in (6). Also, Fig. 8 shows the cumulative distribution function (CDF) of the users achievable rates for

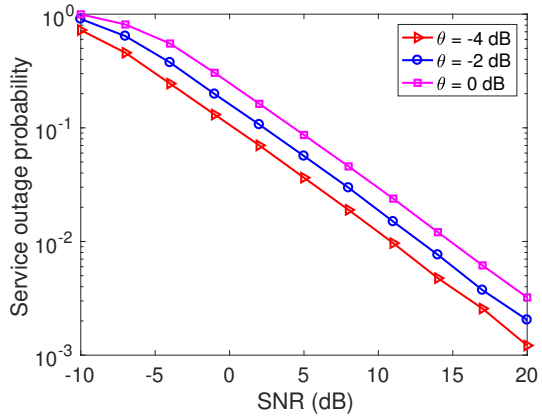


Fig. 9. Service outage probability in the cases optimizing (6), $M = 32$, $N = 8$, $k = 0$, $N_{\text{vec}} = 128$, $\theta = -4, -2, 0$ dB.

different service outage constraints. Here, the results are presented for $N = 8, M = 32, k = 0, N_{\text{vec}} = 128$. Finally, Fig. 9 studies the service outage probability in the cases optimizing (6). As demonstrated in the Figs. 7-8, the service outage constraint affects the end-to-end and the per-user throughput significantly at low SNRs/severe service outage constraints. For instance, with the parameter settings of Fig. 8, more than 50% of the users may receive data with rates less than 1 bpcu, corresponding to $\theta = 0$ dB (see Fig. 8, (6)). However, the effect of the service outage probability decreases as the SNR increases or θ decreases (Figs. 7-9).

TABLE III
AVERAGE NUMBER OF SERVED USERS IN SERVICE OUTAGE-CONSTRAINED SYSTEMS

θ (dB)	-100	-4	-2	0	2
Case1	8.00	7.14	6.82	6.49	6.15
Case2	8.00	7.47	7.15	6.82	6.48

In Table III, we study the average number of served users in Case 1 and 2 where the service outage constrained throughput (6) and the service outage probability (7) are considered as the optimization metric, respectively. Here, the results are presented for $M = 32$, $N = 8$, $k = 0$, $N_{\text{vec}} = 128$. As expected, compared to Case 1 maximizing (6), the average number of served users increases when the goal is to minimize the service outage probability (Case 2). This is indeed at the cost of some end-to-end throughput loss. Finally, the average number of served users increases as θ decreases, i.e., the users minimum required rate decreases.

VI. Conclusion

We studied the performance of initial access beamforming in delay-constrained networks. Considering delay cost of the beamforming procedure, we evaluated the end-to-end throughput as well as the service outage-constrained system performance under different parameter settings. We developed a GA-based beam selection approach which can reach almost the

same throughput as in the exhaustive search-based approach with relatively few iterations. Moreover, the proposed algorithm can be effectively applied for different channel models, performance metrics and beamforming schemes. Therefore, our proposed algorithm is suitable for delay-constrained systems and it can be practically implemented in the future. Then, non-ideal PA affects the system performance remarkably, while its effect decreases at high SNR. Finally, the users' severe service outage constraints affect the end-to-end throughput considerably, while its effect decreases at high SNRs. Comparison between different initial access techniques is an interesting future work on which we are working.

ACKNOWLEDGMENT

The research leading to these results received funding from the European Commission H2020 programme under grant agreement $n^{\circ}671650$ (5G PPP mmMAGIC project), and from the Swedish Governmental Agency for Innovation Systems (VINNOVA) within the VINN Excellence Center Chase.

REFERENCES

[1] T. Bai, A. Alkhateeb, and R. W. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 70–77, Sept. 2014.

[2] G. R. MacCartney, J. Zhang, S. Nie, and T. S. Rappaport, "Path loss models for 5G millimeter wave propagation channels in urban microcells," in *Proc. IEEE GLOBECOM'2013, Atlanta, GA, USA*, Dec. 2013, pp. 3948–3953.

[3] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.

[4] S. Sun, G. R. MacCartney, M. K. Samimi, S. Nie, and T. S. Rappaport, "Millimeter wave multi-beam antenna combining for 5G cellular link improvement in new york city," in *Proc. IEEE ICC'2014, Sydney, Australia*, Jun. 2014, pp. 5468–5473.

[5] M. Cudak, A. Ghosh, T. Kovarik, R. Ratasuk, T. A. Thomas, F. W. Vook, and P. Moorut, "Moving towards mmwave-based beyond-4G (B-4G) technology," in *Proc. IEEE VTC'2013, Dresden, Germany*, Jun. 2013, pp. 1–5.

[6] J. P. Gilb, "IEEE standards 802.15. 3c–part 15.3: wireless medium access control (MAC) and physical layer (PHY) specifications for high rate wireless personal area networks (WPANs) amendment 2: millimeter-wave-based alternative physical layer extension [s]," *IEEE Computer Society, New York*, Aug. 2009.

[7] C. Cordeiro *et al.*, "IEEE P802. 11 Wireless LANs, PHY/MAC Complete Proposal Specification (IEEE 802.11-10/0433r2)," May. 2010.

[8] H. Rate, "GHz PHY, MAC and PALs, Standard ECMA-387, ser. Link: https://www.ecma-international.org/publications/files/ECMA-ST/ECMA-387.pdf," 2010.

[9] C. Cordeiro, D. Akhmetov, and M. Park, "IEEE 802.11 ad: introduction and performance evaluation of the first multi-gbps wifi technology," in *Proc. IEEE mmCOM'2010 Chicago, IL, USA*. ACM, Sept. 2010, pp. 3–8.

[10] V. Desai, L. Krzymien, P. Sartori, W. Xiao, A. Soong, and A. Alkhateeb, "Initial beamforming for mmwave communications," in *Proc. IEEE ACSSC'2014, Pacific Grove, CA, USA*, Nov. 2014, pp. 1926–1930.

[11] J. Palacios, D. De Donno, D. Giustiniano, and J. Widmer, "Speeding up mmwave beam training through low-complexity hybrid transceivers," in *Proc. IEEE PIMRC'2016, Valencia, Spain*, Sept. 2016.

[12] M. Giordani, M. Mezzavilla, C. N. Barati, S. Rangan, and M. Zorzi, "Comparative analysis of initial access techniques in 5G mmwave cellular networks," in *2016 Annual Conference on Information Science and Systems (CISS), Princeton University, Princeton, NJ 08544*, Apr. 2016, pp. 268–273.

[13] Y. Qi and M. Nekovee, "Coordinated initial access in millimetre wave standalone networks," *Millimeter-wave Networking Workshop (mmNet 2016)*, May. 2016.

[14] C. N. Barati, S. A. Hosseini, M. Mezzavilla, T. Korakis, S. S. Panwar, S. Rangan, and M. Zorzi, "Initial access in millimeter wave cellular systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 7926–7940, Dec. 2016.

[15] G. Lee, Y. Sung, and J. Seo, "Randomly-directional beamforming in millimeter-wave multiuser MISO downlink," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1086–1100, Sept. 2016.

[16] J. Singh and S. Ramakrishna, "On the feasibility of codebook-based beamforming in millimeter wave systems with multiple antenna arrays," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2670–2683, Jan. 2015.

[17] J. Choi, "Beam selection in mm-wave multiuser MIMO systems using compressive sensing," *IEEE Trans. Commun.*, vol. 63, no. 8, pp. 2936–2947, Jun. 2015.

[18] C. N. Barati, S. A. Hosseini, S. Rangan, P. Liu, T. Korakis, S. S. Panwar, and T. S. Rappaport, "Directional cell discovery in millimeter wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6664–6678, Jul. 2015.

[19] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, March. 2014.

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

[21] V. Raghavan, J. Cezanne, S. Subramanian, A. Sampath, and O. Koymen, "Beamforming tradeoffs for initial UE discovery in millimeter-wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 543–559, Jan. 2016.

[22] J. Qiao, X. Shen, J. W. Mark, and Y. He, "MAC-layer concurrent beamforming protocol for indoor millimeter-wave networks," *IEEE Trans. Veh. Technol*, vol. 64, no. 1, pp. 327–338, Jan. 2015.

[23] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.

[24] B. Li, Z. Zhou, W. Zou, X. Sun, and G. Du, "On the efficient beamforming training for 60GHz wireless personal area networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 504–515, Feb. 2013.

[25] L. Zhou and Y. Ohashi, "Efficient codebook-based MIMO beamforming for millimeter-wave WLANs," in *Proc. IEEE PIMRC'2012*, Sept. 2012, pp. 1885–1889.

[26] H.-H. Lee and Y.-C. Ko, "Low complexity codebook-based beamforming for MIMO-OFDM systems in millimeter-wave WPAN," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3607–3612, Nov. 2011.

[27] B. Makki, T. Svensson, G. Cocco, T. de Cola, and S. Erl, "On the throughput of the return-link multi-beam satellite systems using genetic algorithm-based schedulers," in *Proc. IEEE ICC'2015, London, UK*, Jun. 2015, pp. 838–843.

[28] S. Payami, M. Shariat, M. Ghoraishi, and M. Dianati, "Effective RF codebook design and channel estimation for millimeter wave communication systems," in *Proc. IEEE ICCW'2015, London, UK*, Jun. 2015, pp. 1226–1231.

[29] L. Wan, X. Zhong, Y. Zheng, and S. Mei, "Adaptive codebook for limited feedback MIMO system," in *Proc. IEEE WOCN'2009, Cairo, Egypt*, Apr. 2009, pp. 1–5.

[30] B. Makki, T. Svensson, T. Eriksson, and M.-S. Alouini, "On the required number of antennas in a point-to-point large-but-finite MIMO system: Outage-limited scenario," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1968–1983, May. 2016.

[31] D. Persson, T. Eriksson, and E. G. Larsson, "Amplifier-aware multiple-input single-output capacity," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 913–919, Jan. 2014.