Antenna Array Synthesis and Beamforming for 5G Applications
An Interdisciplinary Approach

Aslan, Y.

DOI
10.4233/uuid:cd3da8fa-fd23-4d27-823c-b34eb99ad07d

Publication date
2020

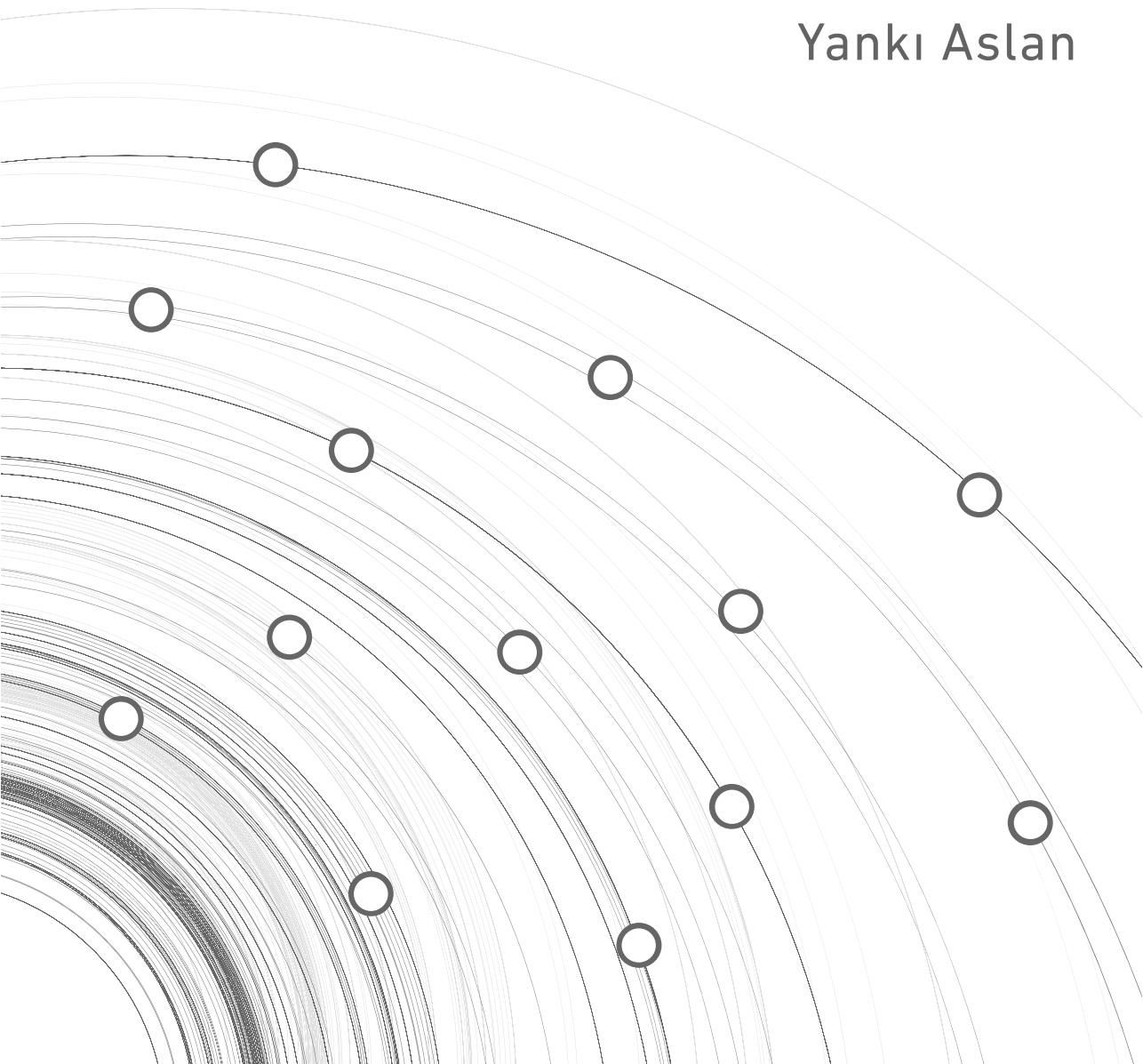
Document Version
Final published version

Citation (APA)
Aslan, Y. (2020). *Antenna Array Synthesis and Beamforming for 5G Applications: An Interdisciplinary Approach*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:cd3da8fa-fd23-4d27-823c-b34eb99ad07d

ANTENNA ARRAY SYNTHESIS AND BEAMFORMING FOR 5G APPLICATIONS

- AN INTERDISCIPLINARY APPROACH -

Yankı Aslan



ANTENNA ARRAY SYNTHESIS AND BEAMFORMING FOR 5G APPLICATIONS

AN INTERDISCIPLINARY APPROACH

ANTENNA ARRAY SYNTHESIS AND BEAMFORMING FOR 5G APPLICATIONS

AN INTERDISCIPLINARY APPROACH

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus, Prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board of Doctorates,
to be defended publicly on
Wednesday 26 August 2020 at 15:00 o'clock

by

Yankı ASLAN

Master of Science in Electrical Engineering,
Delft University of Technology, The Netherlands,
born in Ankara, Turkey.

The dissertation has been approved by the

Promotor: Prof. dr. A.G. Yarovoy

Composition of the doctroral committee:

Rector Magnificus,	Chairman
Prof. dr. A.G. Yarovoy,	Delft University of Technology
Independent members:	
Prof. dr. N. Llombart Juan,	Delft University of Technology
Prof. dr. ir. L.C.N. de Vreede,	Delft University of Technology
Prof. dr. ir. G.J.T. Leus,	Delft University of Technology
Prof. dr. C. Fager,	Chalmers University of Technology, Sweden
Prof. dr. C. Craeye,	Université Catholique de Louvain, Belgium
Other committee members:	
Dr. ir. M. Geurts,	NXP Semiconductors, The Netherlands

This research was supported in part by Netherlands Organisation for Scientific Research (NWO) and in part by NXP Semiconductors N.V. in the framework of the partnership program on Advanced 5G Solutions within the Project 15590 titled “Antenna Topologies and Front-end Configurations for Multiple Beam Generation.”

Keywords: antenna optimization, array synthesis, antenna cooling, beamforming, fifth-generation (5G), integrated arrays, wireless communications.

Printed by: Ipskamp Drukkers B.V., Enschede, Netherlands.

Front & Back: An abstract technology illustration by the artist Majcot, ID: 151625843, downloaded from Shutterstock under the standard license.

ISBN 978-94-6366-297-0

Author e-mail: yankiaslann@gmail.com

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

The DOI of the dataset accompanying the thesis is:
10.4121/uuid:ab0a8412-a3aa-4406-8789-2d7485d3c2d9.

To my family, for their endless love, support, encouragement and sacrifice

CONTENTS

SUMMARY

Realization of the future 5G systems requires the design of novel mm-wave base station antenna systems that are capable of generating multiple beams with low mutual interference, while serving multiple users simultaneously using the same frequency band. Besides, small wavelengths and high packaging densities of front-ends lead to overheating of such systems, which prevents safe and reliable operation. Since the strict cost and energy requirements of the first phase 5G systems favor the use of low complexity beamforming architectures, computationally efficient signal processing techniques and fully passive cooling strategies, it is a major challenge for the antenna community to design multibeam antenna topologies and front-ends with enhanced spatial multiplexing, limited inter-beam interference, acceptable implementation complexity, suitable processing burden and natural-only/radiative cooling.

Traditionally, array design has been performed based on satisfying the given criteria solely on the radiation patterns (gain, side lobe level (SLL), beamwidth etc.). However, in addition to the electromagnetic aspects, multibeam antenna synthesis and performance evaluation in 5G systems at mm-waves must combine different disciplines, including but not limited to, signal processing, front-end circuitry design, thermal management, channel & propagation and medium access control aspects.

Considering the interdisciplinary nature of the problem, the main objective of this research is to develop, evaluate and verify innovative multibeam array techniques and solutions for 5G base station antennas, not yet used nor proposed for mobile communications. The research topics include the investigation of (i) new array topologies, compatible with IC passive cooling, including sparse, space tapered arrays and optimised subarrays, meeting key requirements of 3-D multi-user coverage with frequency re-use and power-efficient side-lobe control, (ii) adaptive multiple beam forming strategies and digital signal processing algorithms, tailored to these new topologies, and (iii) low-cost/competitive and sufficiently generic implementation of the above array topologies and multi beam generation concepts to serve multiple users with the same antenna(s) with best spectrum and power efficiencies.

This doctoral thesis consists of three parts. Part I focuses on the system-driven aspects which cover the system modeling (including the link budget and precoding), propagation in mm-wave channels and statistical assessment of the Quality of Service (QoS). Although separate comprehensive studies exist both in the field of propagation/system modeling and antennas/beamforming, the link between the two disciplines is still weak. In this part, the aim of the study is to bridge the gap between the two domains and to identify the trade-offs between the complexity of beamforming, the QoS and the computational cost of precoding in the 5G multibeam base station arrays for various use cases. Based on the system model developed, a novel quantitative relation between the antenna SLLs/pattern nulls and the statistical QoS is established in a line-of-sight (LoS) dominated mm-wave propagation scenario. Moreover, the potential of using smart (low

in-sector side-lobe) array layouts (with simple beam steering) in obtaining sufficiently high and robust QoS, while achieving the optimally low processing costs is highlighted. For a possible pure non-line-of-sight (NLoS) scenario, the system advantages (in terms of the beamforming complexity and the interference level) of creating a single, directive beam towards the strongest multipath component of a user are explained via ray-tracing based propagation simulations. The insightful system observations from Part I lead to several fundamental research questions: Could we simplify the multiple beamforming architecture while keeping a satisfying QoS? Are there any efficient yet effective alternative interference suppression methods to further improve the QoS? How should we deal with the large heat generation at the base station? These questions, together with the research objectives, form the basis for the studies performed in the remaining parts.

Part II of the thesis focuses on the electromagnetism-driven aspects which include innovative, low-complexity subarray based multibeam architectures and new array optimization strategies for effective SLL suppression. The currently proposed multibeam 5G base stations in the literature for beamforming complexity reduction use either a hybrid array of phased subarrays, which limits the field-of-view significantly, or employ a fully-connected analog structure, which increases the hardware requirements remarkably. Therefore, in the first half of this part, the aim is to design low-complexity hybrid (or hybrid-like) multiple beamforming topologies with a wide angular coverage. For this purpose, two new subarray based multiple beamforming concepts are proposed: (i) a hybrid array of active multiport subarrays with several digitally controlled Butler Matrix beams and (ii) an array of cosecant subarrays with a fixed cosecant shaped beam in elevation and digital beam forming in azimuth. Using the active (but not phased) multiport subarrays, the angular sector coverage is widened as compared to that of a hybrid array of phased subarrays, the system complexity is decreased as compared to that of a hybrid structure with a fully-connected analog network, and the effort in digital signal processing is reduced greatly. The cosecant subarray beamforming, on the other hand, is shown to be extremely efficient in serving multiple simultaneous co-frequency users in the case of a fairness-motivated LoS communication thanks to its low complexity and power equalization capability. Another critical issue with the currently proposed 5G antennas is the large inter-user interference caused by the high average SLL of the regular, periodic arrays. Therefore, in the second half of Part II, the aim is to develop computationally and power-efficient SLL suppression techniques that are compatible with the 5G's multibeam nature in a wide angular sector. To achieve this, two novel techniques (based on iterative parameter perturbations) are proposed: (i) a phase-only control technique and (ii) a position-only control technique. The phase-only technique provides peak SLL minimization and simultaneous pattern nulling, which is more effective than the available phase tapering methods in the literature. The position-only technique, on the other hand, yields uniform-amplitude, (fully-aperiodic and quasi-modular) irregular planar phased arrays with simultaneous multibeam optimization. The latter technique combines interference-awareness (via multibeam SLL minimization in a predefined cell sector) and thermal-awareness (via uniform amplitudes and minimum element spacing constraint) for the first time in an efficient and easy-to-solve optimization algorithm.

Part III of the thesis concentrates on the thermal-driven aspects which cover the ther-

mal system modeling of electronics, passive cooling at the base stations and the role of antenna researchers in array cooling. The major aim here is to form a novel connection between the antenna system design and thermal management, which is not yet widely discussed in the literature. In this part, an efficient thermal system model is developed to perform the thermal simulations. To effectively address the challenge of thermal management at the base stations, fanless CPU heatsinks are exploited for the first time for fully-passive and low-cost cooling of the active integrated antennas. To reduce the size of the heatsinks and ease the thermal problem, novel planar antenna design methodologies are also proposed. In the case of having a low thermal conductivity board, using a sparse irregular antenna array with a large inter-element spacing (such as a sunflower array) is suggested. Alternatively, for the densely packed arrays, increasing the equivalent substrate conductivity by using thick ground planes and simultaneously enlarging the substrate dimensions is proven to be useful.

The performed research presents the first ever irregular/sparse and subarray based antennas with wide scan multibeam capability, low temperature, high efficiency power amplifiers and low level of side lobes. The developed antenna arrays and beam generation concepts could have also an impact over a broad range of applications where they should help overcome the capacity problem by use of multiple adaptive antennas, improve reliability and reduce interference.

SAMENVATTING

Realisatie van toekomstige 5G-systemen vereist het ontwerp van nieuwe mm-golf basisstationantennesystemen die in staat zijn om meerdere bundels te genereren met lage onderlinge interferentie, terwijl ze tegelijkertijd meerdere gebruikers bedienen met dezelfde frequentieband. Bovendien leiden kleine golflengten en hoge verpakkingsdichtheden van front-ends tot oververhitting van dergelijke systemen, wat een veilige en betrouwbare werking verhindert. Aangezien de strikte kosten- en energie-eisen van 5G-systemen van de eerste fase het gebruik van beamforming-architecturen met lage complexiteit, computationeel efficiënte signaalverwerkingstechnieken en volledig passieve koelstrategieën bevorderen, is het een grote uitdaging voor de antennegemeenschap om multibeam-antennetopologieën en front- end met verbeterde ruimtelijke multiplexing, beperkte interbundelinterferentie, acceptabele implementatiecomplexiteit, geschikte verwerkingslast en natuurlijke / stralingskoeling te ontwikkelen.

Traditioneel is het array-ontwerp uitgevoerd op basis van het voldoen aan de gegeven criteria, waarbij de stralingspatronen (versterking, zijlobbeniveau (SLL), bundelbreedte enz.) zijn uitgesloten. Naast de elektromagnetische aspecten moeten multibeam antennesynthese en prestatie-evaluatie in 5G-systemen bij mm-golven verschillende disciplines combineren, waaronder niet alleen signaalverwerking maar ook front-end circuitontwerp, thermisch beheer, kanaal & propagatie en medium toegangscontrole aspecten.

Gezien de interdisciplinaire aard van het probleem, is het hoofddoel van dit onderzoek het ontwikkelen, evalueren en verifiëren van innovatieve multibeam arraytechnieken en -oplossingen voor 5G-basisstationantennes, die nog niet worden gebruikt of voorgesteld voor mobiele communicatie. De onderzoeksthema's omvatten het onderzoek van (i) nieuwe array-topologieën, compatibel met passieve IC-koeling, waaronder spaarzame, taps toelopende arrays en geoptimaliseerde subarrays, die voldoen aan de belangrijkste vereisten van 3D-dekking voor meerdere gebruikers met hergebruik van frequenties en energiezuinig zijlobbenbesturing, (ii) adaptieve strategieën voor het vormen van meerdere bundels en algoritmen voor digitale signaalverwerking, afgestemd op deze nieuwe topologieën, en (iii) goedkope / concurrerende en voldoende generieke implementatie van de bovenstaande array-topologieën en concepten voor het genereren van meerdere bundels om meerdere gebruikers te bedienen met dezelfde antenne (s) met de beste spectrum- en energie-efficiëntie.

Dit proefschrift bestaat uit drie delen. Deel I richt zich op de systeemgestuurde aspecten die betrekking hebben op de systeemmodellering (inclusief het linkbudget en precodering), propagatie in mm-golfkanalen en statistische beoordeling van de Quality of Service (QoS). Hoewel er afzonderlijke uitgebreide studies bestaan, zowel op het gebied van propagatie / systeemmodellering als antennes / beamforming, is de link tussen de twee disciplines nog steeds zwak. In dit deel is het doel van de studie om de kloof tussen de twee domeinen te overbruggen en de afwegingen te identificeren tussen de complexiteit van beamforming, de QoS en de computationele kosten van pre-

codering in de 5G-multibeam-basisstationarrays voor verschillende toepassingen . Op basis van het ontwikkelde systeemmodel wordt een nieuwe kwantitatieve relatie tussen de antenne SLL's / patroonnullen en de statistische QoS vastgesteld in een door het gezichtsveld (LoS) gedomineerd mm-golfvoortplantingsscenario. Bovendien wordt het gebruik van potentieel slimme (lage in-sector zijlobben) array-indelingen (met eenvoudige straalbesturing) om voldoende hoge en robuuste QoS te verkrijgen bij optimaal lage verwerkingskosten benadrukt. Voor een mogelijk puur non-line-of-sight (NLoS) scenario worden de systeemvoordelen (in termen van de bundelvormende complexiteit en het interferentieniveau) van het creëren van een enkele, richtende bundel naar de sterkste multipadcomponent van een gebruiker uitgelegd via op stralen-theorie gebaseerde simulaties. De inzichtelijke systeemwaarnemingen uit deel I leiden tot verschillende fundamentele onderzoeksvragen: kunnen we de meervoudige beamforming-architectuur vereenvoudigen met behoud van een bevredigende QoS? Zijn er efficiënte maar effectieve alternatieve methoden voor het onderdrukken van storingen om de QoS verder te verbeteren? Hoe gaan we om met de grote warmteontwikkeling bij het basisstation? Deze vragen vormen samen met de onderzoeksdoelstellingen de basis voor de onderzoeken die in de overige delen zijn uitgevoerd.

Deel II van het proefschrift richt zich op de door elektromagnetisme aangestuurde aspecten, waaronder innovatieve, op subarray gebaseerde multibeam-architecturen met een lage complexiteit en nieuwe array-optimalisatiestrategieën voor effectieve SLL onderdrukking. De momenteel voorgestelde multibeam 5G-basisstations in de literatuur voor het verminderen van de complexiteit van de bundelvorming gebruiken ofwel een hybride array van gefaseerde subarrays, die het gezichtsveld aanzienlijk beperken, of gebruiken een volledig verbonden analoge structuur, wat de hardwarevereisten opmerkelijk verhoogt. Daarom is het in de eerste helft van dit deel de bedoeling om hybride (of hybride-achtige) meervoudige bundelvormende topologieën met een lage complexiteit te ontwerpen met een brede hoekdekking. Voor dit doel worden twee nieuwe op een subarray gebaseerde concepten voor meervoudige bundelvorming voorgesteld: (i) een hybride array van actieve multipoort-subarrays met verschillende digitaal gestuurde Butler Matrix-bundels en (ii) een array van cosecant-subarrays met een vaste cosecantvormige balk in de hoogte en digitaal straalvorming in azimut. Door gebruik te maken van de actieve (maar niet gefaseerde) multipoort subarrays, wordt de dekking van de hoeksector verbreed in vergelijking met die van een hybride reeks gefaseerde subarrays, wordt de systeemcomplexiteit verminderd in vergelijking met die van een hybride structuur met een volledig verbonden analoog netwerk, waarbij de belasting voor de digitale signaalverwerking sterk wordt sterk verminderd. De cosecant subarray beamforming daarentegen blijkt buitengewoon efficiënt te zijn in het bedienen van meerdere gelijktijdige co-frequentiegebruikers in het geval van een fairness-gemotiveerde LoS-communicatie dankzij zijn lage complexiteit en vermogensegalisatiemogelijkheid. Een ander kritiek probleem met de momenteel voorgestelde 5G-antennes is de grote intergebruikersinterferentie veroorzaakt door de hoge gemiddelde SLL van de reguliere, periodieke arrays. Daarom is het de bedoeling in de tweede helft van deel II computationele en energiezuinige SLL-onderdrukkingstechnieken te ontwikkelen die compatibel zijn met de multibeam-aard van de 5G in een brede hoeksector. Om dit te bereiken, worden twee nieuwe technieken voorgesteld (gebaseerd op iteratieve parameterversto-

ringen): (i) een phase-only controletechniek en (ii) een position-only controletechniek. De phase-only techniek biedt piek SLL-minimalisatie en gelijktijdige patroonnulling, wat effectiever is dan de beschikbare fasevernauwingsmethoden in de literatuur. De position-only techniek levert daarentegen uniforme amplitude, (volledig aperiodieke en quasi-modulaire) onregelmatige planair gefaseerde arrays met gelijktijdige multibeam optimalisatie. De laatste techniek combineert voor het eerst interferentie bewustzijn (via multibeam SLL-minimalisatie in een vooraf gedefinieerde celsector) en thermisch bewustzijn (via uniforme amplitudes en minimale beperking van de afstand tussen elementen) in een efficiënt en eenvoudig op te lossen optimalisatie-algoritme.

Deel III van het proefschrift concentreert zich op de thermisch aangedreven aspecten die betrekking hebben op de thermische systeemmodellering van elektronica, passieve koeling bij de basisstations en de rol van antenne-onderzoekers bij arraykoeling. Het belangrijkste doel hier is om een nieuwe verbinding te vormen tussen het ontwerp van het antennesysteem en het thermisch beheer, dat nog niet algemeen wordt besproken in de literatuur. In dit deel wordt een efficiënt thermisch systeemmodel ontwikkeld om de thermische simulaties uit te voeren. Om de uitdaging van thermisch beheer bij de basisstations effectief aan te pakken, worden ventilatorloze CPU-heatsinks voor het eerst benut voor volledig passieve en goedkope koeling van de actieve geïntegreerde antennes. Om de afmetingen van de heatsinks te verkleinen en het thermische probleem te verminderen, worden ook nieuwe planaire antenne-ontwerpmethodologieën voorgesteld. In het geval van een plaat met een lage thermische geleidbaarheid wordt het gebruik van een spaarzame onregelmatige antenne-array met een grote onderlinge afstand tussen elementen (zoals een sunflower-array) voorgesteld. Als alternatief voor de dicht opeengepakte reeksen is bewezen dat het nuttig is om de equivalente geleidbaarheid van het substraat te verhogen door dikke grondvlakken te gebruiken en tegelijkertijd de afmetingen van het substraat te vergroten.

Het uitgevoerde onderzoek presenteert de eerste onregelmatige / sparse en op subarray gebaseerde antennes met wide scan multibeam-mogelijkheden, lage temperatuur, zeer efficiënte eindversterkers en een laag niveau van zijlobben. De ontwikkelde antennearrays en straalgeneratieconcepten zouden ook een impact kunnen hebben op een breed scala aan toepassingen waar ze zouden moeten helpen het capaciteitsprobleem op te lossen door gebruik te maken van meerdere adaptieve antennes, de betrouwbaarheid te verbeteren en interferentie te verminderen.

1

INTRODUCTION

When wireless is perfectly applied, the whole earth will be converted into a huge brain, which in fact it is, all things being particles of a real and rhythmic whole.

Nikola Tesla

1

We live in an exciting time for the wireless industry, the decade of 5G! Almost every single day, we hear the unveiling of new 5G products, press releases of new field trials and announcements of the upcoming commercial 5G rollouts.

Starting a new business is exciting, but it is also intimidating. It is a brand new trip that takes you into the unknown. As this is the case, the introduction of a new generation communication system brings up many fundamental questions: What is 5G? What will it enable us to do? What will it mean for the society?

In general, the definition of 5G communication is given considering two different views: the hyper-connected vision and the next generation radio access technology. In the hyper-connected view, 5G is seen as a blend of the existing 2G, 3G and 4G technologies achieving far greater coverage and always-on reliability with higher network density in terms of cells and devices. The next-generation radio access technology view, on the other hand, is more traditional in the sense that specific targets for data rates and latency are identified and new technologies proposed for 5G are tested against these criteria. Grouping the aims of both perspectives, a set of expectations from 5G are listed in [1]:

- "1-10 Gbps connections to end points in the field (i.e. not theoretical maximum)"
- "10-100x typical user data rate"
- "1000x mobile data volume per area"
- "1 millisecond (ms) end-to-end round trip delay (latency)"
- "10-100x number of connected devices"
- "(Perception of) 100% availability"
- "(Perception of) 100% coverage"
- "90% reduction in network energy usage"
- "10x battery life for low power devices"

Among these requirements, only two of them are considered as generation-defining aspects that would be unique to 5G: sub-1ms latency and >1 Gbps downlink speed.

The wide-spectrum mobile network services of 5G are classified into three main categories by the International Telecommunication Union (ITU) [2], as shown in Fig. 1.1: *Enhanced Mobile Broadband (eMBB), Ultra-reliable and Low-latency Communications (uRLLC)*, and *Massive Machine Type Communications (mMTC)*.

Among the three categories, *eMBB* studies the increasing demands on the digital lifestyle of individuals for applications such as high definition videos, virtual reality and augmented reality, which require large bandwidth. *uRLLC* targets to satisfy the needs of the future digital industry for applications including assisted/autonomous driving or remote management, for which providing low-latency services is essential, and *mMTC* focuses on the demands of a digital society by considering the IoT services like smart cities that require high connection density.

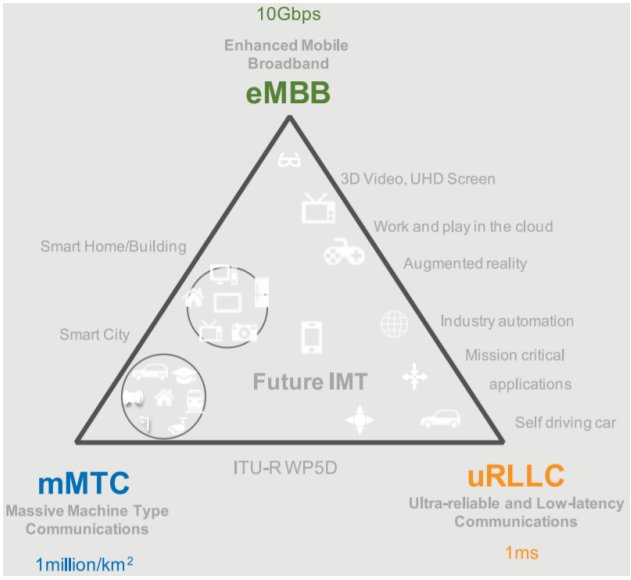


Figure 1.1: 5G mobile network services according to the ITU [2].

Taking into account the listed potential 5G services together with the foreseen challenges and needs of the "information" society in the beyond 2020 timeframe, twelve different use case scenarios are defined within the EU project METIS [3]. Fig. 1.2 presents these use cases (stadiums, open air festivals, shopping malls, highways and so on) whose requirements vary in terms of the throughput, capacity, latency and coverage.

One of the most numerous and costly elements of the communication networks supporting the above-mentioned applications are the base stations. The base stations also have a crucial effect on the end-to-end system performance since, besides the transceiver electronics (amplifiers, duplexers, converters, filters etc.), they include the antenna arrays that are responsible for the beamforming.

To achieve the capacity and throughput requirements of 5G technology, a vast amount of spectrum is necessary. Hence, recently, in addition to the conventional (below 6 GHz)

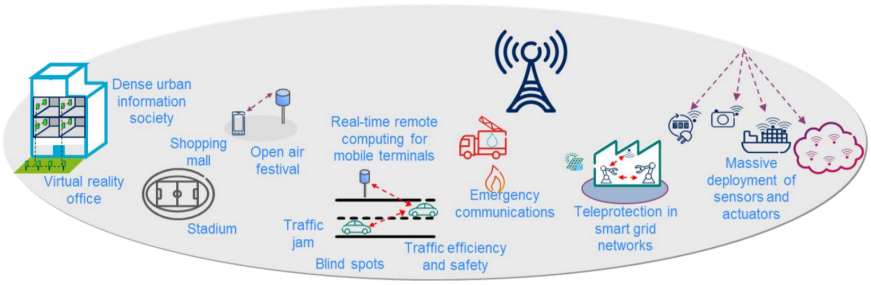


Figure 1.2: Twelve 5G use cases identified within the METIS project [3].

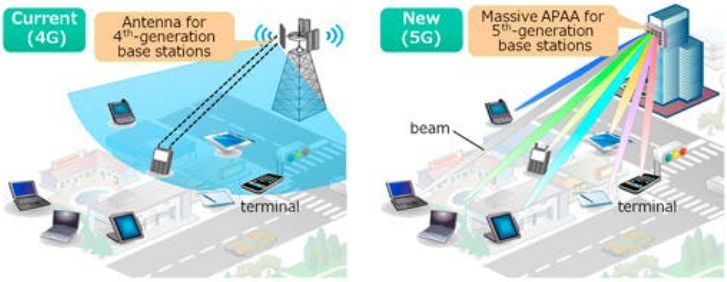


Figure 1.3: Beamforming in 4G vs 5G [4].

frequency bands, millimeter-wave (mm-wave) systems have been aimed for the implementation of 5G. In the antenna engineering community, this brings up the question: What is the suitable beamforming strategy in mm-wave 5G base stations to meet the demanding throughput expectations with low cost?

The fundamental difference in beamforming between the 4G and 5G base stations is as visualized in Fig. 1.3. The base station antennas for the current 4G communication systems form a wide beam with a large angular coverage. Different user terminals within this wide beam are served either at different frequencies (Frequency Division Multiple Access, FDMA) or at different times (Time Division Multiple Access, TDMA). Another alternative is to assign unique orthogonal codes to each user (Code Division Multiple Access, CDMA). All these techniques require additional resources which deteriorate the system optimality, and thus, the system capacity. Therefore, what is desired from the base station arrays in 5G is to efficiently create multiple simultaneous, user-oriented, narrow, co-frequency beams with sufficiently high gain and low inter-beam interference [4]. This is the concept of Space (or Beam) Division Multiple Access, SDMA (or BDMA).

1.1. 5G BASE STATIONS: REQUIREMENTS, CHALLENGES AND DEPLOYMENT SCENARIOS

REQUIREMENTS & CHALLENGES

As highlighted, 5G communication systems are expected to achieve approximately 1000x communication capacity growth and less than 1 ms latency in transferred data stream while supporting massive Internet of Things [5; 6]. Besides, being high volume markets, 5G, or in general mobile communication, systems have strict cost requirements. Therefore, next generation systems demand satisfying the ambitious performance criteria by developing novel and advanced multi-beam antenna arrays integrated with low cost front-ends [7]. By serving users with separate and dedicated antenna beams simultaneously, the aim is to increase the frequency spectrum re-use and thus, the capacity. Since such a system is more interference limited rather than noise limited, sufficient mitigation of the inter-user (or inter-beam) interference becomes very critical.

A major challenge at mm-waves[1] is to establish a reliable 5G system model and a link budget that consider the properties of the transmitters (frequency, bandwidth, transmitted power, antenna gain etc.), receivers (antenna gain, noise figure etc.) & propagation channels (distance, angle, propagation loss etc.) and that can be used to determine requirements or goal functions to be used in the antenna array synthesis.

It is worth to note that the transmissions in the mm-wave bands have significantly less favorable link budgets due to reduced receiving effective aperture, decreased diffraction/dispersion effects and low efficiency of the beamformer ICs [9]. Due to the highly inefficient electronics, there is a large amount of heat generation at the base station antennas, which makes cooling a serious challenge [10]. Ideally, the excess heat should be removed passively via natural convection and heat radiation without using fans or water pumps because of the cost, energy consumption, reliability and maintenance issues. The power amplifier and processing consumption and inefficiency also lead to huge associated electricity costs.

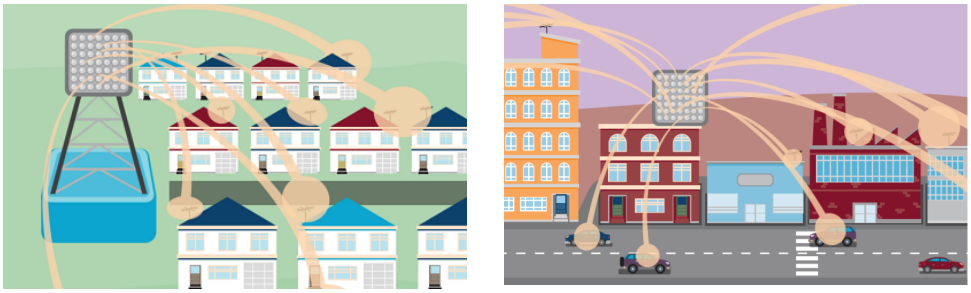
To mitigate the propagation drawbacks at mm-waves and reduce the electricity consumption (with the help of the high antenna gain and low side lobes), large scale antenna arrays with tens or maybe hundreds of elements are expected to be deployed in 5G systems. Using a transceiver behind every antenna element will consume unacceptable amounts of power and will likely be cost prohibitive. This makes RF-oriented (analog) or hybrid (RF and baseband) approaches more attractive compared to the digital beam forming [11]. Yet, the choice of multiple beam generation architecture has also a large impact on performance (spectrum re-use, capacity, quality of service, power consumption), compatibility with non-line of sight and/or MIMO operation, cost and flexibility of the end product. Besides, it has to be tailored to the preferred array and sub-array topologies.

In the longer term, it is probable that fully-digital multiple beam forming will prevail, as it already does now for some space-based mobile communication systems. For instance, in space based communications, after early studies on multi beam arrays [12–14], beam hopping antennas with high performance have been developed in the early 2000's for satellite communications at 20/30 GHz by Boeing Space and Defense for their Spaceway Satellite Program [15]. However, currently proposed satellite antenna solutions have drawbacks in terms of the system cost, complexity (regarding the beamforming architectures, signal processing, cooling etc.), and they cannot handle the 5G's wider field of view as well as its thermal problem without a deeper theoretical revisit of different array topologies & beamforming strategies.

DEPLOYMENT SCENARIOS

When a new wireless technology is developed, it is also very critical to know how it is going to be deployed so as to derive the specific system requirements. In Fig. 1.4, two of the commonly explored 28 and 39 GHz band line-of-sight scenarios today are visualized.

[1] It is reported that the prime spectrum for 5G is set between 24 GHz and 57 GHz considering both the regulation and technology perspectives [8]. Among all the candidate mm-wave frequency bands between 24-57 GHz, currently, 28 GHz and 39 GHz bands are in use or being considered as the primary targets because of the relatively good solid state device capabilities at such relatively low mm-wave frequencies, so that around 10-20 dBm linear output power and more than 10% efficiency can be obtained per power amplifier (PA) by using GaAS, SiGe or even CMOS technology [8].



(a) (b)

Figure 1.4: 5G mm-wave base station deployment scenarios [16]: (a) fixed wireless access in suburban environments, (b) fixed and mobile deployment in dense urban environments.

Fig. 1.4(a) illustrates the Fixed Wireless Access (FWA) suburban environment case where the base station is mounted on a utility pole or a tower and pointing down to deliver high bandwidth data to houses. With fixed users, the antenna system does not require a large vertical scanning range, yet, the required transmit power may be quite high, depending on the cell radius. Fig. 1.4(b), on the other hand, illustrates the dense urban case where the base station is mounted lower to the ground on a rooftop or a lamppost and deliver signals both across the buildings and to the mobile or nomadic ground users. In such a scenario, the base station will require more flexibility in the vertical beam steering, but the transmit power needed may not be so high.

The main motivation of the current 5G research is to find out low-cost solutions in different deployment scenarios that can address all the above-mentioned system challenges (i.e. enhanced spatial multiplexing, acceptable power consumption & processing speed, limited interference, simple passive cooling).

1.2. RESEARCH OBJECTIVES

The objectives of this research are to develop, evaluate and verify for 5G base station antennas innovative multi beam array techniques and solutions, not yet used nor proposed for mobile communications and including:

1) New array topologies, compatible with IC passive cooling, including sparse, space tapered arrays and optimised sub-arrays, meeting key requirements of 3-D multi-user coverage with frequency re-use and power-efficient side-lobe control.

2) Adaptive multiple beam forming strategies and digital signal processing algorithms, tailored to these new topologies.

3) Low-cost/competitive and sufficiently generic implementation[2] of the above array topologies and multi beam generation concepts to serve multiple users with the same antenna(s) with best spectrum and power efficiencies.

[2]The most generic implementation with a single type of base station is what is desired by the chip producers. Yet, such a requirement results in a significant performance loss in different scenarios. A sufficient generalization can be achieved by having a few base station coverage specific antenna topologies (open area, open roof, street canyon etc.) and using the appropriate ones.

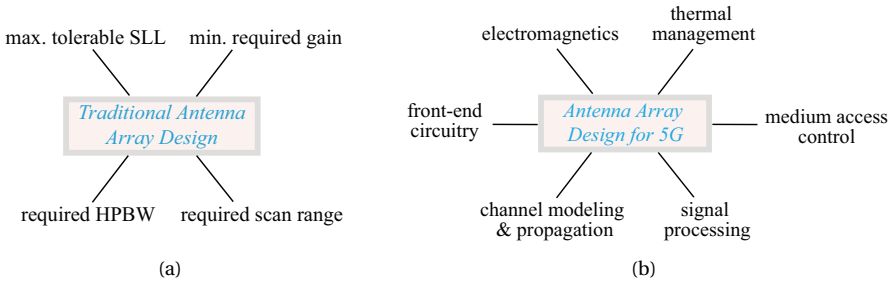


Figure 1.5: Design methodologies for antenna arrays: (a) traditional approach, (b) 5G perspective.

1.3. RESEARCH APPROACH

Traditionally, array design has been performed based on satisfying the given criteria solely on the radiation patterns (gain, SLLs, beamwidth etc.). However, in addition to the electromagnetic aspects, multi-beam antenna synthesis and performance evaluation in 5G SDMA systems at mm-waves must combine different disciplines, including but not limited to, signal processing, front-end circuitry design, thermal management, channel & propagation and medium access control (MAC) aspects (see Fig. 1.5). The reasons why the traditional approach is not sufficient any more and why the listed complementary research domains must be considered in 5G antenna system design are explained in Table 1.1.

Among the listed study areas, the signal processing aspects consist of the study on linear, nonlinear, narrowband, wideband, switched or adaptive beamforming algorithms. In this field, it is desired to achieve low computational complexity and high processing speed. The front-end circuitry and thermal management aspects are related to the power output, linearity & efficiency of the transceiver chips and the heat dissipation capabilities. It is important to keep the amplifiers operating at the same, efficient working point and ensure sufficient cooling of the beamforming chips at the integrated antennas. The channel modeling & propagation aspects deal with deciding on the propagation parameters and an appropriate link budget depending on the particular environment and beamforming strategy. Lastly, the MAC layer aspects cover the design of the medium access periods and rules when setting up an efficient link between the base station and users.

Although separate comprehensive studies exist in each field, the concept of interdisciplinary work is relatively new in the antenna engineering community. To the best of the authors' knowledge, there is no prior work that jointly considers all the above-mentioned complementary research domains simultaneously. This thesis, on the other hand, approaches the problem by combining different disciplines and aims to highlight some of the system design trade-offs that will be faced when designing a complete 5G antenna array system. The main focus of the thesis is on getting generic system solutions for 5G, rather than the design of a particular array.

A summary of the thesis content with respect to each complementary research domain is given in Table 1.2.

Table 1.1: Motivations behind the interdisciplinary 5G study

Complementary Research Domain	Motivation
Channel modeling & propagation	The requirement that a narrow beam has to be pointed to a user (especially in NLOS scenarios) and has zeroes towards other users, while a wide beam will always cover the user
Signal processing	Large amount of control parameters in arrays with limited processing power and speed
Front-end circuitry	Compact system integration challenges, component (and cost) reduction requirements
Thermal management	Low efficiency of PAs, high spatial density of the heat sources, much higher heat generation than the conventional systems
Medium access control (MAC)	The need to manage multiple data streams to simultaneous users, where each user receives a dedicated co-frequency beam

Table 1.2: Thesis content with respect to the research topics

Research Topic	Thesis Content
Electromagnetics	Multibeam architectures, high gain, low sidelobe antennas
Channel modeling & propagation	LoS SDMA system model & mm-wave link budget in different BS deployment scenarios, impact of multipath on the system performance
Signal processing	Several beamforming algorithms with/without complexity reduction
Front-end circuitry	Low-complexity beamforming architectures, power-efficient array synthesis
Thermal management	Thermal system modeling, passive cooling methods
Medium access control (MAC)	User selection & scheduling approaches

1.4. NOVELTIES AND MAIN RESULTS

This thesis proposes a new multidisciplinary metaparadigm for the antenna array synthesis and beamforming for future wireless communications. In particular, aperiodic arrays have been introduced as an innovative approach to 5G mm-wave systems, and novel methods for irregular / sparse array topology optimization, beamforming and thermal management have been investigated.

The array topology problem is formulated as a system-based optimization procedure for the first time, and a new system-driven array synthesis approach is proposed. The presented approach is superior to the existing techniques in terms of 5G compatibility since it (simultaneously) achieves: (i) remarkable SLL suppression in wide angle scanning, (ii) uniform-amplitude excitations with optimum power efficiency, (iii) sparsity for extra space for electronics and cooling, (iv) modularity for design and fabrication simplicity. As a result, when compared to a referenced fully populated equally spaced array in a rectangular topology, the proposed aperiodic sparse arrays can improve the total statistical SINR within the same scanning area and decrease the maximal temperature of the beamforming chips.

Furthermore, pursuing the given inter-disciplinary approach (shown in Table 1.2) and addressing the aforementioned research objectives (listed in Section 1.2), the system demands of the 5G base station antennas are derived. During this study, the following main original methods are proposed:

- Subarray based hybrid beamforming architectures (an array of active multiport subarrays, an array of cosecant subarrays) for beamforming complexity reduction with a superior radiation performance than the existing hybrid configurations,
- Effective inter-user interference suppression via low-complexity and power-efficient optimization techniques based on phase-only and position-only iterative perturbation algorithms,
- Efficient thermal system performance evaluation of the chip-integrated antenna arrays via compact thermal models,
- The exploitation of CPU heatsinks in the passive-only thermal management of 5G arrays,
- The use of antenna arrays (sparse topologies and thick ground plane designs) for passive cooling extension, and thus heatsink size reduction.

The above-mentioned novel ideas presented in this thesis lead to the following main outcomes:

- A multi-user LoS 5G SDMA system model has been presented to investigate, for the first time, the trade-offs between the statistical quality-of-service, computational burden of precoding and beamforming complexity for various use cases and number of users with different beamforming algorithms and array layouts. Based on the model, a novel quantitative relation between the antenna side lobe levels (SLLs)/pattern nulls and the quality of service (QoS) is established. Considering the comparative analysis, smart array layouts with low in-sector average SLLs are suggested to be used with simple beam steering in order to achieve sufficiently high and robust QoS while achieving the optimally low processing costs.
- Considering a more realistic mm-wave multipath propagation, the effect of forming a single-lobe beam towards the strongest multipath direction and forming a multi-lobe beam towards the few strongest multipaths is studied for the first time in urban outdoor environments. It is observed that, optimally, one could use multiple simultaneous single lobed beams and ports per user and recombine signals coherently after time synchronization of each. That should provide the optimal gain, but it limits greatly the same frequency re-use by others and increases the beamforming complexity. Therefore, if it satisfies the minimum SNR requirement, it is preferable to form a single beam with one lobe to the strongest ray.
- For beamforming complexity reduction, two new subarray based multiple beamforming concepts have been proposed that are based on (i) a hybrid array of active multiport subarrays with several digitally controlled Butler Matrix beams covering

a wide angular sector and (ii) an array of cosecant subarrays with a fixed cosecant shaped beam in elevation and digital beam forming in azimuth. Using the active multiport subarrays, the angular sector coverage is widened as compared to that of a hybrid array of phased subarrays, and the system complexity is decreased as compared to that of a hybrid structure with a fully-connected analog network. The cosecant subarray beamforming, on the other hand, is shown to be efficient in serving multiple SDMA users in the case of a fairness-motivated LoS communication thanks to its low complexity and power equalization capability.

- A new phase-only peak SLL minimization and simultaneous pattern nulling algorithm has been proposed for low-cost and low-complexity inter-user interference mitigation. The proposed approach outperforms the existing methods when radiation pattern nulling in certain sectors is required and achieves at least comparative results (in terms of the maximal SLL, beamwidth and tapering efficiency) in cases without nulling.

- An extended-feature, 5G SDMA system-driven convex algorithm has also been presented for the synthesis of uniform-amplitude, (fully-aperiodic & quasi-modular) irregular planar phased arrays with simultaneous multi-beam optimization. The presented technique is the first-ever optimization scheme with the following joint capabilities required for the 5G antennas: (i) minimized maximum SLL everywhere in the field of view (or within a sector) for a beam scanned freely inside a given cell sector, which creates the interference-awareness and with (ii) uniform excitation amplitudes for optimum power efficiency, (iii) a restricted aperture size for direct integration with the existing passive cooling systems, (iv) a predefined minimum inter-element spacing to prevent element overlapping and ease the thermal problem caused by the temperature increase due to the too-close elements, which creates the thermal-awareness.

- To effectively address the challenge of thermal management at the base stations, fanless CPU coolers have been exploited for the first time for fully-passive and low-cost cooling of the active integrated antennas. An efficient thermal system model has been developed to perform the thermal simulations and the simulation results have been verified by the experiments. To the best of the author's knowledge, this is the first study that quantifies the thermal properties of 5G antennas with a compact, fast and easy-to-solve model.

- To reduce the size of the heatsinks and ease the thermal problem, novel planar antenna design methodologies have been proposed for a better dissipation of the heat flowing towards the electromagnetic radiators. If the thermal conductivity of the board is small, sparse antenna arrays with large inter-element spacing are used, which yields up to 15% reduction in the maximum temperature of the ICs. Alternatively, the arrays can be kept dense while increasing the substrate conductivity by using thick ground planes and enlarging the substrate dimensions simultaneously. Such an approach provides up to 30% reduction in the maximum IC temperature.

In summary, based on the topologies and beam generation strategies resulting from the research contents, **the first ever irregular / sparse and sub-array based antennas with wide scan multi-beam capability, low temperature, high efficiency power amplifiers and low level of side lobes have been designed**. The developed antenna arrays and beam generation concepts could have also an impact over a broad range of applications where they should help overcome the capacity problem by use of multiple adaptive antennas, improve reliability, reduce interference and mitigate fading problems. Such applications could be in wireless sensing networks and in Machine-to-Machine communication. Other application domains could be automotive radar or defence and security, where intelligent antenna systems can help limit interference and electronic jamming.

The research presented in this thesis was supported in part by NWO and in part by NXP Semiconductors in the framework of the partnership program on "Advanced 5G Solutions" within the Project 15590 titled "Antenna Topologies and Front-end Configurations for Multiple Beam Generation". The outcome of the thesis matches well with the objectives and milestones stated in the project proposal [17]. Throughout the project, the antenna array and beamforming topologies with the algorithms developed have been yearly discussed with and approved by the project partners from NWO, NXP Semiconductors, TU Eindhoven, TU Delft and TU Twente.

1.5. OUTLINE OF THE THESIS AND CHAPTER ABSTRACTS

The remaining of the thesis is organized in three parts which are related to system-, electromagnetism- and thermal-driven aspects of 5G antennas and beamforming. As for the system-driven aspects in the first part (Chapter 2-3), a communication system model is developed, different propagation phenomena are discussed and the statistical quality-of-service is assessed by using different precoding algorithms. After identifying the major system problems for 5G base station antennas, in the second part (Chapter 4-5-6), new array optimization strategies and innovative hybrid multibeam architectures are presented, which forms the electromagnetism-driven aspects. Finally, in the third part (Chapter 7-8-9), the thermal-driven aspects are studied by covering the thermal system modeling of electronics, passive cooling strategies at the base stations and the role of antenna researchers in array cooling. The flow of the thesis is summarized in Fig. 1.6. The content of the chapters to follow are briefly described as follows:

Chapter 2 - *5G System Simulations and Performance Assessment in LoS*

In this chapter, a multi-user line-of-sight (LoS) SDMA system model is developed with a meaningful link-budget analysis so as to investigate the trade-offs between the QoS, computational burden in precoding and complexity of beamforming for various use cases and number of users. The QoS at the user ends is rated by assessing the statistical signal-to-interference-plus-noise ratios (SINRs). Two beamforming algorithms, namely Conjugate Beamforming (CB) and Zero-Forcing (ZF), are considered and compared. Considering the candidate 5G base station deployment scenarios, rotated version of the currently proposed square array layouts are proposed to be used in CB with the least computational complexity while providing relatively good QoS. Different reduced-complexity ZF algorithms are also introduced as a compromise between the SINR performance and computational burden. The effects of several practical factors such as ap-

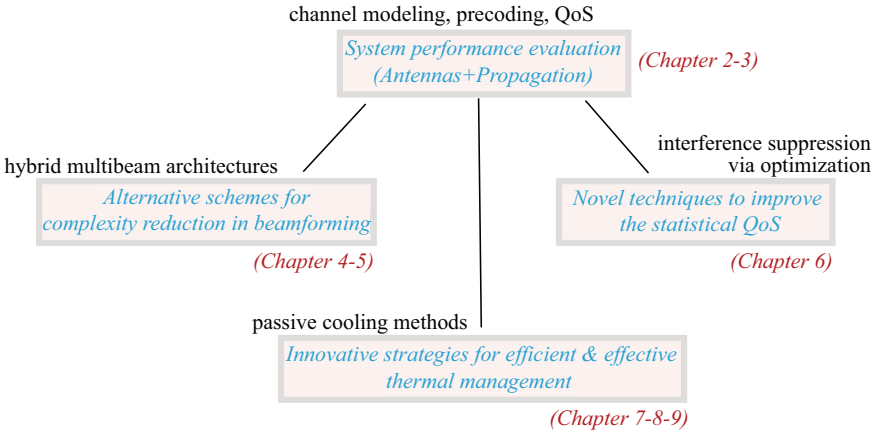


Figure 1.6: Thesis flow scheme.

proaches in user scheduling, errors in channel state information estimation and quantization in excitation amplitudes and phases are studied as well.

Chapter 3 - *mm-Wave Propagation in Urban Outdoor Environments - LoS vs NLoS*

In this chapter, the effect of forming single/multiple and multi-lobe beam patterns at mm-wave base station antennas on the received signal strength, beamforming complexity and co-channel interference is studied for mm-wave urban outdoor environments. A sample, simplified urban city model is used with randomly selected user positions. Ray tracing simulations are performed to analyze the channel's directional characteristics towards the test users. Depending on the number of dominant paths, single or multiple main lobes are created in the appropriate directions. Through the simulations, it is observed that for optimal gain, multiple beams should be used with one lobe and one port each, where signals from the same stream can be time and amplitude weighted before combining into one port/signal with increased SINR. Yet, this approach restricts the frequency re-use significantly and increases the beamforming complexity. Therefore, due to its key spectrum efficiency and simplicity, it is suggested to steer a single-lobe beam towards the dominant path of a user, whenever it provides a sufficient received signal level.

Chapter 4 - *Multibeam Antenna Array Architectures for 5G - An Overview*

In this chapter, the existing multiple beam generation strategies, traditionally developed and used in radar and space applications, are revisited. The advantages and disadvantages of each topology are discussed, with a primary focus on the multi-user 5G aspects.

Chapter 5 - *Subarray Based Beamforming Concepts*

In this chapter, two low-complexity beamforming concepts are studied which are based on active multiport subarrays and cosecant subarrays. In the first half of the chapter, a novel hybrid beamforming architecture based on a phased array of active multiport subarrays is proposed for multi-user 5G applications. The subarrays have multiple

simultaneous fixed beams which are smartly combined at the user locations with relatively large gains towards the intended co-frequency users and sufficiently low inter-user interference levels. The presented architecture has a wider angular coverage than the existing subarray-based hybrid beamforming schemes and reduces the signal processing complexity significantly as compared to the fully digital beamforming. In the second half, with the cosecant subarrays, a 5G base station phased array antenna concept is presented which is based on the power flux equalization in the elevation plane. Furthermore, novel SIW and microstrip line feeding based base station antenna designs have been proposed that have an approximate cosecant-square pattern in elevation and achieve flexible beam scanning in azimuth.

Chapter 6 - *Array Optimization for Inter-User Interference Minimization*

In this chapter, two new interference suppression techniques are presented: a phase-only control technique and a position-only control technique. The capabilities, limits, advantages and drawbacks of each method are explained in detail. In the first half of the chapter, with the phase-only control technique, minimization of the maximum sidelobe level for a given array geometry, amplitude distribution, and nulling sectors by phase-only adjustment of the element coefficients is studied. Nonlinear optimization problem for phase distribution is solved using a novel iterative convex optimization algorithm, which includes mutual coupling effects and exploits small phase perturbations at each step. Superiority of the algorithm in terms of the peak SLL and nulling depth achieved over several optimization methods reported in the most relevant literature is demonstrated in several case studies. Finally, a case study is performed to demonstrate the added value of the algorithm for mm-wave 5G application with phase-only radiation pattern forming. In the second half, with the position-only control technique, an extended-feature, system-driven convex algorithm for the synthesis of uniform-amplitude, irregular planar phased arrays with simultaneous multi-beam optimization for mm-wave 5G base station applications in multi-user scenarios is presented. The inter-user interferences are suppressed by minimizing the SLL for a beam scanned freely inside a given sector. A minimum guaranteed inter-element spacing in the final layout is predefined, which prevents element overlapping, eases the thermal problem and helps reduce the effects of high mutual coupling. The optimized array results show that, compared to their regular counterparts, significant reduction in the SLLs is achieved for a beam scanned inside the defined sector, which improves the statistical QoS significantly. The effect of mutual coupling on the results is also investigated via full-wave simulations and it is explained how embedded element patterns can potentially be included in the optimization. Superior capabilities of the proposed method are illustrated by comparing the algorithm output to those reported in the state-of-the-art literature. It is also explained how the algorithm can be used to synthesize quasi-modular arrays with reduced design/fabrication/calibration complexity.

Chapter 7 - *Cooling Strategies for Active Integrated Phased Arrays - An Overview*

In this chapter, various antenna system thermal management approaches studied in the literature are revisited. Considering the impact of the selected array architecture on the possible cooling options, first, the two configurations of modern phased arrays, i.e. Active Electronically Scanned Arrays (AESAs), are presented: Traditional AESAs and Pla-

nar AESAs. Moreover, a critical discussion on their advantages/disadvantages is given with the assessment criteria on 5G compatibility. Second, thermal management in these two types of array topologies are reviewed by providing relevant state-of-the-art examples. Third, the growing challenges with excessive heat generation at mm-waves are illustrated. Last, the role of antenna engineers/researchers in array thermal control is discussed. Under this topic, an unconventional antenna design concept, i.e. heatsink antenna, is considered and several existing prototypes are shown.

Chapter 8 - *Passive Cooling in 5G Base Station Antennas - Why and How?*

The strict cost and heat removal requirements of high volume 5G base station market favor the use of low-cost, passive cooling strategies via the use of heat sinks, heat spreaders, heat pipes or thermal interface materials (TIMs) instead of fans or forced liquids that require the use of electricity. In this chapter, the mechanisms of heat transfer phenomenon are described and the challenge of passive cooling in mm-wave chip-integrated base station antenna arrays is addressed. Several approaches in thermal modeling of electronics are revisited and discussed. Due to its computational efficiency and low-complexity, the two-resistor compact thermal model is applied to predict the junction temperatures of the beamformer chips. Thermal simulations are performed for NXP's Ka- band 8x8 fabricated array which has 16 quad channel ICs. Two commercial passive CPU heatsinks with different capabilities are modeled and attached to the chips. The simulation results are validated through experiments using the ICs' temperature sensor readings. The reasons of discrepancies between the simulated and measured results are explained. Furthermore, the system limitations (in terms of the maximum number of SDMA users and minimum number of antenna elements) due to fully passive array cooling at the base station antennas are discussed.

Chapter 9 - *Alternative Ways for Cooling Enhancement*

Although the CPU heat sinks are demonstrated as effective cooling modules for fully passive heat removal in mm-wave integrated 5G base station antenna arrays, they are bulky and heavy, which limits their use. In order to relax the requirements on the CPU coolers, reduce the size of the heat sinks and miniaturize the 5G units, it is necessary to provide additional (and preferably passive) cooling paths in the system. A way to achieve this is to make use of the antenna itself. The concept of "heat sink antennas" is revisited in Chapter 7 and it is shown that despite being a promising concept, the existing techniques are all based on complex 3-D antenna designs which are not desirable in the first-phase of low-cost 5G products. Therefore, in this chapter, the planar alternatives are investigated and two original approaches are presented. The first approach is based on optimizing the array layout, in other words, finding out the best positions of the heat sources (or the antenna elements). Since such an approach requires an interdisciplinary approach, the 2-D heat conduction dominated scenarios are first studied, for which several works exist in the thermal engineering domain. A novel connection between the field of antenna synthesis and heat source position optimization is established and an efficient sparsity-based convex layout optimization algorithm is proposed to be used for the "thermal-concern-only" applications. Then, the applicability of the method to the 5G antennas is discussed and the approach is adopted to the 3-D heat convection dominated scenarios, which is effective in applications where both thermal

and electromagnetic (referred to as dual) performances are concerned. The second approach is also proposed to satisfy the dual performance criteria and is based on increasing the thermal conductivity and the size of the antenna substrate simultaneously. The presented study in this chapter highlights the similarities and the differences between the antenna and thermal engineering disciplines and combines their strenghts to design dual-performance 5G arrays.

Chapter 10 - *Conclusion and Future Work*

This final chapter summarizes the main achievements presented in this dissertation and provides recommendations for future research.

I

COMMUNICATION SYSTEM DRIVEN ASPECTS

The value of a telecommunications network is proportional to the square of the number of connected users of the system.

Robert Metcalfe

2

5G System Simulations and Performance Assessment in LoS

Parts of this chapter have been published as:
[J-1] Y. Aslan, J. Puskely, A. Roederer and A. Yarovoy, "Trade-offs between the quality of service, computational cost and cooling complexity in interference-dominated multi-user SDMA systems," *IET Comm.*, vol. 14, no. 1, pp. 144-151, Jan. 2020.
[C-1] Y. Aslan, S. Salman, J. Puskely, A. Roederer and A. Yarovoy, "5G Multi-User System Simulations in Line-of-Sight with Space-Tapered Cellular Base Station Phased Arrays," in *Proc. 13th EuCAP*, Krakow, Poland, Apr. 2019.

2.1. INTRODUCTION

The concept of space (or beam) division multiple access (SDMA), which was first introduced in [18] and has been used in space applications for decades, has recently regained attention with the development of mm-wave 5G technologies having demanding criteria on the channel capacity [19]. Using SDMA at the multi-beam 5G base station (BS) antennas, the system capacity of the communication network can be increased by reusing the same time-frequency resource in several different (and preferably well separated) beams. Furthermore, SDMA can be straightforwardly combined with the well-established code division multiple access (CDMA) technique to further enhance the capacity, especially for the users that are too close to each other and cannot be resolved by the antenna beam [20].

This chapter combines multiple disciplines in 5G SDMA by addressing and discussing all the aspects mentioned in Fig. 1.5(b). A multi-user system model with a meaningful link budget is formulated considering the early deployment line-of-sight (LoS) only propagation environments via Monte Carlo simulations of random user locations. Two different application scenarios are investigated considering either ground-only users or users/relays at different heights. An interference-aware user scheduling algorithm is proposed, which guarantees an almost-uniform power distribution across the array. Being the most power and computation efficient beamforming algorithm, Conjugate Beamforming (CB) is investigated with different array topologies. The performance of CB is compared with the complexity-reduced Zero-Forcing (ZF) algorithms (i.e. ZF with a limited number of zeroes, only towards the most interfering users [21]) that are able to strongly reduce or cancel out the inter-user interferences.

Using the system model developed, the trade-offs between the complexity of beamforming, the Quality of Service (QoS) and the computational cost of precoding are studied for the 5G multi-beam BS antenna arrays. The statistical system QoS performance evaluation is given in terms of the cumulative distribution function (CDF) of signal-to-interference-plus-noise ratio (SINR) at the user ends, while the computational cost is given in terms of floating point operations per second (FLOPS).

The major contributions of this chapter are listed as follows:

(i) A novel LoS SDMA system model that incorparates a 5G mm-wave link budget is proposed.

(ii) A novel and generalized classification of the 5G use cases (depending on the BS deployment scenario and user locations) is performed.

(iii) From the antenna radiation pattern perspective, an original and intuitive simultaneous, co-frequency user selection technique in SDMA is proposed.

(iv) The 5G BS array rotation effect on the statistical user SINRs is investigated for the first time.

(v) A novel reduced complexity ZF algorithm is introduced which is based on a reference interference-to-noise-ratio (INR) threshold.

(vi) The pairwise trade-offs between the QoS, computational burden and beamforming complexity in 5G BS arrays are analyzed for the first time, with useful insights and system design recommendations.

2.2. SYSTEM MODEL

Let us consider an isolated cell sector in which a BS with M antenna elements is serving K single, omnidirectional antenna users in pure LoS simultaneously in the same narrow frequency sub-band using SDMA. It is worthy of note that the BS will potentially serve K', K'', etc. users in the other sub-bands at the same time. However, in this paper, we focus on the system simulations only in a single sub-band to analyze the SDMA performance in a single time-frequency resource block. Let $\boldsymbol{q}$ contain the symbols intended for the users ($E\{|q_k|^2\} = 1$), $\boldsymbol{\rho}$ is proportional to the maximal user SNRs, $\boldsymbol{n}$ consist of the unit-variance Additive White Gaussian Noise (AWGN) at each user and $\boldsymbol{y}$ be the received signal vector, such that $\{\boldsymbol{q}, \boldsymbol{\rho}, \boldsymbol{n}, \boldsymbol{y}\} \in \mathbb{C}^{Kx1}$.

The precoded signal vector $\boldsymbol{x} \in \mathbb{C}^{Mx1}$ is simply given by

$$\boldsymbol{x} = \boldsymbol{W}\boldsymbol{q} \tag{2.1}$$

where $\boldsymbol{W} \in \mathbb{C}^{MxK}$, is the precoding matrix satisfying

$$\sum_{m=1}^{M} |W_{m,k}|^2 = 1 \text{ for } \forall k \in \{1, \ldots, K\} \tag{2.2}$$

Furthermore, the received signal vector $\boldsymbol{y}$ can be computed by

$$\boldsymbol{y} = \sqrt{\boldsymbol{\rho}} \cdot (\boldsymbol{H}\boldsymbol{x}) + \boldsymbol{n} \tag{2.3}$$

where $\boldsymbol{H} \in \mathbb{C}^{KxM}$ denotes the downlink channel matrix with $E\{|H_{k,m}|^2\} = 1$ whose entries are formulated as [22]

$$H_{k,m} = \beta_{k,m} G_m(\hat{\boldsymbol{r}}_{km}) \frac{e^{-j\frac{2\pi}{\lambda}|\boldsymbol{r}_k - \boldsymbol{r}_m|}}{|\boldsymbol{r}_k - \boldsymbol{r}_m|} \tag{2.4}$$

where $G_m(\hat{\boldsymbol{r}}_{km})$ is the far-field function of the m^{th} BS antenna element in the direction $\hat{\boldsymbol{r}}_{km}$ towards the k^{th} user. $\beta_{k,m}$ is the normalization constant. For simplicity, let us assume a common cosine shaped embedded element pattern, i.e. $G(\hat{\boldsymbol{r}}_k) = G_m(\hat{\boldsymbol{r}}_{km}),\ \forall m \in \{1, \ldots, M\}$.

Consequently, the SINR of the transmission from the BS array to the k-th user is given by

$$SINR_k = \frac{\rho_{k,k} |\boldsymbol{H}_{k,:}\boldsymbol{W}_{:,k}|^2}{\sum_{j \neq k}^{K} \rho_{k,j} |\boldsymbol{H}_{k,:}\boldsymbol{W}_{:,j}|^2 + 1} \tag{2.5}$$

where $\rho_{k,j}$ is given by

$$\rho_{k,j}(dB) = P_j(dBm) - 20\log_{10}[f_c] - 20\log_{10}[\frac{4\pi}{c}] - 20\log_{10}[|\boldsymbol{r}_k|] + G(\hat{\boldsymbol{r}}_k)(dB) - N_{th}(dBm) \tag{2.6}$$

where f_c is the carrier frequency, N_{th} is the thermal noise power and P_j is the average adaptive transmit power to the j^{th} user with equalized SNRs which is calculated by

$$P_j = P_{\max} \frac{(d^2_{user,j})}{\max(\boldsymbol{d}^2_{user})} \frac{\cos(\min(\boldsymbol{\theta}_{sc}))}{\cos(\theta_{sc,j})} \tag{2.7}$$

where $\boldsymbol{d}_{user}$ is a vector containing the distances between the users and the BS and $\boldsymbol{\theta}_{sc} \in \mathbb{C}^{Kx1}$ contains the scan angles towards the users.

Regarding the precoding algorithms, two commonly applied techniques, CB and ZF are studied for which the precoding matrix $\boldsymbol{W}$ is given by

$$\boldsymbol{W} = \begin{cases} \boldsymbol{H}^{\dagger} & for\ CB \\ \boldsymbol{H}^{\dagger}(\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1} & for\ ZF \end{cases} \tag{2.8}$$

where † denotes the Hermitian transpose. In CB, the beam is steered towards the intended user with maximum gain (thus with maximal SNR) and no dynamic control on the inter-user interference, whereas, in ZF, the interference towards the unintended users is canceled out, but the gain towards the aimed user will not be optimal.

It is well-known that in an interference-dominated application scenario such as the 5G multi-user SDMA communication systems, ZF (or a ZF-like precoding algorithm) provides better statistical SINR performance than the CB. However, as a trade-off, ZF demands more accurate channel state information (CSI) and more processing power.

2.3. SIMULATION SCENARIOS AND SETTINGS

We first assume that an M-element $\lambda/2$-spaced regular square-grid multibeam array is serving K randomly located simultaneous co-frequency users within a predefined sector in user azimuth & elevation, with a maximum range of $r_{\max}$. This scenario is visualized in Fig. 2.1(a). While serving a user, the side lobes of the beam generated by an array will cause interference (whose strength depends on the user distance) towards the others, as visualized in Fig. 2.1(b). This will reduce the SINR and thus, the throughput at the user ends.

Such a simulation scenario requires full flexibility in the formation of the multiple beams if the best attainable system performance in terms of the statistical SINR is aimed for. Fig. 2.2 shows two potential multibeam phased array architectures that are compatible with the presented system studies. An analog, fully-connected array[1] is given in Fig. 2.2(a), while its digital counterpart is visualized in Fig. 2.2(b). More information on these beamforming structures, together with several other multiple beam generation strategies, will be provided later in Chapter 4. For now, it is worth to notice that both schemes have a large design and processing complexity.

To realize comprehensive system simulations, a generalized classification of the 5G mm-wave deployment scenarios is performed. We consider two different use cases depending on the BS position and accordingly, the user locations (i.e. ground-only users and users/relays at different heights) with the common simulation parameters listed in Table 2.1.

[1] Note the significant RF combining loss in the fully-connected configuration, such that 90% of the power is lost for 10 beams! (which has to be compensated by PAs after the combiners), unless a Butler-like combining, only producing agile orthogonal beams, is used. More details on front-end configurations are given in Chapter 4.

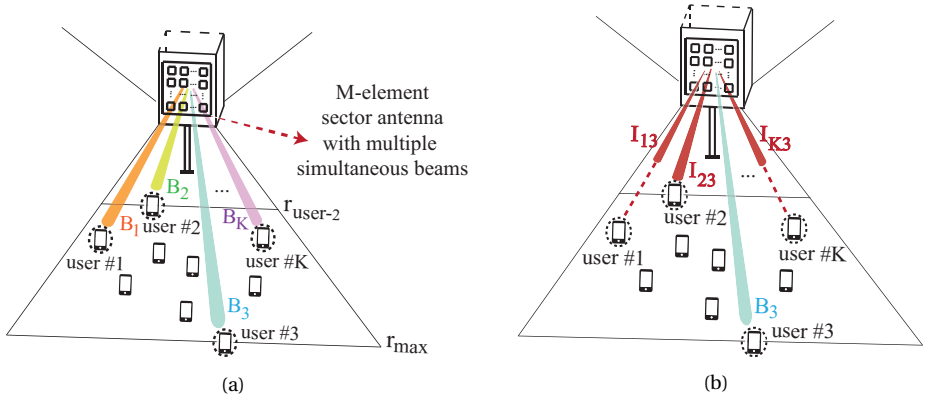


Figure 2.1: An illustration of the simulation scenario: (a) a sector antenna with M elements serving K users in the same time-frequency source block (side lobes are not shown for simplicity), (b) interference towards the other simultaneous, co-frequency users for beam #3 (note that beam #3 creates the largest interference towards the other users since user #3 is located further from the base station than the other users).

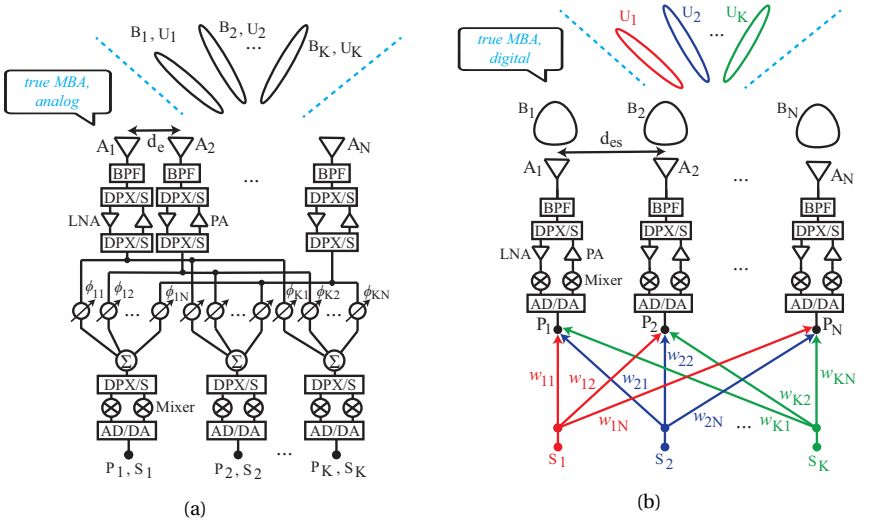


Figure 2.2: Potential multi-beam phased array (MBA) architectures compatible with the system studies: (a) fully-connected analog array, (b) fully-digital array (A: antenna element, Ar: antenna array, P: antenna port, S: user stream, B: antenna beam, U: user beam)

It is worthy to note that the two uses cases investigated here represent the possible 5G millimeter-wave deployment scenarios that are currently discussed [16] (see Figure 1.4). The ground-only user case (Use Case A) mimics the scenario where the base station will be on a tower at a certain height, which does not require a large amount of vertical scanning range. The use case with users/relays at different heights (Use Case B), on the other hand, mimics the scenario where the base station is mounted lower

Table 2.1: A list of sector simulation parameters

Parameter definition	Symbol	Value
Center frequency (GHz)	f_c	28
Number of antennas at the BS	M	64
Maximum cell range (m)	$r_{\max}$	200
Directivity of each antenna element at Tx (dB)	$D_{tx,el}$	8
Directivity of each user antenna at Rx (dB)	D_{rx}	0
Maximum average BS transmit power per user (dBm)	$P_{\max}$	26
Bandwidth (MHz)	BW	100
Thermal Power Spectral Density (dBm/Hz)	PSD	-170
Noise figure (dB)	NF	10
Thermal noise (dBm)	N_{th}	-80
Number of random user location realizations	N_{sim}	10,000

to the ground, which will require vertical scanning ability to deliver signals across the buildings in elevation as well as to the users on the ground. The impact of the two use cases on the distribution and scheduling of the SDMA users and on the statistical system performance will be explained in more detail in Section 2.4. In principle, these two scenarios require very different performance specifications, not necessarily best served by the same generic antenna.

The coordinate system used in the studies is shown in Fig. 2.3. The relation between the conventional spherical coordinate angles and the user elevation/azimuth angles are also provided in (2.9). The given definitions are important to understand the distribution of users in the angular domain and the sector definition in the $u(\sin\theta\cos\phi) - v(\sin\theta\sin\phi)$ plane later in Section 2.4.

$$\begin{gathered} x = R\sin(\theta)\cos(\phi),\ \ y = R\sin(\theta)\sin(\phi),\ \ z = R\cos(\theta), \\ \varphi_{az} = \arctan(x/z) = \arctan(\tan(\theta)\cos(\phi)), \\ \varphi_{el} = \arctan(y/z) = \arctan(\tan(\theta)\sin(\phi)), \\ \phi = \arctan(\tan(\varphi_{el})/\tan(\varphi_{az})),\ \ \theta = \arctan(\sqrt{\tan(\varphi_{el})^2 + \tan(\varphi_{az})^2}) \end{gathered} \tag{2.9}$$

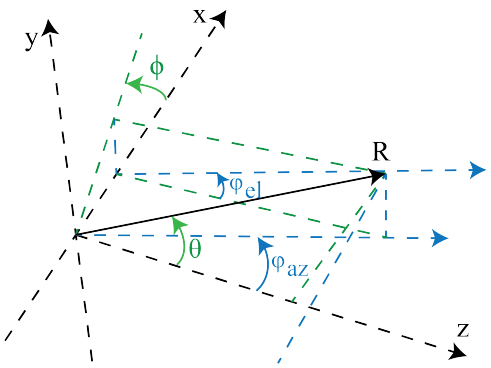


Figure 2.3: Coordinate system: conventional spherical coordinates and user elevation & azimuth angles.

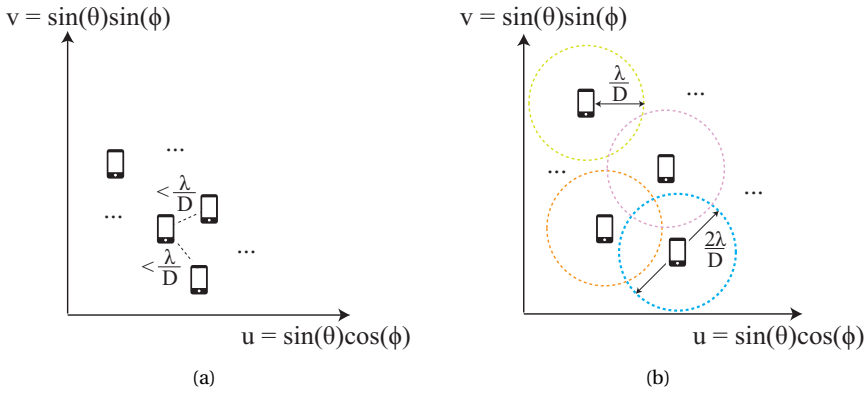


Figure 2.4: Illustration of random scheduling of K concurrent co-frequency users among all the users in a sector, (a) interference-unaware, (b) interference-aware.

To keep the random nature of the Monte Carlo simulations, K random user positions are to be selected among all the potential SDMA users at each realization of the communication system. The assigned user positions at a time should be controlled in the MAC layer scheduling. One option is to select completely random positions, which is commonly applied in papers focusing only on the signal processing aspects. This approach will be referred as the interference-unaware scheduling in this paper (see Fig. 2.4(a)). A smarter approach is to consider the main lobe-width of the antenna in which no other co-channel users are allowed to enter. This (at this point strictly for LoS) will be called as the interference-aware approach since, when there are two users in the main beam of an array, either the interference between them will be very large (in the case of CB) or the user SNRs will be very low (in the case of ZF). In this work, we keep the minimum angular separation at λ/D radians for the interference-aware approach as seen in Fig. 2.4(b) (which is approximately the half power beamwidth) where D is the side length of the array.

Throughout the simulations, mutual coupling is not taken into account and the embedded pattern of a patch element is assumed to be equal to $\sqrt{\cos\theta_{sc}}$, where θ_{sc} is the scan angle in spherical coordinates which is different from the user elevation angle, φ_{el}. The losses in the antenna system (which depend on a particular design, element structure, PCB material, etc.) are not included either, so the system model is ideal in this sense. In future studies, all types of losses can be included in a more detailed link budget, which will have a negative impact on the operational cell range and user SINRs unless compensated by the amplifiers.

2.4. SIMULATION RESULTS AND DISCUSSION

In this section, the system simulation results are provided and discussed for the two simulation scenarios explained in Section 2.3: ground users (Use Case A) and users/relays at different heights (Use Case B).

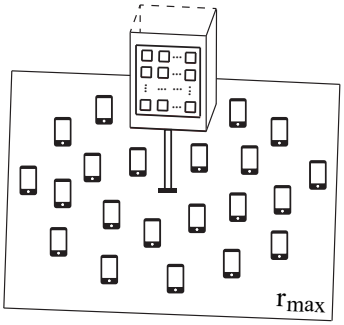


Figure 2.5: Uniform user distribution in a cell around the base station.

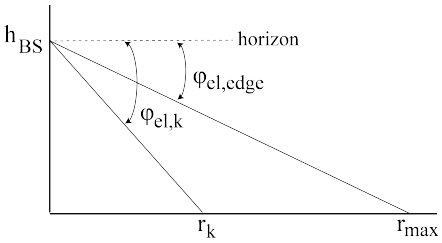


Figure 2.6: Definition of the edge angle ($\varphi_{el,edge}$), scan angle in elevation ($\varphi_{el,k}$) and corresponding distances from the base station.

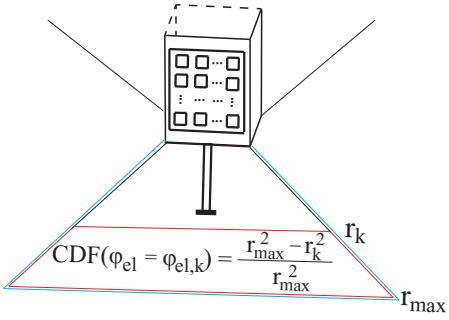


Figure 2.7: Visualization of the area covered when φ_{el} is scanned from $\varphi_{el,edge}$ to $\varphi_{el,k}$.

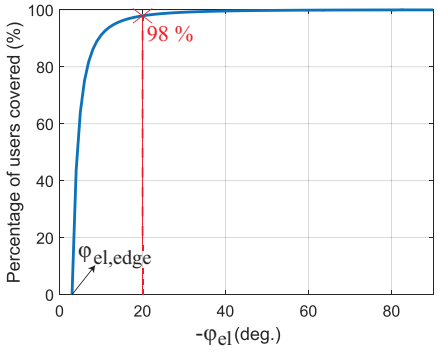


Figure 2.8: Percentage of user coverage with respect to $\varphi_{el,k}$ in a cell with uniformly distributed users over the area for h_{BS} = 10.5 m, $r_{\max}$ = 200 m.

2.4.1. GROUND USERS

In this situation, the users are uniformly distributed on the ground in the cell sector as shown in Fig. 2.5 and it is assumed that the BS is located on a street lamp pole at a height of 10.5 m from the users. The relation between the user elevation angle corresponding to the cell edge ($\varphi_{el,edge}$), base station height (h_{BS}), scan angle in elevation ($\varphi_{el,k}$) and corresponding range (r_k) is shown in Fig. 2.6. The area covered when the elevation angle is scanned from $\varphi_{el,edge}$ to $\varphi_{el,k}$ is computed as in Fig. 2.7 in terms of a cumulative density function (CDF). The result is plotted in Fig. 2.8 which shows that 98% of users are in the first 20 degrees below the horizon.

In the first study case, we assume that 4 random user positions are to be selected according to Fig. 2.8 in elevation and from a uniform distribution in ±45 degrees in azimuth. Note that the scan angle in elevation towards the cell edge $\varphi_{el,edge}$ is equal to 3 degrees below the horizon since $r_{\max}$ = 200 m. The histograms of the selected user positions for 10,000 random realizations for the two scheduling approaches (interference-unaware vs interference-aware) are shown in Fig. 2.9.

Four different approaches are investigated that differ in the scheduling technique,

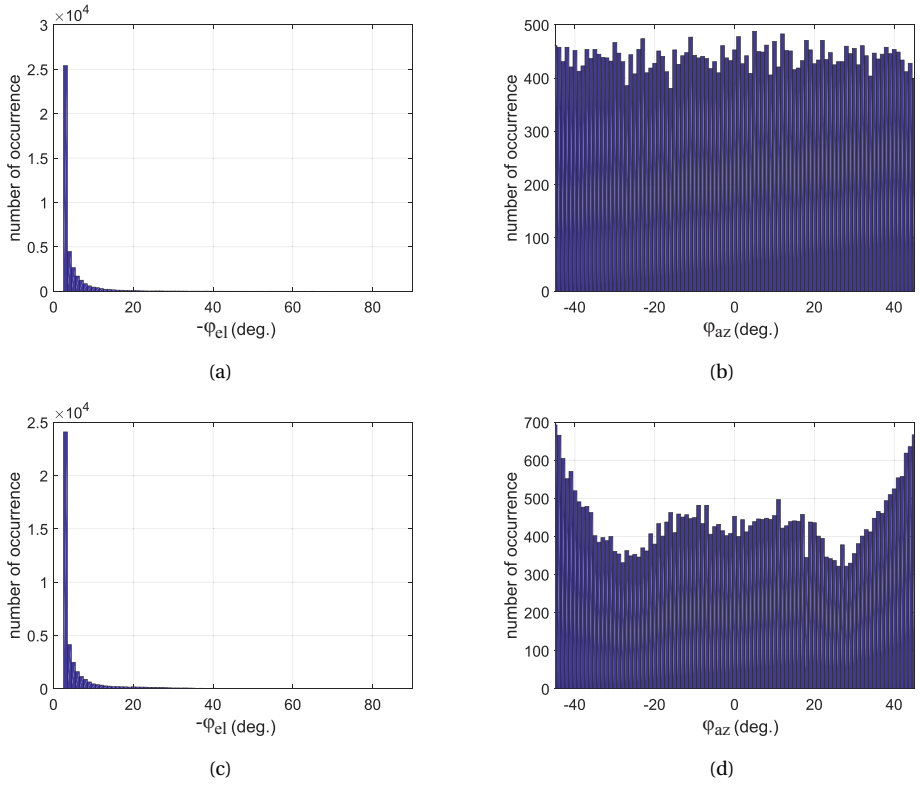


Figure 2.9: Histogram of angular positions for 4 users for 10,000 random realizations in Use Case A, (a) interference-unaware, φ_{el}, (b) interference-unaware, φ_{az}, (c) interference-aware, φ_{el}, (d) interference-aware, φ_{az}.

errors in CSI estimation and quantization in antenna excitations. Two precoding approaches mentioned before in Section 2.2 are studied for 0.5λ-spaced regular square arrays: CB and ZF. The statistical performance is provided as a graph of CDF of SNR, INR and SINR. In such graphs, it is a common practice to define a horizontal line at a certain CDF value that represents the percentage of occurrence (for example, a line at a CDF value of 0.1 represents 90% of the cases while 0.05 represents 95% of the cases). The intersection point of this horizontal line with the related curve will give the minimum SNR/INR/SINR value for the chosen percentage of the total occurrences.

Approach A: Interference-unaware scheduling, perfect CSI, non-quantized excitations

Fig. 2.10 provides the results for interference-unaware scheduling with perfect CSI and non-quantized excitations in terms of the CDF of the signal-to-noise ratio (SNR), interference-to-noise ratio (INR) and signal-to-interference-plus-noise ratio (SINR) at all users for all random realizations. It is seen that very low SNRs can be observed in ZF, while INRs can be quite large in CB. Both effects occur when the angular inter-spacing of the users are smaller than the main beamwidth of the base station antenna, which results in very low SINRs.

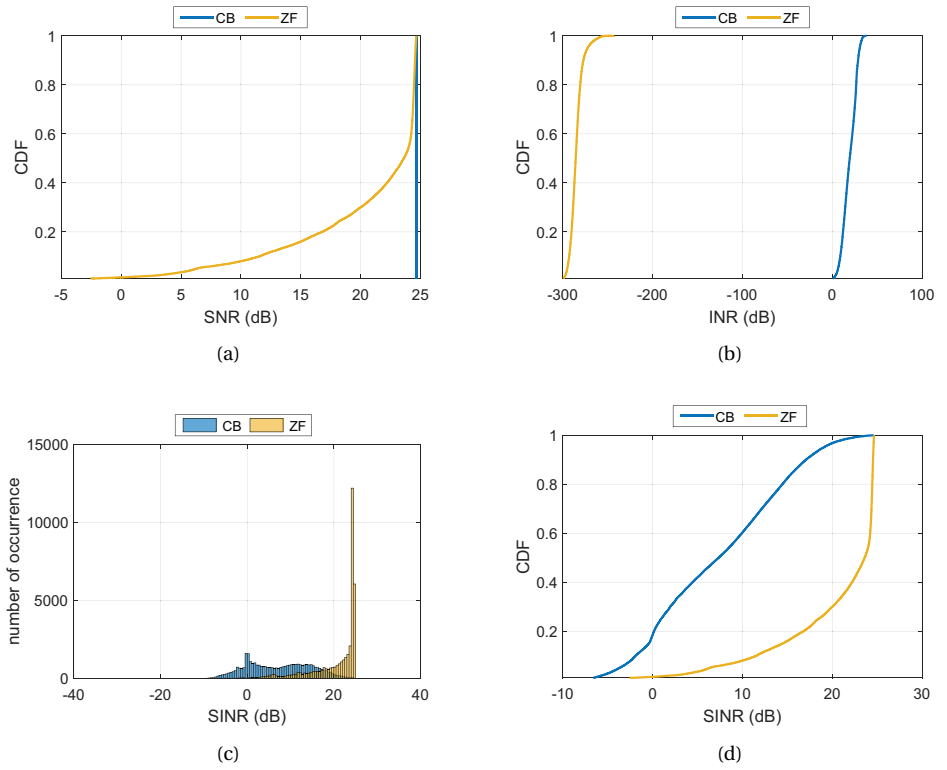


Figure 2.10: Results of Approach A in Use Case A: Interference-unaware scheduling, perfect CSI, non-quantized excitations, (a) CDF of SNR, (b) CDF of INR, (c) histogram of SINR, (d) CDF of SINR.

Approach B: Interference-aware scheduling, perfect CSI, non-quantized excitations

In this case, the user scheduling approach is changed to guarantee a minimum angular separation among the users that is larger than the half main beamwidth. The results are given in Fig. 2.11, which show significant improvement compared to the results in Fig. 2.10. It is seen that both approaches can achieve very good SNR, while the interference is nearly zero in ZF and relatively large in CB, which is expected due to the high SLLs of the regular square array. Fig. 2.11(d) shows that the SINR in 90% of the cases is larger than 7.23 dB and 24.31 dB for the CB and the ZF, respectively. Note that although it occurs rarely, the side lobes of faraway users in CB may create very large interference towards the closeby users. Therefore, it is preferable to treat the closeby users with a separate sub-band and possibly a separate antenna (or a separate part/portion of the same antenna), which is not implemented in this thesis.

Approach C: Interference-aware scheduling, imperfect CSI, non-quantized excitations

The effect of CSI errors on the results in Fig. 2.11 is studied here. The impurities in CSI is realized by adding normally distributed random complex numbers in every element of $\boldsymbol{H}$ within a certain maximum error range. Fig. 2.12 provides the CDF of SINR for two different maximum error values. It is seen that if the error is smaller than 0.001,

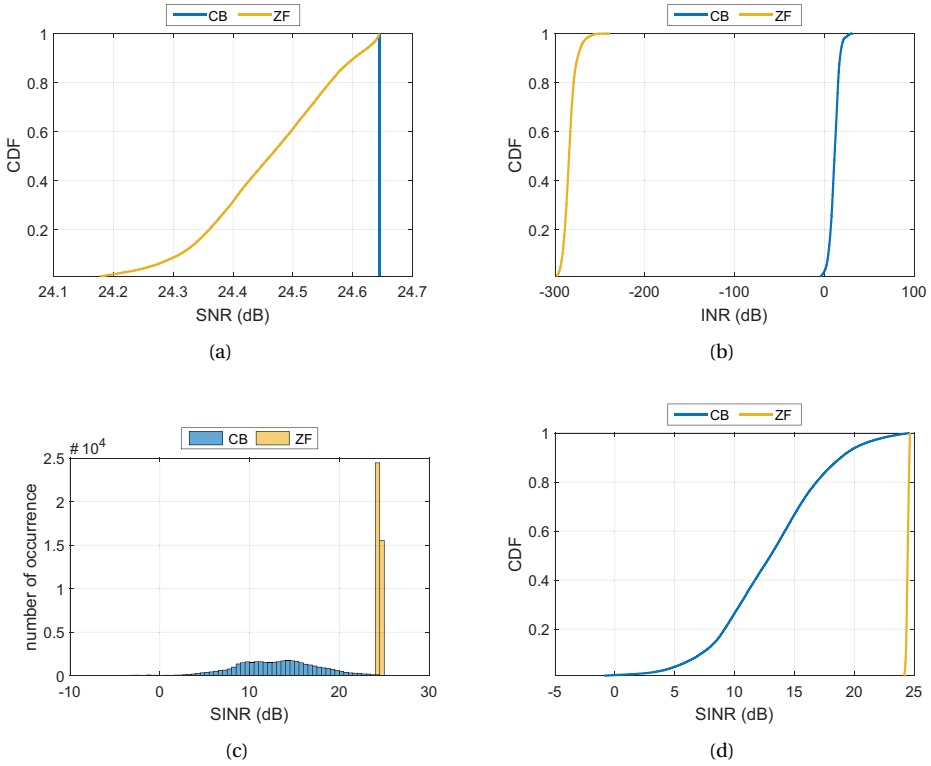


Figure 2.11: Results of Approach B in Use Case A: Interference-aware scheduling, perfect CSI, non-quantized excitations, (a) CDF of SNR, (b) CDF of INR, (c) histogram of SINR, (d) CDF of SINR.

the result in Fig. 2.11(d) is completely recovered. However, increasing the errors deteriorates the performance, as expected. Fig. 2.12(b) shows that for a maximum error of 0.01, the SINR in 90% of the cases is larger than 6.08 dB and 17.96 dB for the CB and the ZF, respectively. This shows that CB is more robust to imperfect CSI (at relatively small errors) compared to ZF, which relies on placing perfect nulls at interferer positions.

Approach D: Interference-aware scheduling, perfect CSI, quantized excitations

Effect of amplitude and phase quantization of the antenna weights on the user SINR performance is studied here by rounding the ideal weights to the practically available ones in the case of analog beamforming. Since CB does not apply amplitude tapering and ZF yields very small excitation amplitude ranges in the interference-aware scheduling (only a few dBs), the main focus is on the phase shifters while keeping the amplitude steps at 0.5 dB. In general, phase shifters with low number of bits are relatively low cost, low loss and wideband since the number of integrated solid state devices and layers is reduced. However, the deteriorated phase resolution may lead to directivity loss, high side lobes and filled nulls.

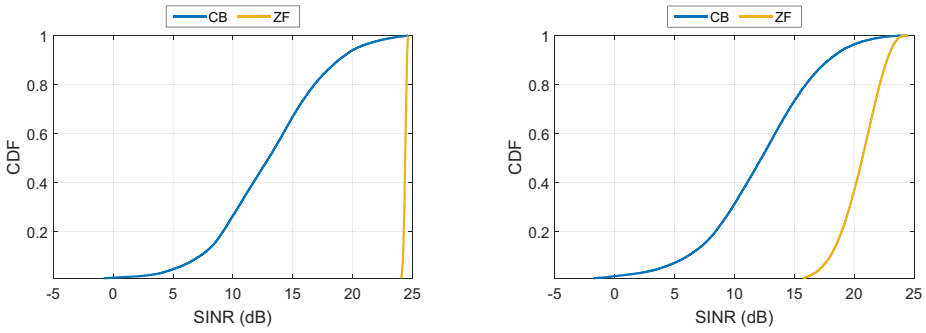


Figure 2.12: SINR results of Approach C in Use Case A: Interference-aware scheduling, imperfect CSI, non-quantized excitations, (a) max error of 0.001, (b) max error of 0.01.

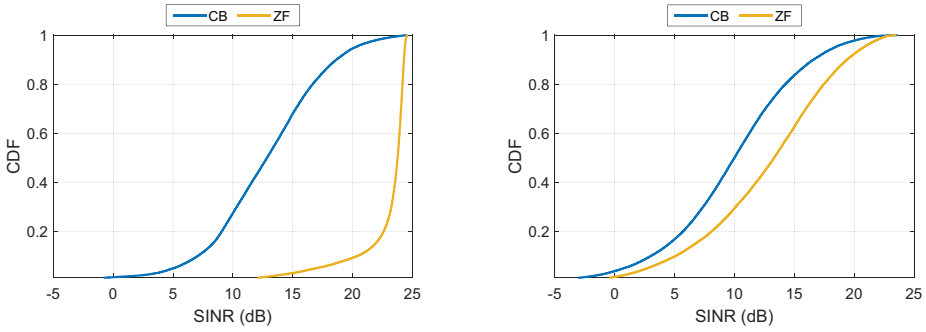
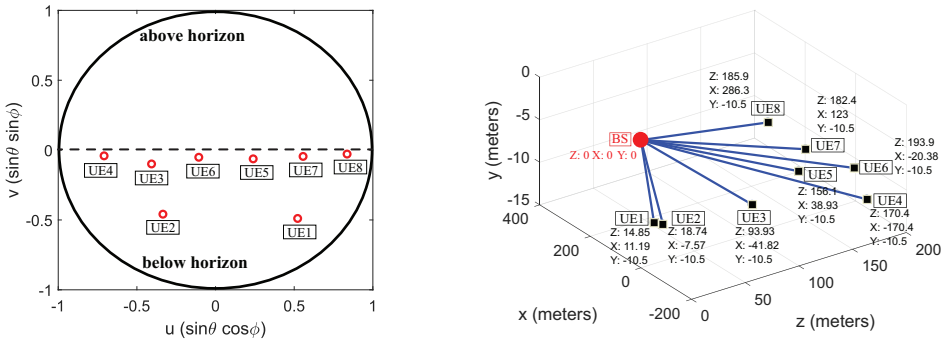


Figure 2.13: SINR results of Approach D in Use Case A: Interference-aware scheduling, perfect CSI, quantized excitations, (a) 0.5 dB amplitude step, 4-bit phase, (b) 0.5 dB amplitude step, 2-bit phase.

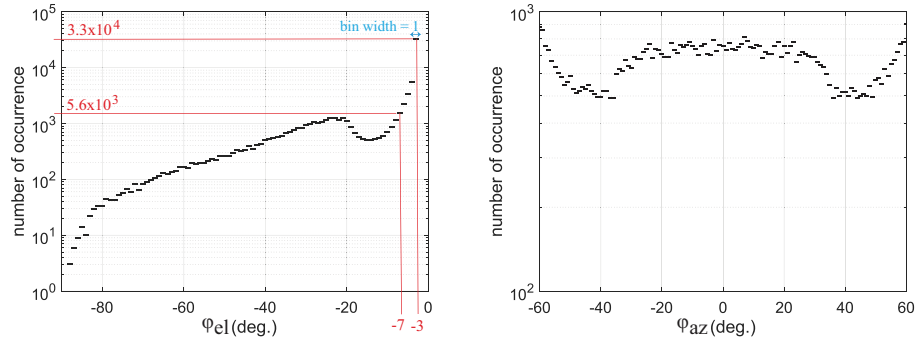
The statistical SINR results with 4-bit and 2-bit phase shifters are given in Fig. 2.13. It is seen that for 4-bit phase shifters, the results in Fig. 2.13(a) become similar to the ones in Fig. 2.12(b). However, Fig. 2.13(b) shows that for 2-bit phase shifters, the SINR in 90% of the cases is larger than only 3.12 dB and 5.15 dB for the CB and the ZF, respectively. This shows that CB is more robust to phase quantization than ZF and that CB performance approaches to ZF statistically, when the number of phase bits is small.

In the second study case, the number of SDMA co-channel users is increased from 4 to 6 and to 8. In order to keep the randomness of Monte Carlo simulations with a larger number of user ends, the sector coverage in azimuth is increased to ±60 degrees. Having seen its advantage, the interference-aware scheduling technique is applied in the MAC layer. A sample distribution of locations for 8 users is given in Fig. 2.14. After repeating the random location distribution for 10,000 times for 8 users, a histogram of user elevation (φ_{el}) and azimuth (φ_{az}) angles is obtained, which is given in Fig. 2.15. It is seen that very small elevation angles below horizon occur the most, such that the



(a) (b)

Figure 2.14: A sample distribution of 8 random ground users with interference-aware selection in Use Case A in (a) uv-plane, (b) Cartesian coordinates. (Note that the users end up close to the peaks of the sidelobes in the CB case for the square generic arrays.)



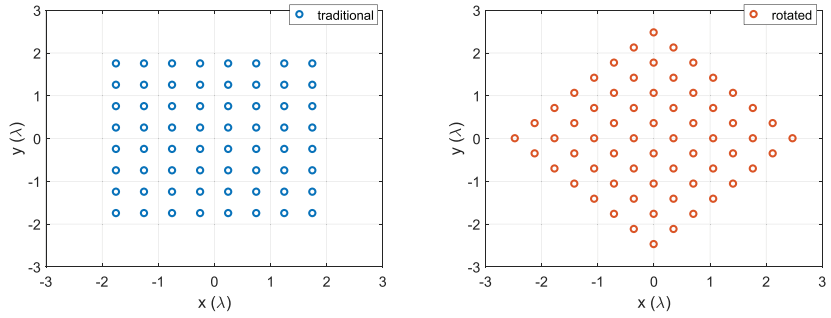
(a) (b)

Figure 2.15: Statistical distribution of 8 ground users with interference-aware selection for 10,000 random realizations in Use Case A: (a) user elevation (φ_{el}), (b) user azimuth (φ_{az}).

number of occurrences starting from -3 deg. at the cell edge, up to -7 deg. covers nearly 60% of the total occurrences. This is expected since, similarly to the first study case, most of the ground sector area can be covered by scanning φ_{el} by a few degrees, which makes the use of 2-D scanning questionable for the ground-only users.

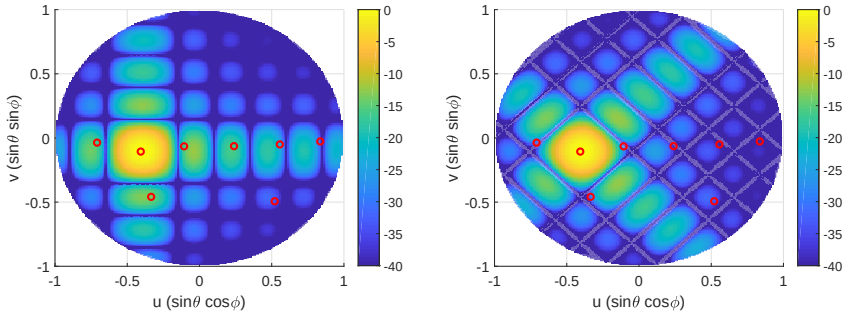
In this study case, we investigate the QoS performance and computational complexities of the two precoding techniques using two different array topologies having 64 elements: the traditional 0.5λ-spaced regular square array (as used in the first study case) and the 45-degree rotated version of it (see Fig. 2.16). At this point, one may fairly ask what the motivation is for examining a rotated array.

In Fig. 2.14, it is seen that most of the users are located next to each other in the uv-plane at very small v values. In such a scenario, using the traditional array frequently results in having the high side lobes at the neighboring users, as illustrated in Fig. 2.17(a).



(a) (b)

Figure 2.16: Investigated array layouts: (a) traditional 0.5λ-spaced regular square array, (b) 45-degree rotated version of it.



(a) (b)

Figure 2.17: Normalized radiation patterns in Use Case A while serving UE3 in Fig. 2.14 with CB using (a) traditional array, (b) rotated array.

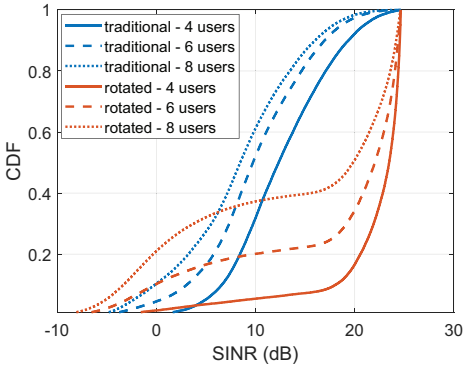


Figure 2.18: CDF of SINR in Use Case A - Approach B using CB with the traditional and rotated arrays.

However, by simply rotating the array by 45 deg., it is possible to shift the higher side lobes to the diagonal axis where, in most of the time, there are no users (see Fig. 2.17(b)). This analysis provides an original and intuitive approach to increase the statistical QoS via rotation of the traditional arrays placed on a square grid.

The QoS comparison for the two layouts is given in Fig. 2.18 for different number of users in terms of CDF of SINR. From Fig. 2.18, by looking at the intersection points of the curves, it is seen that the minimum guaranteed SINR becomes larger for the rotated array as compared to the traditional array when the pre-defined percentage of the total occurrences is less than 97%, 83% and 66% for $K = 4$, $K = 6$ and $K = 8$, respectively.

2.4.2. USERS/RELAYS AT DIFFERENT HEIGHTS

In this situation, the sector is assumed to be formed by ±15 deg. in elevation and ±60 deg. in azimuth (which is used as the 5G cell definition in some of the current systems

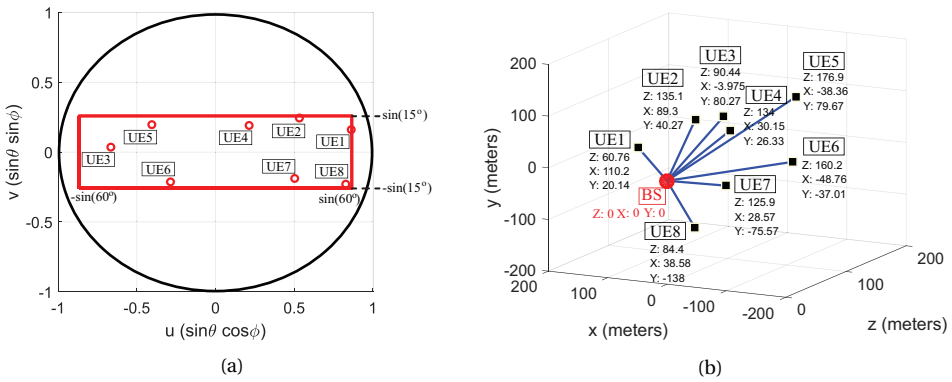


Figure 2.19: A sample distribution of 8 random users/relays at different heights with interference-aware selection in Use Case B in (a) uv-plane, (b) Cartesian coordinates.

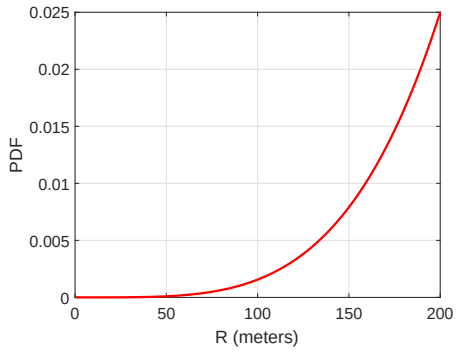


Figure 2.20: PDF of the Beta distribution with the design parameters α=5 and β=1, scaled in an interval of R = [0, 200 m].

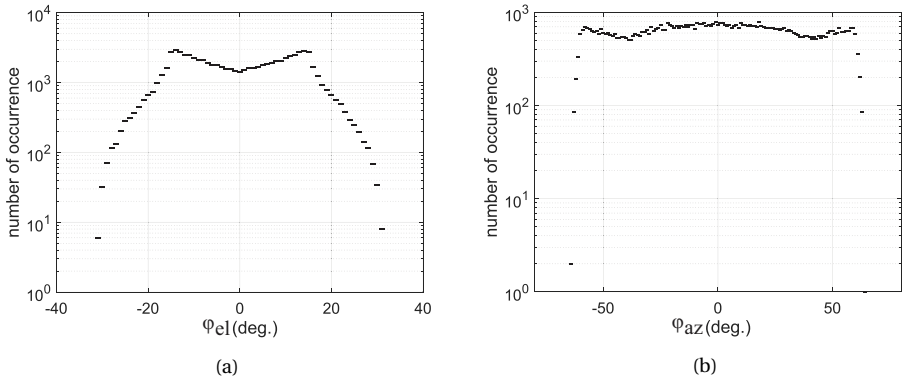


Figure 2.21: Statistical distribution of 8 random users/relays at different heights with interference-aware selection for 10,000 random realizations in Use Case B: (a) user elevation (φ_{el}), (b) user azimuth (φ_{az}).

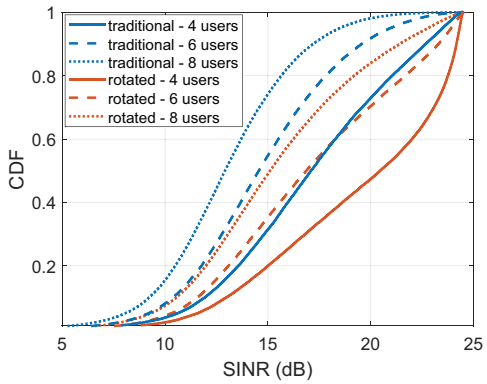


Figure 2.22: CDF of SINR in Use Case B - Approach B using CB with the traditional and rotated arrays.

[23] that cover both the ground users and the buildings), as shown in Fig. 2.19(a). The users' locations are randomly picked from uniformly distributed points in the uv-plane using the interference-aware user selection algorithm. Moreover, the distances from each user to the BS is randomly picked from a scaled Beta distribution with α=5 and β=1 in a range of [0, 200 m] to mimic the realistic scenario with higher probability for larger distances. The probability density function (PDF) of the Beta distribution used in this section is given in Fig. 2.20. Fig. 2.19 shows a sample selection of 8 SDMA users in this case. Similarly to the ground-only users, a statistical distribution of φ_{el} and φ_{az} for 10,000 random realizations is provided in Fig. 2.21, which shows more uniformity in the distribution of angles as compared to Fig. 2.15.

Note that according to Fig. 2.21, $|\varphi_{el}|$ can be larger than 15 deg. (can reach up to 30 deg. with low probability) and $|\varphi_{az}|$ can be slightly larger than 60 deg. (up to 64 deg.), which is observed around the corners of the rectangular-shaped cell which is a

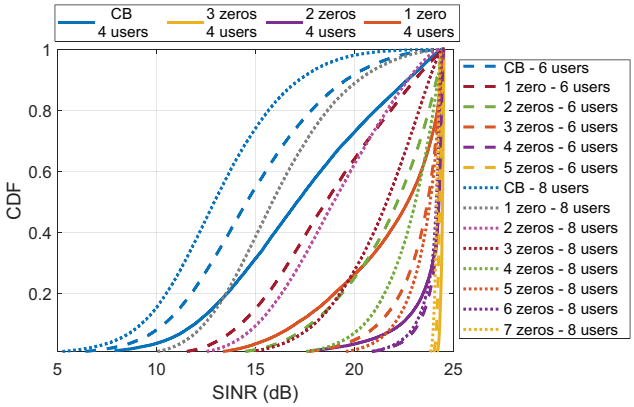


Figure 2.23: CDF of SINR for reduced complexity ZF based on the number of zeros in Use Case B - Approach B using the square grid array.

straightforward approximation of the elliptical cell in the uv-plane formed using (2.9).

First, we evaluate the statistical QoS performance of traditional and rotated arrays in the case of CB. The motivation behind studying the rotated array is similar to the one for the ground-only users since the sector has a relatively narrow angular width in elevation. The CB results in this case, given in Fig. 2.22, shows that in 90% of the total occurrances, the minimum SINR is 1.0 dB, 1.2 dB and 1.4 dB larger in the rotated array for $K = 4$, $K = 6$ and $K = 8$, respectively, as compared the array with the traditional layout.

As seen previously in Section 2.4.1, ZF is superior to CB in terms of QoS since it can cancel out the interferences. However, the computational complexity of ZF is much larger than CB. Therefore, two different reduced complexity ZF algorithms are investigated here in Use Case B for the traditional, square-grid array.

Firstly, only the strongest N_c interferers are cancelled for each user as done in [21]. According to [21], the number of FLOPS in this case ($\#\mathscr{F}_{ZF}$) of ZF (for $N_c > 0$) is given by

$$\#\mathscr{F}_{ZF} = K(24N_cM^2 + 48{N_c}^2M + 54{N_c}^3 + 6M) \tag{2.10}$$

On the other hand, for CB (with no zeros), the number of FLOPS ($\#\mathscr{F}_{CB}$) is equal to

$$\#\mathscr{F}_{CB} = K(14M - 2) \tag{2.11}$$

SINR results for different number of zeros per user (N_c) is given in Fig. 2.23. The number of FLOPS corresponding to diffrent N_c and number of user (K) combinations is plotted in Fig. 2.24. It is seen that CB can provide a three-order of magnitude reduction in FLOPS as compared to the ideal ZF, but it has 8.8 dB, 10.6 dB and 11.9 dB less minimum guaranteed SINR in 95% of the total occurrances for $K = 4$, $K = 6$ and $K = 8$, respectively. Therefore, reduced complexity ZF algorithm with a pre-defined N_c (depending on the SINR requirements) can provide a good compromise between the QoS and computational complexity.

Secondly, instead of cancelling a constant number of interferers for each user, an adaptive cancellation scheme is proposed where the interferers whose power is larger

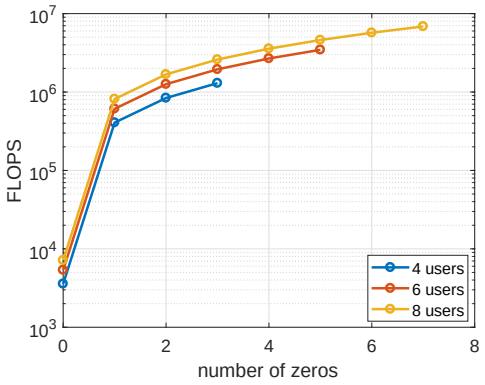


Figure 2.24: Computational complexity in reduced-complexity ZF (based on the number of zeros, with no zeros representing the CB) for a single realization of the random user distribution.

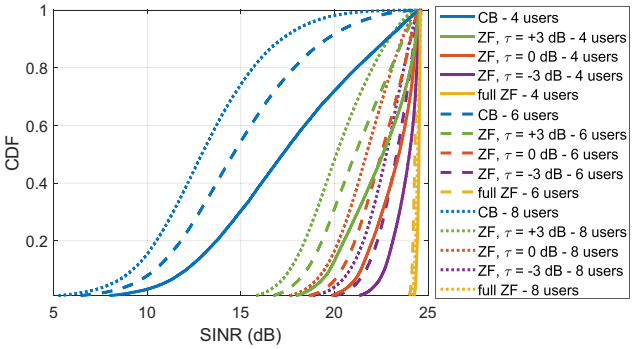


Figure 2.25: CDF of SINR for reduced complexity ZF based on different INR threshold values (τ) in Use Case B - Approach B using the square grid array.

than a certain threshold τ (with respect to the noise power) are cancelled out. The statistical SINR results in this approach for τ = +3 dB, 0 dB and -3 dB are given in Fig. 2.25.

It is worth to note that since the users are already much closer to zero levels with the 45-degree rotated square array (see Fig. 2.17(b)), ideal ZF (and simplified ZF) would require less perturbation from the CB and thus yield more equal amplitudes.

2.5. CONCLUSIONS

In this chapter, ***a multi-user LoS SDMA system model has been presented to investigate, for the first time, the trade-offs between QoS, computational burden of precoding and beamforming complexity for various use cases and number of users with different beamforming algorithms and array layouts.***

From the system discussions and the simulation results, the following major observations have been made:

1. To obtain full flexibility in the formation of the multiple beams for the best attainable SINR performance, fully-connected analog or fully-digital phased array architectures are used, for which each beam has its own beamforming structure (analog or digital). Both topologies have a large complexity, high cost and extremely large heat generation at mm-wave frequencies.

2. A certain minimum angular spacing among the simultaneous co-frequency users has to be satisfied in MAC layer user scheduling to guarantee high SNR values in zero-forcing and low interference levels in conjugate beamforming. Therefore, **a novel co-channel user selection algorithm for MAC has been proposed that is based on sufficient user isolation**.

3. The signal processing in CB is very simple and more robust to channel impurities as compared to the ZF which relies on the pseudo-inverse of the channel matrix. Therefore, **the idea of achieving close performance to ZF by CB using smart antenna layouts has been proposed for the first time**, which is very important to reduce the complexity and computational burden on the mm-wave 5G systems.

4. For the ground-only users with homogeneous user distribution in a cell area and for users/relays at different heights (with a narrow angular coverage in elevation), it has been seen that most of the users are located in a limited angular sector in elevation. In these cases, even a simple rotation of the array can increase the QoS and the uniformity of the excitation amplitudes, due to its low average SLLs inside the cell sector. Therefore, **45-degree rotated version of the generic square 5G arrays have been proposed to be used at 5G base stations, for the first time**, with CB, ZF or simplified ZF.

5. ZF precoding has been observed to provide the best QoS (ideally with zero interference), but with the most computational complexity (three-order of magnitude larger than CB). Therefore, **reduced complexity ZF algorithms (with up to an order of magnitude less complexity than the ideal ZF) have been proposed** that are based on cancelling only a certain number of interferers for each user instead of all. It has been seen that such algorithms perform better than CB in terms of the statistical SINR even with the cancellation of only the strongest interferer, which makes them a good compromise between the computational burden and QoS.

Based on the system analyses performed in this chapter, the following main questions arise, which are discussed in detail and answered rigorously in the upcoming chapters:

1. How should the LoS system model be modified to handle a more realistic NLoS downlink scenario at mm-wave frequencies? Is it possible to use a similar LoS-like model with the lowest system complexity by considering the single strongest multipath component or is it needed to take into account several multipath components? (Chapter 3).

2. How could the complexity of multiple beamforming architectures be reduced while keeping relatively good system performance? (Chapter 4-5).

3. What are the possible efficient and effective ways to suppress the inter-beam interference and improve the statistical QoS performance further? (Chapter 6).

4. How could safe and reliable communications be maintained in the presence of a large heat generation within the array? What are the energy-efficient and effective thermal management strategies for 5G antennas? (Chapter 7-8-9).

3

mm-Wave Propagation in Urban Outdoor Environments - LoS vs NLoS

Parts of this chapter have been published as:
[C-2] Y. Aslan, J. Puskely, A. Roederer and A. Yarovoy, "Performance comparison of single- and multi-lobe antenna arrays in 5G urban outdoor environments at mm-waves via intelligent ray tracing," in *Proc. 14th EuCAP*, Copenhagen, Denmark, Mar. 2020.

3.1. INTRODUCTION

Approaches to base station (BS) antenna synthesis and beamforming (BF) in future communication systems are expected to be capacity-driven, rather than directivity-driven which has been traditionally used with free-space based performance evaluation metrics [24]. Therefore, it is crucial to incorporate the channel and propagation aspects in the process of selecting the optimal BF strategy in 5G.

Ray tracing is one of the most commonly used methods to identify the directional characteristics of a channel and predict the received signal level for coverage and interference assessments [25]. However, the accuracy of ray phases and amplitudes in ray tracing is limited due to simplifications in the building database. Still, the major interest is in the spatially averaged received power over a small area, which is acceptably approximated in ray tracing by the sum of the ray powers [26].

In [27], using ray tracing, three major BF strategies were studied for mm-wave indoor applications: radial, single and multi BF. In radial BF, the environment is ignored and a beam is created considering the known positions of the BS and the user. As expected, such a strategy has the worst performance among the three, since the LoS path might be blocked in many cases. In single BF, the most dominant ray direction of a user is found via ray tracing with an omnidirectional BS antenna and a sharp, single-lobe beam with the optimal directivity is steered accordingly. On the other hand, in multi BF, multiple main lobes are formed (towards the first few strongest ray directions) and equalized to focus the power at the user. From the simulations in [27], it was inferred that single BF is the most efficient solution since there is a single dominant path in most of the cases. It was also claimed that multi BF performs as well as single BF in the case of multiple dominant paths and provides additional robustness against sudden changes in the channel (blockage, movement etc.). However, the higher co-channel interference caused by the multiple lobes and the corresponding decrease in the link quality was not discussed. Later in [28], using the single BF approach, an interference-aware beam-searching solution was proposed to have the highest possible minimum SINR in a multi-user indoor scenario. The price to pay was reported as the large overhead and time-consuming signaling that occur because of the more challenging CSI estimation procedure.

An important point that is commonly ignored by the propagation community is that the beamforming strategy should also be tailored to the front-end architecture of the multibeam antenna. In fact, there are three main beamforming strategies which can be used to serve a user in a NLoS scenario:

(S-1) Single beam with one lobe to strongest ray *(referred to as single BF in the literature)*.

(S-2) One multi-lobe beam and one single port where several slightly time shifted signals from the same stream are superposed and might be realigned by processing of the time spread signal *(referred to as multi BF in the literature)*.

(S-3) Multiple beams with one lobe and one port each, where signals from the same stream can be time and amplitude weighted before combining into one port/signal with increased SINR.

Among the three options, (S-3) has the most potential for maximizing the “gain”, but it limits re-use of the frequency band since a user occupies multiple beam ports and a

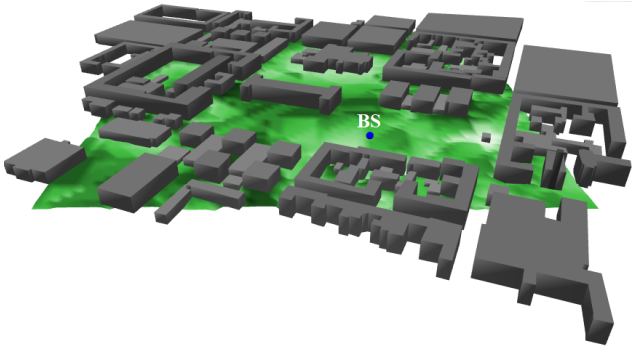


Figure 3.1: 3-D city view with the buildings and the base station (BS).

high interference is generated towards multiple directions. Considering this aspect, and the direction of current interest of the propagation community on mm-wave beamforming strategies [24; 27; 28], we focus on the comparison of approaches (S-1) and (S-2) in this chapter. To the authors' knowledge, there is no prior study that compares the performance of single- and multi-lobe BF in an urban outdoor environment. In this work, we compare the two candidate BF methods in a sample city environment with a focus on their performance regarding the intended received power and the unwanted co-channel interference.

3.2. SIMULATION SETTINGS

The propagation environment, simulation scenario, applied settings and assumptions are summarized as follows:

- Altair HyperWorks WinProp - Propagation Modeling tool [29] is used with the Student License for the three-dimensional rigorous intelligent ray tracing simulations.
- The hybrid (deterministic + statistical) ray tracing solver of WinProp is run in the noncoherent[1] (power related, no consideration of phase) mode to plot the received power distributions. Later in the related tables, both the coherent and unchoherent signal summation results are provided.
- CST MWS is used to compute the antenna patterns.
- An in-house developed array synthesis technique (see Section 5.2.2) is used in MATLAB in order to find out the excitation coefficients of the shaped (multi-lobed) beam patterns.
- The contribution of rays are given by Fresnel coefficients (for reflection losses) and GTD/UTD (for diffraction losses), based on permittivity, permeability, conductivity and thickness of the materials.

[1] In the noncoherent mode, the phase of the carrier is lost, but the modulating signal, at much lower frequency, is preserved. Thus, there is a 3dB loss with respect to the coherent mode.

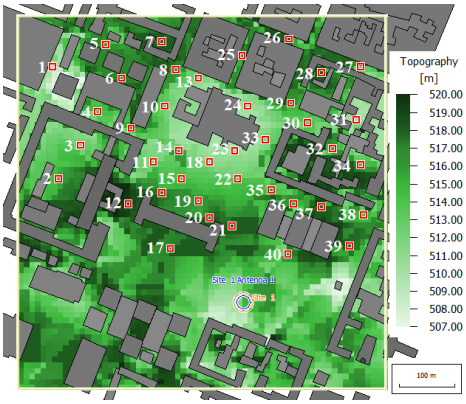


Figure 3.2: Topography and the positions of test users.

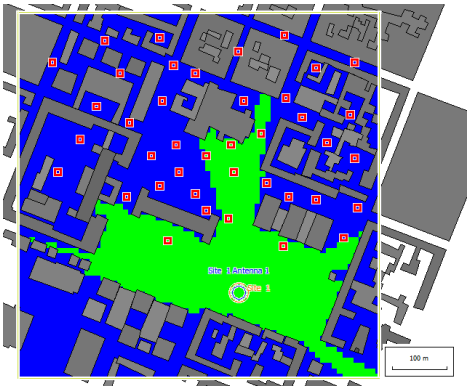


Figure 3.3: LoS area.

- Reflection from the ground is not taken into account, only the buildings at the fixed positions with the fixed shapes are considered, which is the sample city model "*city.oib*" provided by WinProp (see Fig. 3.1).
- Isotropic antenna at the users is assumed. (The system is not true single user MIMO compatible.)
- The frequency is set to 28 GHz.
- The resolution of the prediction results is 10 meters. In other words, the computation grid is set to 10 meter by 10 meter pixels (defined by the model, cannot be modified).
- Users are located at 2 meter height above the ground with the given topography (the topography and the value of 2 meters are dictated by the model and cannot be changed). For investigation purposes, 40 test users are randomly distributed over the city (see Fig. 3.2).
- Base station is at 15 meters above the ground. There is no tilting in elevation.
- Assumed polarization is vertical.
- In the power distribution visuals, the propagation results are filtered using a smoothing filter of order 3.
- The buildings are made of concrete with permittivity of 5.31 and conductivity of 0.48 S/m at 28 GHz [30].
- In all simulations, the transmit power is set to 1.6 Watts, assuming 20 dBm power amplifier output per element in a linear, $\lambda/2$-spaced 16 element array. The value of 20 dBm is set considering the state-of-the-art low-cost, silicon-based analog beamforming IC of NXP Semiconductors. Note that in this study, we are interested

in the relative changes in the received power at the users for different beamforming strategies at the BS. The true value of the chosen transmit power does not play an important role here.

- Path loss exponent is taken as 2 / 2.4 for LoS / NLoS before the breakpoint and 3.6 / 3.6 for LoS / NLoS after the breakpoint. The breakpoint is set to 500 meters. Note that these values are rough estimates for our test case and the real values should be computed by fitting the ray tracing results to the experiments as performed in [31].

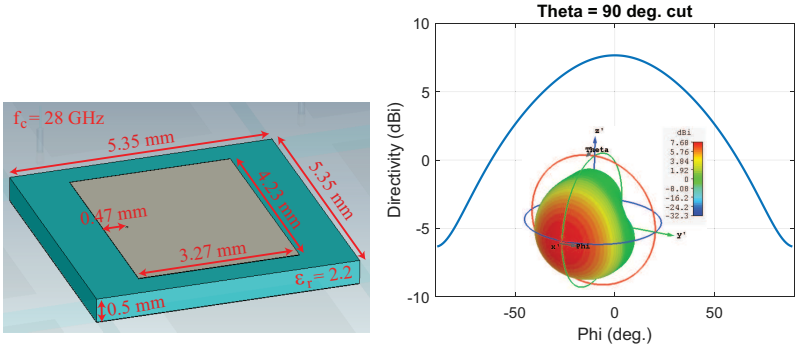
- Possible interaction types are reflections and diffractions.

- Maximum path loss of contributing rays is 200 dB and the maximum dynamic range per pixel is 100 dB.

3.3. SIMULATION RESULTS AND DISCUSSION

The LoS and NLoS areas are shown in Fig. 3.3. First, an omnidirectional antenna (with 0 dBi gain) is used at the BS and the received powers of the first 5 strongest rays towards each user are computed. The results are given in Table 3.1. Following the assumption in [27], it is assumed that if the difference between 1st and 2nd ray is larger than 5 dB, there is a single dominant path. Thus, it can be seen that most of the users (30 out of 40) have a single dominant path.

A pin-fed patch element is designed in CST MWS to be used in the array simulations. The design parameters and simulation results of the single antenna element are given in Fig. 3.4. In this section, the NLoS simulations will be grouped into two categories depending on the number of dominant paths and results will be given for several sample user positions.



(a) (b)

Figure 3.4: Single patch antenna element in CST: (a) design parameters, (b) directivity pattern.

Table 3.1: The received powers from the first 5 strongest rays towards each of the 40 test users in the case of an omnidirectional BS antenna. *(N/A means that the ray path loss is higher than 200 dB. The explanation of color coding is as follows: Green : Single dominant path - LoS. Yellow : Single dominant path - NLoS. Red : Two or more comparable main paths - NLoS.)*

User #	1st ray (dBm)	2nd ray (dBm)	3rd ray (dBm)	4th ray (dBm)	5th ray (dBm)
1	-131.75	N/A	N/A	N/A	N/A
2	-114.01	-156.26	N/A	N/A	N/A
3	-122.41	N/A	N/A	N/A	N/A
4	-122.28	-160.87	N/A	N/A	N/A
5	-127.80	N/A	N/A	N/A	N/A
6	-123.26	N/A	N/A	N/A	N/A
7	-138.84	N/A	N/A	N/A	N/A
8	-114.48	-114.70	-116.95	-120.25	-151.30
9	-127.03	-130.33	-142.92	-143.32	-144.34
10	-107.74	-112.16	-112.64	-115.34	-115.42
11	-107.84	-108.73	-113.98	-114.49	-114.87
12	-101.99	-104.97	-112.02	-117.87	-120.13
13	-109.24	-138.68	-142.39	-155.02	-159.91
14	-106.47	-108.48	-110.62	-114.80	-118.58
15	-98.15	-103.58	-104.22	-106.94	-107.73
16	-103.16	-108.42	-108.64	-108.93	-113.26
17	-74.26	-88.86	-99.47	-99.66	-100.82
18	-77.51	-97.36	-102.01	-102.24	-104.53
19	-103.89	-106.16	-106.44	-107.29	-109.03
20	-102.77	-103.43	-104.73	-116.39	-117.79
21	-72.88	-94.12	-98.00	-99.93	-103.09
22	-76.58	-100.70	-101.07	-105.42	-118.50
23	-77.97	-96.06	-101.83	-104.55	-106.83
24	-117.30	-140.09	-143.52	-148.80	-149.87
25	-121.56	-137.31	-141.12	-144.70	-156.83
26	-134.30	N/A	N/A	N/A	N/A
27	-137.41	N/A	N/A	N/A	N/A
28	-126.68	N/A	N/A	N/A	N/A
29	-115.97	-117.37	-120.18	-120.36	-124.43
30	-127.84	-155.90	-157.36	-159.10	-161.95
31	-131.07	N/A	N/A	N/A	N/A
32	-119.25	N/A	N/A	N/A	N/A
33	-78.66	-99.12	-113.89	-120.35	-130.61
34	-125.68	N/A	N/A	N/A	N/A
35	-90.49	-96.84	-99.38	-101.01	-103.59
36	-99.72	-102.33	-107.44	-107.77	-112.96
37	-96.18	-101.24	-110.48	-119.59	-123.05
38	-95.19	-103.29	-108.90	-135.05	-135.45
39	-93.96	-101.97	-103.22	-104.74	-110.30
40	-71.03	-88.79	-92.35	-96.69	-99.84

NLOS SIMULATIONS IN THE CASE OF HAVING A SINGLE DOMINANT PATH – USERS #4, #13, #37

For Users #4, #13 and #37, there is a single dominant path shown in Fig. 3.5(a). In this case, a single-lobe beam is formed towards this path using progressive phase shifts in a 16-element $\lambda/2$ spaced array of the designed patch antenna. The results are provided in Table 3.2. When compared to Table 3.1, it can be seen that the magnitude of the 1st ray is increased by the amount of the array gain, as expected.

Table 3.2: Angle of main lobes, array directivities and received powers from the first 5 strongest rays for User #4, #13 and #37.

Angle of main lobe(s) (degrees)	Array directivity (dBi)	1st ray (dBm)	2nd ray (dBm)	3rd ray (dBm)	4th ray (dBm)	5th ray (dBm)
-37	17.38	-105.25	-161.33	N/A	N/A	N/A
-11	17.75	-91.50	-123.63	-125.68	-158.16	-161.33
66	16.43	-80.16	-85.22	-107.15	-108.10	-112.97

Table 3.3: Angle of main lobes, array directivities and received powers from the first 5 strongest rays for User #8.

Angle of main lobe(s) (degrees)	Array directivity (dBi)	1st ray (dBm)	2nd ray (dBm)	3rd ray (dBm)	4th ray (dBm)	5th ray (dBm)	Power sum (dBm)	Coherent sum (dBm)
-3	17.80	-96.95	-99.42	-112.47	-134.86	-136.38	-94.92	-91.18
-15	17.71	-97.13	-115.41	-117.89	-120.78	-141.23	-97.01	-94.95
-3 & -15	15.74 & 15.06	-99.42	-99.78	-101.89	-131.27	-138.15	-95.46	-90.64

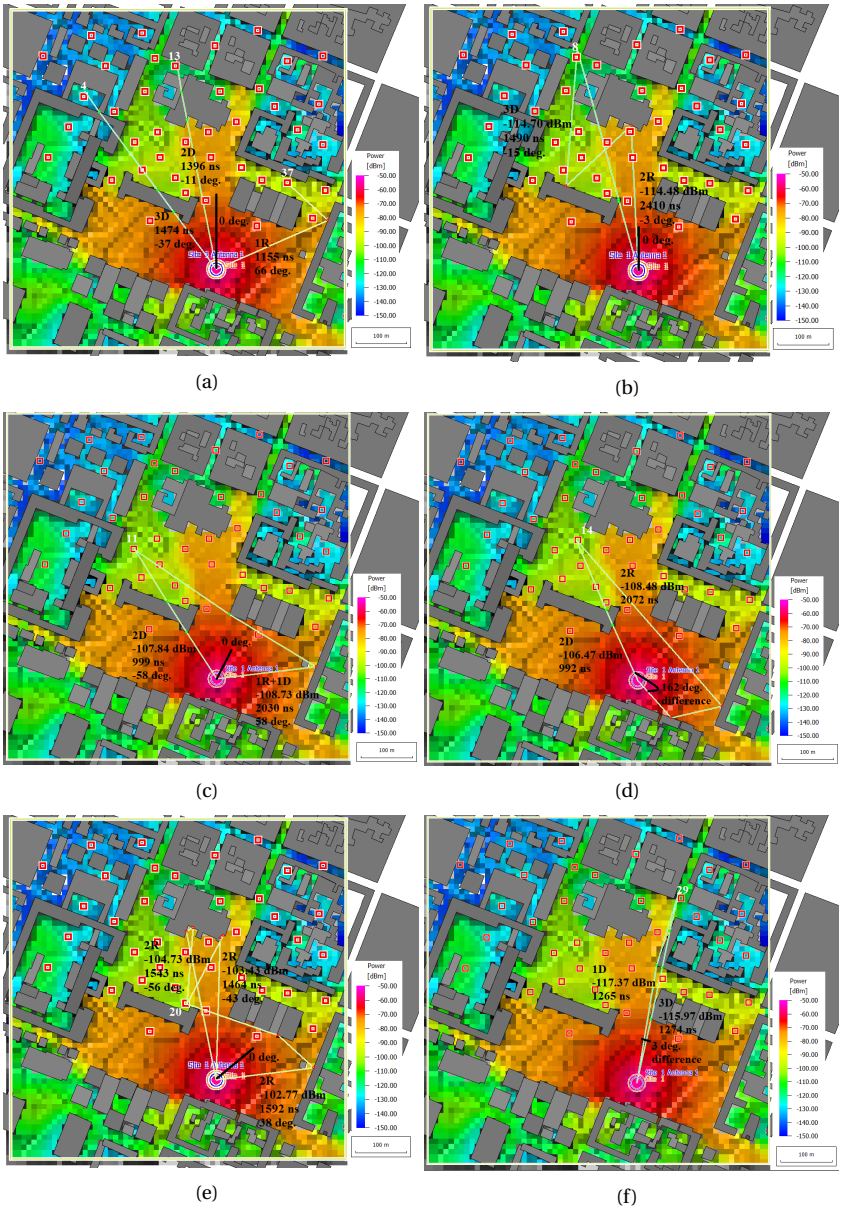
Table 3.4: Angle of main lobes, array directivities and received powers from the first 5 strongest rays for User #11.

Angle of main lobe(s) (degrees)	Array directivity (dBi)	1st ray (dBm)	2nd ray (dBm)	3rd ray (dBm)	4th ray (dBm)	5th ray (dBm)	Power sum (dBm)	Coherent sum (dBm)
-58	16.87	-91.63	-109.78	-112.12	-116.02	-118.79	-91.50	-89.20
58	16.87	-92.30	-98.53	-109.33	-116.08	-119.23	-91.28	-87.50
-58 & 58	14.09 & 14.13	-94.25	-94.77	-101.00	-115.54	-116.12	-91.00	-86.06

NLOS SIMULATIONS IN THE CASE OF HAVING TWO OR MORE DOMINANT PATHS – USERS #8, #11, #14, #20, #29

Here, we will study the effect of using single- and multi-lobe beams (in the 16-element $\lambda/2$ spaced array) on the received power of the users having more than one dominant path.

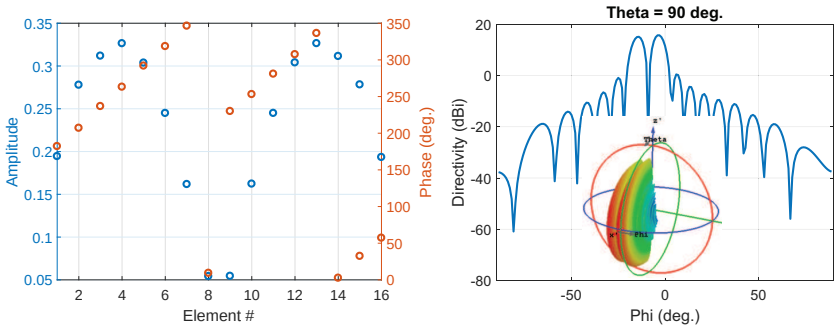
For User #8, there are two main paths positioned at -3 and -15 degrees off-broadside (see Fig. 3.5(b)). The single-lobe beams are formed using progressive phase shifts, while beam shaping is applied to form the two-lobe pattern. Fig. 3.6 provides the excitation coefficients and radiation pattern of the two-lobe beam. The simulation results for User #8 are summarized in Table 3.3. We see that forming a single beam at -3 degrees performs the best in the non-coherent mode, which is followed by the two-lobe beam. This is due to the gain drop in the two-lobe case and the non-significant contribution from the 2nd ray in the single-lobe case, which is also in the direction of -3 degrees with diffraction instead of reflection. In the fully-coherent mode, the two-lobe beam increases the re-



(a) (b) (c) (d) (e) (f)

Figure 3.5: Interaction type and number, time delay and angle of departures in the case of an omnidirectional BS antenna for Users (a) #4, #13, #37, (b) #8, (c) #11, (d) #14, (e) #20, (f) #29.

ceived power only by 0.5 dB as compared to the closest single beam option. This level of increase depends on the changes in the propagation scenario for different BF strategies. In this case, the difference is not so significant since a single lobe beam towards -3



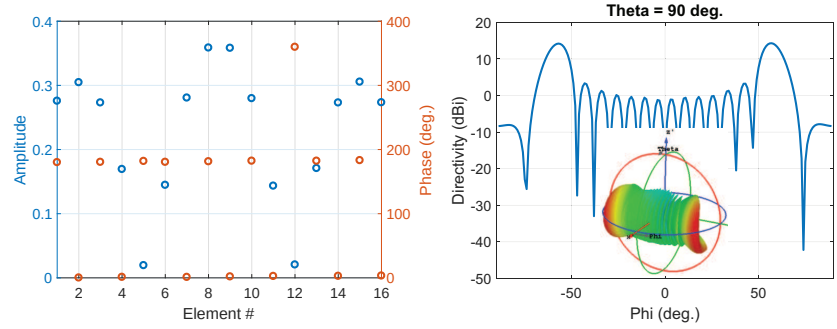
(a) (b)

Figure 3.6: Two main-lobe 16-element array synthesis for User #8 with beams at -3 and -15 degrees: (a) amplitudes and phases, (b) radiation pattern.

degrees gives a strong second ray contribution from the same direction, which is seen as the third ray in Table 3.1.

For User #11, the BS antenna is mechanically rotated in azimuth by 25 degrees to steer the broadside and the two main paths are oriented towards ±58 degrees (see Fig. 3.5(c)). Fig. 3.7 provides the excitation coefficients and radiation pattern of the two-lobe beam. The simulation results for User #11 are summarized in Table 3.4. In this case, the two-lobe beam pattern performs slightly better than the single-lobe pattern in the non-coherent mode. The gain drop in the two-lobe beam is compensated by the addition of the ray intensities originated from the two main lobes. In the fully-coherent mode, the two-lobe beam increases the received power by 1.4 dB as compared to the closest single beam option.

For User #14, two strongest paths have a wide angular separation (see Fig 3.5(d)). In such cases, two (probably out of three or four) sectors can cooperate to serve the user.



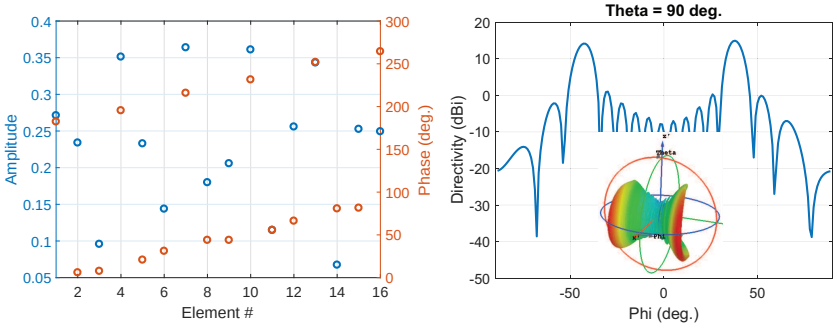
(a) (b)

Figure 3.7: Two main-lobe 16-element array synthesis for User #11 with beams at -58 and 58 degrees: (a) amplitudes and phases, (b) radiation pattern.

For User #20, the BS antenna is mechanically rotated in azimuth by 45 degrees to steer the broadside and the main paths are oriented towards 38, -43 and -56 degrees (see

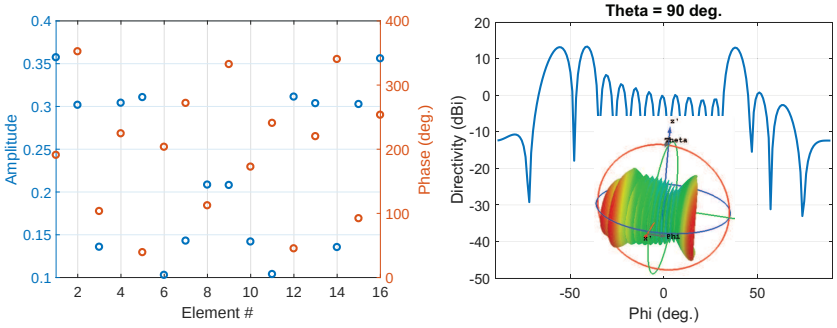
Table 3.5: Angle of main lobes, array directivities and received powers from the first 5 strongest rays for User #20.

Angle of main lobe(s) (degrees)	Array directivity (dBi)	1st ray (dBm)	2nd ray (dBm)	3rd ray (dBm)	4th ray (dBm)	5th ray (dBm)	Power sum (dBm)	Coherent sum (dBm)
38	17.37	-85.52	-109.64	-110.69	-112.66	-112.91	-85.47	-83.91
-43	17.28	-86.30	-101.65	-107.22	-110.34	-117.71	-86.12	-83.69
-56	16.94	-88.29	-101.08	-104.65	-110.74	-111.44	-87.93	-84.62
38 & -43	14.93 & 14.14	-88.01	-89.36	-104.89	-109.62	-112.13	-85.54	-81.38
38, -43 & -56	12.93, 13.27 & 13.14	-90.09	-91.00	-91.90	-105.01	-114.22	-86.10	-80.68



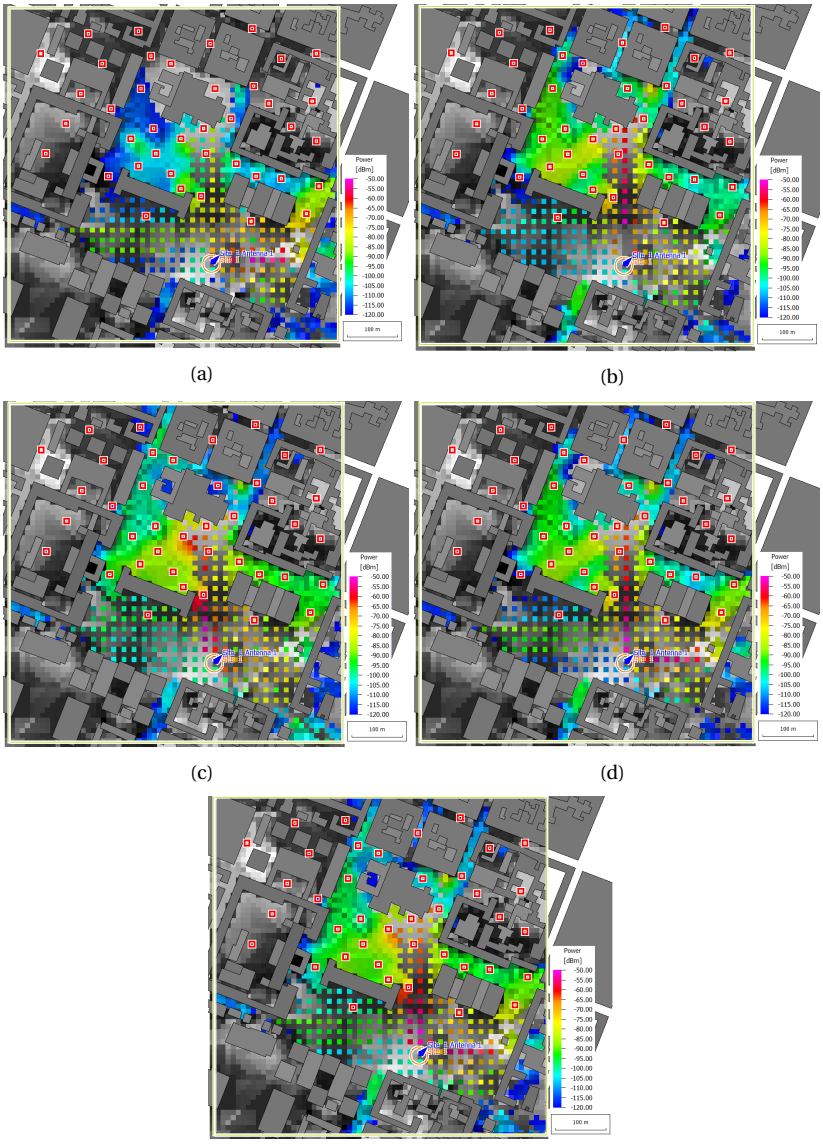
(a) (b)

Figure 3.8: Two main-lobe 16-element array synthesis for User #20 with beams at 38 and -43 degrees: (a) amplitudes and phases, (b) radiation pattern.



(a) (b)

Figure 3.9: Three main-lobe 16-element array synthesis for User #20 with beams at 38, -43 and -56 degrees: (a) amplitudes and phases, (b) radiation pattern.



(a) (b) (c) (d) (e)

Figure 3.10: The distribution of received power at the other (potentially co-channel) NLoS users while serving User #20 with: (a) a single lobe at 38 degrees, (b) a single lobe at -43 degrees, (c) a single lobe at -56 degrees, (d) two lobes at 38 and -43 degrees, (e) three lobes at 38, -43 and -56 degrees *Note: In the case of importing a directional beam, there is a bug in the software which does not allow to see a smooth power distribution in the LoS region.*

Fig. 3.5(e)). A two-lobe (38 and -43 degrees) and a three-lobe (38, -43 and -56 degrees) beams are formed. Fig. 3.8 and Fig. 3.9 provide the excitation coefficients and radia-

tion pattern of the two-lobe and three-lobe beams, respectively. The simulation results for User #20 are summarized in Table 3.5. In this case, steering the beam towards the strongest ray direction (38 degrees) provides the best result in the non-coherent mode. However, the total received power values are very close to each other, especially in the case of having a single-lobe at 38 degrees and two-lobes at 38 and -43 degrees. In the fully-coherent mode, the three-lobe beam increases the received power by 3 dB as compared to the closest single beam option.

Note that the advantage of the fully-coherent mode in increasing the received power is obtained for constant Tx power. If the Tx power is tapered to get the desired element weights in the multi-lobed beams, the advantage is lost due to the EIRP loss.

Furthermore, the interference towards the other users is different in each case. Therefore, the selection of beam position should be made depending on the other users sharing the same time/frequency slot. The received power distribution for each beam position in Table 3.5 are given in Fig 3.10. It is seen that the multiple lobes increase the inter-user interference and in practice, forbids the re-use of the same frequency around the main lobes. Besides, multilobe beams are not good for MIMO situations.

For User #29, two strongest paths are not resolvable for the 16 element array (see Fig. 3.5(f)). In such a case, a single-lobe beam can cover both paths.

For completeness, the simulations with the two-lobe beams were repeated using the coherent mode of WinProp which takes into account the uncompensated ray phases [32]. It was seen that compared to the non-coherent mode, the received power is 3.9 dB less, 5.4 dB less and 1.5 dB more in the coherent mode for User #8, User #11 and User #20, respectively. This study shows that the result depends very much on the relative phases of the dominant paths. Ideally, with the accurate channel state information at the receiver or transmitter, all the incoming rays should be equalized in time and phase by employing a Rake (at receiver) or Time Reversal (at transmitter) architecture, which yields the fully-coherent field sums in Tables 3.3-3.5. However, such a case increases the complexity of the system and, as already stated in Section 3.1, does not comply with the limited-accuracy ray tracing simulations.

3.4. CONCLUSIONS

In this chapter, ***the intended received power and unwanted co-channel interference performances of single-lobe and (adaptively-combined) multi-lobe beam arrays have been studied in a 5G urban outdoor environment at mm-waves***. A sample urban city model has been considered with randomly selected user locations. Ray tracing simulations have been performed (at 28 GHz) to compute the ray directions, interactions and intensities. An omnidirectional pattern has been used at the BS to identify the characteristics of the dominant paths towards each user. It has been observed that, **in most of the cases, there is only a single dominant path for which the single-lobe beam with the optimal directivity can be steered** using a phased array by applying progressive phase shifts. For the NLoS users having more than one dominant path, the following three strategies are considered:

(S-1) Single beam with one lobe to strongest ray.

(S-2) One multi-lobe beam and one single port where several slightly time shifted

signals from the same stream are superposed and might be realigned by processing of the time spread signal.

(S-3) Multiple beams with one lobe and one port each, where signals from the same stream can be time and amplitude weighted before combining into one port/signal with increased SINR.

One could use multiple simultaneous single lobed beams and ports per user and recombine signals coherently after time synchronization of each. That should provide the optimal gain, but limits greatly same frequency re-use by others.

As for (S-1) and (S-2), it has been concluded that, unless the receiver (or transmitter) has the accurate channel state information and a complex Time Reversal[2] (or Rake) architecture to equalize the ray phases, **steering a single-lobe beam towards the dominant path could yield comparable or better received powers than the multi-lobe ones which suffer from directivity loss**. Besides, the single-lobe beam option generally has a key spectrum efficiency and flexibility advantage over the multi-lobe because it has optimum interference-cancellation capability towards the other co-channel users. On the other hand, **the multi-lobe practice automatically forbids re-use of the same band in and around the other lobe directions**. The major advantage of multi-lobe pattern is its spatial diversity that can be helpful in sudden, unexpected blockage of the main path in the direction of the single-lobe beam. Such an advantage is only relevant when the CSI updating interval is relatively long. It is worth to mention that the CSI updating frequency of a communication system is generally set considering the trade-off between the CSI estimation accuracy and the overhead of CSI acquisition [33].

In summary, **among the three NLoS beamforming strategies listed, (S-1) seems the best from the system perspective**, if it satisfies the minimum SNR threshold. The approach (S-2) has less system gain than that of (S-1) and (S-3), and has a frequency re-use problem due to interference as in (S-3), but less strong. If the signal level remains below the threshold with (S-1), option (S-3) could be preferred over (S-1) and (S-2).

[2] Although with an omnidirectional user antenna, and a fortiori with a single port multi lobe beam at the BS (i.e. approach (S-2)), Time Reversal would generate a one stronger cumulated signal, lots of unwanted small signals will also be observed. The undesirable signals are filtered out when multiple beams with one lobe and one port each (i.e. approach (S-3)) are used instead of the multiple lobes.

II

Electromagnetism Driven Aspects

It is electromagnetism in all its many forms
that has been so basic, that haunts us and guides us.

Nick Holonyak

4

Multibeam Antenna Array Architectures for 5G - An Overview

Parts of this chapter have been published as:
[C-3] Y. Aslan, J. Puskely, A. Roederer and A. Yarovoy, "Active multiport subarrays for 5G communications," in *Proc. IEEE APWC*, Granada, Spain, pp. 298-303, Sept. 2019.

4.1. INTRODUCTION

In multi-user applications, using distributed single beam antennas (SBAs) is currently proposed for 5G (Massive) MIMO base stations by Nokia Bell Labs [34], Ericsson [23], IBM [35] and NXP [C-7], which results in a multiplicity of single beam antennas, generally acive phased arrays with analog beam forming and digital MIMO processing. The more ambitious alternative is to generate all the agile beams simultaneously from a single multiple beam antenna (MBA). The IEEE definition of MBA is as follows [36]: "An antenna capable of creating a family of major lobes from a single nonmoving aperture through the use of a multiport feed, with one-to-one correspondence between input ports and member major lobes, the latter characterized by having unique main-beam pointing directions". Despite being relatively new in the domain of cellular communications with the introduction of 5G, MBAs have been used since decades for space applications. For example, with the Boeing Spaceway system [15], the same 500 MHz band is reused 24 times in 24 simultaneous hopping beams around 20 GHz (they hop into pre-programmed cells with about -1dB crossover between cells).

The multiple beam antennas (MBAs) can roughly be grouped into two categories: (i) passive/active multiple fixed beam antennas (MFBAs), (ii) passive/active multiple beam phased array antennas (MBPAAs). The phrase 'active' means that the LNA and PA are placed just behind each antenna element or subarray, which helps compensate the losses in the beamforming network and generate larger RF power.

MFBAs are mainly based on reflectors [37], lenses [38] or beamforming circuits (transmission lines, directional couplers etc.) [39]. At the output, a fixed number of beams are created only in specific directions. Unless there is a large number of highly overlapped beams, such a limitation is not compatible with the demanding 5G performance criteria that require flexible beamforming and versatility. The passive and active MBPAAs, on the other hand, have the ability to scan the simultaneously created beams. As indicated in [40], active MBPAAs can achieve much better power efficiency and system linearity than the passive MBPAAs[1]. Therefore, it can be inferred that active MBPAAs are the most suitable candidates for simultaneous multi-user 5G communication.

In MBAs, the beams can be generated in various ways depending on the beamforming strategy. The possible options are based on fully-analog, fully-digital and hybrid schemes. Among these options, fully-analog MBAs are not as flexible, versatile and robust-to-failure as the fully-digital ones. They also suffer from combining losses that must be compensated by the amplifier gain. On the other hand, fully-digital MBAs suffer from high design cost and complexity. Therefore, being an attractive compromise between the performance and complexity, several hybrid beamforming strategies have recently been proposed in the 5G literature [41]. In the past few years, hybrid topologies received a lot of attention, especially in the signal processing community [42; 43], due to the need for highly-complex beamforming algorithms. Using elementary antenna elements, it has been shown that baseband processing and RF beamforming can be efficiently combined, which comes at the expense of a radiation pattern performance degradation. Therefore, for the antenna community, it is still a challenge to make use of

[1] Passive multiple beam arrays are in practice never used because there is nothing to compensate the combining losses, $10\log(N_{beam})$dB at each element, where N_{beam} denotes the number of simultanous beams. So, with 10 beams one gets roughly 10% of the efficiency that would be achieved with an active array.

smart antenna designs that help reduce the design complexity, computational burden and required resources, while keeping sufficiently low inter-beam interferences.

In this chapter, the existing beamforming architectures for the 5G base station antennas are reviewed with examples of the state-of-the-art prototypes both from the industry and academia. The advantages / disadvantages of each topology are discussed with multiple performance criteria in terms of the radiation pattern properties (gain, side lobes etc.), design / fabrication complexity, signal processing burden and cost.

4.2. MULTIPLE BEAM GENERATION STRATEGIES - PROS AND CONS

In this section, first, some of the existing active 5G beamforming architectures will be covered and the advantages/disadvantages of each will be explained, taking into account the current status of the industry. Later, a novel active multiport subarray based array topology will be introduced considering the trade-offs between the design complexity, radiation performance and digital signal processing (DSP) burden. Interested readers are referred to [7] for a deeper discussion on the possible antenna technologies for 5G.

In 5G, ideally, each stream needs one own beam. Otherwise, if several streams share the same beam, as was the case in the previous generations, the system will not be optimum. Currently, the industrial baseline 5G antennas are mostly based on the classical analog MIMO array (possibly with additional baseband processing) which is shown in Fig. 4.1(a). In this configuration each stream only uses a single array (with P elements) among many (= Q) closely located or distributed single beam arrays. Therefore, the full potential of the overall antenna array is not exploited, which results in relatively wide beams with low gain unless each array is large enough. Besides, the power sharing flexibility and efficiency between all beams and streams is not at all as good as with a true MBA where all the power can be put in a single stream, if needed. Several single beam array examples from the industry are provided in Fig. 4.2. The arrays could also be hybrid single beam arrays, i.e. with sub-arrays of elements each with a T/R module with analog phase shifter and true time delay at sub-array level (to improve the bandwidth performance), as used for radar applications. Note that the distribution/separation of the single beam arrays might also bring some potential MIMO improvement.

Fig. 4.1(b) shows an alternative structure which will be called as the true multi-beam fully-connected analog array. The massive advantages of this topology as compared to classical MIMO arrays with the same total number of elements are: (i) each stream benefits from the full gain (and if needed full power) of all the array elements and not only from the gain of one “sub-array", in other words, the gain is multiplied by Q, (ii) the power needed per stream is divided by Q for the same EIRP, (iii) since the beams are Q times narrower, it is possible to re-use the same frequency Q times more thus, in theory, multiply the spectrum efficiency by Q. However, in true multi-beam fully-connected arrays, there is a large number of phase shifters and adders (with combining losses[2]) and

[2] A substantial loss ($10\log(N_{beam})$dB) occurs at combiners into element chains. The loss has to be compensated in the PA's, which is not a big issue in space MBA's, but it is a big problem with CMOS PA's. Therefore, beamforming and combining at intermediate frequency (IF) has been proposed and implemented, which is shown to be a lot easier with lower level power amplification at IF before up-conversion [7].

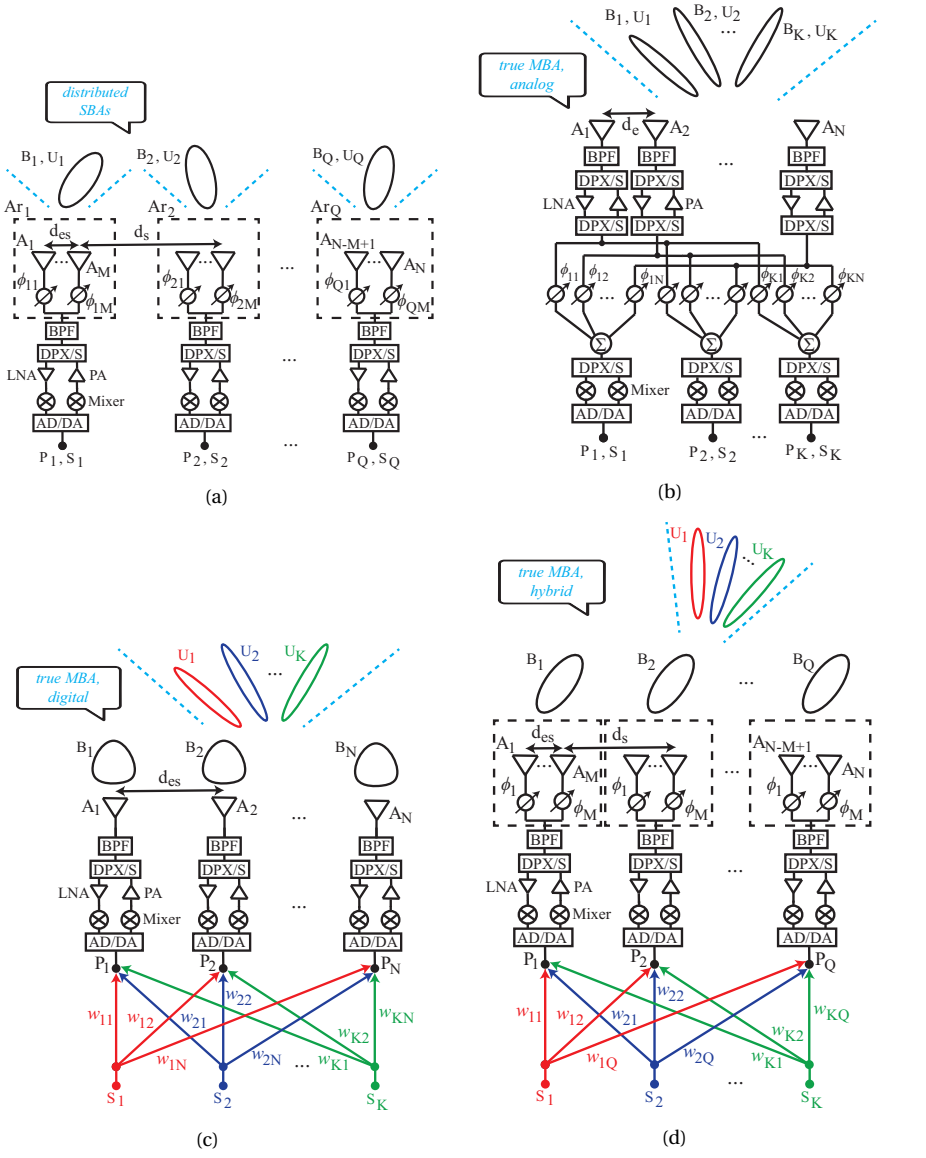
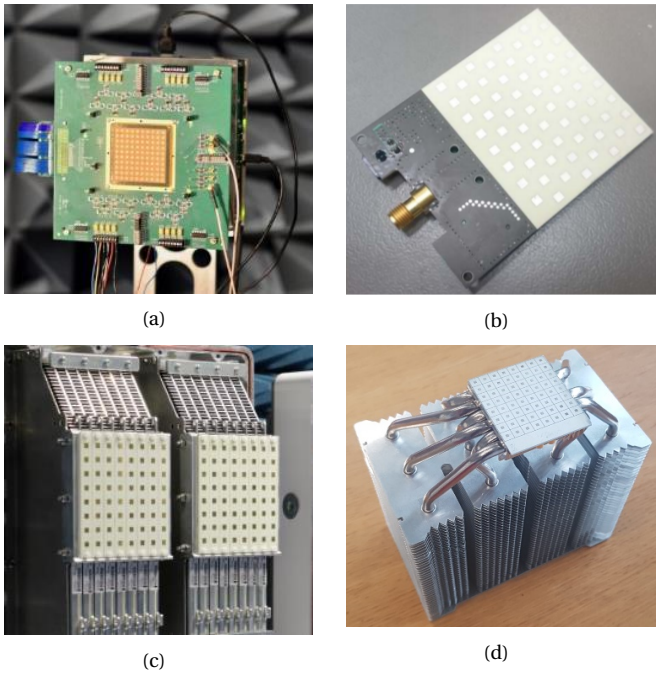


Figure 4.1: Potential active multi-beam generation architectures proposed in the 5G literature [C-3]: (a) classic analog MIMO array, (b) fully-connected analog array, (c) fully-digital array, (d) array of phased subarrays. (A: antenna element, Ar: antenna array, P: antenna port, S: user stream, B: antenna beam, U: user beam)

each beam has its own beamforming network. Due to the system complexity, the literature is not rich regarding the implemented analog multi-beam arrays. Some examples can be found in radio astronomy [44], satellite communications [45; 46], radar defense [47] areas.



(a) (b) (c) (d)

Figure 4.2: 5G state-of-the-art active integrated phased base station antenna arrays from the industry, (a) IBM [35], (b) Nokia Bell Labs [34], (c) Ericsson [23], (d) NXP Semiconductors [C-7].

Fig. 4.1(c) presents the fully digital architecture which is considered as the final goal since it is able to provide the most flexible, accurate and versatile performance via beamforming in the baseband. Each stream uses all the antenna elements and each element has a separate RF chain, but no phase shifters or adders are used. The major issues with the fully digital architectures are the cost and complexity, which increase significantly with the number of simultaneous beams and the number of array elements. Earlier designs with fully-digital arrays were mainly used in military radar applications [48] and satellite communications [49]. Later, with the development of advanced DSP chips, the application domains have been extended to personnel imaging [50], automotive radar [51] and so on. Very recently, the first 5G fully-digital array hardware was presented in [52]. However, the industrial high-volume 5G market is still far from that due to many practical factors such as cost, design complexity, cooling and computational burden.

Considering the drawbacks of the fully-digital arrays, hybrid architectures have been commonly proposed for 5G base stations. Due to having less number of RF chains, hybrid beamforming can lower the system cost and complexity. However, the price to pay is the formation of high side/grating lobes. The most commonly used hybrid architecture, namely the array of phased subarrays is given here in Fig. 4.1(d). In this case, the covered sector can be steered using the phase shifters in the subarrays. However, since the

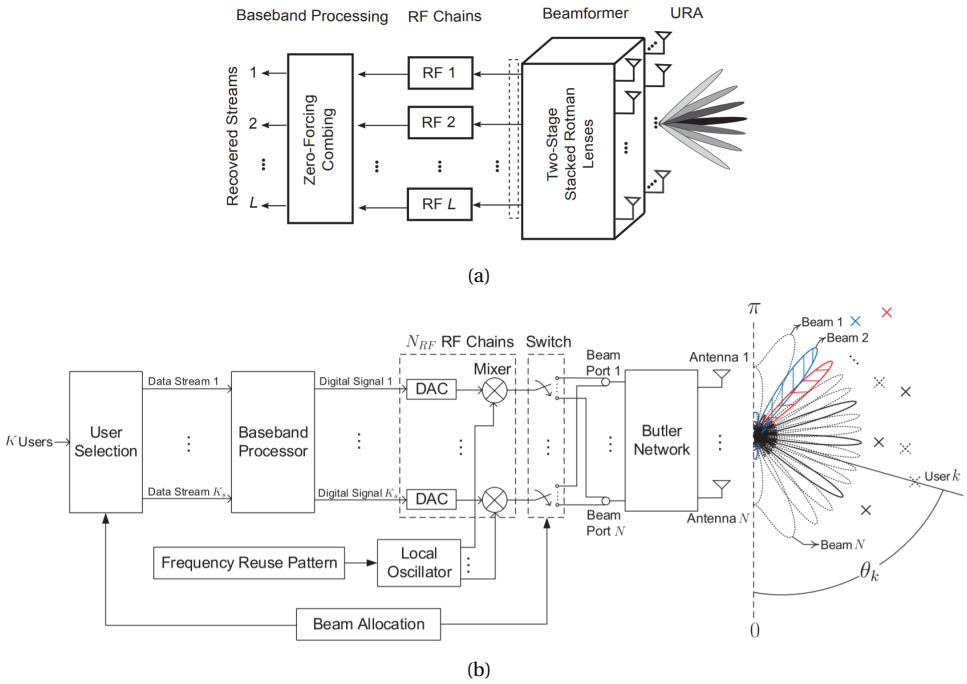


Figure 4.3: Single passive multiple fixed beam antenna array topologies in the recent literature: (a) an uplink scheme with a Rotman lens [53], (b) a downlink scheme with a Butler Matrix [54].

angular coverage range is limited by the subarray pattern, only a narrow sector[3] (defined by the number of subarray elements, M) can be covered by multiple beams. Besides, because of the large inter-spacing (d_s) between the subarrays, grating lobes might occur within the sector, especially for relatively large scan angles [7].

Very recently, hybrid beamforming schemes using a single passive multiple fixed beam antenna (Rotman lens [53] or Butler matrix [54]) have also been proposed to achieve more simplicity through exploiting multiple fixed analog beams covering a wide sector and controlling them digitally in the baseband. However, it is well-known that in 5G (especially at mm-waves), sharp beams with large gains are needed to satisfy the link budget and allow sufficient frequency reuse. This may lead to large (N x N where $N \geq 16$) multiple fixed beam antenna structures that can result in bulky designs with high design complexity, insertion loss, power inefficiency and parasitic radiation. In fact, such passive configurations will never work in practice at mm-wave frequencies unless active T/R amplifiers are added to power each element. Also, there is 3dB (1-D) to 5dB (2-D) crossover loss and need for switching. For completeness, Fig. 4.3 shows sample schemes for

[3]The sector width limitation here is given by considering relatively small-sized subarrays, as reported in [7]. The angular coverage can be increased by generating a fat or hat-like subarray sector pattern with the phase shifters, but then many elements are required per subarray to achieve the beam shaping, there is gain loss and the grating lobe problem becomes more significant.

Table 4.1: Comparison of the existing multiple beam generation strategies for 5G base station antennas. *(Note that it is desired to have a narrow beamwidth (to increase frequency re-use), wide angular scanning range (to reduce the number of array panels), high array gain (to compensate the propagation loss and reduce the electricity consumption), high number of SDMA users (to maximize the capacity), and low RF & digital complexity with a low implementation cost.)*

	Distributed single beam analog/hybrid arrays	Analog, fully-connected array	fully-digital array	Hybrid, array of phased subarrays	Hybrid, passive multiple fixed beam arrays
Beamwidth	Wide	Narrow	Narrow	Narrow	Narrow
Scanning range	Wide	Wide	Wide	Narrow	Wide
Array gain	Low	Moderate	High	High	Moderate
Number of SDMA users	Low	Moderate	High	Moderate	Moderate
RF complexity	Low	High	High	Moderate	High
Digital complexity	Low	-	High	Moderate	Low
Cost	Low	High	High	Moderate	Moderate

the single passive multiple fixed beam antennas found in the recent literature.

As previously mentioned, hybrid architectures are getting a lot of attention from the signal processing experts, but the implemented 5G hybrid beamforming prototypes are still very limited [55; 56].

4.3. CONCLUSIONS

In this chapter, ***the existing simultaneous multiple beam generation strategies for high-capacity multi-user 5G SDMA applications have been revisited.*** From the comparative analyses of different architectures, the following main features have been reported:

1. Using distributed single beam arrays results in relatively wide beams with low gain and resolution since the total array potential is not used. In other words, only a small part of the total available array elements is used for each stream with reduced gain and angular resolution.

2. In the analog, fully connected arrays, each stream benefits from the full gain of the array elements. Yet, the complexity and losses are high due to the need for a large number of phase shifters and combiners.

3. The fully digital arrays provide the most flexible, accurate and versatile performance. However, the cost, implementation & processing complexity and heat generation are extremely high at the current technology level[4].

4. The hybrid arrays of phased subarrays use reduced number of RF chains (except the same number of T/R amplifiers at the elements) with lower system cost and complexity. Yet, the angular coverage is narrow and grating lobes might occur in the sector.

[4] It is predicted that, with the advance of technology, fully digital beamforming will win in the medium to long term because hardware complexity (i.e. T/R modules, up/down conversion chains, A/D, D/A converters do not increase much with the number of simulaneous streams. Also, digital beamforming is much better for MIMO and broadband operation with true time delays instead of phase shifts.

5. The hybrid, passive multiple fixed beam antennas can also reduce the system cost. However, since sharp beams with large gains are required in 5G, the array design becomes complex (to the point that this option can probably not be considered) and bulky, with high insertion loss, power inefficiency and parasitic radiation.

Table 4.1 provides a performance summary of the multiple beam generation strategies that have been discussed throughout this chapter. By simply counting the red points in the table, it can be observed that **on average, the hybrid multibeam architectures provide the best performance/cost trade-off**. Nevertheless, the array of phased subarrays suffer from narrow angular coverage, while the single passive multiple fixed beam antennas suffer from high implementation complexity and losses. Therefore, the main question is the following: "**Can we keep a hybrid (or hybrid-like) scheme, achieve wide coverage and have low implementation & processing complexity at the same time?**". This question provides the basis for the next chapter.

5

SUBARRAY BASED MULTIPLE BEAMFORMING CONCEPTS

Parts of this chapter have been/will be published as:
[C-3] Y. Aslan, J. Puskely, A. Roederer and A. Yarovoy, "Active multiport subarrays for 5G communications," in *Proc. IEEE APWC*, Granada, Spain, pp. 298-303, Sept. 2019.
[C-4] J. Puskely, Y. Aslan, A. Roederer and A. Yarovoy, "SIW based antenna array with power equalization in elevation plane for 5G base stations," in *Proc. 12th EuCAP*, London, UK, Apr. 2018.
[J-2] J. Puskely, T. Mikulasek, Y. Aslan, A. Roederer and A. Yarovoy, "5G SIW based phased antenna array with cosecant-squared shaped pattern," *IEEE Trans. Antennas Propag.*, 2020, under review.
[J-8] J. Puskely, Y. Aslan, A. Roederer and A. Yarovoy, "Analog beamforming cosecant squared pattern active phased array antennas for 5G cellular systems," *IEEE Antennas Propag. Mag.*, 2020, in preparation.

5.1. HYBRID ARRAY OF ACTIVE MULTIPORT SUBARRAYS

5.1.1. INTRODUCTION

Hybrid and active beamforming started in **single beam** radar systems. This is because the use of subarrays of M active elements each equipped with an analog T/R module with phase shifter, and possibly variable gain, has the following advantages:

1) To divide by M the number of up/down converters & of D/A and A/D converters needed for fully digital beam forming,

2) To divide by M the numbers of digital weights and true time delays generated with fully digital beam forming,

3) Less frequency dependent beam forming/pointing than with a fully analog array,

4) No subarray scan loss/grating lobes with RF variable phase shifters.

The disadvantages, on the other hand, are as follow:

1) Same number of analog PA's and LNA's with in addition RF variable phase shifters and some (but simpler) digital beam forming,

2) RF phases still need to be computed.

5

In this section, we propose the active **multibeam** multiport subarray phased array antenna (AMSPAA) architecture[1] shown in Fig. 5.1 which combines several (= Q) multiport subarrays with M elements and backed by M x M Butler Matrices (BMs) (or Butler-like Matrices that can be configured to the coverage sector, suppress one or two BM port(s) etc.) where M is relatively small so that the overall design is less lossy, more compact and power-efficient. The major advantage of the active multiport subarrays over the array of phased subarrays is the increased angular coverage via the use of multiple Butler beams. Besides, compared to their fully-digital counterparts, AMSPAAs greatly reduce the processing complexity since instead of all the antenna elements, the digital weights are applied only to the BM ports corresponding to the user positions.

The major advantages (A-i–A-iv) / disadvantages (D-i–D-iv) of the AMSPAAs can be summarized as follows:

(A-i) AMSPAAs can be straightforwardly integrated with 1D arrays or efficient 2D beamforming can be achieved with planar antenna layouts (see the simulation results).

(A-ii) The DSP complexity is remarkably less as compared to the fully-digital beamforming. For example, if we assume conjugate-beamforming (CB) or zero-forcing (ZF) pre-coding at the baseband, the number of floating point operations ($\mathscr{F}_{CB}, \mathscr{F}_{ZF}$) is given by [21]

$$\mathscr{F}_{CB} = K(14N-2)$$

$$\mathscr{F}_{ZF} = K(24(K-1)N^2 + 48(K-1)^2N + 54(K-1)^3 + 6N)$$

where K is the number of simultaneous users and N is the number of antennas in fully-digital arrays which is replaced by Q, the number of subarrays, in AMSPAAs. If we assume an array with $N = 32$, $Q = 8$, $K = 4$, DSP complexity of the AMSPAA becomes 25% and 10% of the fully-digital beamforming in the case of CB and ZF precoding, respectively. For an

[1] Phase shifters at the element level might also be needed in the end for the required trimming/calibration needed with PA temperature differences.

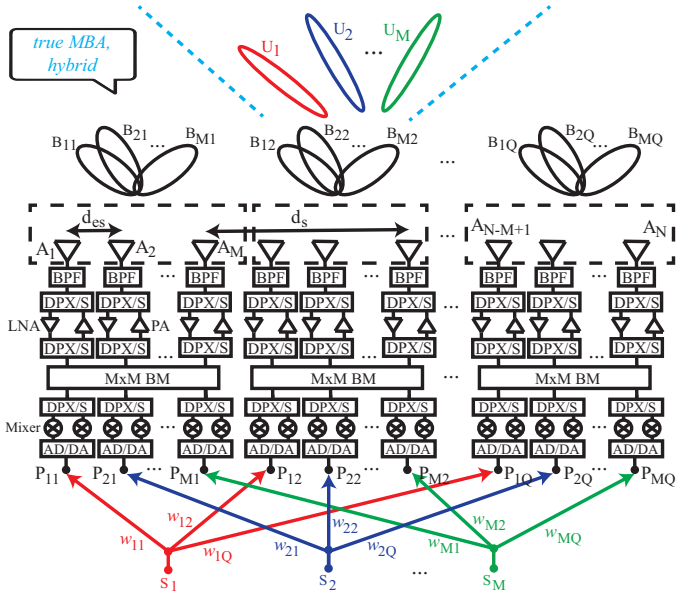


Figure 5.1: Proposed AMSPAA architecture with fixed-beam multiport subarrays.

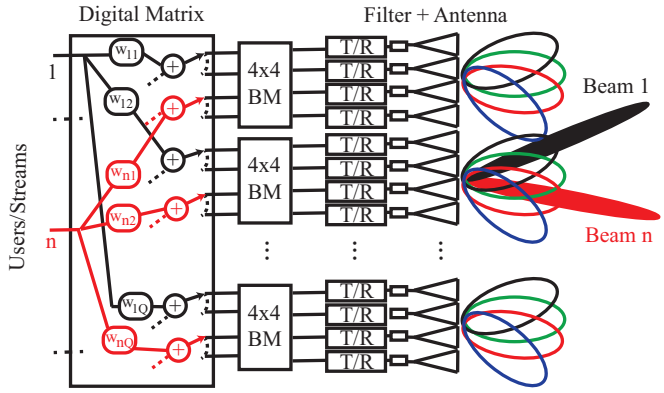


Figure 5.2: A sample use of multiport sub-arrays of 4 elements fed by 4x4 Butler Matrix to reduce DSP and hardware requirements.

array with $N = 128$, $Q = 16$, $K = 8$, the DSP complexity of the AMSPAA is reduced to only 12% and 3% of the fully-digital array for CB and ZF, respectively.

(A-iii) Optimally, M well-separated far-away users can be simultaneously served by making use of all the subarrays for each user (as shown in Fig. 5.1), with all the available gain in order to save power.

(A-iv) In the case of having a few well-separated simultaneous co-frequency users, it is possible to send the same user stream to multiple BM ports as in [57] and reduce the side lobe level with no additional DSP algorithms such as ZF. However, this comes at the expense of the increased beamwidth and reduced gain that occurs because of the

amplitude tapering in the subarrays.

(D-i) The hardware requirements of AMSPAAs become comparable to the fully-digital arrays if PAs are used at each antenna port to compensate the BM losses. However, in the case of having low loss matrices, the array could be powered at the subarray ports only, which reduces the number of amplifiers by a factor of M. In addition, there can be cases where all the matrices/sub-arrays input ports are not used at the same time (particularly in a 2D array) or where the number of simultaneous users is restricted. Then, it might be worth having less power greedy A/D and D/A converters than sub-array ports by introducing an RF (or IF) 'String to Sector Port' switching matrix. This concept is visualized in Fig. 5.2 where two users switch between the beam ports associated with the upper and lower angular sectors.

(D-ii) It is highly probable that two (or more) users positioned in the same Butler beam have to be served at different time slots or sub-bands due to the relatively wide beamwidth and thus, the large inter-beam interference. Besides, the number of simultaneously served far-away users (that need the gain from all the subarrays at the same frequency sub-band) is restricted by the number of BM ports. However, in [J-1], it was shown that the number of users cannot be very large ($K, M < 8$ for the current chip technologies) due to the thermal management problems which is preferred to be handled passively via natural convection and radiative cooling [J-6].

(D-iii) Due to large spacing between the subarrays, grating lobes (GLs) will occur in the field of view which may create very large interference if not suppressed. Besides, ZF algorithm does not work for two users located near each other's grating lobes [J-6]. However, placing the subarrays smartly can help suppress the GLs and allow us to use CB or ZF [J-6] with sufficiently low interference. An alternative way to dissolve the GLs is to use overlapped subarrays [58] (with a more complex analog beamformer) instead of the contiguous ones, which is not discussed in this section.

(D-iv) There will be a significant scan loss up to 3dB in linear and 5dB in planar BMs. This can be improved using variable orthogonal Butler-like matrices instead of fixed ones, at the expense of increased complexity, which is listed as an item in the recommended future research lines in Section 10.2.

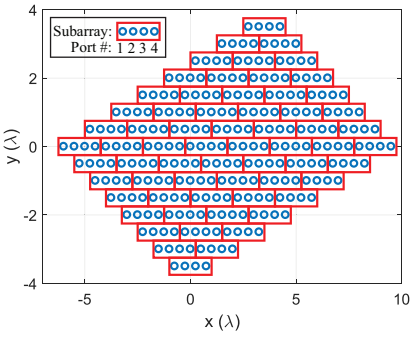


Figure 5.3: The proposed AMSPAA layout.

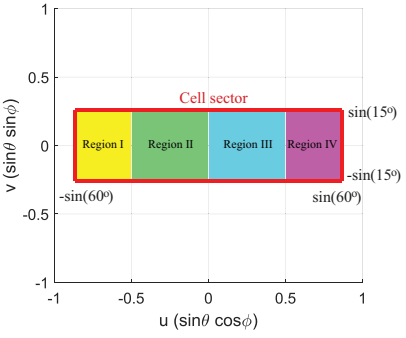


Figure 5.4: Division of angular regions corresponding to the multiport subarray beams in a typical 5G cell.

Table 5.1: Progressive phase shifts in a 4-element subarray fed by a 4x4 BM

Excited subarray port #	Progressive phase shift in the subarray (in degrees)
1	+135
2	+45
3	-45
4	-135

5.1.2. SIMULATION RESULTS AND DISCUSSION

In this section, we propose a planar array version of AMSPAA with 0.5λ-spaced triangular lattice at element level and grating lobe dissolving row shifting in a diamond like configuration. The elements are grouped in horizontal subarrays which are staircase shifted to reduce grating/side lobes over the ±60 degree (in azimuth) by ±15 degree (in elevation) field of view, which is currently defined as the typical 5G angular sector in urban micro-cell environments [23].

The array layout is given in Fig. 5.3. In total, 256 elements are used with 64 4x4 BMs. The cell sector is visualized in Fig. 5.4 in the u-v plane. The sector is divided into four regions. Each Butler beam is responsible for serving a separate region.

The progressive phase shifts in a 4x4 BM for excitation of each port is provided in Table 5.1. By exciting the first port of each BM and applying the proper phase shifts, the beam is scanned in Region I. The same principle applies to the other BM ports (2, 3, 4) and the corresponding regions (II, III, IV).

Next, the CB pre-coding radiation pattern results for two beams scanned in Region I and Region II are shown. It is worthy of note that due to the symmetry in the layout, the beams in Region III and Region IV will have the symmetrical pattern properties as compared to the results given for Region II and Region I, respectively.

Fig. 5.5 shows the array directivity when a beam is scanned in Region I by applying the proper progressive phase shift to the first ports of the 64 subarrays. An isolated element pattern of $\sqrt{\cos\theta}$ is assumed with a directivity equal to 6 dBi. It is seen that in majority of the cases, the average interference within the sector is sufficiently low (side/grating lobes are all below -15 dB and most below -30 dB even for the worst scan condition). A high interference level (-8 to -10 dB) is only observed at specific areas in Region IV when $u = -\sin(60^o)$ and when $u = -\sin(30^o)$, $v = \sin(15^o)$.

Similarly, Fig. 5.6 shows the pattern results for a beam scanned in Region II for the excitation of the second BM ports. In this case, compared to Fig. 5.5, much cleaner beams are obtained. The only high interference is observed at the lower part of the boundary between Region III and Region IV when the beam is scanned towards the top left of Region II, where $u = -\sin(30^o)$, $v = \sin(15^o)$. This is expected since the pattern must be in line with the results given in Fig. 5.5 at the boundary shared between Region I and Region II.

Overall, competitive radiation pattern results to the fully-connected and fully-digital arrays with 0.5λ-spacing are observed using the proposed AMSPAA topology in the defined cell sector. In the case of having the channel state information or user positions available, it is also possible to further decrease the inter-user interference in AMSPAAs by

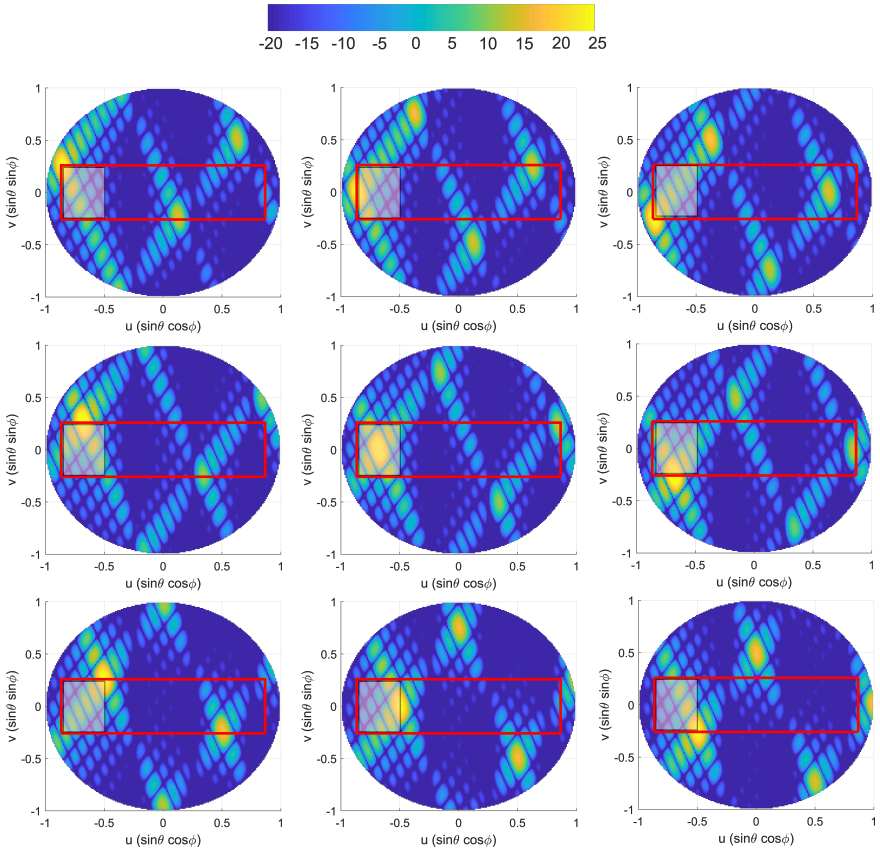


Figure 5.5: Array directivity (in dBi) for a beam scanned inside Region I.

applying ZF pre-coding, while maintaining relatively low side/grating lobes everywhere inside the sector. As previously mentioned in (D-i), further reduction in the processing and hardware (LNA/PAs, AD/DAs, mixers, DPXs) requirements can be achieved with the use of only one or two of four BM ports at a time with a switching network (see Fig. 5.2).

5.1.3. CONCLUSION

In this section, an original 5G multi-user hybrid beamforming architecture has been presented. The proposed technique is based on relatively small-sized active subarrays with multiple ports generating multiple fixed beams that are digitally controlled in the baseband.

Remember that the state-of-the-art 5G active multi-beam phased array architectures have been reviewed in Chapter 4. Here, in Table 5.2, those topologies are qualitatively compared with the proposed scheme. It has been shown that as compared to existing analog and hybrid multiple beamforming methods, active multiport subarrays can decrease the design and implementation complexity and / or increase the angular sector

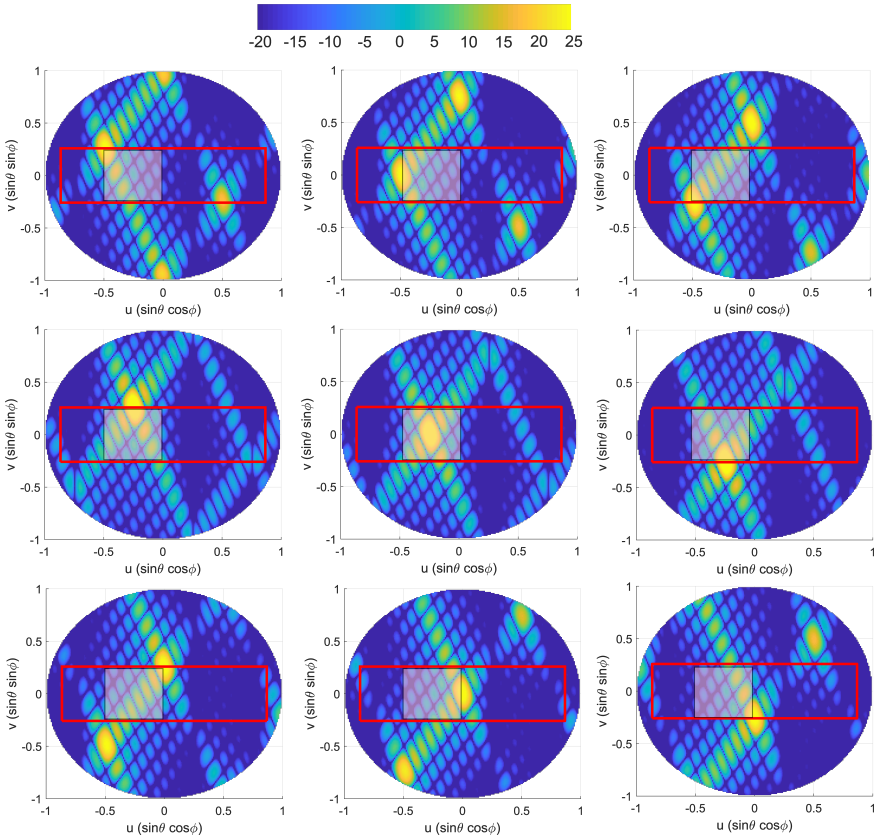


Figure 5.6: Array directivity (in dBi) for a beam scanned inside Region II.

coverage. Besides, by introducing a smart 2D array layout of the subarrays, competitive radiation performance to the fully digital beamforming has been obtained, while reducing the processing complexity significantly. It has also been suggested that hardware simplification in multiport subarrays can be achieved by scheduling use of only one or two sub-array ports/beams at the same time with a switching arrangement.

It is important to emphasize here that true hybrid beamforming is simply defined as RF beamforming at element level with DBF at the subarray port level. Yet, some in the literature do not refer to the combination of digital BF and analog BF at different radiation planes as hybrid beamforming. Therefore, we have focused on a commonly-agreed hybrid architecture in which analog beamforming is combined with digital precoding in the same radiation plane. Yet, an effective and well-known (in communication and radar domains) approach in hybrid-like beamforming is to form a single (or multiple) analog beam(s) in the elevation plane and digitally control the beams in the azimuth plane. In fact, the BMs in AMSPAAs can be oriented vertically, providing multiple fixed analog beams in elevation and digital beamforming in azimuth can be applied on the densely packed vertical BM ports. Moreover, reconfigurability can be added to the matrices for gain improvement, which is not discussed here.

Table 5.2: A qualitative comparison of the proposed AMSPAA with the existing multiple beam generation strategies studied in Chapter 4. *(Note that it is desired to have a narrow beamwidth (to increase frequency re-use), wide angular scanning range (to reduce the number of array panels), high array gain (to compensate the propagation loss and reduce the electricity consumption), high number of SDMA users (to maximize the capacity), and low RF & digital complexity with a low implementation cost.)*

	Distributed single beam analog/hybrid arrays	Analog, fully-connected array	fully-digital array	Hybrid, array of phased subarrays	Hybrid, passive multiple fixed beam arrays	Hybrid, active multiport subarrays
Beamwidth	Wide	Narrow	Narrow	Narrow	Narrow	Narrow
Scanning range	Wide	Wide	Wide	Narrow	Wide	Wide
Array gain	Low	Moderate	High	High	Moderate	High
Number of SDMA users	Low	Moderate	High	Moderate	Moderate	Moderate
RF complexity	Low	High	High	Moderate	High	Moderate
Digital complexity	Low	-	High	Moderate	Low	Low
Cost	Low	High	High	Moderate	Moderate	Moderate

Apart from the multiport subarrays, an intriguing, yet simple approach for the above-mentioned hybrid-like beamforming topology is to use vertical subarrays with a cosecant-squared shape pattern in elevation. This brings many advantages at the 5G base stations in terms of the cost and complexity, which is discussed in detail next.

5.2. ARRAY OF COSECANT SUBARRAYS

5.2.1. INTRODUCTION

The cosecant-square elevation pattern is applied as a means of achieving a uniform signal strength at the users with varying distances, but with a constant height relative to the base station array. Such a strategy is approriate and practical when the base station is on a tower at a certain height and transmits signals down towards the receiver units with a clear line-of-sight (LoS) propagation path, which does not require a large amount of vertical scanning range. In fact, **there is no clear gain in multiple beam operation in the vertical plane in this situation**. Remember that this is one of the use cases of 5G that is to be used for fixed wireless access in suburban environments [16] (see Figure 1.4(a)). Despite being well established in other applications, the introduction of cosecant-based hybrid beamforming in millimeter wave, multi-user 5G applications was first presented in [C-4].

Besides decreasing the cost and complexity significantly, the cosecant shape in elevation is particularly interesting in 5G since,

1. The power flux and power per user is optimally equalized (in LoS) for all users, which is key for PA performance and for providing more fairness with a predictable minimum SNR.

2. Overshadowing of weak signals by the stronger ones (i.e. the near-far problem) observed in the proximity of the base station is prevented (crucial for additional CDMA [59]).

3. A lower number of chips[2] are used to achieve the same EIRP with the 2-D arrays

[2] Less, thus larger amplifiers can be used with the cosecant subarrays to increase the signal quality or the cell range. However, this poses challenges in terms of the availability in the CMOS technology and on the thermal control, which is addressed later in Part III.

thanks to the gain of the vertical subarrays. As a result, the number of control points, design complexity, mismatches and errors are reduced.

4. The cosecant-squared pattern subarray has a similar gain to a pencil beam subarray, and covers the foot of the tower better.

5. The safety issues related with the vertical compliance distance [60] is automatically resolved.

Many algorithms exist in the literature which can be used to find out the element excitation coefficients and synthesize cosecant square shaped array factors. Some examples include Iterative Sampling Method [61], Alternating Projection Method [62], Convex Optimization [63], Genetic Algorithm [64], Differential Evolution [65], Particle Swarm Optimization [66] and Tabu Search Algorithm [67].

A wide range of topologies have also been proposed for the realization of cosecant squared pattern antennas. Among them, reflector antennas can yield a high gain and broad bandwidth, but they are large in size, heavy and costly to fabricate [68; 69]. Reflectarrays need feeding horns with high precision [70; 71]. On the other hand, planar microstrip array antennas can achive a lightweight and low profile product [72; 73]. However, microstrip-line feeding causes strong parasitic radiation and large losses, especially at high frequencies [74]. To improve the efficiency of the feeding network, the substrate integrated waveguide (SIW) technique with low cost and high integration capabilitiy is commonly adopted [75–79].

Within the project, we propose novel SIW and microstrip line feeding based base station antenna designs that have an approximate cosecant-square pattern in elevation and achieve flexible beam scanning in azimuth. The designs are realized and published by Dr. Jan Puskely *et al.* in [C-4; J-2]. The main features of the proposed designs are summarized here in Section 5.2.2 and the antenna performances are compared next in Section 5.2.3 for the sake of completeness of this part. Due to the limited scope of the Ph.D. thesis, the detailed model descriptions, simulation results and experiments are provided in Appendix A.

It is also helpful to mention here that the LoS system performance of the designed array of cosecant subarrays was studied as a MSc. thesis project with the help of a MSc. student, S. Salman, in which the system impact of the synthesized non-ideal cosecant-square shaped patterns is investigated and the proposed approach is compared with the alternative state-of-the-art hybrid beamforming techniques in 5G. Interested readers are referred to [T-1, C-10] for more details.

5.2.2. SUBARRAY DESIGN STRATEGIES

In this section, a short summary of the design approaches of the novel aperture coupled microstrip patch antenna (AC-MPA) arrays studied within the project framework is given. In total, three different strategies are discussed: (i) horizontally (transverse to the waveguide axis) polarized AC-MPA array fed by SIW, (ii) vertically (along the waveguide axis) polarized AC-MPA array fed by SIW, and (iii) vertically polarized AC-MPA array fed by microstrip line. The methods presented here have been used to synthesize a cosecant squared shape radiation pattern. However, they can be exploited in any series-fed array realizing any amplitude and phase distribution along the line.

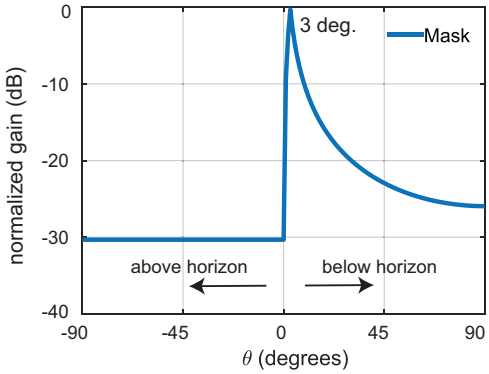


Figure 5.7: Cosecant-square shaped radiation pattern mask used in pattern matching optimization. *(Note: The computed and measured cosecant-square shaped patterns based on the given mask can be found in Appendix A.)*

An in-house developed technique[3] is used to get the excitation coefficients and synthesize the desired cosecant shaped radiation pattern with the optimization goal of matching the mask given in Fig. 5.7. Considering a cell radius of 200 meters and a base station height of 10.5 meters as done in Chapter 2, the maximum gain is directed to the cell edge, i.e. 3 degrees below the horizon.

Depending on the input parameters (the number of elements and spacing between the elements), the desired pattern is mimicked by the optimization algorithm. After obtaining the aimed element amplitudes and phases, the goal is to match the excitation coefficient of each element in the antenna design as closely as possible. To simplify the overall design by reducing the variety of the elements, a clever design strategy might be to group similarly excited (in amplitude and/or phase) elements together in a single set of amplitude and/or phase values.

In the presented studies, a sub-array of 12 radiators is chosen for the final design since it is observed to provide a good trade-off between the array gain and the subarray length.

In the first design, a series SIW-fed horizontally polarized cosecant shaped-beam AC-MPA sub-array is designed as a basic element for the active phased array antenna. The basic antenna element consists of two dielectric layers, a SIW layer and a patch layer, with a prepreg layer between them. The rectangular microstrip patch is placed on the top surface of the patch layer and helps achieve higher gain and much wider operational bandwidth. The microstrip patch is aperture coupled through the longitudinal slot integrated to the SIW layer. At the operating frequency, the coupling slots are centered at the peaks of the guided wave and has an offset from the centre line of the SIW. The centre of the last slot is placed a quarter guided wavelength from the end wall. The patch is placed symmetrically to the centre of the slot. The via is used as tuning element in the final cosecant subarray design.

[3]The technique is developed by Dr. Antoine Roederer and is based on a method of power synthesis which iteratively updates the phase assigned to the wanted amplitude pattern before the next "projection", with the phase law of the realized pattern from the previous iteration.

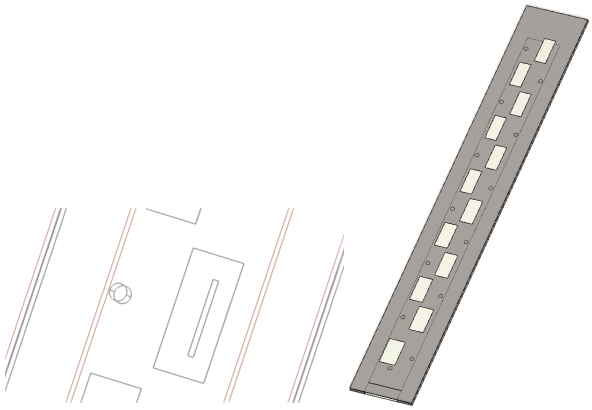


Figure 5.8: Illustration of the SIW-fed horizontally polarized AC-MPA subarray: (a) a single element in the array, (b) the complete subarray.

The illustration of the first design technique is shown in Fig. 5.8.

To simplify the antenna modelling, the SIW is replaced by an equivalent rectangular waveguide with solid walls [80]. This approach significantly reduces the simulation runtime. The fundamental TE_{10} mode is excited from the edge of the SIW.

An approach based on Elliott's design procedure for the slotted waveguide arrays [81] is used for the design of an AC-MPA array fed by a SIW. The self-admittance, Y, of an isolated antenna element is extracted using a two port structure with the antenna element placed in between. The self-admittance normalized to the characteristic conductance, G_0, of the waveguide is obtained by optimizing the slot length & offset, the patch length and the via position compensating the capacitive character of the radiator and it is given by [81]

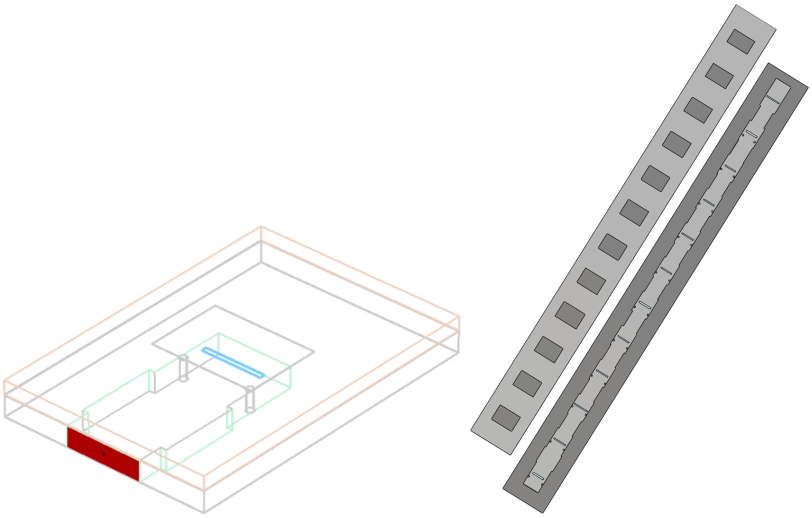
$$Y/G_0 = G/G_0 + jB/G_0 = -2S_{11}/(1+S_{11}). \tag{5.1}$$

The sub-array of AC-MPAs is impedance matched through the desired frequency band when the sum of normalized conductance, G/G_0, of each element in the subarray is equal to 1 at the operating frequency. Additionally, the sum of the normalized susceptances, B/G_0 should be zero.

The detailed design procedure, element amplitudes & phases, dimensions and simulation results of the SIW-fed horizontally polarized AC-MPA array can be found in [C-4].

The problem with the first design is that due to the relatively small spacing (one-half the guide wavelength) between the longitudinal slots, it is not possible to realize the optimal phase distribution. Therefore, in [C-4], except the first element, a single phase value is used for all the subarray elements, which deteriorates the pattern matching performance. In addition, due to high mutual coupling between the slots in the array of subarrays, the angular scan range in the horizontal axis becomes limited to ±45 degrees.

In the second design, SIW-feeding is used with transversal slots for vertical polarization. Transversal slots in the SIW which are placed at the maximum of the stand-



(a) (b)

Figure 5.9: Illustration of the SIW-fed vertically polarized AC-MPA subarray: (a) a single element in the array, (b) the complete subarray.

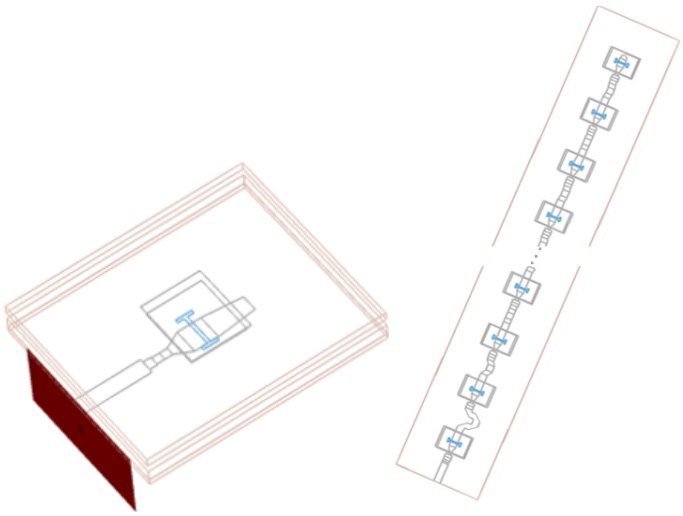
ing wave in the waveguide allows to have larger spacing among elements (around one guided wavelength) and to easily realize the synthesized optimal phase distribution. The slots are complemented by a pairs of vias (as an alternative to the reflection-canceling slot pairs [77]) and a phase element (by varying the SIW width) to get a tapered distribution both in amplitude and phase to provide a nice cosecant radiation pattern shape. Moreover, the transversal slots are coupled to microstrip patches to improve impedance bandwidth and gain.

Similarly to the first design, the subarray consists of two dielectric layers, a SIW layer and a patch layer, with a prepreg layer between them. The illustration of the second design technique is shown in Fig. 5.9.

The detailed design procedure, element amplitudes & phases, dimensions and simulation results of the SIW-fed horizontally polarized AC-MPA array can be found in [J-2]. An important point to note here is that thanks to the reduced mutual coupling with the transversal slots, the scan range in azimuth is increased to ±65 degrees.

Despite their many advantages, the common problem with the SIW-fed subarrays is the requirement of side wall vias which is cumbersome in the manufacturing process. To eliminate the need for vias, **in the third alternative design**, a series microstrip line fed aperture coupled stacked patch sub-array antenna is proposed. The drawbacks in this case are the higher loss in the feeding line at mm-waves and the increase in the number of layers due to the stacked patch structure.

In this design, the subarray consists of three dielectric layers, one for the microstrip feed line, one for the lower patch and one for the upper patch. Two prepreg layers are needed in between. The lower patch is aperture coupled through a transversal slot (with vertical polarization) integrated to the slot plane in between the two substrates support-



(a) (b)

Figure 5.10: Illustration of the microstrip line fed vertically polarized AC-MPA subarray: (a) a single element in the array, (b) the complete subarray.

ing the feed line and the lower patch. The upper patch is used as a parasitic element to further increase the gain and bandwidth.

The illustration of the third design technique is shown in Fig. 5.10.

Considering their superior performance, the second and third designs (vertically polarized AC-MPA arrays fed by SIW and microstrip line, respectively) were extended to an array of 16 subarrays, which were then fabricated and measured. The simulation and measurement results of them are provided in this thesis in Appendix A. The results include the performance of the isolated antenna panel, as well as the chip-integrated antenna system that uses NXP's 4-channel mm-wave analog beamforming chip for demonstration purposes. The chip structure is as shown in Fig. 5.11. It is possible to control the amplitude and phase in each channel (or each subarray in our case) by interfacing the chip to a computer using a NI USB-845x module that provide I2C and SPI connectivity.

5.2.3. CONCLUSION

In this section, we have proposed novel subarray design strategies to realize a 5G base station phased array antenna concept based on power flux equalization in the elevation plane for complexity reduction of the antenna system via one-dimensional (horizontal-only) beamforming.

The presented hybrid beamforming technique is especially useful in clear LoS propagation scenarios that do not require a large amount of vertical scanning range.

A desired cosecant shaped pattern has been formed considering a typical deployment scenario. The antenna element excitation coefficients have been synthesized following an iterative numerical optimization procedure.

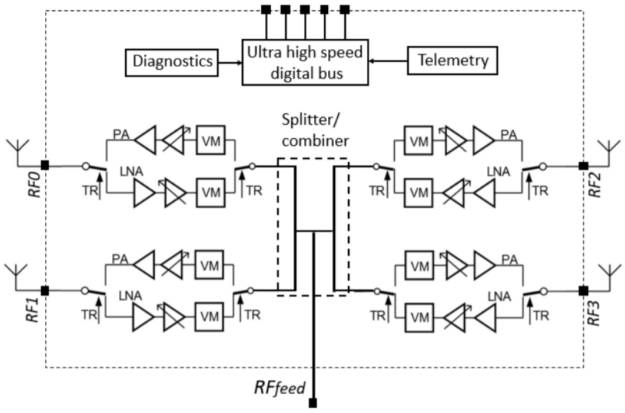


Figure 5.11: Structure of the 4-channel analog mm-wave beamforming chip of NXP Semiconductors.

5

Several implementation strategies have been investigated to realize the synthesized cosecant shaped pattern, including the SIW and microstrip line feeding. In all designs, aperture-coupled microstrip patches have been used to achieve sufficient gain and bandwidth.

From the simulation results with an array of 16 subarrays, it has been observed that all antenna arrays can achieve around 10% bandwidth and more than 24 dB realized gain. For the SIW-fed aperture-coupled patch subarrays, it has been concluded that vertically polarized slots has lower coupling among subarrays and allow a scan range of ±65 degrees in azimuth. On the other hand, the longitudinal slots with horizontal polarization has resulted in ±45 degree scan range.

Microstrip line fed aperture-coupled patch subarrays have also been proposed considering their low cost fabrication process. With the vertically polarized slots, a ±60 degree azimuthal scan range has been observed.

The vertically polarized antenna panels, as well as the chip-integrated antenna systems, have been fabricated and tested. The measurement results and comparisons with the simulations are reported in Appendix A. The communication system studies for the proposed array of cosecant subarrays have been performed within a complementary MSc. thesis project [T-1] based on the array simulation results. The work is in progress for analyzing the system impact of the measured patterns of the fabricated cosecant arrays with respect to the ideal/simulated ones.

5.3. CONCLUSIONS

After revisiting the currently proposed multiple beam generation strategies in 5G and listing their disadvantages for multiuser SDMA applications, the major question posed in the end of Chapter 4 was the following:

"Can we keep the hybrid (or hybrid-like) beamforming to decrease the cost while increasing the angular coverage and reducing the processing complexity further?"

In this chapter, we have addressed this question and ***proposed novel 5G hybrid mul-***

tiple beamforming concepts *based on: (i) active multiport subarrays with several digitally controlled Butler Matrix beams covering a wide sector and (ii) cosecant subarrays with a fixed cosecant shaped beam in elevation and digital beam forming in azimuth.*

In the first half of the chapter, we introduced the **hybrid array of active multiport subarrays** which is an array of small-sized active subarrays with multiple ports generating multiple fixed beams that are digitally controlled in the baseband. The pros and cons (with possible solutions to alleviate them) of the proposed approach have been explained in a comprehensive way. It has been shown that the major advantages of this method over the current multibeam generation methods are; (1) **wider scanning range** as compared to the hybrid array of phased subarrays, (2) **lower design complexity and higher efficiency** as compared to the fully-connected arrays and multiple fixed beam arrays, and (3) **significantly reduced processing complexity** as compared to fully-digital array.

In the second half, being one of the most promising innovative concepts with its low-cost and low-complexity, yet high-performance, a **cosecant subarray based base station antenna concept** has been introduced. The concept is realized by having a single analog beam in elevation with a cosecant shape and beam steering is achieved in azimuth. It has been explained that such a strategy is **especially useful in a power-fairness motivated, clear LoS communication scenario with very limited field-of-view in elevation**. Thanks to the cosecant-squared radiation pattern in elevation, the near-far problem, i.e. overshadowing of weak signals by the stronger ones that are observed near the base station, is solved. Furthermore, since less number of chips are used to achieve the same EIRP with the 2-D arrays because of the high gain of the subarrays, **the design complexity, mismatch, errors are reduced, less heat is generated and signal processing complexity is significantly decreased**. Few major drawbacks of the proposed approach have also been reported: the users within the same cosecant beam are very likely to be served using different time/frequency/code, and there will be more multipath to be dealt with in comparison with 2-D beamforming since the beam is relatively wide in elevation.

In the end, **three different cosecant subarray design strategies** have been shortly discussed: (i) horizontally (transverse to the waveguide axis) polarized AC-MPA array fed by SIW, (ii) vertically (along the waveguide axis) polarized AC-MPA array fed by SIW, and (iii) vertically polarized AC-MPA array fed by microstrip line. The **SIW-fed arrays have been proposed to achieve lower losses at mm-waves**, while the **microstrip line feeding alternative has been proposed to decrease the fabrication efforts and cost**. Considering their relatively low mutual coupling levels and the resulting wide azimuthal scan range, arrays of vertically polarized SIW-fed and microstrip line fed cosecant subarrays have been fabricated. Within the NWO-NXP partnership project framework, the complete antenna systems have been prototyped, as well as their sub-systems (power divider, chip board, antenna board etc.) for testing. The corresponding results are provided in Appendix A.

6

ARRAY OPTIMIZATION FOR INTER-USER INTERFERENCE MINIMIZATION

Parts of this chapter have been/will be published as:
[J-3] Y. Aslan, J. Puskely, A. Roederer and A. Yarovoy, "Phase-only control of peak sidelobe level and pattern nulls using iterative phase perturbations," *IEEE Antennas Wirel. Propag. Lett*, vol. 18, no. 10, pp. 2081–2085, Oct. 2019.
[J-4] Y. Aslan, J. Puskely, A. Roederer and A. Yarovoy, "Multiple beam synthesis of passively cooled 5G planar arrays using convex optimization," *IEEE Trans. Antennas Propag.*, vol. 68, no. 5, pp. 3557–3566, May 2020.
[C-5] Y. Aslan, J. Puskely, A. Roederer and A. Yarovoy, "Synthesis of multiple beam linear arrays with uniform amplitudes," in *Proc. 12th EuCAP*, London, UK, Apr. 2018.
[C-6] Y. Aslan, M. Candotti and A. Yarovoy, "Synthesis of multi-beam space-tapered linear arrays with side lobe level minimization in the presence of mutual coupling," in *Proc. 13th EuCAP*, Krakow, Poland, Apr. 2019.
[J-1] Y. Aslan, J. Puskely, A. Roederer and A. Yarovoy, "Trade-offs between the quality of service, computational cost and cooling complexity in interference-dominated multi-user SDMA systems," *IET Comm.*, vol. 14, no. 1, pp. 144-151, Jan. 2020.
[C-1] Y. Aslan, S. Salman, J. Puskely, A. Roederer and A. Yarovoy, "5G Multi-User System Simulations in Line-of-Sight with Space-Tapered Cellular Base Station Phased Arrays," in *Proc. 13th EuCAP*, Krakow, Poland, Apr. 2019.
[J-5] Y. Aslan, J. Puskely, A. Roederer and A. Yarovoy, "Synthesis of quasi-modular circularly polarized 5G base station antenna arrays based on irregular clustering and sequential rotation," *Microw. Opt. Technol. Lett.*, 2020, under review.
[J-10] Y. Aslan, A. Roederer and A. Yarovoy, "Synthesis of optimal 5G array layouts with wide-angle scanning and zooming ability for efficient link setup and high-QoS communication," *IEEE Antennas Wirel. Propag. Lett.*, vol. 19, no.9, Sep. 2020.

6.1. A PHASE-ONLY CONTROL TECHNIQUE

6.1.1. INTRODUCTION

In array synthesis with given element locations, interference suppression by controlling the complex excitation weights (amplitudes, phases) at each antenna element provides the most degree of freedom and best pattern performances. However, it is also the most expensive strategy due to the need of both a phase shifter and a variable gain amplifier per element. Considering this drawback, phase-only control has been introduced, which exploits the phase shifters used for beam steering in phased arrays also to obtain low sidelobes and nulls, while maintaining a relatively simpler feed network and typically higher power efficiency as compared to the comparable amplitude-tapered arrays [82]. Such an approach is more attractive for the 5G market in which system cost is of extreme importance.

Being inherently nonlinear, phase-only pattern design problem has been studied in the literature by employing a large variety of synthesis strategies. Phase-only pattern shaping with a pre-specified mask was studied using deterministic methods [83], numerical approaches [84; 85] and population-based optimization techniques like Genetic Algorithm [86]. In [87], best common amplitude distribution was searched for several beam shapes (flat-topped, cosecant, pencil) in reconfigurable arrays with phase-only control using the Intersection Approach. The same problem was addressed in [88; 89] using Convex Optimization and in [90] using Vector Projection Approach.

Considering multi-user communication aspects, statistical interference suppression via peak sidelobe level minimization [J-1] with phase-only optimization was also studied using Steepest Descent Method [82], Iterative Fast Fourier Transform [91] and Iterative Projection Method [92]. Furthermore, it was shown that, phase optimization is effective in pattern nulling of both narrow [93] and wide [94] angular sectors. Recently, in [95], a phase-only method based on Successive Alternating Projections was used in uniform-amplitude arrays to create Gaussian-shaped null regions for 5G applications.

In this section, we propose a novel phase-only, mutual coupling (MC) aware peak sidelobe level (SLL) minimization and (simultaneous) pattern nulling technique that is based on Iterative Convex Optimization. The nonlinear problem is linearized by introducing small phase perturbations at each iteration. The algorithm performance is compared with several techniques reported in the literature and its superiority is shown by examples. Finally, a case study is performed to demonstrate added value of the algorithm for mm-wave 5G application with phase-only radiation pattern forming.

6.1.2. PROBLEM FORMULATION

For an array with configuration given in Fig. 6.1, the far field, $f^{i,s}$, at the i^{th} iteration of the algorithm for a scanned beam, s, is given by

$$f^{i,s}(\theta,\phi) = \sum_{n=1}^{N} f_n(\theta,\phi)\, w_n^{i,s}\, e^{jk_0(x_n \sin\theta\cos\phi + y_n \sin\theta\sin\phi)} \tag{6.1}$$

where f_n is the complex far-field of the n^{th} element when the field origin is at the element center, k_0 is the wavenumber, (x_n, y_n) denotes the position of the n^{th} element and

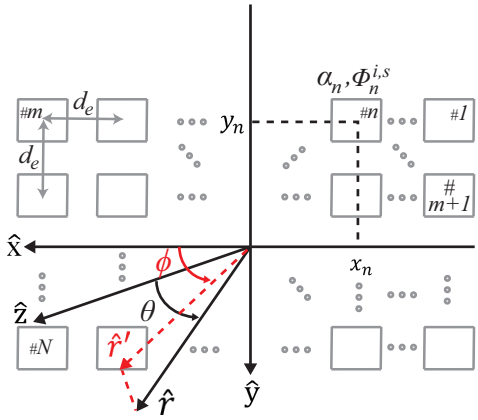


Figure 6.1: Schematic of a uniform planar array of N patch antennas. θ is the elevation angle defined as the angle between the observation direction $\hat{r}$ and $\hat{z}$. The unit vector $\hat{r}'$ is the projection of $\hat{r}$ onto the $\hat{x}\hat{y}$ plane. ϕ is the azimuth angle defined as the angle between $\hat{r}'$ and $\hat{x}$.

$w_n^{i,s}$ is the excitation weight of the n^{th} element at the i^{th} iteration for a scanned beam, s. The excitation weights are given by

$$w_n^{i,s} = \alpha_n e^{j(\Phi_n^{i-1,s} + \Phi_n^{i,s})} \tag{6.2}$$

where α_n is the pre-given amplitude and $\Phi_n^{i,s}$ is the phase variation of the n^{th} element at the i^{th} iteration (with respect to the element phase at the previous iteration) for a scanned beam, s.

Let us introduce the u-v coordinates as

$$\begin{aligned} u &= \sin\theta\cos\phi, \quad & u_s &= \sin\theta_s\cos\phi_s, \\ v &= \sin\theta\sin\phi, \quad & v_s &= \sin\theta_s\sin\phi_s \end{aligned} \tag{6.3}$$

where θ_s, ϕ_s (and the corresponding u_s, v_s) represents the desired beam pointing direction for a scanned beam, s.

Assuming $|\Phi_n^{i,s}| \ll 1$, and using the first-order Taylor expansion, (6.1) can be approximated and written as

$$f^{i,s}(u,v) = \sum_{n=1}^{N} f_n(u,v)\alpha_n e^{j\Phi_n^{i-1,s}}(1 + j\Phi_n^{i,s})e^{jk_0(x_n u + y_n v)} \tag{6.4}$$

The initial phases for the beam s are given by

$$\Phi_n^{0,s} = -k_0(x_n u_s + y_n v_s) \tag{6.5}$$

Overall, the optimization problem at the i^{th} iteration of the algorithm becomes

$$\min_{\mathbf{\Phi}^{i,s}} \rho^i, \;\; s.t. \begin{cases} |f^{i,s}(\{\boldsymbol{u},\boldsymbol{v}\}_{\boldsymbol{SL},\mathrm{s}})| \leq \rho^i \\ f^{i,s}(u_s, v_s) = 1 \\ |f^{i,s}(\{\boldsymbol{u},\boldsymbol{v}\}_{\boldsymbol{NR},\mathrm{s}})| \leq \delta \\ |\Phi_n^{i,s}| \leq \mu, \;\; \forall n \in \{1,2,\cdots,N\} \end{cases} \tag{6.6}$$

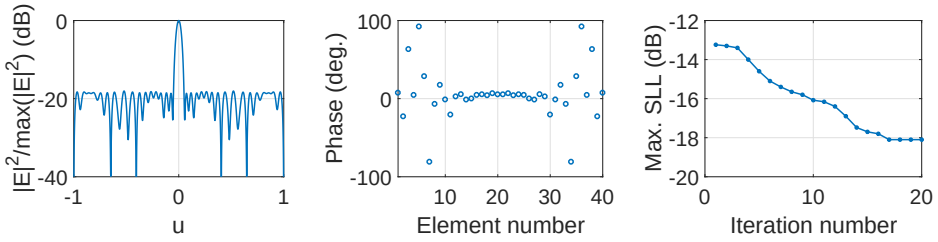
where $\{\boldsymbol{u},\boldsymbol{v}\}_{\boldsymbol{SL},\mathrm{s}}$ and $\{\boldsymbol{u},\boldsymbol{v}\}_{\boldsymbol{NR},\mathrm{s}}$ define the sidelobe and null regions, respectively. ρ^i is the maximum SLL to be minimized at the i^{th} iteration. δ is the suppression level in the null region. μ defines the upper-bound of the phase perturbations such that $|\Phi_n^{i,s}| \ll 1$, which is needed to properly apply the Taylor expansion in (6.4).

The optimization problem presented in (6.6) is a second-order cone program (SOCP) problem where a linear function is minimized over the intersection of an affine set and the product of quadratic cones [96]. For a comprehensive introduction to SOCP and its applications, the interested readers are referred to [96; 97]. Many aprroaches exist in the literature to efficiently solve the SOCP problems. Some examples include Interior-Point Method (IPM) [96; 98], Reduced-Augmented-Equation Approach [99], Pivoting Method [100] and Parametric Approach [101]. Today, SOCP problems can be easily solved in polynomial time by available convex programming toolboxes [102; 103] using IPM, which is commonly exploited by antenna researchers [104–107]. In this study, CVX, a MATLAB-based modeling system for convex optimization, is used to formulate and solve the problem in (6.6).

6.1.3. SIMULATION RESULTS AND DISCUSSION

In this section, the proposed algorithm's performance is evaluated by comparative case studies presented in the related literature with additional realistic, mutual coupling embedded patterns and 5G-oriented examples. In the simulations, it is assumed that $\mu = \pi/3$ and $\delta = 0.0001$. All numerical computations have been carried out on an Intel(R) Core(TM) i7-4710HQ 2.5GHz CPU, 16GB RAM computer. Each iteration takes about a few seconds in the linear arrays and an hour in the considered 64-element planar array for a uniform discretization step of 0.01 in the u-v plane. It is also worthy to note that in this work, we do not adopt a stop condition. Instead, we define a maximum number of iterations and observe the behavior of the maximum SLL to study its convergence. However, such a condition can be specified considering the relative change in the peak SLL in a few successive iterations so that the iterative process stops when the maximum SLL no longer diminishes. An example is to stop when $20\log(\rho^i \;/\; \rho^{i-1}) \leq 0.01$, as applied in [106]. If there is a peak SLL targeted for a system, it is also possible to stop the iterations as soon as the aim is reached.

Case-1: First, the results of the presented convex optimization method are compared with the phase-only synthesis technique presented in [82]. Using the Steepest Descent Method, the authors of [82] tried to minimize the peak SLL, which is the same goal function that is used in this work. As done in [82], the synthesis is performed on the array factor (AF) where α_n is the same for all n and the phases are even-symmetric. Note that the even-symmetry can be easily enforced using an additional constraint, $\Phi_n^{i,s} = \Phi_{N-n+1}^{i,s}$ for even N, in (6.6). The inter-element spacing, d_e, is equal to $0.5\lambda_0$. The maximum iteration number is set to 20. The comparison is given in terms of the maximum SLL, first-null



(a) (b) (c)

Figure 6.2: Results of Case-1 for N=40, (a) normalized array factor (at iteration number 20), (b) even-symmetric element phases (at iteration number 20), (c) iterative trend of the maximum SLL.

Table 6.1: Performance comparison between the Steepest Descent Method in [82] and this method in the case of phase-only array synthesis with uniform amplitudes.

Number of elements (N)	**Steepest Descent Method in [82]**			**This method**		
	Peak SLL (dB)	**FNBW HPBW (deg.)**	**Eff. (%)**	**Peak SLL (dB)**	**FNBW HPBW (deg.)**	**Eff. (%)**
20	-15.8	12.5 5.4	81.8	-16.1	12.2 5.2	83.6
40	-17.9	6.3 2.6	76.6	-18.1	6.4 2.7	74.7
80	-20.1	3.4 1.4	69.6	-20.3	3.4 1.4	67.3

beamwidth (FNBW), half-power beamwidth (HPBW) and array efficiency (which is defined as the ratio of the peak power density of the phase (and/or amplitude) controlled array to the peak power density of the uniformly excited array with progressive phase shifts defined by the scan angle). Table 6.1 provides a summary of the results for the linear array with N = 20, 40 and 80. It can be seen that the convex optimization provides competitive or better (especially for the relatively smaller array with N = 20) results. The synthesized array pattern, element phases and maximum SLL convergence for N = 40 with the presented method are also given in Fig. 6.2 for comparison with *Case-2*, which is studied next.

Case-2: The AF-only optimization in *Case-1* does not consider the mutual coupling effects and may result in unreliable patterns. In *Case-2*, full-wave simulations are performed using a conventional pin-fed patch antenna (see Fig. 6.3) in a 40-element $\lambda_0/2$-spaced E-plane array at a candidate 5G frequency of 28 GHz. A sample set of embedded far E-field magnitudes is given in Fig. 6.4. Upon using equi-amplitude element excitations, due to the difference in the embedded patterns, the AF-only optimized phases lead to a modified final pattern (compare Fig. 6.5(a) with Fig. 6.2(a)), a maximum SLL of -17.7 dB instead of -18.1 dB and an efficiency of 74.1% instead of 74.7% as reported in Table 6.1. By inserting the embedded patterns in the presented MC-aware method, we are able to synthesize a reliable pattern with a maximum SLL of -19.0 dB, but with an efficiency of 70.6%. The maximum iteration number is now set to 25. The E-field pattern,

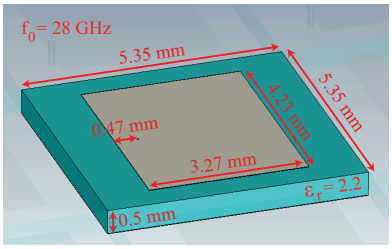


Figure 6.3: Pin-fed patch antenna dimensions used in the full-wave simulations.

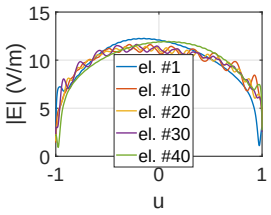
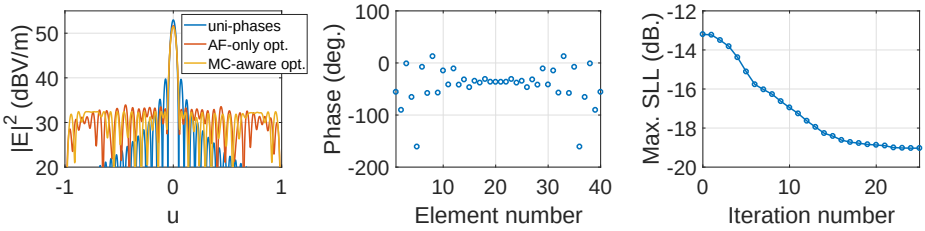


Figure 6.4: Sample set of embedded electric far-fields in the 40-element E-plane linear patch array.



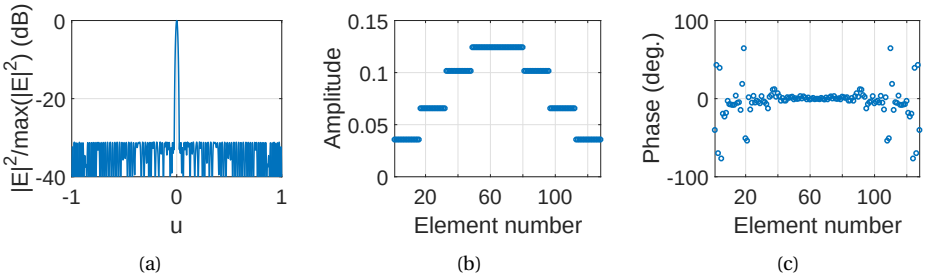
(a) (b) (c)

Figure 6.5: Results of Case-2 for N=40, (a) far E-field pattern (at iteration number 25), (b) even-symmetric element phases (at iteration number 25), (c) iterative trend of the maximum SLL.

even-symmetric element phases and convergence results of the MC-aware optimization with a 40-element E-plane patch array are shown in Fig. 6.5.

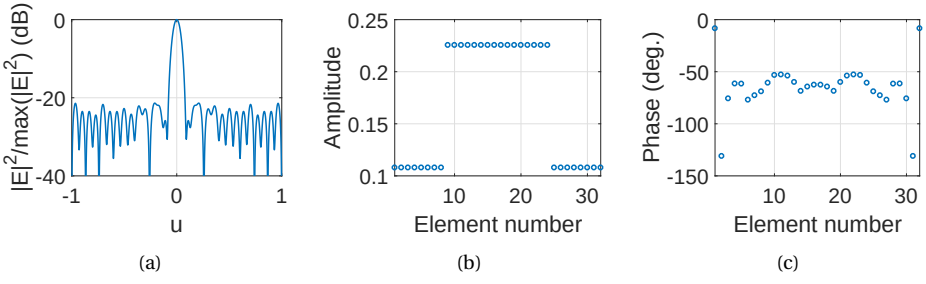
Case-3: In this case, the aim is to compute the phases of the array elements for a fixed subarray amplitude weighting in order to minimize the maximum SLL, as studied in [92]. The N-element array is partitioned into Q uniform and contiguous subarrays with a Taylor taper. Two cases are studied here for direct comparison with the results in [92]: i) N = 128, Q = 8 and ii) N = 32, Q = 4. It is assumed that the elements are isotropic and separated uniformly by $\lambda_0/2$. The phases are forced to be even-symmetric. For completeness, the AF's, element amplitudes and phases for (N, Q) = (128, 8) and (N, Q) = (32, 4) are given here in Fig. 6.6 and Fig. 6.7, respectively. A comparison with the Iterative Projection Method in [92] is provided in Table 6.2. It can be seen that the presented technique outperforms the method in [92] in terms of the maximum SLL and efficiency in the case of the large array with N = 128 elements. For the smaller array, the efficiency is 3.7% larger, while the maximum SLL increases only by 0.3 dB.

Case-4: Many pattern control methods in the literature study pattern-nulling with pre-specified amplitude tapering for SLL reduction and phase perturbation for creating the zeroes. Here, we focus on pattern-nulling by phase-control for narrow-sector interferers as studied in [93]. For fair comparison, we consider N = 20 isotropic elements with $0.5\lambda_0$ regular spacing. Three nulls are placed at u = {-0.5, 0.4, 0.61}. In [93], the Levenberg-Marquardt algorithm was used by targeting the array pattern as the enve-



(a) (b) (c)

Figure 6.6: Results of Case-3 for N=128, Q=8, (a) normalized array factor, (b) even-symmetric element amplitudes, (c) even-symmetric element phases.



(a) (b) (c)

Figure 6.7: Results of Case-3 for N=32, Q=4, (a) normalized array factor, (b) even-symmetric element amplitudes, (c) even-symmetric element phases.

Table 6.2: Performance comparison between the Iterative Projection Method in [92] and this method in the case of uniform subarrayed array antennas.

Number of elements (N)	**Number of sub-arrays (Q)**	**No phase optimization**			**Iterative projection method in [92]**			**This method**		
		Peak SLL (dB)	**HPBW (deg.)**	**Eff. (%)**	**Peak SLL (dB)**	**HPBW (deg.)**	**Eff. (%)**	**Peak SLL (dB)**	**HPBW (deg.)**	**Eff. (%)**
128	8	-25.1	1.0	84.3	-29.2	1.1	78.5	-31.1	1.1	79.2
32	4	-18.6	3.9	86.3	-21.7	4.2	76.6	-21.4	4.0	80.3

lope of the 30 dB Dolph-Chebyshev amplitude taper. At the output, -80 dB null levels were obtained with a maximum SLL of -25 dB. First, we assume a 30 dB Chebyshev windowing and optimize the phases to minimize the maximum SLL while nulling out the interferences. The results in this case are summarized in Fig. 6.8. It is seen that using the proposed method, the null levels become -92 dB and the maximum SLL is reduced to -28 dB. Second, we consider a phase-only control with equi-amplitude element excitations for the same array topology and optimization goal, which is more preferable for reduced design complexity and increased efficiency. For the uniform-amplitude counterpart, similar to the Chebyshev tapered array, -92 dB nulls are observed. However, the maximum SLL becomes -15.6 dB.

Case-5: The problem of broad-sector nulling and simultaneous SLL minimization is considered in this part. The proposed algorithm's performance is illustrated using the

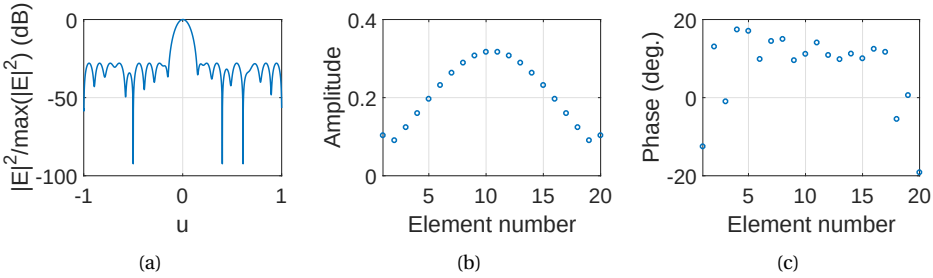


Figure 6.8: Results of Case-4 for N=20 with phase-only control, (a) normalized array factor, (b) element amplitudes, (c) element phases.

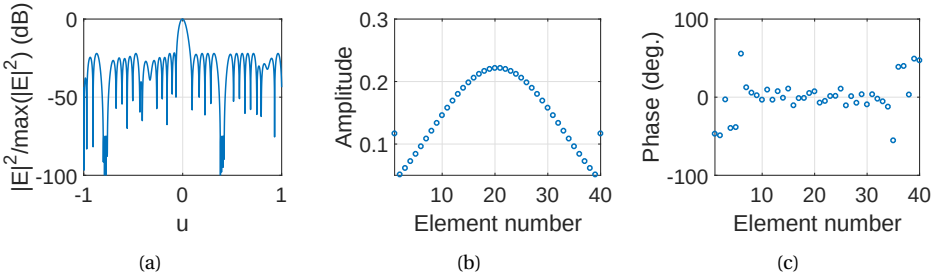
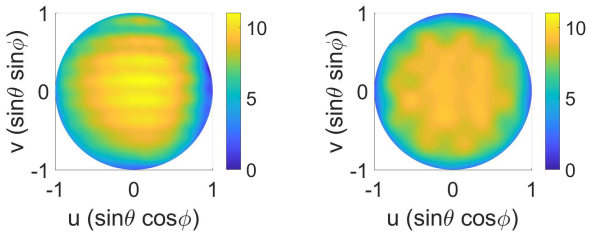


Figure 6.9: Results of Case-5 for N=40 with phase-only control, (a) normalized array factor, (b) element amplitudes, (c) odd-symmetric element phases.

numerical example in [94] with $N = 40$ isotropic, $\lambda_0/2$-spaced elements. Two sector nulls at u = [-0.8 -0.76] and u = [0.38 0.42] are desired. The phases are assumed to be odd-symmetric to reduce the number of calculations as done in [94]. Using a 30-dB Chebyshev initial amplitude taper, sector depths around -70 dB and -80 dB were obtained in the linear programming technique in [94] while having a maximum SLL around -19 dB. For the same initial settings, our method provides sector depths of -75 dB while keeping the maximum SLL at -22 dB. The results for this case are summarized in Fig. 6.9. For phase-only control with a uniform amplitude array, our method is able to decrease the maximum SLL to -16.2 dB, while keeping the sector nulls at -75 dB.

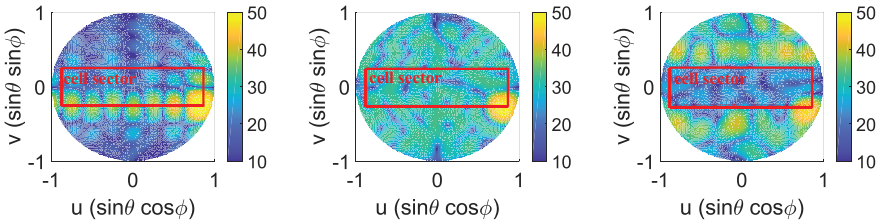
Case-6: In the last case, we study SLL minimization for an 8x8 $\lambda_0/2$-spaced uniform-amplitude patch antenna array at 28 GHz considering a typical 5G angular cell sector ($\pm$15 deg. in elevation and $\pm$60 deg. in azimuth) [23] and using our phase-only control algorithm. The embedded far field pattern of each patch element is computed to include the impact of mutual coupling. Sample embedded far field patterns for the patch element in Fig. 6.3 are given in Fig. 6.10 to show this impact. Note that the edge effects seen in Fig. 6.10 will be much weaker in sparse arrays than in regular dense arrays due to spatial tapping and large distance between neighboring antenna nodes.

The array is fed with uniform amplitudes and the beam is steered towards the cell edge, i.e. $u_s = \sin(\pi/3)$, $v_s = -\sin(\pi/12)$. The maximum iteration number is set to 25. Fig. 6.11(a) shows the pattern for the standard progressive phase shifts, for which the maximum SLL is -9.1 dB. For SLL minimization everywhere in the visible region, the pattern



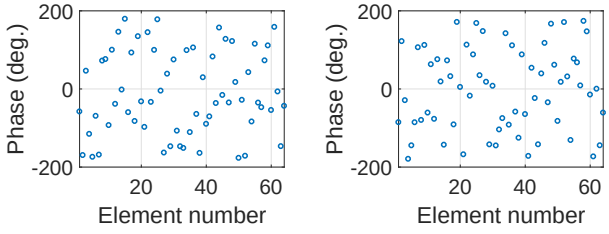
(a) (b)

Figure 6.10: Sample embedded electric far-fields (in V/m) in the 8x8 patch array for the element in the (a) corner (element #1), (b) center (element #29).



(a) (b) (c)

Figure 6.11: Far E-field patterns (in dBV/m) of the 8x8 patch array in Case-6 for a beam scanned towards the 5G cell sector edge u_s=sin(π/3), v_s=-sin(π/12) for (a) progressive phases, (b) phases optimized for minimum SLL everywhere in the visible region, (c) phases optimized for minimum SLL in the sector.



(a) (b)

Figure 6.12: Element phases at iteration number 25 with $\mu = \pi/3$ in Case-6 with optimization for minimum SLL (a) everywhere in the visible region, (b) only in the sector.

in Fig. 6.11(b) and the phases in Fig. 6.12(a) are observed, which yields a maximum SLL of -15.7 dB and efficiency of 55%. For completeness, the convergence of the maximum SLL for varying μ values is plotted in Fig. 6.13. It can be seen that the chosen value of $\mu = \pi/3$ provides a stable convergence within a reasonable number of iterations (= 5). If the optimization is done to minimize the interference only in the sector, the pattern in Fig. 6.11(c) is obtained with the phases given in Fig. 6.12(b). In this case, the maximum SLL in the sector becomes as low as -18.5 dB, but the efficiency drops to 30%.

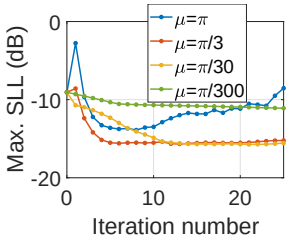


Figure 6.13: Maximum SLL trend with μ for optimization everywhere in the visible region.

6.1.4. CONCLUSION

This section presents a new phase-only peak SLL minimization and simultaneous pattern nulling algorithm. The nonlinear optimization problem is linearized using an iterative procedure by introducing small phase perturbations on the element excitation coefficients at each iteration. Embedded element patterns which are obtained via full-wave simulations can be integrated into the optimization procedure. In terms of the peak SLL reduction, the proposed approach outperforms the existing methods when radiation pattern nulling in certain sectors is required and achieves at least comparative results in cases without nulling, as illustrated via the case studies.

The advantages of the proposed phase-only control method are:

(i) better efficiency than the amplitude-tapered arrays

(ii) lower SLL compared to progressive phase shifts

(iii) no increase in the array size

(iv) easy integration of the embedded element patterns for mutual coupling awareness

(v) simpler feed network than amplitude or position tapering

(vi) possibility to create ZF-like orthogonal beams based on users' angular location information and not on the full-CSI as required by the ZF precoding

Because of the advantages listed, the proposed phase-only radiation pattern control method is promising for the initial, low-cost 5G base station antennas serving multiple SDMA users (with accurately estimated angular locations) at mm-waves. However, there are several disadvantages and limitations too, which are:

(i) much worse power efficiency than position-only-tapered arrays

(ii) the need for different set of phases optimized for each scan angle (a phase codebook is required)

(iii) higher SLL than the amplitude or position-tapered arrays

6.2. A POSITION-ONLY CONTROL TECHNIQUE

6.2.1. INTRODUCTION

As already stated in the preceding chapters, high expectations for 5G require sufficient simultaneous frequency re-use through formation of multiple beams (SDMA) with low interference and thus with low side lobes. The maximum SLL of the regular layouts (which is around -13 dB) contradicts with this requirement. On the other hand, such regular arrays are currently proposed in the very recent 5G (Massive) MIMO base stations by the leading organizations such as IBM, Nokia Bell Labs, Ericsson, and NXP Semiconductors (see Fig. 4.2).

In Section 6.1, phase optimization was proposed as an option to reduce the average SLL within the field of view, while keeping the dense layout of the array. However, it was seen that because of the nonlinearity in the phase shifts, the power efficiency might drop significantly, especially when a strong in-sector SLL suppression is intended. An alternative way to achieve optimal power efficiency is to keep the conventional progressive phase shifts at the uniform-amplitude elements, but to create irregularity in the array layout. It is worth to state that under a design criterion on the minimum inter-element distance, creating irregularity means moving the elements further from each other. In other words, the SLL improvement comes at the expense of a slightly increased array size. Nevertheless, in Section 9.1.2 and [J-6], it is shown that irregular arrays with larger than half-wavelength inter-element spacing can bring further passive cooling due to the increased substrate-to-air surface area around each element. On top of the better SLL suppression capability, this provides an additional superiority of irregular layouts with controlled minimum element spacing over the traditional regular half a wavelength spaced arrays.

In the literature on antenna (or chip) array topology optimization for SLL minimization, being motivated by the practicality and optimum power efficiency of uniformly excited arrays [108], many uniform amplitude array synthesis techniques have been proposed that are based on global optimization [109–111] and deterministic [112; 113] methods. It was shown that computational burden in global optimization increases rapidly with the number of array elements and the optimality of the solution is not guaranteed. Although deterministic methods are much more efficient, it was found out that they require the a priori knowledge of a reference (and ideally optimal) continuous source to be emulated. Recently, iterative convex optimization has been introduced for the synthesis of equi-amplitude arrays as a relatively efficient and powerful alternative [104; 106].

Multi-beam optimization (MBO) was recently introduced in [111; 114; 115] and it was shown that simultaneous parameter optimization for the broadside and the outermost scanned beams is necessary to guarantee that the SLL everywhere in the visible region will be less than the minimized SLL when the beam is scanned in a defined sector. In 5G, an array topology optimized via MBO will perform statistically the best in terms of the inter-user interference among the SDMA users sharing the same time and frequency resources. However, in the presented MBO literature, amplitude tapering with a large dynamic range was applied and there was no control on the spacing between the elements, which are two undesired features resulting from the sparsity-based nature of the proposed algorithms.

Considering this gap in the literature, we proposed a convex iterative position per-

turbation technique for multibeam optimization of uniform-amplitude linear arrays in [C-5]. Later, we extended the method to include mutual coupling into the design procedure via embedded element pattern simulations [C-6]. However, in convex optimization for planar array synthesis, defining a minimum inter-element spacing for the final layout is more challenging compared to the linear case and, if ignored, the designed array elements may have quite small spacings (<0.3λ), as obtained in [105; 106]. In addition to the thermal problem, this may lead to very high mutual coupling (MC) levels or even an unrealizable design. An effort was made in [107] to force a minimum inter-spacing by defining a circle around each element in which the total selection weight of the elements cannot exceed one. However, since any element could be selected with a weight less than 0.5, there is no guarantee that the spacing would be kept at the desired value. In [104], the minimum spacing control was applied in between the concentric circular rings, which can actually be represented as a 1-D constraint similarly to the linear array case. In [116], convex optimization was combined with the method of perturbed compressive sampling (PCS) to give a constraint to the inter-element spacing in a free 2-D space with no forced symmetry. Nevertheless, the synthesized arrays were power-inefficient due to the large-scale amplitude tapering. In [117], Chaotic PSO technique was efficiently applied to synthesize uniformly excited arrays with a pre-defined minimum inter-spacing. However, the aim was to maximize the power collected by a receiving rectenna and the maximum allowed SLL was defined as a constraint, which is not minimized. Besides, the effect of beam scanning was not taken into account during the optimization in the above-mentioned methods.

In this section, we propose the first optimization technique which simultaneously (i) formulates the array topology optimization problem in an advance way (i.e. system-based optimization to achieve certain system performance within the angular sector statistically, while systematically approaching the thermal issues); (ii) uses the same (or more advanced) optimization algorithm than any previously published paper; (iii) applies more advanced optimization constraints than any previously published paper; (iv) results in 2-D array topologies with electromagnetic and thermal performances better than any previously published; (v) is able to take into account mutual coupling of the elements in natural way; (vi) is computationally efficient.

The superiority of the proposed approach as compared to the state-of-the-art literature is summarized in Table 6.3. As per authors' knowledge, none of the earlier studies have combined interference-awareness (via multi-beam SLL minimization in a predefined cell sector) and thermal-awareness (via uniform amplitudes and minimum element spacing) with an efficient and easy-to-solve optimization algorithm.

The major contributions of this work are listed as follows:

1. Practical, 5G SDMA application specific integrated system design challenges and limitations in array topology optimization are discussed for the first time, with useful insights and recommendations.

2. A novel, extended-feature array synthesis method with possiblities to have uniform-amplitude excitations, simultaneous multi-beam optimization, inter-element spacing and aperture size restriction and mutual coupling inclusion via embedded element pattern simulation is introduced using a low analytical complexity, easy-to-

Table 6.3: Characteristics of the proposed approach in comparison to the state-of-the-art methods

	Uniform amplitude excitation	Pre-defined min. element spacing	Multi-beam optimization	Computation efficiency	5G SDMA compatibility
Ref. [110]	✓	✓	×	Low	×
Ref. [111]	×	✓	✓	Low	×
Ref. [112]	✓	×	×	High	×
Ref. [113]	✓	✓	×	High	×
Ref. [106]	✓	×	×	Moderate	×
Ref. [104]	✓	✓	×	Moderate	×
Ref. [107]	×	✓	×	Moderate	×
Ref. [116]	×	✓	×	Moderate	×
Ref. [117]	✓	✓	×	Low	×
Ref. [114]	×	×	✓	Moderate	×
Ref. [115]	×	×	✓	Moderate	×
This method	✓	✓	✓	Moderate	✓

solve and relatively efficient convex optimization technique.

3. A novel convex formulation of the minimum inter-element spacing constraint in planar arrays with no forced symmetry is proposed.

4. The radiation pattern performance of the optimal space-tapered arrays in a 5G SDMA system is studied for the first time under a typical 5G base station deployment scenario.

5. As compared to the outcome of the techniques presented in the state-of-the-art literature, competitive or better radiation pattern and, in general, system performances are obtained using the proposed method with increased 5G SDMA compatibility.

6.2.2. PROBLEM FORMULATION

Let us consider an N-element uniformly excited planar array, as shown in Fig. 6.14 with the coordinate system and notations used throughout the paper.

If the magnitude of the field is equated to one towards broadside as in [106], the far field radiated by such an array is given by

$$f(\theta,\phi) = \frac{1}{N}\sum_{n=1}^{N} E_n(\theta,\phi)\, e^{jk(x_n \sin\theta\cos\phi + y_n \sin\theta\sin\phi)} \tag{6.7}$$

where $E_n(\theta)$ is the embedded pattern of the n^{th} element. If the same isolated pattern, $E(\theta,\phi)$ is assumed for each element for simplicity, the far field, $f(\theta,\phi)$ becomes

$$f(\theta,\phi) = \frac{E(\theta,\phi)}{N}\sum_{n=1}^{N} e^{jk(x_n \sin\theta\cos\phi + y_n \sin\theta\sin\phi)} \tag{6.8}$$

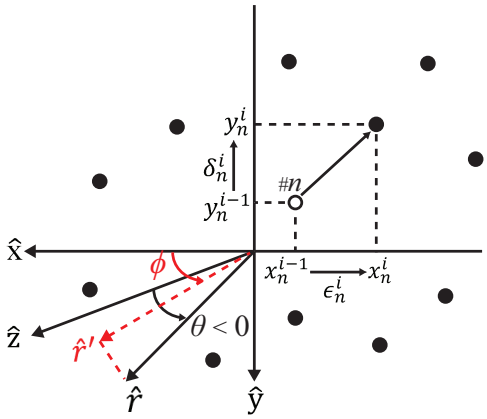


Figure 6.14: Schematic of the planar array, coordinate system and notations. θ is the elevation angle defined as the angle between the observation direction $\hat{r}$ and $\hat{z}$. The unit vector $\hat{r}'$ is the projection of $\hat{r}$ onto the $\hat{x}\hat{y}$ plane. ϕ is the azimuth angle defined as the angle between $\hat{r}'$ and $\hat{x}$.

As proposed in [106], the iterative optimization technique is applied by starting from a uniformly distributed array and moving the n^{th} element by ϵ_n^i in the $\hat{x}$ and δ_n^i in the $\hat{y}$ direction at the i^{th} step of the algorithm, which gives

$$x_n^i = x_n^{i-1} + \epsilon_n^i, \quad y_n^i = y_n^{i-1} + \delta_n^i \tag{6.9}$$

As previously realized in [106], the far field expression in (6.8) can be linearized around the element locations using the Taylor expansion when the following relations hold

$$|k\sin\theta\cos\phi(\epsilon,\delta)_n^i| \ll 1, \text{i.e.,} \; |(\epsilon,\delta)_n^i| \ll \lambda/2\pi = 0.16\lambda \tag{6.10}$$

In this case, the far field at the i^{th} iteration can be approximated by

$$f_{\epsilon_n,\delta_n}^i(\theta,\phi) \approx \frac{E(\theta,\phi)}{N}\sum_{n=1}^{N} e^{jk(x_n^{i-1}\sin\theta\cos\phi + y_n^{i-1}\sin\theta\sin\phi)} (1 + jk\epsilon_n^i\sin\theta\cos\phi)(1 + jk\delta_n^i\sin\theta\sin\phi) \tag{6.11}$$

If the sufficiently small high-order terms in (6.11) are ignored to keep the convexity of the problem, the far field becomes

$$f_{\epsilon_n,\delta_n}^i(\theta,\phi) \approx \frac{E(\theta,\phi)}{N}\sum_{n=1}^{N} e^{jk(x_n^{i-1}\sin\theta\cos\phi + y_n^{i-1}\sin\theta\sin\phi)} (1 + jk\epsilon_n^i\sin\theta\cos\phi + jk\delta_n^i\sin\theta\sin\phi) \tag{6.12}$$

Assume that the beam is scanned at p different angles where $(\theta_{s_m},\phi_{s_m})$ represents the direction of maximum radiation for the scanned beam $s_{m=1,2,\ldots,p}$. In such a scenario, the phase shift of the n^{th} element for the scan angle $(\theta_{s_m},\phi_{s_m})$ at the i^{th} iteration is given by

$$\Phi_{n,s_m}^i = e^{-jk(x_n^i\sin\theta_{s_m}\cos\phi_{s_m} + y_n^i\sin\theta_{s_m}\sin\phi_{s_m})} \tag{6.13}$$

Following similar steps, the far field expression in (6.12) can be modified to take into account the effect of scanning via multiplication with (6.13). Thus, the far field of a scanned beam s_m at the i^{th} iteration can be computed as follows

$$f^{i,s_m}_{\epsilon_n,\delta_n}(\theta,\phi) \approx \frac{E(\theta,\phi)}{N}\sum_{n=1}^{N} e^{jkx_n^{i-1}(\sin\theta\cos\phi-\sin\theta_{s_m}\cos\phi_{s_m})} e^{jky_n^{i-1}(\sin\theta\sin\phi-\sin\theta_{s_m}\sin\phi_{s_m})} (1+jk\epsilon_n^i(\sin\theta\cos\phi-\sin\theta_{s_m}\cos\phi_{s_m}) + jk\delta_n^i(\sin\theta\sin\phi-\sin\theta_{s_m}\sin\phi_{s_m})) \quad (6.14)$$

If the u-v coordinates are introduced as

$$\begin{aligned} u &= \sin\theta\cos\phi, \quad u_{s_m} = \sin\theta_{s_m}\cos\phi_{s_m}, \\ v &= \sin\theta\sin\phi, \quad v_{s_m} = \sin\theta_{s_m}\sin\phi_{s_m} \end{aligned} \quad (6.15)$$

the field expression in (6.14) can be rewritten as

$$f^{i,s_m}_{\epsilon_n,\delta_n}(u,v) \approx \frac{E(u,v)}{N}\sum_{n=1}^{N} e^{jk(x_n^{i-1}(u-u_{s_m})+y_n^{i-1}(v-v_{s_m}))} (1+jk\epsilon_n^i(u-u_{s_m})+jk\delta_n^i(v-v_{s_m})) \quad (6.16)$$

The vectors of parameters in the algorithm are defined as

$$\begin{aligned} \boldsymbol{x}^i &= \begin{bmatrix} x_1^i & \cdots & x_N^i \end{bmatrix}^T, \quad \boldsymbol{y}^i = \begin{bmatrix} y_1^i & \cdots & y_N^i \end{bmatrix}^T, \\ \boldsymbol{\epsilon}^i &= \begin{bmatrix} \epsilon_1^i & \cdots & \epsilon_N^i \end{bmatrix}^T, \quad \boldsymbol{\delta}^i = \begin{bmatrix} \delta_1^i & \cdots & \delta_N^i \end{bmatrix}^T, \\ \boldsymbol{u_s} &= \begin{bmatrix} u_{s_1} & \cdots & u_{s_p} \end{bmatrix}, \quad \boldsymbol{v_s} = \begin{bmatrix} v_{s_1} & \cdots & v_{s_p} \end{bmatrix}, \\ \boldsymbol{(U,V)_{SL,s}} &= \begin{bmatrix} \boldsymbol{(u,v)_{SL,s_1}} & \cdots & \boldsymbol{(u,v)_{SL,s_p}} \end{bmatrix} \end{aligned} \quad (6.17)$$

In (6.17), $\boldsymbol{x}^i$ and $\boldsymbol{y}^i$ contain the element locations. $\boldsymbol{\epsilon}^i$ and $\boldsymbol{\delta}^i$ contain the position shifts in the $\hat{x}$ and $\hat{y}$ directions at the i^{th} iteration, respectively. Scan angles used in the optimization form $\boldsymbol{u_s}$ and $\boldsymbol{v_s}$. $\boldsymbol{(U,V)_{SL,s_m}}$ is a vector containing the (u,v) values forming the side lobe region for each scan angle, s_m. These regions are determined according to a pre-specified main lobe radius, r_b, such that

$$(u,v) \in \boldsymbol{(u,v)_{SL,s_m}} \; if \; (u-u_{s_m})^2 + (v-v_{s_m})^2 > r_b^2 \quad (6.18)$$

Furthermore, in order to guarantee a pre-defined minimum inter-element spacing $d_{\min}$ in the final layout, the spacing between each element pair (α,β) is forced to be larger than $d_{\min}$ at each iteration

$$(x_\alpha^i - x_\beta^i)^2 + (y_\alpha^i - y_\beta^i)^2 \geq d_{\min}^2 \quad (6.19)$$

The constraint in (6.19) is non-convex since a convex expression on the left-hand-side of the inequality is forced to be larger than a real constant. Recall that the positions of the n^{th} element at the i^{th} iteration of the algorithm are given by (6.9). Considering

that $(\epsilon^i_{\alpha,\beta})^2$, $(\delta^i_{\alpha,\beta})^2$, $(\epsilon^i_\alpha \epsilon^i_\beta)$ and $(\delta^i_\alpha \delta^i_\beta)$ are sufficiently small, (6.19) can be convexly approximated to a linear programming (LP) constraint [96] as

$$(\epsilon^i_\alpha - \epsilon^i_\beta)(2x^{i-1}_\alpha - 2x^{i-1}_\beta) + (\delta^i_\alpha - \delta^i_\beta)(2y^{i-1}_\alpha - 2y^{i-1}_\beta) + (x^{i-1}_\alpha - x^{i-1}_\beta)^2 + (y^{i-1}_\alpha - y^{i-1}_\beta)^2 \geq d^2_{\min} \quad (6.20)$$

Note that it is also possible to modify this constraint to define an elliptical keep-out region that may have unequal dimensions in horizontal and vertical axes, which will be useful in the case of having subarrays as the smallest moving blocks instead of single antenna elements.

Finally, if the aperture size is limited, an additional condition should be forced on the element positions at each iteration so that no element can go beyond the pre-defined aperture region. Assuming a square aperture with maximum edge length $L_{\max}$, this condition is represented by

$$|\boldsymbol{x}^{i-1} + \boldsymbol{\epsilon}^i| \leq L_{\max}/2, \quad |\boldsymbol{y}^{i-1} + \boldsymbol{\delta}^i| \leq L_{\max}/2 \quad (6.21)$$

Overall, the convex problem to be solved at the i^{th} iteration of the algorithm is formulated as follows

$$\min_{\boldsymbol{\epsilon}^i, \boldsymbol{\delta}^i} \rho, \ \ s.t. \begin{cases} \max |f^{i,\boldsymbol{u_s},\boldsymbol{v_s}}_{\boldsymbol{\epsilon}^i,\boldsymbol{\delta}^i}((\boldsymbol{U},\boldsymbol{V_{SL,s}})| \leq \rho, \\ |\boldsymbol{\epsilon}^i| \leq \mu, \quad |\boldsymbol{\delta}^i| \leq \mu, \\ (6.20), (6.21) \text{ hold for } \forall(\alpha,\beta) \end{cases} \quad (6.22)$$

where ρ is the maximum SLL which is simultaneously minimized for all the defined scan angles, $\boldsymbol{u_s}, \boldsymbol{v_s}$. $|\boldsymbol{\epsilon}^i|$ and $|\boldsymbol{\delta}^i|$ are upper-bounded by a user-defined constant μ, as in (6.10). The last two constraints guarantee that the minimum inter-element spacing at each iteration is larger than or equal to a desired value, $d_{\min}$, and the maximum aperture size is limited to $L_{\max}$ x $L_{\max}$.

The optimization problem presented in (6.22) is a nonlinear convex problem, namely a second-order cone program [96], which, again, can be efficiently solved using interior point methods by available solvers such as CVX [102].

6.2.3. SIMULATION RESULTS AND DISCUSSION

A STUDY CASE AND COMPARISONS WITH THE LITERATURE

The algorithm is first tested for a study case using a 64 element array (N = 64) with a minimum inter-element spacing $d_{\min}$. A regularly spaced array (with spacing also equal to $d_{\min}$) is used at the algorithm input and an isolated element pattern $E(\theta,\phi) = \sqrt{\cos(\theta)}$ is assumed. At this point, it is worth to mention that the square equispaced layout, that is used as the benchmark initial array topology in the paper, has an impact on the optimized topologies. However, as stated in [115], the effect of using different starting element positions on the final results can be investigated by the interested users simply by modifying the element locations at algorithm initialization.

The MBO is performed for the broadside beam ($u_{s_1} = v_{s_1} = 0$) and five other beam positions (u_{s_2} = -sin(π/4), v_{s_2} = -sin(π/9), $u_{s_3} = 0$, v_{s_3} = -sin(π/9), u_{s_4} = sin(π/4), v_{s_4} =

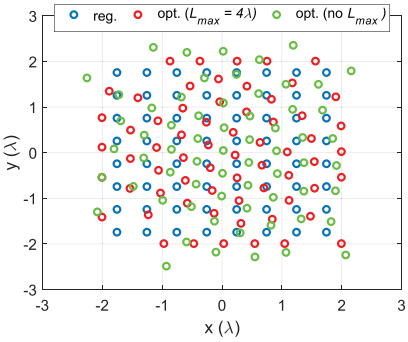


Figure 6.15: Antenna locations for the test case (N = 64, $d_{\min} = 0.5\lambda$, $r_b = 0.28$, $\mu = 0.08\lambda$).

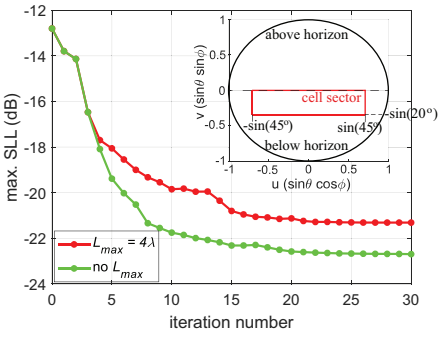


Figure 6.16: Convergence of the maximum SLL with respect to the broadside gain for the test case ($N = 64$, $d_{\min} = 0.5\lambda$, $r_b = 0.28$, $\mu = 0.08\lambda$) in the sector defined by a 20 degree range in elevation below horizon and a 90 degree range in azimuth.

-sin(π/9), u_{s_5} = -sin(π/4), v_{s_5} = 0, u_{s_6} = sin(π/4), v_{s_6} = 0), covering a sector defined by a 20 degree range in elevation below horizon (where, statistically, almost all the ground users are located for the urban base station towers as previously seen in Fig. 2.8) and a 90 degree azimuth coverage. The u-v plane is discretized in steps of 0.01. The upper bound for the position shifts $\mu = 0.08\lambda$ is used since it has been experimentally noticed that it provides a stable convergence in a reasonable time with small transient oscillations.

The initial settings $d_{\min} = 0.5\lambda$ and $r_b = 0.28$ are used. Two cases are studied: with a constraint on the aperture size ($L_{\max} = 4\lambda$) and without (no $L_{\max}$). Fig. 6.15 shows the element locations of the initial regular array and the final optimized arrays. The convergence of the maximum SLLs are shown in Fig. 6.16, together with the cell sector.

Fig. 6.17, Fig. 6.18 and Fig. 6.19 provide the multi-beam radiation patterns when the beam is at broadside and is scanned towards the sector edges, in the case of the regular array, optimized array with $L_{\max} = 4\lambda$ and optimized array with no $L_{\max}$, respectively. It is seen that the optimized arrays provide 9.9 dB and 8.5 dB reduction in the maximum SLL for no $L_{\max}$ and $L_{\max} = 4\lambda$, respectively, when the beam is freely scanned inside the defined sector. The improvement in SLLs causes an increase in the aperture size in each dimension (14.3% for $L_{\max} = 4\lambda$) when compared to the regular array with 0.5λ spacing.

All numerical computations for the synthesized 64-element arrays have been carried out on an Intel(R) Core(TM) i7-4710HQ 2.5GHz CPU, 16GB RAM computer using MATLAB and CVX. Each iteration takes about two hours. The simulation time mainly depends on the number of elements, the number of scan directions that is used in simultaneous MBO, the step size in the u-v plane grid and if applied, the forced symmetry order in the layout.

In order to see the effect of MC on the radiation patterns, full-wave EM simulation results for the optimized layout with $d_{\min} = 0.5\lambda$ and $L_{\max} = 4\lambda$ are provided in Fig. 6.20. The simulations have been performed in CST Microwave Studio (MWS) for a pin-fed microstrip patch antenna array operating at the center frequency of 28.5 GHz with the patch length and width equal to 0.3λ, where λ is the free space wavelength at the center

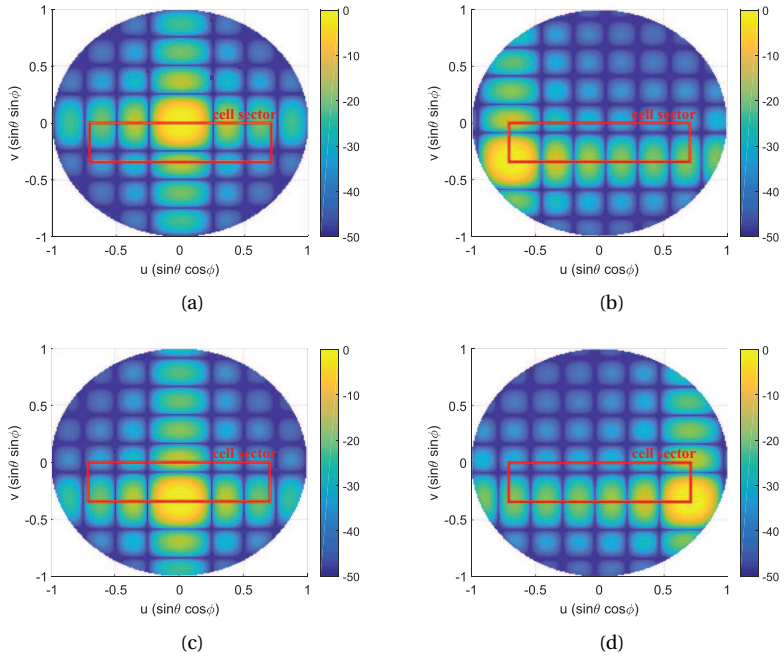


Figure 6.17: Multi-beam radiation patterns (in dB, normalized) for the test case with the regular array for a beam scanned towards (a) $u_{s_1} = v_{s_1} = 0$, (b) $u_{s_2} = -\sin(\pi/4)$, $v_{s_2} = -\sin(\pi/9)$, (c) $u_{s_3} = 0$, $v_{s_3} = -\sin(\pi/9)$, (d) $u_{s_4} = \sin(\pi/4)$, $v_{s_4} = -\sin(\pi/9)$.

frequency, on a 0.508 mm thick substrate with $\epsilon_r = 2.2$. The losses are included. A square aperture is used with an edge length equal to 5λ. It is seen that due to the relatively high MC (up to -13 dB), the radiation patterns given in Fig. 6.18 are modified and the maximum SLL increases up to -18 dB, which is seen on a circle around the main lobe. This analysis shows that, to have a better accuracy in the optimization results, embedded element patterns (EEPs) should be used at each iteration of the algorithm, as done in [104; 118–120], at the expense of increased optimization times and resources.

In [C-6], by using linear arrays for simplification, we demonstrated the applicability of embedded pattern inclusion in our iterative convex optimization technique. We also introduced a fast and efficient two-step approach where in Step #1, the average EEP (obtained from the full-wave simulation of the initial regular array with spacing d_{ini}) is used to optimize the element locations. Using these locations as an initial guess, embedded far-field of each element is simulated at each iteration in Step #2 and integrated into the optimization routine to obtain the final array layout. The procedure is given in Fig. 6.21. Interested readers are referred to [C-6] and the references therein for a detailed discussion on techniques to include MC in array optimization. It is worth here also to mention that, alternatively, MC reduction techniques, such as [121–123], might be exploited in the antenna design to minimize the unpredictability of the MC-unaware optimization results.

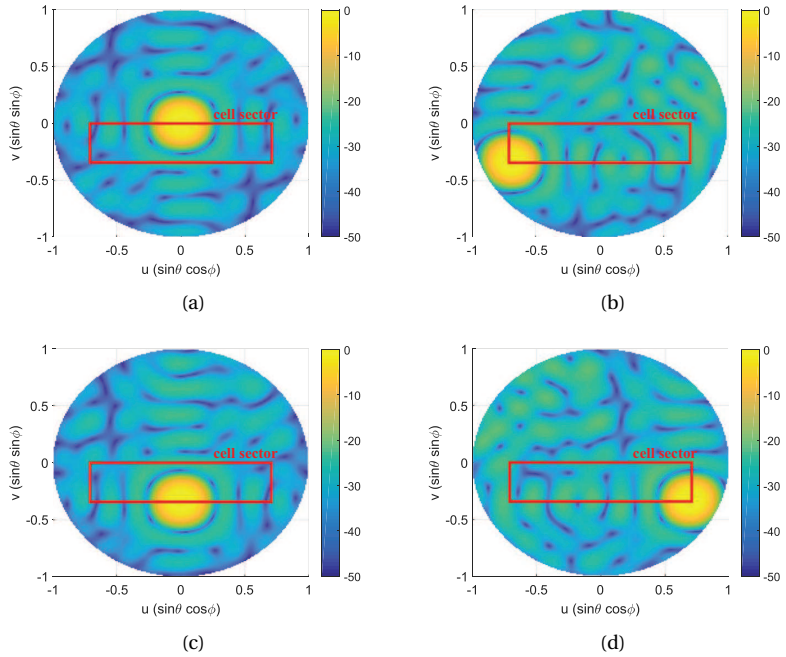


Figure 6.18: Multi-beam radiation patterns (in dB, normalized) for the test case with the optimized array with $L_{\max} = 4\lambda$ for a beam scanned towards (a) $u_{s_1} = v_{s_1} = 0$, (b) $u_{s_2} = -\sin(\pi/4)$, $v_{s_2} = -\sin(\pi/9)$, (c) $u_{s_3} = 0$, $v_{s_3} = -\sin(\pi/9)$, (d) $u_{s_4} = \sin(\pi/4)$, $v_{s_4} = -\sin(\pi/9)$.

Next, the effectiveness of the proposed method is shown by adjusting the optimization parameters considering the results reported in [107; 115; 117] and comparing the outcomes. The performance metrics used for the comparison are the maximum SLL, dynamic range of excitation amplitudes and beam collection efficiency (BCE) which, according to [124], can be expressed as

$$\text{Max. SLL (dB)} = 10\lg\frac{\max_{u,v\notin\Omega}|f(u,v)|^2}{\max_{u,v}|f(u,v)|^2} \tag{6.23}$$

$$\text{BCE} = \frac{P_\Omega}{P_\Gamma} = \frac{\int_\Omega |f(u,v)|^2 du\,dv}{\int_\Gamma |f(u,v)|^2 du\,dv} \tag{6.24}$$

where P_Ω and P_Γ denote the power radiated over the angular rectenna region Ω and the total transmitted power over the whole visible space Γ, respectively. Amplitude dynamic range is the range between the smallest and largest output levels of the amplifiers feeding the elements.

In [107], a 35-element array with isotropic elements was synthesized with $d_{\min} = 0.6\lambda$ and $L_{\max} = 5\lambda$. Both element positions and excitation weights were optimized for the broadside beam, which resulted in 8.01 dB dynamic range in weights. Here, only the element positions in a uniformly excited 35-element array are optimized using the

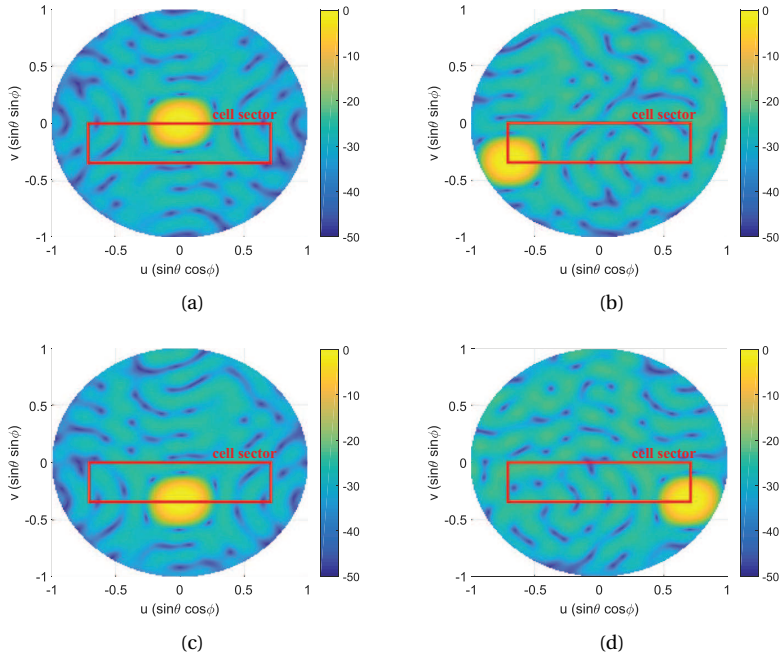


Figure 6.19: Multi-beam radiation patterns (in dB, normalized) for the test case with the optimized array with no $L_{\max}$ for a beam scanned towards (a) $u_{s_1} = v_{s_1} = 0$, (b) $u_{s_2} = -\sin(\pi/4)$, $v_{s_2} = -\sin(\pi/9)$, (c) $u_{s_3} = 0$, $v_{s_3} = -\sin(\pi/9)$, (d) $u_{s_4} = \sin(\pi/4)$, $v_{s_4} = -\sin(\pi/9)$.

same constraints in [107] for the broadside beam and for multiple beams (broadside and scanned towards $u_s = v_s = 0.5$ to show the superiority of the proposed algorithm). The resulting element locations are given in Fig. 6.22(a). Fig. 6.23 provides the corresponding radiation patterns (a circular rectenna region Ω with radius $r_b = 0.2$ is defined for BCE calculations). Comparison of results is shown in Table 6.4. It is seen that the broadside optimization with the proposed method yields decreased maximum SLL and increased BCE. It is also seen that since the optimization in [107] was performed only for broadside radiation, scanning the beam leads to the formation of grating lobes in the visible region, which decreases the BCE significantly. However, the proposed MBO approach is able to guarantee low side lobes, and thus much better BCE, for both the broadside and the scanned beam simultaneously.

In [115], a 48-element patch antenna array was synthesized to achieve less than -20 dB SLL outside a circular main lobe region, Ω, which is defined by a radius of $r_b = 0.25$. Element positions and excitation weights were optimized for the broadside beam and the beam scanned towards $u_s = 0.5$, $v_s = 0$, which resulted in a 6.90 dB dynamic range in weights. Here, using the proposed multi-beam optimization method, a new, uniformly-fed 48-element array is synthesized in order to minimize the maximum SLL while using the same settings with [115]. The resulting element locations are given in Fig. 6.22(b). Fig. 6.24 shows the corresponding radiation patterns. Comparison of results is

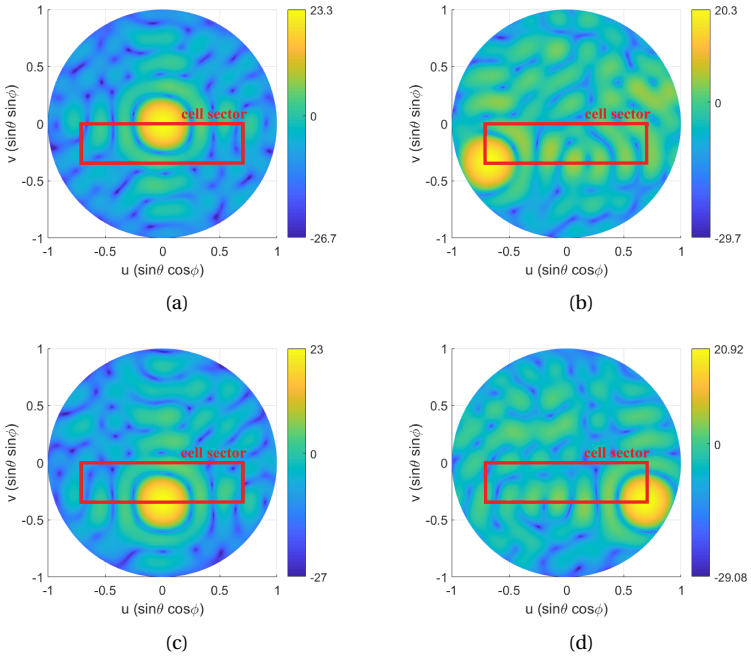


Figure 6.20: Full-wave simulation multi-beam realized gain patterns for the test case with the optimized array with $L_{\max} = 4\lambda$ for a beam scanned towards (a) $u_{s_1} = v_{s_1} = 0$, (b) $u_{s_2} = -\sin(\pi/4)$, $v_{s_2} = -\sin(\pi/9)$, (c) $u_{s_3} = 0$, $v_{s_3} = -\sin(\pi/9)$, (d) $u_{s_4} = \sin(\pi/4)$, $v_{s_4} = -\sin(\pi/9)$.

summarized in Table 6.5. It is seen that the proposed method is able to provide improved maximal SLL and BCE for both the broadside and scanned beam, while increasing the power-efficiency by equi-amplitude excitations.

In [117], the work in [124] was improved by synthesizing a uniformly-fed 100 element

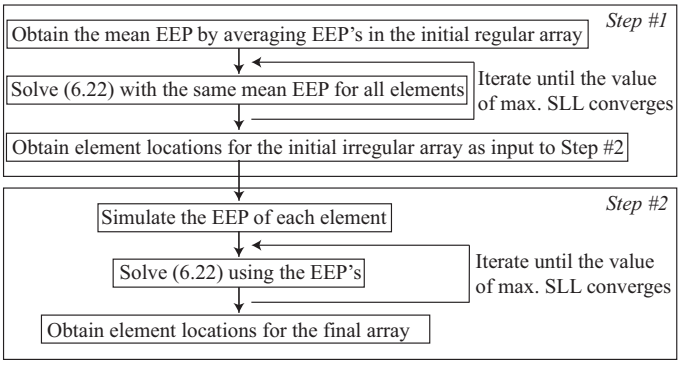


Figure 6.21: An efficient two-step optimization scheme for array optimization in the presence of strong mutual coupling (adapted from [C-6]).

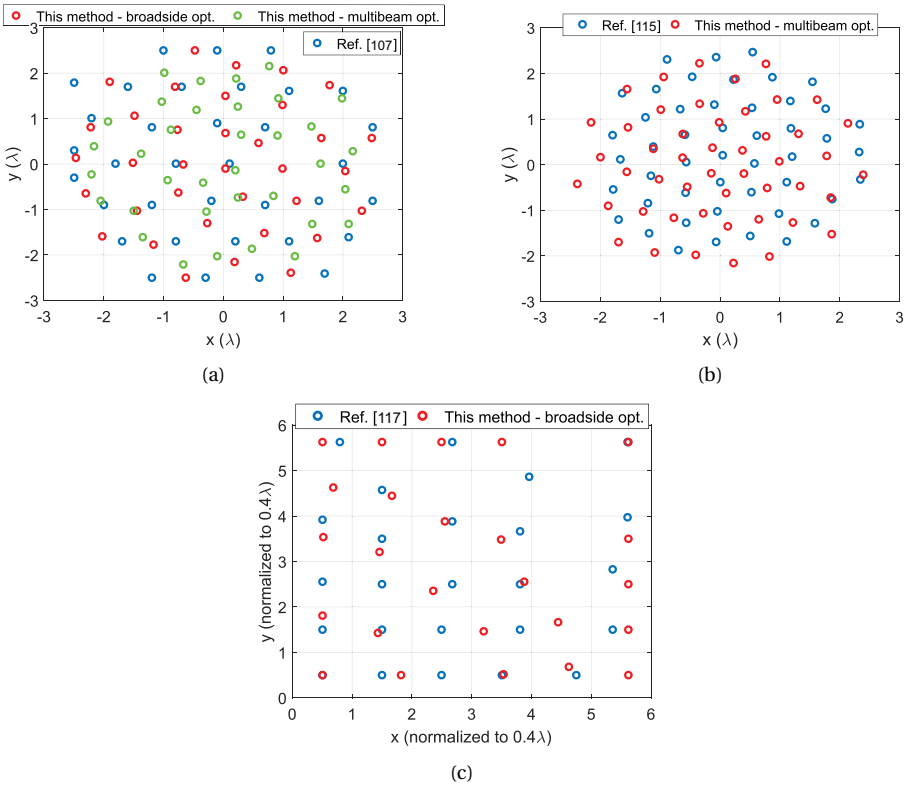


Figure 6.22: Antenna locations obtained using the proposed algorithm and locations in (a) Ref. [107], (b) Ref. [115], (c) Ref. [117].

array while limiting the maximum SLL and maximizing the BCE in a circular rectenna region Ω with radius 0.2. It was given that $d_{\min} = 0.4\lambda$ and $L_{\max} = 4.5\lambda$. Positions of 25 elements were optimized in the first quadrant and symmetry about the x-axis and y-axis was assumed. The resulting element locations in [117] and locations given by the proposed method are shown in Fig. 6.22(c). Fig. 6.25 provides the radiation patterns at broadside for the both methods. The performance comparison is given in Table 6.6. It is seen that the maximum SLL can be suppressed further by the proposed method. Hence, the interference among the users located at these side lobes and the main lobe becomes lower. However, due to the increase in the average level of side lobes, BCE is lower for the proposed method. In other words, the proposed algorithm yields a lower maximum SLL, but the number of side lobes reaching this maximum value is larger. Therefore, the selection between these two methods depends on the particular system requirements.

COMPARATIVE SYSTEM STUDIES

In this section, the performance of space-tapered multi-beam arrays with minimized side lobe levels is statistically evaluated in a LoS propagation environment within a cell sector in terms of the SINR at the multiple user ends. Comparative analyses are per-

Table 6.4: Comparison of the maximum SLL, amplitude range and BCE obtained using the proposed algorithm with [107]

	Ref. [107]	This method -broadside opt.-	This method -multibeam opt.-
Max. SLL (dB) -broadside-	-17.50	-19.51	-15.43
Max. SLL (dB) -scanned (u_s=v_s=0.5)-	-1.30	-5.50	-15.43
Excitation amplitude range (dB)	8.01	0	0
Min. inter-element spacing (λ)	0.60	0.60	0.60
BCE (%) -broadside-	54.91	68.96	52.02
BCE (%) -scanned (u_s=v_s=0.5)-	27.68	34.99	50.23

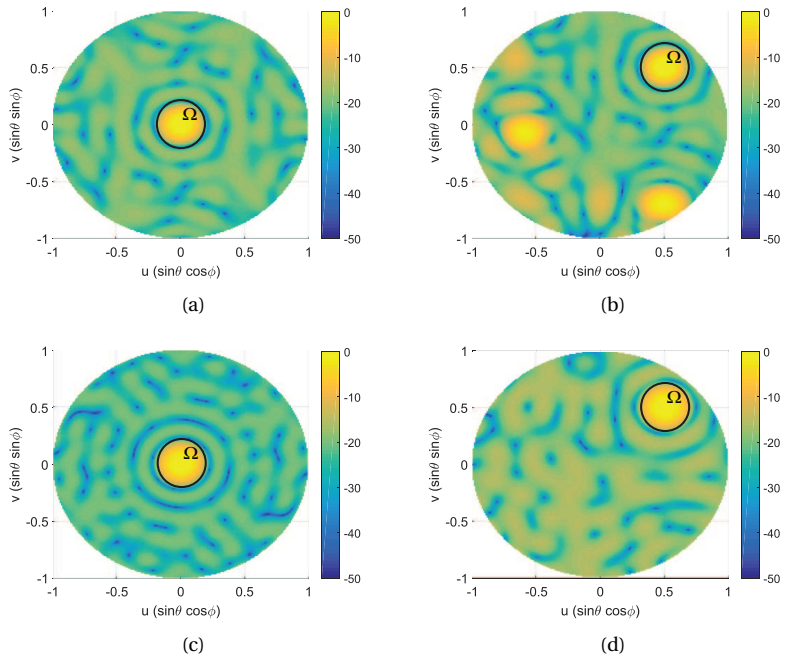


Figure 6.23: Radiation patterns (in dB, normalized) for the proposed algorithm and the method in [107], (a) from [107], broadside, (b) from [107], scanned towards u_s=v_s=0.5, (c) proposed method, broadside, (d) proposed method, scanned towards u_s=v_s=0.5.

formed to examine the advantages of space-tapered, irregular arrays over the conventional 0.5λ-spaced regular array layouts which were studied previously in Section 2.4.

First, we will focus on the ground users situation with the first study case in Use Case

Table 6.5: Comparison of the maximum SLL, amplitude range and BCE obtained using the proposed algorithm with [115]

	Ref. [115]	This method -multibeam opt.-
Max. SLL (dB) -broadside-	-22.13	-23.73
Max. SLL (dB) -scanned (u_s=0.5, v_s=0)-	-21.19	-23.07
Excitation amplitude range (dB)	6.90	0
Min. inter-element spacing (λ)	0.53	0.50
BCE (%) -broadside-	88.70	91.17
BCE (%) -scanned (u_s=0.5, v_s=0)-	86.44	88.97

Table 6.6: Comparison of the maximum SLL and BCE obtained using the proposed algorithm with [117]

	Ref. [117]	This method -broadside opt.-
Max. SLL (dB)	-12.30	-18.29
Excitation amplitude range (dB)	0	0
Min. inter-element spacing (λ)	0.40	0.40
BCE (%)	89.96	79.93

A examined in Section 2.4.1. In that case, it was seen in Fig. 2.8 which shows that 98% of users are in the first 20 degrees below the horizon. At each realization of the system simulation, 4 random user positions were selected according to Fig. 2.8 in elevation and from a uniform distribution in ±45 degrees in azimuth. Here, considering the sector definition, we will add the performance of the array layout optimized in the study case in Section 6.2.3. For completeness, the investigated layouts are given here as well in Fig. 6.26. The corresponding results for the beamforming approaches B, C, D described in Section 2.4.1 are given in Fig. 6.27. For *Approach B* with interference-aware scheduling, perfect CSI, non-quantized excitations, it is seen in Fig 6.27(a) that the SINR in 90% of the cases is larger than 7.23 dB, 10.43 dB and 24.31 dB for the regular array with CB, space-tapered array with CB and regular array with ZF, respectively. For *Approach C* with interference-aware scheduling, imperfect CSI (for a maximum error of 0.01), and non-quantized excitations, it is seen in Fig. 6.27(b) that the SINR in 90% of the cases is larger than 6.08 dB, 8.54 dB and 17.96 dB for the regular array with CB, space-tapered array with CB and regular array with ZF, respectively. For *Approach D* with interference-aware scheduling, perfect CSI, quantized excitations (2-bit phase shifters), Fig. 6.27(c) shows that the SINR in 90% of the cases is larger than only 3.12 dB, 5.14 dB and 5.15 dB for

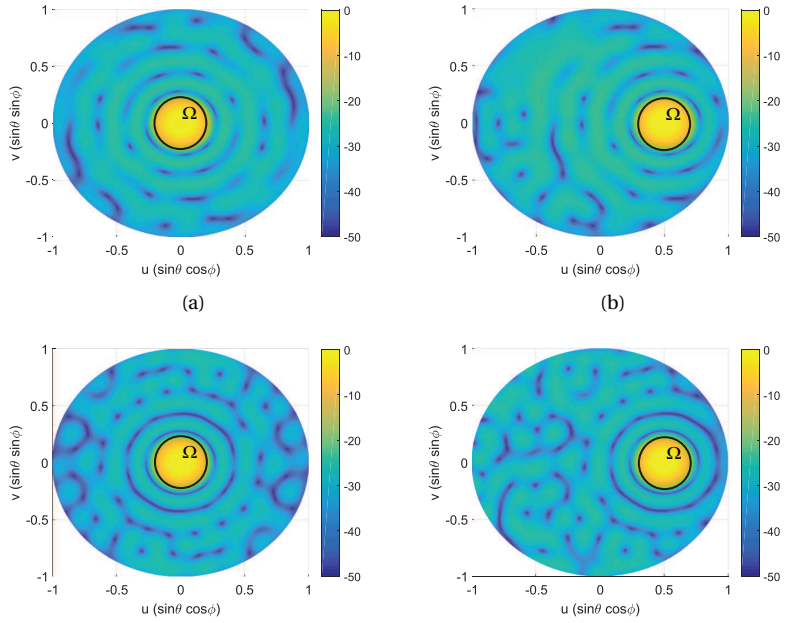


Figure 6.24: Radiation patterns (in dB, normalized) for the proposed algorithm and the method in [115], (a) from [115], broadside, (b) from [115], scanned towards u_s=0.5, v_s=0, (c) proposed method, broadside, (d) proposed method, scanned towards u_s=0.5, v_s=0.

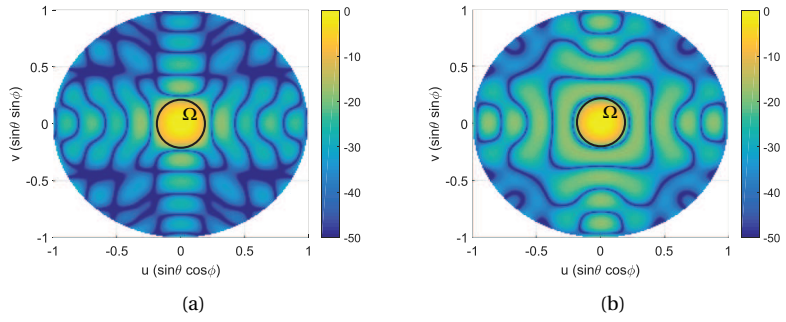


Figure 6.25: Broadside radiation patterns (in dB, normalized) for the proposed algorithm and the method in [117], (a) from [117], (b) proposed method.

the regular array with CB, space-tapered array with CB and regular array with ZF, respectively. These results indicate that CB with space-tapered arrays results in better performance than the traditional counterparts, provides a good compromise between the QoS and signal processing complexity and is relatively robust to system impurities such that it even performs equally well as ZF statistically, when the number of phase bits is small.

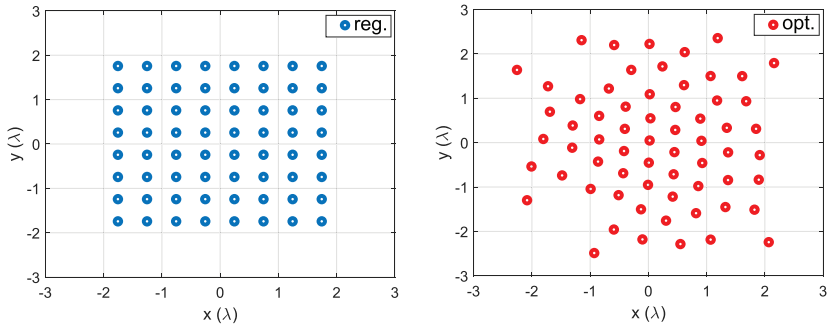


Figure 6.26: Element locations of the traditional regular and space-tapered array used in the comparative system studies for the ground-only user scenario (Use Case A).

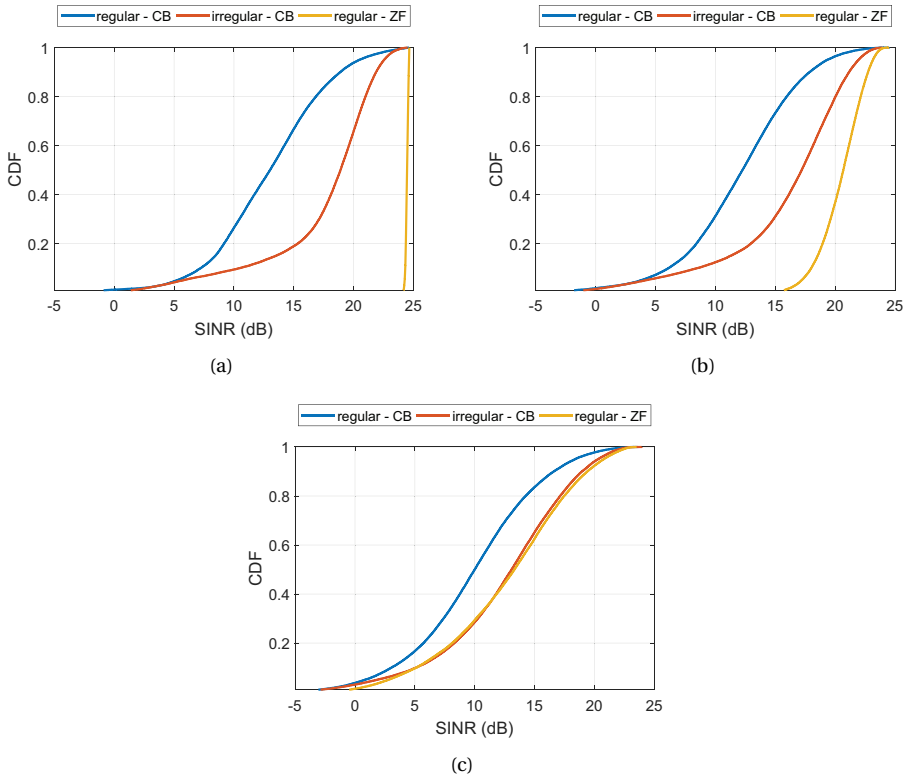


Figure 6.27: CDF of SINR for different array topologies and precoding techniques in Use Case A for: (a) Approach B, (b) Approach C, (c) Approach D. (Go back to Section 2.4.1 for the detailed description of each beamforming approach).

Next, we revisit the situation with users/relays at different heights (Use Case B) which was previously discussed in Section 2.4.2. In that section, the sector was assumed to

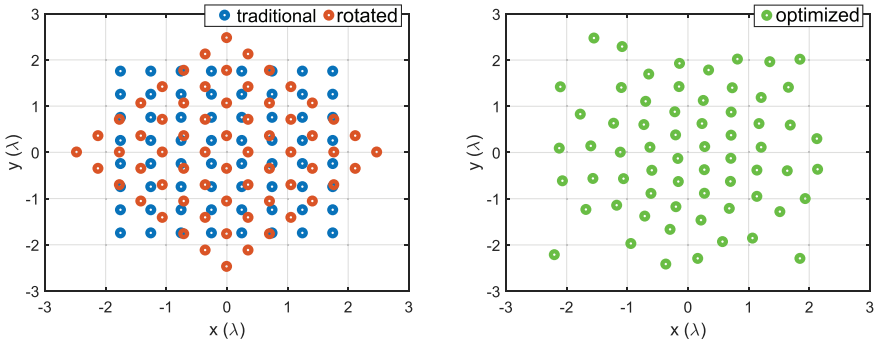


Figure 6.28: Element locations of the traditional regular, rotated and space-tapered array used in the comparative system studies for the users/relays at different heights (Use Case B).

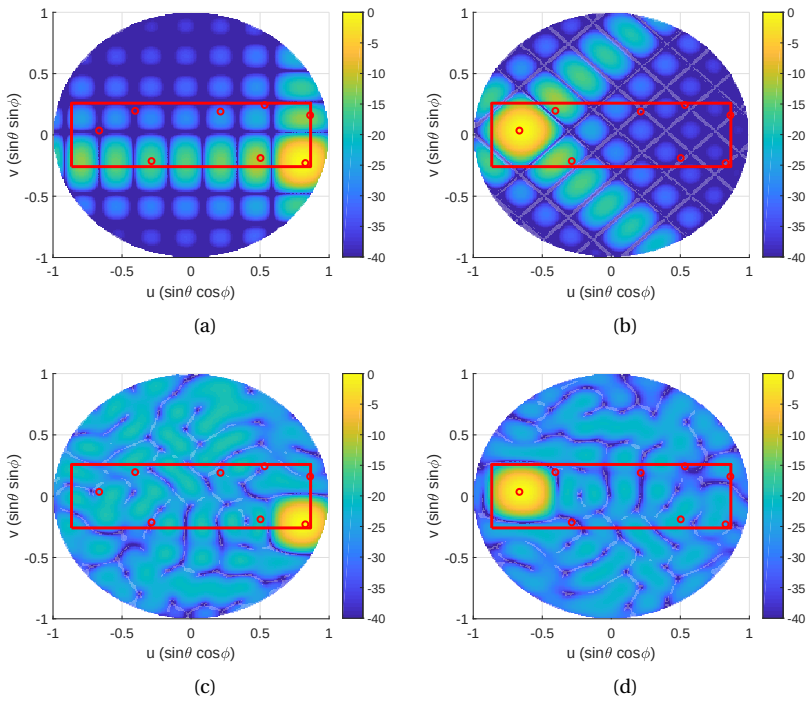


Figure 6.29: Normalized radiation patterns in Use Case B using CB and while serving: (a) UE8 with the square grid array, (b) UE3 with the rotated array, (c) UE8 with the optimized array, (d) UE3 with the optimized array. (Revisit Fig. 2.19 to see the corresponding sample user distribution graph).

be formed by ±15 deg. in elevation and ±60 deg. in azimuth. The CB and reduced-complexity ZF performances of traditional 0.5λ-spaced square array were investigated and the advantage of using the rotated version of array in statistically reducing the inter-user interference was shown. Here, we investigate the performance of an array with the

optimized layout. The array layouts used in this study are given in Fig. 6.28. It is worth to mention that the optimized layout with the ±15 deg. by ±60 deg. sector yields -22 dB maximum SLL with respect to the broadside beam for multiple scanned beams. Considering the sample distribution of 8 users previously given in Fig. 2.19 in Section 2.4.2, the reason behind expecting an improved statistical QoS for the optimal irregular arrays is visualized in Fig. 6.29 where sample cases are shown supporting the reduced inter-user interference of the optimized array. The CB results in this case, given in Fig. 6.30, show that in 95% of the total occurrances, the minimum SINR is 4.9 dB, 4.7 dB and 4.6 dB larger in the optimized array for $K = 4$, $K = 6$ and $K = 8$, respectively, as compared the array with the traditional layout and 3.9 dB, 3.6 dB and 3.4 dB larger for $K = 4$, $K = 6$ and $K = 8$, respectively, as compared the 45-degree rotated array. Last, the superior statistical QoS performance of the optimized arrays in the case of reduced-complexity ZF precoding is shown in Fig. 6.31, which clearly shows the movement of the curves towards the large SINR region with the optimized array in Fig. 6.31(b), as compared to the results from the traditional array in Fig. 6.31(a) (which was also given previously in Fig. 2.23).

SYNTHESIS OF QUASI-MODULAR ARRAYS

The major issue with the fully aperiodic integrated antennas is that they are hard to realize and use in practice since the circuit routing and array calibration may become very complex. Therefore, to maintain the feasibility of array fabrication and reliability of operation, the industry prefers to have more modular layouts, which comes at the expense of increased SLLs. One way to achieve modularity is to assume several different subarray shapes (such as polyominoes) which, when tiled with each other, can mimic the fully aperiodic array [125–127]. Another, and a more straightforward, way to achieve modularity is to enforce a layout symmetry in the beginning of the optimization procedure [114].

In the latter case, the modularity can be based on an N-th order rotational symmetry for which only a slice of $1/N$-th of the total array is to be optimized. Besides the advan-

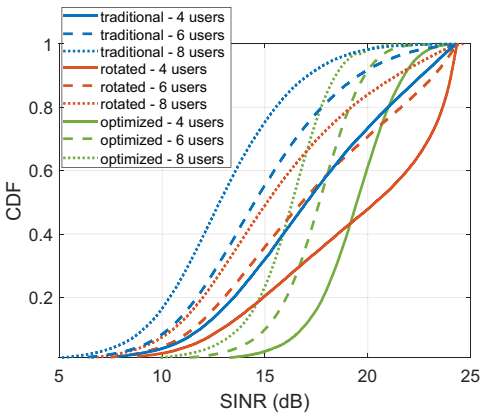


Figure 6.30: CDF of SINR in Use Case B - Approach B using CB with the traditional, rotated and optimized arrays for different number of SDMA users.

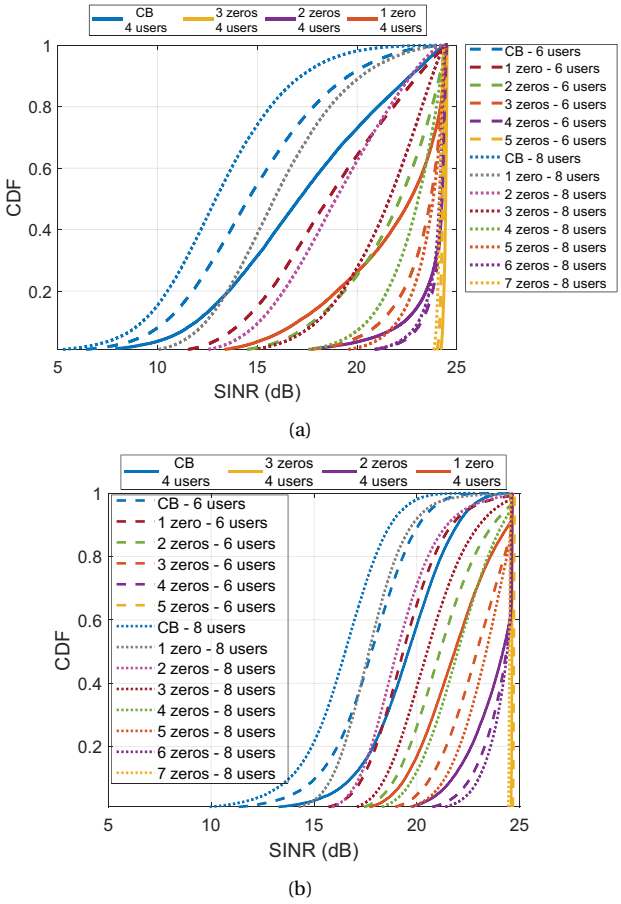


Figure 6.31: CDF of SINR for reduced complexity ZF based on the number of zeros in Use Case B - Approach B using: (a) the square grid array (same with Fig. 2.23), (b) the space-tapered array.

tage regarding the reduced computational efforts, sequentially rotated arrays allow us to synthesize circularly polarized (CP) arrays in an elegant way (using either CP or densely packed linearly polarized (LP) elements) with low cross-polarization in the main beam, reduced mutual coupling and wide axial ratio (AR) bandwidth [128]. Note that CP solutions may even be more preferable in some mm-wave wireless communication systems since LP alternatives would need accurate polarization alignment between the base station and the users [129].

It is also worth to mention here that the sequential rotation technique was first introduced in mid-80's in [130], using elliptically polarized notched patches. The use of LP elements was later proposed in [131]. As evidenced in [132], this requires element spacings much below 0.5λ because, otherwise, with the polarization mismatch at LP element level, only half of the incoming CP power is accepted and, in the transmit mode, essentially half of the power is radiated in four cross polarized lobes. Sequential rotation

has been widely used since then [133; 134].

Here, we combine, for the first time, the irregular array layout design approach with the sequential rotation technique in order to synthesize novel quasi-modular integrated multibeam CP 5G base station antennas with the optimal power efficiencies and relatively low SLLs.

In this study, the element positions are optimized by applying the proposed iterative convex position perturbation technique proposed. The optimization goal this time is to minimize the maximum SLL inside a pre-defined angular sector for a beam which is freely scanned within that sector. The sector is again formed by a ±15 / ±60 degree range in elevation / azimuth. A regular, $\lambda/2$-spaced, 8x8 square grid array is used as an initial layout. The minimum allowed element spacing is set to $\lambda/2$. The main lobe radius in the $u-v$ plane is chosen as 0.2. Depending on the order of rotation (= N), only $1/N$-th of the 8x8 array is used in the optimization and the optimal slice is rotationally repeated over 360 degrees, which gives the total number of elements (= M).

Furthermore, considering the Ka-band quad channel analog beamformer chip of NXP Semiconductors (revisit Fig. 5.11), 4-element subarray clusters are formed within the optimized slice. Since the aim is to have each cluster as compact as possible (to prevent losses) and to place the chips as far as possible (to have enough space for routing and to reduce coupling), the “equal-size-k-means" algorithm from ELKI [135] is used. As the result depends on the algorithm initialization, the chip positions and formed clusters are sub-optimal. The ultimate selection has to be made by considering the design limitations. Note that the state-of-the-art chip of NXP can only generate a single beam at a time. However, the proposed arrays will also work fine with the 5G's multibeam architectures presented in Chapter 4.

***Fully-Irregular 64-Element Array*:** First, to have a benchmark, no symmetry is enforced in the layout optimization, as previously performed in this section. In this case, the maximum SLL (w.r.t. the broadside gain) inside the sector becomes -25 dB (see Fig. 6.32). However, as shown in Fig. 6.33, such a topology requires 16 unique subarray clusters.

***Quasi-Modular 64-Element Array with N = 4*:** In this case, 1/4-th of the initial 8x8 element array (with 16 elements in the first quadrant) is optimized and sequential rotation is used. In this layout, the maximum SLL w.r.t. the broadside gain increases to -19.5 dB, as shown in Fig. 6.34. However, only 4 different clusters are required (see Fig. 6.35).

***Quasi-Modular 72-Element Array with N = 6*:** In the last case, only 1/6-th of the initial 8x8 element array (12 elements in the first slice) is optimized and sequential rotation is applied. In this layout, the maximum SLL w.r.t. the broadside gain becomes -20.3 dB, as shown in Fig. 6.36. The element number increases from 64 to 72. Similarly, the number of chips increases from 16 to 18. However, now, only 3 different clusters are needed (see Fig. 6.37).

Next, the radiation performance of the sequentially rotated arrays is validated via full-wave simulations in CST. Following the work in [130; 132], a pin-fed notched circular patch antenna, shown in Fig. 6.38, is used as the array element. The corresponding design parameters for the center frequency of 27.4 GHz are listed in Table 6.7. Note that the chosen element has a narrow axial ratio (AR) bandwidth and used only for demonstration. In the final prototype, more complex designs with wider AR bandwidth (e.g.

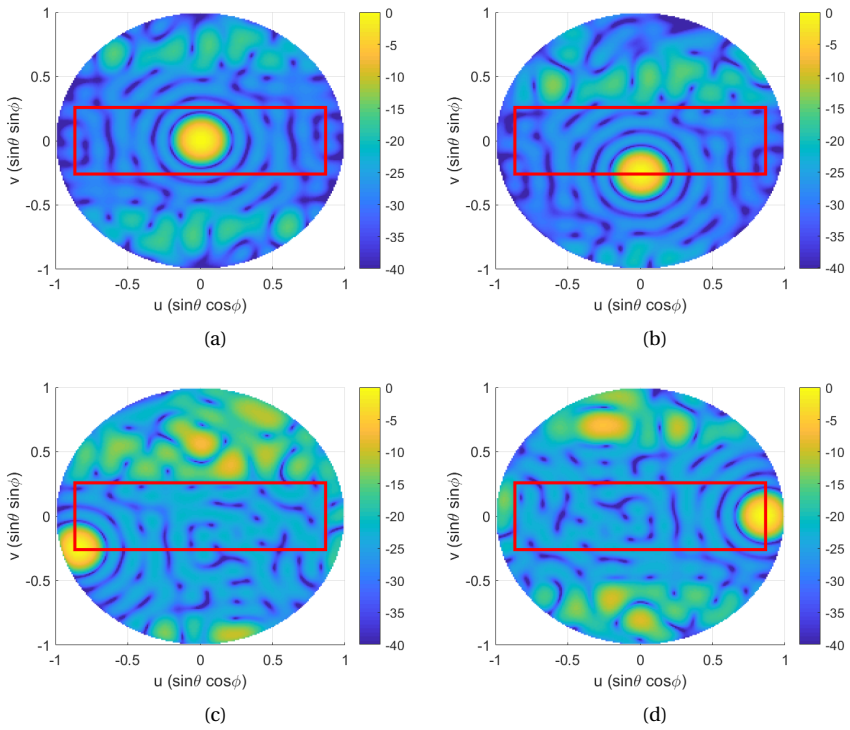


Figure 6.32: Multibeam radiation patterns (in dB, normalized w.r.t. the broadside beam) of the optimized fully-irregular 64-element array: (a) $u = 0$, $v = 0$, (b) $u = 0$, $v = -\sin\pi/12$, (c) $u = -\sin\pi/3$, $v = -\sin\pi/12$, (d) $u = \sin\pi/3$, $v = 0$.

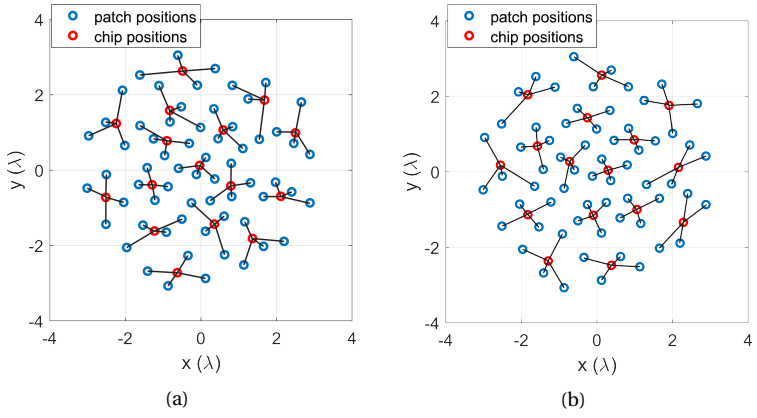


Figure 6.33: Sample sub-optimal subarray clusters for the fully-irregular 64-element array: (a) option #1, (b) option #2.

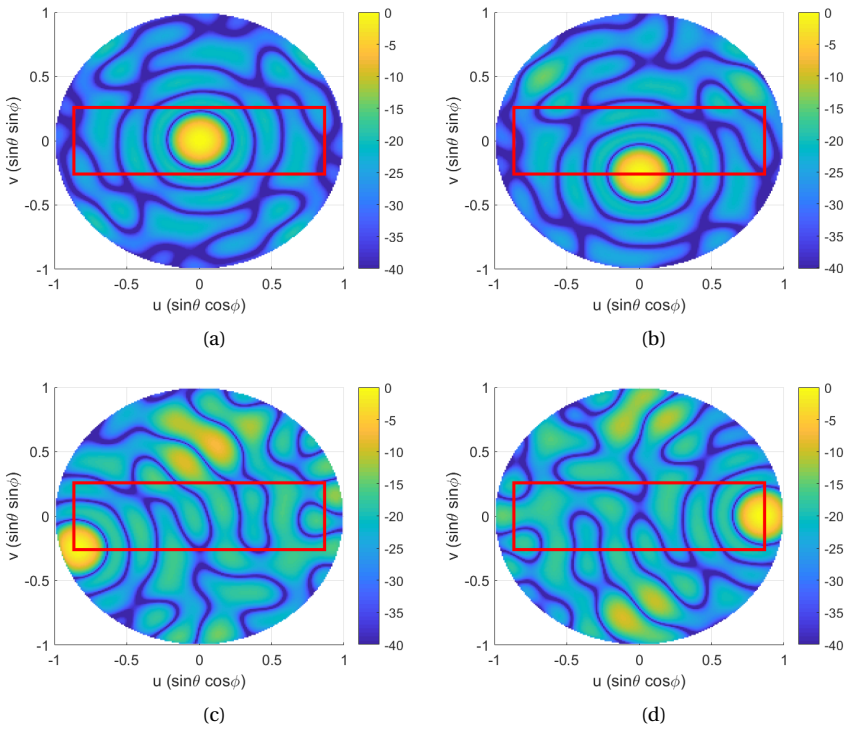


Figure 6.34: Multibeam radiation patterns (in dB, normalized w.r.t. the broadside beam) of the optimized quasi-modular 64-element array: (a) $u = 0$, $v = 0$, (b) $u = 0$, $v = -\sin\pi/12$, (c) $u = -\sin\pi/3$, $v = -\sin\pi/12$, (d) $u = \sin\pi/3$, $v = 0$.

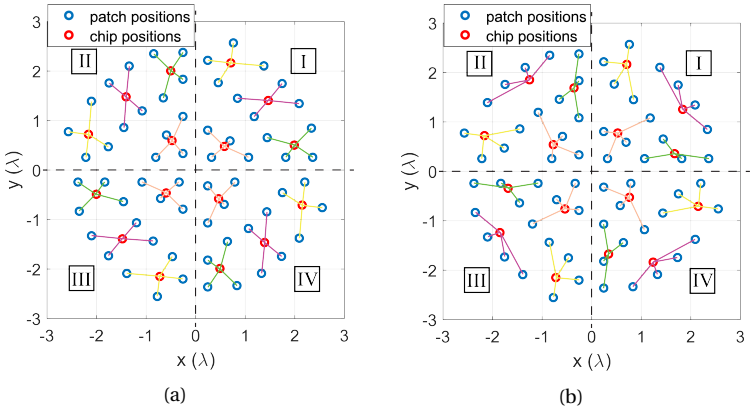


Figure 6.35: Sample sub-optimal subarray clusters for the quasi-modular 64-element array: (a) option #1, (b) option #2.

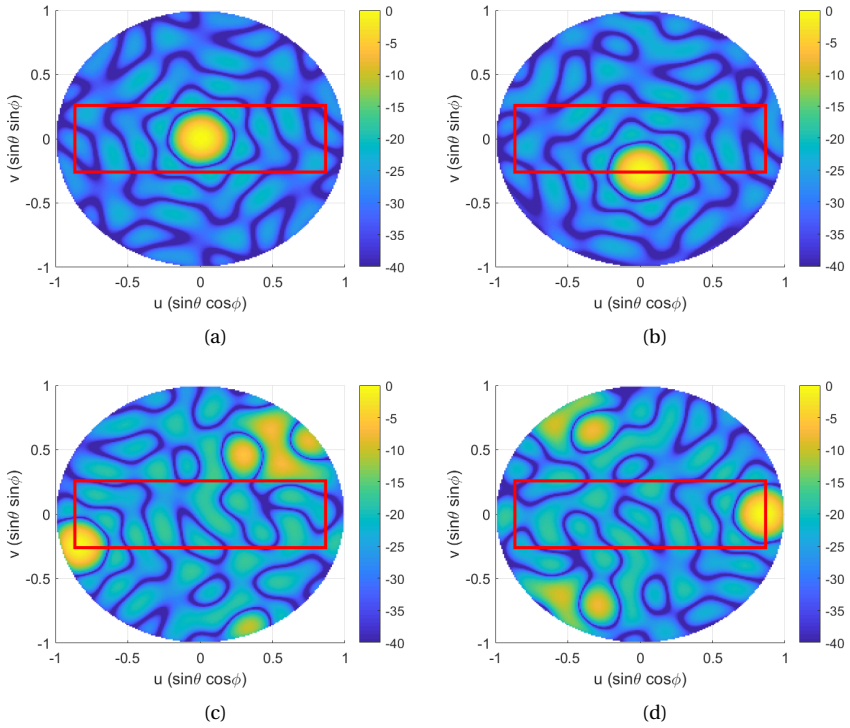


Figure 6.36: Multibeam radiation patterns (in dB, normalized w.r.t. the broadside beam) of the optimized quasi-modular 72-element array: (a) $u = 0$, $v = 0$, (b) $u = 0$, $v = -\sin\pi/12$, (c) $u = -\sin\pi/3$, $v = -\sin\pi/12$, (d) $u = \sin\pi/3$, $v = 0$.

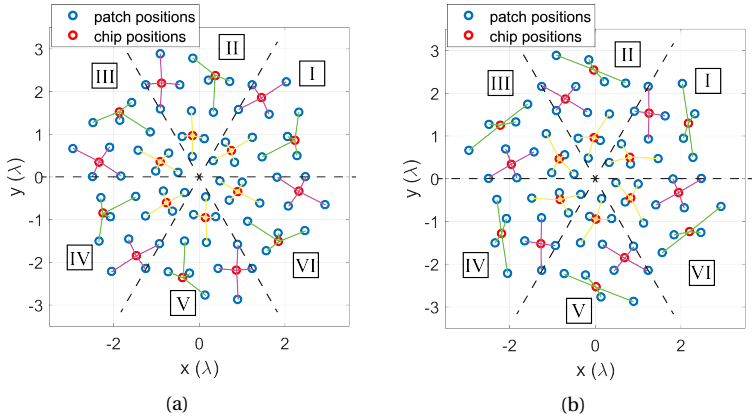


Figure 6.37: Sample sub-optimal subarray clusters for the quasi-modular 72-element array: (a) option #1, (b) option #2.

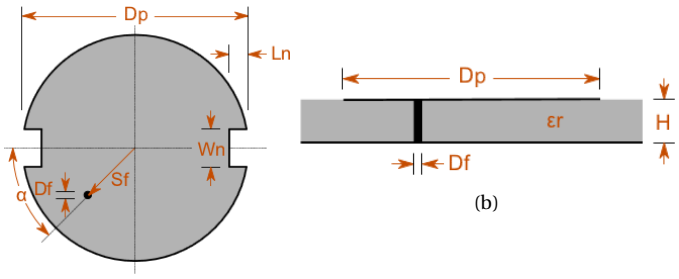


Figure 6.38: Pin-fed notched circular patch antenna: (a) top view, (b) side view.

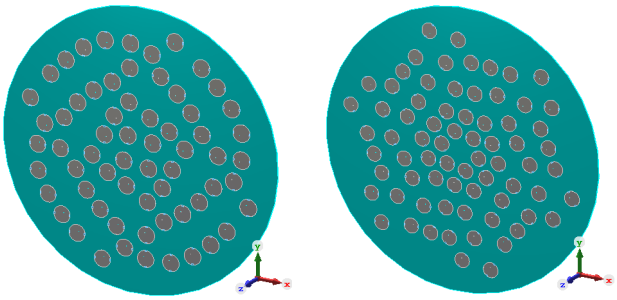


Figure 6.39: Quasi-modular antenna arrays designed in CST for full-wave simulations.

[136]) may be preferable.

Fig. 6.39 shows the optimized 64- and 72-element sequentially rotated arrays designed in CST. On top of the progressive phase shifts used for beam steering, the element phases are increased in 90 degree steps from Region I to Region IV in the 64-element array, and in 60 degree steps from Region I to Region VI in the 72-element array (see Fig. 6.35 and Fig. 6.37 for the denoted regions). The corresponding right-handed CP radiation patterns at broadside and at the cell edge are provided in Fig. 6.40 and Fig. 6.41

Table 6.7: CP antenna element design parameters

Symbol	Definition	Value
D_p	patch diameter	4.202 mm
L_n	notch length	0.268 mm
W_n	notch width	0.536 mm
D_f	feed pin diameter	0.107 mm
S_f	feed pin offset	0.654 mm
α	feed pin rotation angle	45 deg.
ϵ_r	substrate relative permittivity	2
H	substrate height	0.227 mm

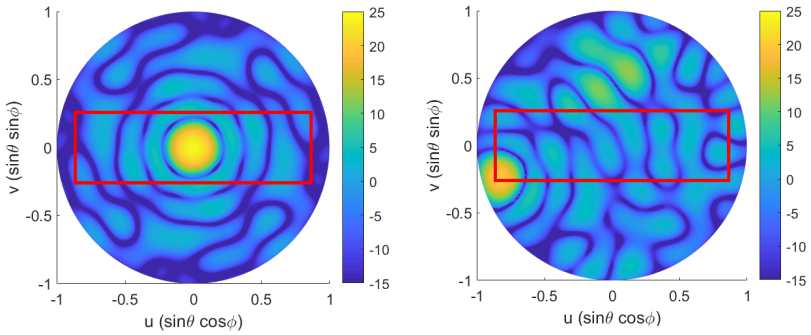


Figure 6.40: Array directivity pattern (right-handed CP, in dBi) using full-wave simulation for the 64-element quasi-modular array: (a) $u = 0$, $v = 0$, (b) $u = -\sin\pi/3$, $v = -\sin\pi/12$.

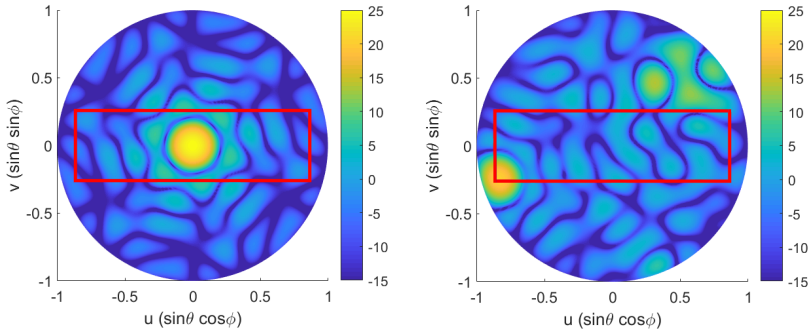


Figure 6.41: Array directivity pattern (right-handed CP, in dBi) using full-wave simulation for the 72-element quasi-modular array: (a) $u = 0$, $v = 0$, (b) $u = -\sin\pi/3$, $v = -\sin\pi/12$.

Table 6.8: Array directivities (in dBi) at the center frequency of 27.4 GHz

Array type & Beam position	Co-Polarization	Cross-Polarization
64-element, broadside	24.4	-1.3
64-element, sector edge	20.1	3.4
72-element, broadside	25	1.6
72-element, sector edge	20.2	2

for the 64-element array with $N = 4$ and the 72-element array with $N = 6$, respectively. It is seen that the results are very similar to the ones previously shown in Fig. 6.34 and Fig. 6.36, except for the slightly higher (about 1-2 dB) first side lobe along the symmetry axes. The difference originates from the effect of mutual coupling which is ignored during the optimization to reduce the computational efforts. For completeness, the co- (right-handed CP) and cross- (left-handed CP) polarization directivities at the center frequency are listed in Table 6.8.

6.2.4. CONCLUSION

A novel, system-driven array layout synthesis technique has been proposed for multi-beam phased arrays that are suitable for mm-wave 5G base station applications in simultaneous multi-user scenarios. The proposed technique achieves the optimal statistical system performance within a cell sector, while approaching the thermal management issue in the active inteagrated antenna arrays in a systematic way, which is discussed in more details in the following chapters.

The presented algorithm combines multiple system-related issues in a single optimization procedure with low analytical complexity, compactness and easy solvability. It is based on an efficient iterative convex optimization scheme with joint capabilities, namely, (i) minimized maximum SLL everywhere in the field of view for a beam scanned freely inside a given cell sector, which creates the interference-awareness and with (ii) uniform excitation amplitudes for optimum power efficiency, (iii) a pre-defined minimum inter-element spacing to prevent element overlapping and ease the thermal problem caused by the temperature increase due to the too-close elements, (iv) an increased layout sparsity via irregularity to obtain further passive cooling, which imply the thermal-awareness.

The algorithm's superior performance in comparison with the state-of-the-art methods ([104; 106; 107; 110–117]) has also been demonstrated both qualitatively and quantitatively.

Full-wave EM simulations have shown that high mutual coupling levels may affect the radiation performance and lead to unreliable optimization results. In such cases, either mutual coupling reduction techniques can be exploited in the antenna design or embedded radiation patterns can be straightforwardly integrated into the proposed optimization procedure at the cost of increased computation time and resources.

The proposed method has been applied for the synthesis of a 64 element irregular active integrated base station arrays to be used in different 5G use case scenarios. From the simulation results, it has been observed that, by using the optimized layouts, SLLs lower than -20 dB can be achieved in the field of view, which leads to a superior system performance (in terms of the QoS and processing complexity) as compared to that of the traditional regular arrays.

Finally, the synthesis of quasi-modular, circularly polarized multibeam 5G base station antennas with relatively low SLLs has been studied to ease the fabrication and circuit routing requirements of the fully aperiodic arrays. Only a small portion of the whole array has been used in the iterative convex layout optimization routine, which has been combined with (i) the equal-size-k-means clustering algorithm to form the sub-optimal subarray clusters and (ii) the sequential rotation technique to achieve modularity. Note that, with the proper patch center adjustments, the proposed irregular layout optimization algorithm can also be straightforwardly used to synthesize low-sidelobe linearly polarized arrays having a rotationally symmetric feed network.

By using an interleaved-shared layout, it is also possible to extend the proposed layout optimization technique to SLL minimization for multiple steerable beams having distinct beamwidths. The major advantages of this approach are:

(i) The cooling challenge in the integrated 5G antennas is relaxed via power-efficient uniform-amplitude excitations and array subset activations (i.e. thinning) correspond-

ing to the desired beam resolution.

(ii) By applying a smart irregularity in the array subset layouts, the SLLs are significantly reduced jointly for multiple wide-angle scanning varying-resolution beams.

(iii) Since the smaller-sized arrays generating wider beamwidths share their elements with the larger arrays, the total array aperture is efficiently used.

Having optimal array subset layouts with wide-angle scanning and zooming ability is crucial for efficient link setup and high-QoS communication. Interested readers are referred to [J-10] and the references therein for more details.

6.3. CONCLUSIONS

The system analysis in Chapter 2 has shown the importance of effective and efficient SLL suppression on the statistical QoS performance in 5G SDMA applications. There, it was indicated that CB precoding algorithm provides optimal power efficiency and the least computational complexity, but its SINR performance is, in general, not satisfying, especially when it is used with the currently proposed $\lambda/2$-spaced square-grid arrays having a relatively high maximal SLL. The ZF precoding, on the other hand, can cancel out the inter-beam interference in the current systems, but it works well only under ideal system conditions in the case of having full and accurate channel state information. Therefore, in this section, we have proposed ***two novel effective SLL suppression approaches that are based on: (i) a phase-only control technique, (ii) a position-only control technique***.

An **iterative convex optimization method** has been used in both techniques such that small phase (or position) perturbations are applied to each element at each iteration, which allow us to linearize the non-linear far-field expression around the unknown phase (or position) parameters. The resulting SOCP optimization problem has been solved using the free commercial convex programming solver CVX. The convergence in the minimized maximal SLL has been observed after several iterations.

In the **phase-only control**, it has been shown that the existing phase shifters in phased arrays can be exploited to **reduce the SLLs and put narrow/wide nulls towards the unintended users** (with known angular locations) while scanning the beam, but at the expense of reduced power efficiency (whose extent depends on the array topology). The main advantages of the proposed technique have been reported as **lower SLL compared to progressive phase shifts, no increase in array size, easy integration of embedded patterns and simpler feed network than amplitude or position control**. The disadvantages, on the other hand, have been listed as much worse power efficiency than non-optimized arrays, the need for different set of phases optimized for each scan angle and SLLs higher than that of the amplitude or position-optimized arrays.

In the **position-only control**, it has been explained that the positions of the antenna elements can be optimized to **reduce the SLLs with maximal energy efficiency (no amplitude/phase control used), relax the thermal requirement and have more space for the electronics**. The main advantages of this technique have been mentioned as its **practical interest, optimal power efficiency, simultaneous optimization capability for multiple beams and the possibility to control the minimum element spacing/sparsity** (space for the electronics, cooling). The system advantages (in terms of the QoS and processing burden) of the synthesized irregular arrays over their regular counterparts have

also been pointed out via comparisons with the study cases previously used in Section 2. On the other hand, the disadvantages of position-tapered arrays have been listed as the difficult integration of the embedded element patterns (and thus, mutual coupling), increase in array size and irregular feed network with increased need for calibration. To **ease the design, fabrication and calibration challenges of the fully-aperiodic arrays, quasi-modular arrays with rotationally symmetric feed networks** have also been proposed, where the sequential rotation has been exploited to synthesize circularly polarized arrays with improved polarization characteristics (wide axial ratio beamwidth and low cross-polarization in the main beam).

III

THERMAL DRIVEN ASPECTS

Those who cannot understand how to put their thoughts on ice should not enter into the heat of debate.

Friedrich Nietzsche

7

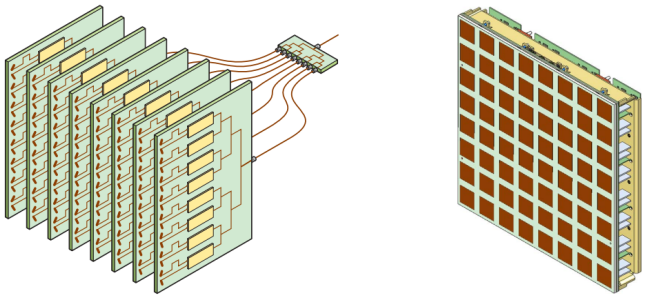
Cooling Strategies for Active Integrated Phased Arrays - An Overview

Parts of this chapter have been published as:
[J-6] Y. Aslan, J. Puskely, J. H. J. Janssen, M. Geurts, A. Roederer and A. Yarovoy, "Thermal-aware synthesis of 5G base station antenna arrays: an overview and a sparsity-based approach," *IEEE Access*, vol. 6, pp. 58868-58882, Oct. 2018.

7.1. TRADITIONAL VS PLANAR AESAS

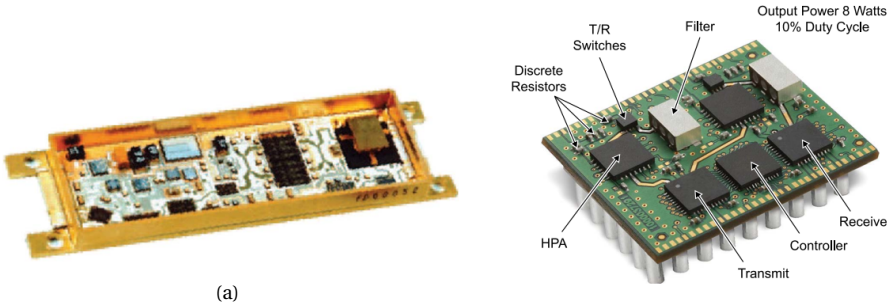
The development and use of modern phased arrays (i.e. Active Electronically Scanned Arrays, AESAs) roots back to more than 50 years of valuable research and implementation, exclusively in space and defense applications [137; 138]. An example of traditional structure of AESAs is given in Fig. 7.1(a). In such constructions, a number of circuit cards are placed orthogonal to the antenna array. Each card feeds a row of antenna elements with individual transmit/receive (T/R) modules mounted on them. The advantage of this approach can be stated as providing a large surface area for the T/R modules and for the thermal load. A major disadvantage is the need for a large number of RF boards and cabling for signal routing. Moreover, in the case of having large size T/R modules, it could be difficult to make a compact design.

Fig. 7.1(b) visualizes an alternative approach that, in this paper, is referred to as planar AESA. In this approach, a single multi-layer RF board is used to integrate the antenna elements and RF beamformers. The radiating elements are placed on one side of the board while all the beamforming electronics are mounted directly onto the back of the board. Compared to the traditional approach, this alternative has the advantage of reducing the area of RF boards and the number of connectors/cables. Therefore, pla-



(a) (b)

Figure 7.1: Illustration of AESA construction [137], (a) traditional approach, (b) planar approach.



(a) (b)

Figure 7.2: Sample AESA T/R modules [137] for, (a) a traditional AESA, (b) a planar AESA.

nar AESAs can provide low-cost and low-profile apertures, which makes handling and mounting of the arrays more flexible. Due to these advantages, planar AESAs are good candidates for the high-throughput, but low-cost antennas in the first phase of digital 5G systems.

Despite their superiority in compactness, planar AESAs have limited space to fit all the beamforming electronics. In the traditional approach, a separate T/R module with a discrete low noise amplifier (LNA), power amplifier (PA), converters, control components etc. is required for every radiating element. Using a similar design strategy in planar arrays would result in an unrealizable layout where the required electronics do not fit in the allowable space on the board. With increased frequency, since the physical spacing between the radiating elements becomes even smaller, the problem of excessive heat dissipation gets more serious.

For better understanding and visualization of the described concepts, an active X-band T/R module in a traditional AESA developed by Raytheon for missile defense applications [137] and a low-profile planar AESA T/R module developed for a multi-function phased array radar [137] are shown in Fig. 7.2.

7.2. COOLING IN TRADITIONAL AESAS

In the traditional AESAs, no cooling module can be placed to the bottom of the array because of the space occupied by the electronics. The same restriction holds for the upper region as well, since otherwise the antenna pattern would be significantly influenced. Therefore, the cooling has to be achieved from the sides of the array, where the space for sufficient heat dissipation exists.

In [139], the thermal problem in traditional AESA's was discussed by investigating the effect of PA efficiency on the operating temperature of the PAs. The severity of the problem was shown using the PA in [140] as the baseline for LTE signals, which provides around 9% PA efficiency. An 16 x 8 array was assumed around which an ideal heat sink with 30 degree constant temperature is maintained. The resulting temperature distribution is given here in Fig. 7.3 which clearly shows that the heat gets trapped in the center of the array, resulting in an unacceptable temperature rise. That is why thermal engineers often include (possibly liquid-cooled) cold plates in the array assembly as shown in Fig. 7.4 for rapid heat transfer towards the sides where a secondary cooling system is placed.

Although serious challenges with thermal management occurred recently in the wireless communication area with the introduction of 5G and its intended performance re-

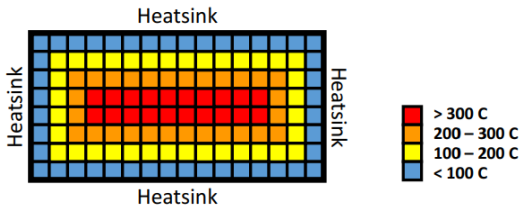


Figure 7.3: Temperature distribution across a traditional 8x16 AESA [139].

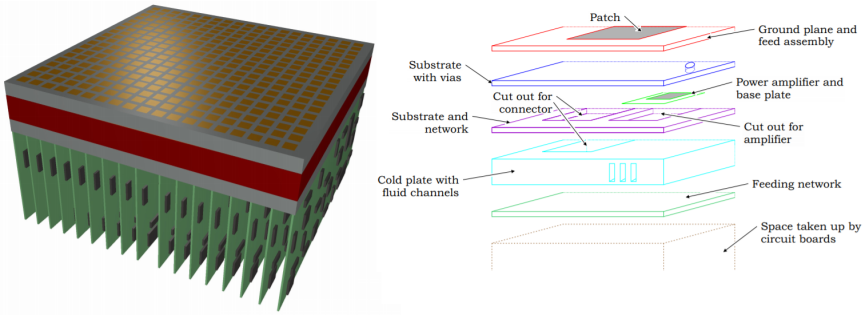


Figure 7.4: Sketch of an AESA array with the layer structure showing one element. The position of the cold plate is indicated by the red layer (by Deutsches Zentrum für Luft- und Raumfahrt (DLR), [147]).

(a) (b)

Figure 7.5: An airborne phased array radar antenna [146], (a) array aperture, (b) a row of microwave modules on a slat.

7

quirements, which has created more interest in array cooling [139; 141; 142], excessive heat generation has always been a critical issue for phased array antenna design in such systems as military electronics [143; 144] and space applications [145]. As an example, an X-band airborne phased array antenna [146] is given here in Fig. 7.5 which shows many microwave modules having transmit and receive circuitry with GaAs power amplifiers that are attached to liquid cooled slats.

7.3. COOLING IN PLANAR AESAS

It was seen that, the heat generated from a traditional AESA gets trapped in the middle region of the array since the cooling is done at the array edges where the heat sink is placed. Therefore, for such arrays, it is necessary to transfer the generated heat towards the edges via cold plates or fluid channels. Thermal management in planar AESAs is different than the traditional counterparts due to the placement of the beamforming chips. Since the ICs are now located at the opposite side of the substrate compared to the radiating elements, external heat sinks can be attached to the chips, without affecting the radiation performance. As seen in Fig. 7.6, thermal interface materials (TIMs), thermal glues or some flat plates with large thermal conductivities are used to enhance the thermal conduction and heat spreading from the chip to the lid and the heat sink [148].

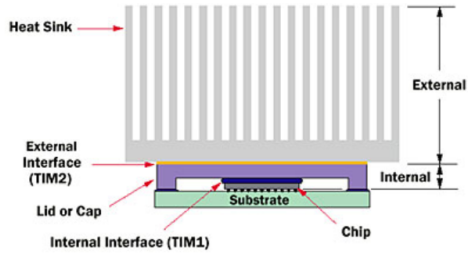


Figure 7.6: Chip package attached to a heat sink [148].

Figure 7.7: Package mounted on evaluation board with heat sink [142].

For better visualization, a sample design with a package mounted on an evaluation board with heat sink is shown here in Fig. 7.7, which presents a low-cost 60 GHz antenna-in-package (AiP) phased array [142].

7.4. ARRAY COOLING AT MM-WAVES

To achieve the demanding throughput requirements of 5G, a vast amount of spectrum is necessary, which is available in the mm-wave bands. However, despite the advantage of having a large bandwidth, the system efficiency of the base station arrays at mm-waves gets worse due to highly inefficient silicon-based PAs (or transceiver ICs in general) [149; 150], with high linearity requirements. As a result of the low efficiency of the mass-produced chips at high frequencies, more heat is generated in the system which makes the thermal management even more challenging.

Therefore, this section is allocated to the existing cooling strategies for mm-wave arrays. The techniques applied today are mainly active cooling methods exploiting either forced air or liquid cooling. It is worth to mention here that such active systems require the use of electricity and therefore results in higher costs and maintenance requirements.

Now, let us go through some examples that shows the problem more clearly.

In [147], cooling systems for a Ka-band (operating at 30 GHz) transmit aircraft antenna were designed. TriQuint's GaAs MMIC power amplifier [152] was used to feed each antenna, for which the power dissipation could reach up to 3.8 W in the worst case. Two active cooling strategies were employed; forced liquid cooling and forced air cooling (which is easier to set up than liquid cooling). The designs were demonstrated for a 4 x 4 array of amplifiers. In Fig. 7.8(a), the forced air cooling design is shown where the heat is spread through the copper cold plate and fans are used to blow air towards the heat sinks. In Fig. 7.8(b), the liquid is pumped through the channels of the cold plate in order to transfer the heat towards the outside of the array, where it is then dissipated via heat exchange to a different liquid or air.

Another liquid cooling design was made for a 3x16 element digital beamforming Substrate Integrated Waveguide (SIW) based transmitter array at 30 GHz for mobile satellite communication application at Ka-band [151], which is shown here in Fig. 7.9. Together with the designs given in Fig. 7.8, this example highlights the challenge of thermal management in mm-wave antennas.

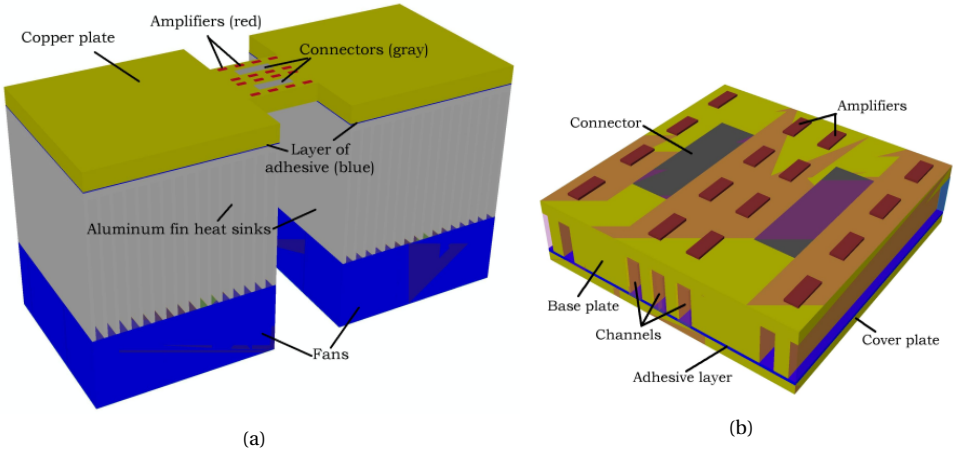


Figure 7.8: Cooling systems for a 4x4 Ka-band transmit antenna array [147], (a) forced air cooling, (b) forced liquid cooling.

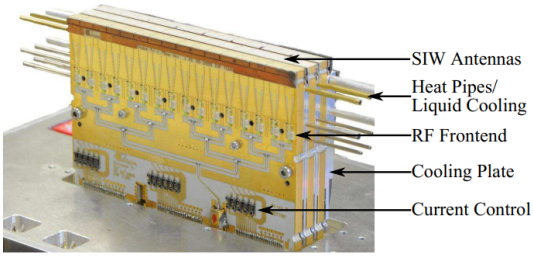


Figure 7.9: Cooling of a 3x16 element digital beamforming transmitter array at 30 GHz [151].

7.5. HEAT SINK ANTENNAS

In the previous sections, several electronics cooling strategies were explained and illustrated by focusing on the removal of the heat which flows from the ICs towards the cool-

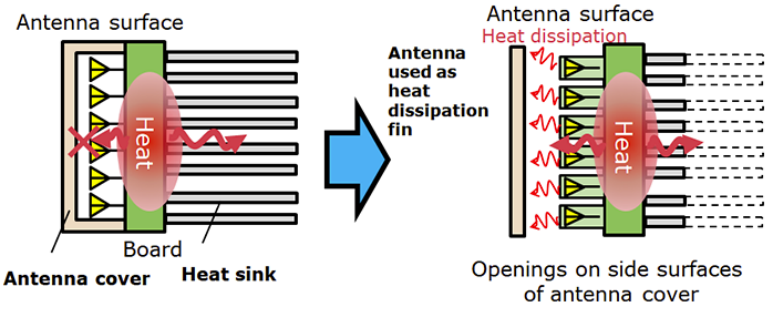


Figure 7.10: The concept of "heat sink antenna": a method to ease the thermal problem in planar AESAs [153].

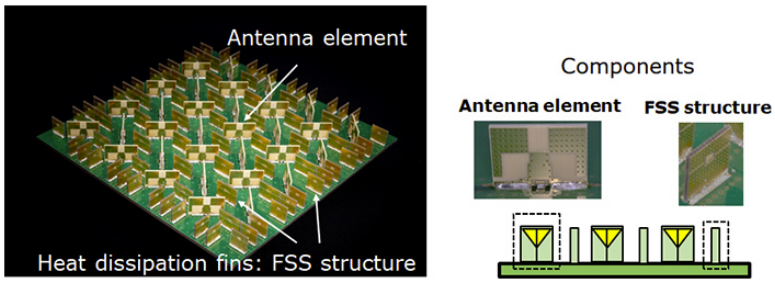


Figure 7.11: The heat dissipation antenna array from NEC Corporation [153].

ing system (fin heat sinks, fans, water pumps etc.). However, one should not neglect that there is a second side of the coin, the antenna element itself ! As shown in Fig. 7.10, in planar double-sided AESAs, part of the heat goes from the ICs to the heat sink at the back of the board, while the remaining reaches to the antenna elements on the other side. In conventional array designs, the antenna surface is not made to dissipate heat. On the other hand, in mm-wave 5G, using the antennas as a heat sink can potentially ease the thermal problem by providing an additional cooling path and decreasing the stringent requirements (size, energy consumption etc.) on the cooling systems. In this section, a number of "heat sink antenna" concepts presented in the literature are briefly covered and their suitability for the 5G systems are discussed.

The first example is a very recent 5G heat sink antenna array from NEC Corporation, which can be seen in Fig. 7.11. In this array, the antenna elements function as heat dissipation fins. Besides, the elements employ a Split-Ring structure for size reduction and additional metal plates are placed in between the sparsely distributed elements. These plates are made electromagnetically transparent to the antennas at the operating frequency by employing a Frequency Selective Surface (FSS) on them. This way, the radiation properties of the conventional array was preserved.

Although the concept of heat dissipation antenna was brought back to 5G arrays as a new technology by NEC, similar approaches were in fact studied by few others in the field more than a decade ago. For example, an active heat sink patch antenna was proposed in [154]. The idea was to create a vertical heat spreader connecting the IC to the patch,

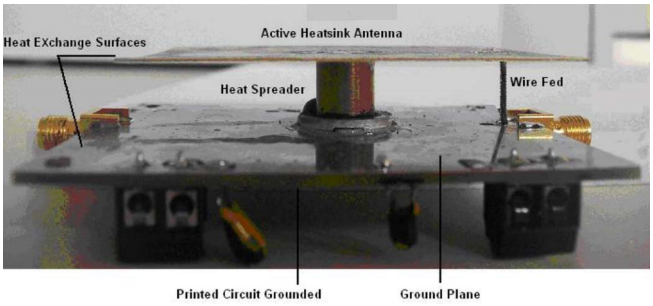


Figure 7.12: RF transmitter with the heat-sink patch antenna using a heat spreader [154].

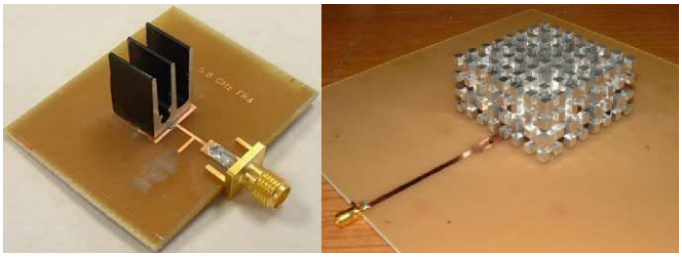


Figure 7.13: Finned and fractal heat-sink antennas [155; 156].

as seen in Fig. 7.12. It was found that such an antenna is able to provide a desirable radiation and thermal performance at the same time. Some other techniques concentrate more on the finned or fractal element structures for dual (electromagnetic and thermal) functionality [155; 156]. Examples of such antennas are illustrated in Fig. 7.13.

Despite proven to be effective in terms of the heat removal, the presented heat sink antennas employ a complex 3-D antenna element structure which is hard and costly to fabricate. Since the 5G industry today prefers to use fully-planar antennas with low-cost mass production via printed circuit board technology, the applicability of the existing techniques is very limited and almost non-existent in the first phase of the commercial products.

7.6. CONCLUSIONS

In this chapter, ***various thermal management strategies in AESAs*** have been discussed through the help of many illustrative examples from the state-of-the-art literature. Firstly, the evolution of modern phased array architectures and its impacts on both design and cooling challenges have been explained from a 5G system perspective. It has been shown that traditional AESAs consist of large number of RF boards, need cabling for signal routing and result in incompact designs, which makes them unsuitable for the mm-wave 5G communication systems. Besides, it has been stated that active cooling methods (fans or forced liquids) are commonly used in such configurations in order to transfer the heat trapped in the middle of the array to the array edges. On the other hand, being **compact, low-cost and low-profile, planar AESAs** have been presented as favorable candidates in 5G. The thermal challenge in planar topologies has also been reported as very serious due to the **limited space for the electronics and extreme heat generation in this limited space**, especially at mm-waves. Finally, **the role of antenna researchers in array cooling** has been discussed and a relatively **new concept to 5G, heatsink antenna**, has been revisited. It has been seen through existing prototypes that when meticulously designed in 3-D, the antenna itself can be used as a heatsink that helps remove the heat reaching to the antenna side. Such antennas can **ease the thermal problem and relax the cooling requirements on the chip side**. However, the applicability of them for the first-phase of 5G systems has been claimed to be very limited in practice since today, the industry prefers fully-planar and easy-to-fabricate array options. Therefore, later in Chapter 9, we introduce passive cooling & cooling enhancement by element topology optimization and enlarged/thick array base plate.

8

PASSIVE COOLING IN 5G BASE STATION ANTENNAS - WHY AND HOW?

Parts of this chapter have been published as:
[J-6] Y. Aslan, J. Puskely, J. H. J. Janssen, M. Geurts, A. Roederer and A. Yarovoy, "Thermal-aware synthesis of 5G base station antenna arrays: an overview and a sparsity-based approach," *IEEE Access*, vol. 6, pp. 58868-58882, Oct. 2018.
[C-7] Y. Aslan, C. E. Kiper, A. Biggelaar, U. Johannsen and A. Yarovoy, "Passive cooling of mm-wave active integrated 5G base station antennas using CPU heatsinks," in *Proc. 16th EuRAD*, Paris, France, pp. 121-124, Oct. 2019.
[C-8] Y. Aslan, J. Puskely, A. Roederer and A. Yarovoy, "Effect of element number reduction on inter-user interference and chip temperatures in passively-cooled integrated antenna arrays for 5G," in *Proc. 14th EuCAP*, Copenhagen, Denmark, Mar. 2020.

8.1. INTRODUCTION

As previously discussed in Chapter 2, because of the relatively small package sizes and low-efficiency chips at mm-waves, the 5G industry is currently facing serious thermal challenges. As a rule of thumb, it is commonly accepted that every 10^oC rise in the integrated circuit (IC) junction temperature reduces the average life of the chip by 50% (through the Arrhenius equation [157]). Moreover, the average junction temperature should be kept below 125^oC under typical working conditions at steady state for a safe and reliable device operation [158].

Heat transfer phenomenon can be grouped into three categories: conduction, convection and radiation. In most cases, there is no single dominant heat transfer mechanism and more than one mechanism is simultaneously effective. Therefore, it is important to understand the physics behind each category and sort out the most powerful one(s) in terms of cooling, depending on the scenario at hand.

Heat conduction is the transfer of heat through physical contact for which the heat flows from hot to cold regions. Thus, it occurs in solid, liquid or gaseous media when there is a temperature gradient. Heat convection, on the other hand, is due to a moving gas or liquid. Convection mechanism generally consists of natural (or free) and forced methods. Natural convection occurs passively via the gravity. To be more exact, the buoyancy of the fluid created by the temperature gradients causes pressure and density differences such that the fluids become less dense when heated and float up. In the forced (active) methods, the fluid flow is initiated and maintained by an external appliance such as a pump or a fan. Despite being effective in removing the heat, active cooling systems with fans or pumps require the use of electricity and make the system more complex and harder to maintain. On the other hand, passive thermal management is a relatively cheaper and more energy-efficient solution. However, since the heat is removed only passively via natural convection by utilizing only heat spreaders or heat sinks, it is not easy to achieve thermal performance similar to that of the active counterparts unless there is sufficient heat sink surface area that is in contact with outside. Since gravitation plays an important role, the orientation of the heat sink becomes critical as well.

Apart from the conduction and convection, there is a third mechanism, namely heat transfer by means of electromagnetic radiation. In fact, no additional medium is necessary for thermal radiation since it occurs due to the thermal motion of particles in matter. Every object emits thermal radiation whose spectrum and power depends on the object properties and its absolute temperature. Now, for easy demonstration of the effectiveness of radiative cooling, let us consider a special kind of object, a blackbody, which is assumed to emit maximal thermal radiation. The power (in W) emitted by a blackbody is given by

$$P = \sigma A T^4 \tag{8.1}$$

which is called as the Stefan-Boltzmann law, where σ is the Stefan-Boltzmann constant given as 5.67 x 10^{-8} $Js^{-1}m^{-2}K^{-4}$, A is the total radiating area in m^2, and T is the absolute temperature in Kelvin. If we assume an IC that has a temperature of 80^oC with a surface area of 5 mm x 5 mm ($\lambda/2$ x $\lambda/2$ at 30 GHz) around it, the dissipated power via thermal radiation is calculated through (8.1) as 22 mW from one surface and 44 mW if

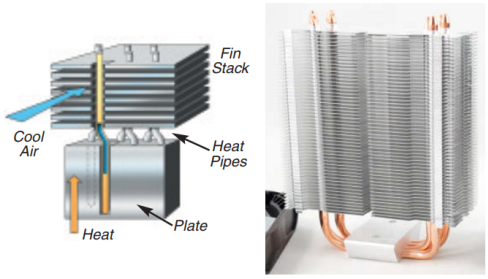


Figure 8.1: CPU cooling via heat pipes and a remote fin stack (taken from [160; 161]).

we consider the surfaces on both sides. Such figures on the order of tens of mW indicate that only a few percent of the total generated heat (on the order of a few Watts per IC [152]) can be removed via radiation, unless there is a large cooling surface. Therefore, conduction and convection are the dominating heat transfer and cooling mechanisms in highly packed mm-wave integrated antennas. The contribution from radiative cooling increases with extended heat exchange surface designs, and becomes more noticeable in a low-convection scenario, such as cooling with a low-airflow non-finned flat plate.

At the Ka- frequency band, in which most of the current 5G mm-wave work is conducted, the severe cooling issue is handled by active cooling strategies employing fans or water pumps [147]. However, high-volume and low-cost communication market demands fully-passive thermal management. For this purpose, we proposed, for the first time, to make use of the commercial fanless CPU coolers [J-6]. Indeed, CPU coolers (without fans) can potentially be used in planar AESA's, for which the heat sink can be attached to the chips at the opposite side of the radiating elements. Although they are bulky and heavy, the heat sink modules for CPU coolers are able to provide competitive results in heat removal when compared to the active systems [159]. This claim will be further supported in the next sections via the simulation results and experiments.

Here, for better visualization, Fig. 8.1 is given, which shows how the heat can be absorbed from the chips by the heat pipes (which are totally passive with low thermal resistance and very efficient heat transfer [160]) and transported to the fin stack which is cooled via natural (or forced) convection, and a sample CPU cooler fin stack from FSP Windale [161].

The rest of the chapter is organized as follows. Thermal system modeling approaches in electronics design are revisited in Section 8.2. The thermal simulation settings and the results for the CPU heat sink integrated antennas are given in Section 8.3. Section 8.4 presents the experimental validation and discussions. The system limitations dictated by the passive-only thermal control are explianed in Section 8.5. The chapter is concluded in Section 8.6.

8.2. THERMAL SYSTEM MODELING IN ELECTRONICS DESIGN

In order to assess the performance of the cooling systems and investigate the effect of different design parameters (heat sink size, shape, power per IC, thermal characteristics of the IC and the board etc.) on the junction temperature of the beamformer chips, it is

very important to have an appropriate thermal model of the overall system.

The thermal models in electronics design can be grouped into two categories: detailed thermal models (DTMs) and compact thermal models (CTMs) [162]. DTMs present an almost-exact physical model of the chips with the actual package geometry. Despite being very accurate, DTMs are not suitable for system-level studies with a large number of ICs due to their high computational burden. CTMs, on the other hand, do not try to mimic the package geometry and material properties. Instead, they aim to accurately predict the temperature only at a few critical points of the chip package (junction, case, etc.), mostly by using an equivalent thermal resistor network. Because of their high computational efficiency, CTMs are much more suitable for investigating system-level designs and exploring what-if scenarios.

Today, the most commonly used CTMs are the two-resistor [163] and DELPHI [164] models, which are shortly discussed next. It is worthy to note that both models are standardized by the JEDEC Solid State Technology Association.

8.2.1. TWO-RESISTOR CTM

The two-resistor model of JEDEC consists of three nodes, as seen in Fig. 8.2. The thermal problem is re-formulated using a thermal resistance network which is similar to electrical circuit modeling. The junction node represents the junction of the IC where the heating power, P_h, is applied. The case node represents the top of the package. The equivalent thermal resistance between the junction and the top of the case is given by θ_{JCtop}. The board node represents the contact of the package with the PCB. θ_{JB} is the equivalent thermal resistance between the junction and the board. It is worthy of note that, in this model, it is assumed that the heat flows through the case node and the board node. Therefore, the heat flow through the sides of the package is not taken into account.

In planar AESAs, the top of the package is in contact with a thermal interface material and a heat sink, which is equivalent to a case-to-ambient resistance. The printed circuit

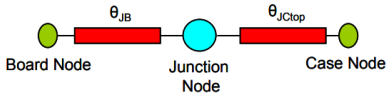


Figure 8.2: Two-resistor thermal model of JEDEC (taken from [163]).

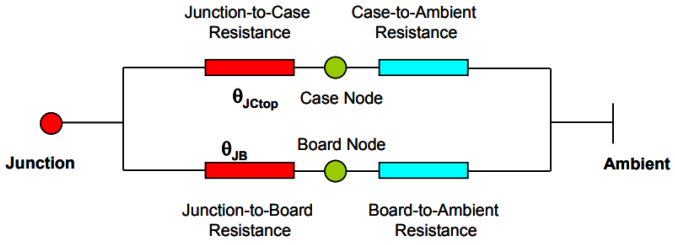


Figure 8.3: Complete equivalent thermal resistance representation of the two-resistor model (taken from [163]).

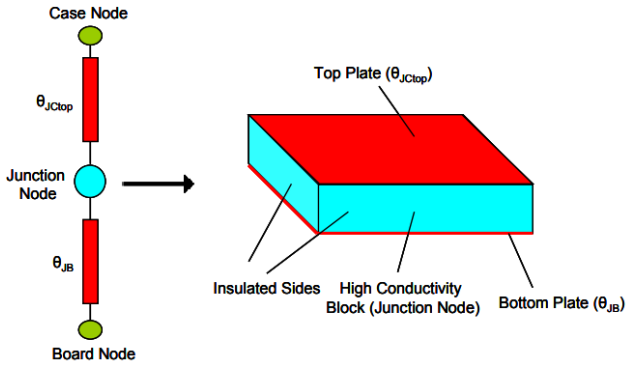


Figure 8.4: Representation of two-resistor model in a simulation environment via a block-and-surface resistance approach (taken from [163]).

board (PCB) has several layers including the ground plane and the radiating elements on the opposite sides. The overall topology of PCB and the radiating elements can be included in an equivalent board-to-ambient resistance. Fig. 8.3 presents the complete network of such a thermal model.

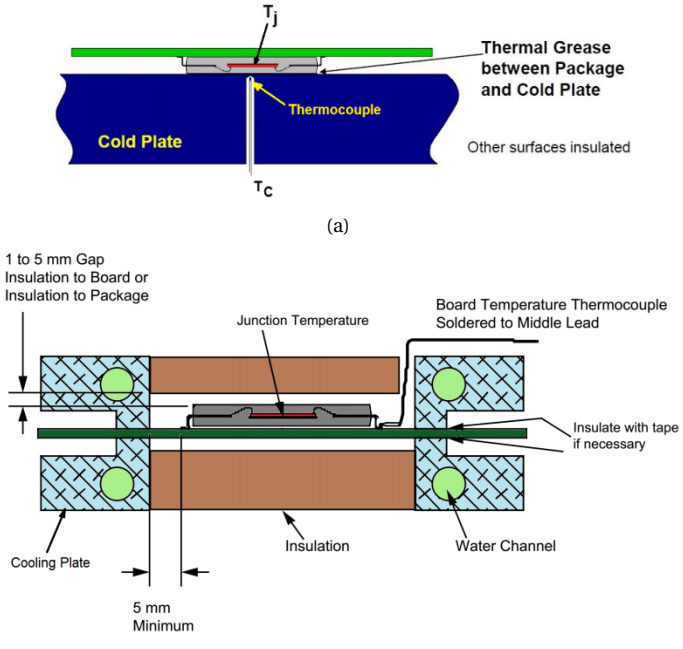


Figure 8.5: Thermal resistance measurement setups from [165] for, (a) junction-to-case resistance, (b) junction-to-board resistance.

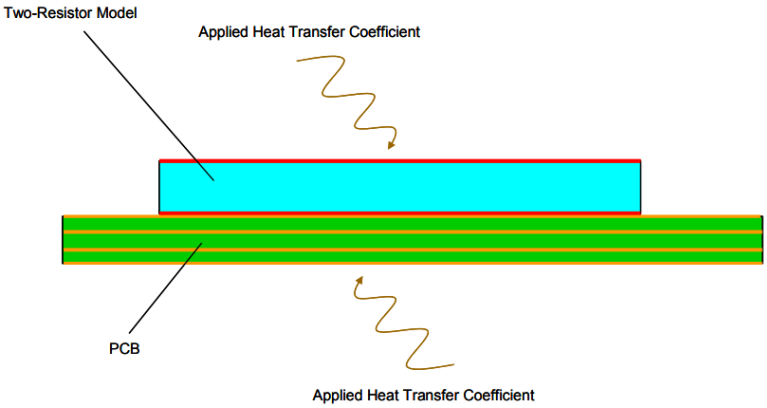


Figure 8.6: Application of two-resistor model in conduction-based simulation tools (taken from [163]).

To perform thermal simulations within the environment, it is needed to represent the two-resistor model in 3-D. Fig. 8.4 explains how this is achieved. The junction node is represented by a block formed by a high conductivity material (such as aluminum), while the equivalent thermal resistances are replaced with surface resistances on the top and bottom of the block for θ_{JCtop} and θ_{JB}, respectively. The sides of the block are insulated to prevent the heat flow through the sides.

While calculating the junction-to-case resistance, almost all the heat is forced to be dissipated through a single surface of the chip (case top or bottom) [165]. A cold plate is attached to the top (or bottom) of the chip and works as a heatsink. The case temperature is measured via a thin thermocouple wire, as shown in Fig. 8.5(a). Similarly, for calculation of the junction-to-board resistance, it is ensured that almost all of the heat flows through the substrate. The chip top and bottom are covered with a thermal insulation material. A circular cooling plate holds the board (see Fig. 8.5(b)).

As for the simulation tools, the 3-D thermal solvers are divided into two main categories; conduction-based tools and computational fluid dynamics (CFD) tools. In the conduction-based tools, the solid portions in the design are handled with the governing equations for conduction heat transfer, but unlike the CFD tools, the airflow effects are not solved directly. Instead, they are represented by equivalent heat transfer coefficients (H.T.C.s) at the solid-air interfaces, as shown in Fig. 8.6. The approximate H.T.C.s depend on the cooling regime and the typical H.T.C. for natural air convection is on the order of 10 Watts per meter-square Kelvin (W/m^{2}K), while it can reach values up to the order of 1000 W/m^{2}K for forced air convection and 10,000 W/m^{2}K for forced water convection [163].

In general, the prediction accuracy of the two-resistor CTM is not very high, but the accuracy will be much higher in the cases where most of the heat flows through the heatsink (as in the CPU cooler integrated arrays) and/or the board.

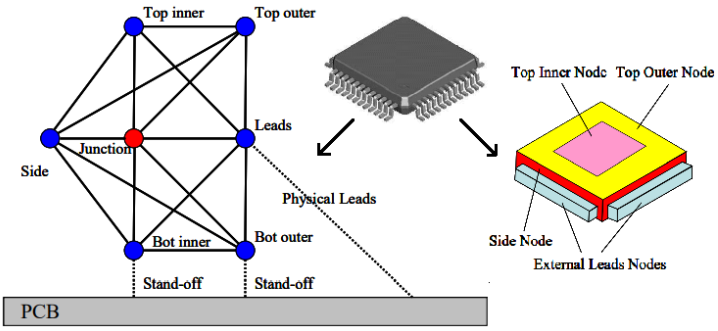


Figure 8.7: DELPHI CTM representation from [164]

Table 8.1: A qualitative comparison of the CTMs

Two-resistor CTM	DELPHI CTM
Intuitive & Simple	Complex
Test-based	Simulation-based
Artifacts from the test environment	No environmental artifacts
No error analysis	Statistical optimization with low errors
Less accurate	More accurate
Well-developed (aerospace / defense applications etc.)	Not adopted universally
Lack of CTM data by the semiconductor manufacturers	

8.2.2. DELPHI CTM

As shown in Fig. 8.7, the DELPHI CTM consists of several thermal resistors connecting the junction node to the surface nodes. The resistor values are computed by a detailed simulation procedure followed by a statistical optimization process that minimizes the errors in the prediction of the junction temperature through averaging over a wide range of environmental conditions (free convection, fan cooling, cold plate, high/low conductivity substrate etc.).

A comparison between the two models is provided in Table 8.1. The DELPHI CTM has a relatively high accuracy with a reasonable computational efficiency, which makes it more advantageous than the two-resistor CTM. Yet, it is advisable to use the two-resistor CTM in the initial design stage to simply obtain the first predictions on the chip temperatures and efficiently observe the outcome of different what-if scenarios.

8.3. THERMAL SIMULATION RESULTS WITH CPU HEAT SINKS

Having introduced the potential of the CPU coolers in mm-wave array cooling and the thermal models for performance assessment, we will now present the thermal simulations. It is worthy to note here that throughout the thesis, CST MultiPhysics (MPS) Studio's conduction-focused steady state solver is used for thermal simulations. Furthermore, the two-resistor CTM is used in the simulations, with the thermal properties of the Ka-band quad channel analog beamformer chip of NXP. The thermal model param-

Table 8.2: Thermal model parameters of the Ka band quad channel analog beamformer chip of NXP

Heat produced per chip in idle mode	0.36 W
Heat produced per chip in Tx mode	2.3 W
Heat produced per chip in Rx mode	1.3 W
Chip dimensions	4.3 x 3.5 x 0.5 mm
IC junction-to-case resistance	10 W/K
IC junction-to-board resistance	14 W/K
H.T.C. at the air interfaces	10 W/m^2K
Surface emissivity	0.9
Ambient temperature	25^oC

Table 8.3: NXP's integrated antenna array properties

Operating frequency	28 GHz
Board material	LTCC
Board dimensions (length x width x height)	40.6 x 47.5 x 1.4 mm
Board thermal conductivities (length x width x height)	11.53 x 11.53 x 3.06 W/mK
Number of chips	16
Spacing of chips (along length x along width)	10 x 10 mm
Number of patches	64 (8 by 8)
Patch dimensions	1.5 x 1.5 mm
Inter-patch spacings	5 mm

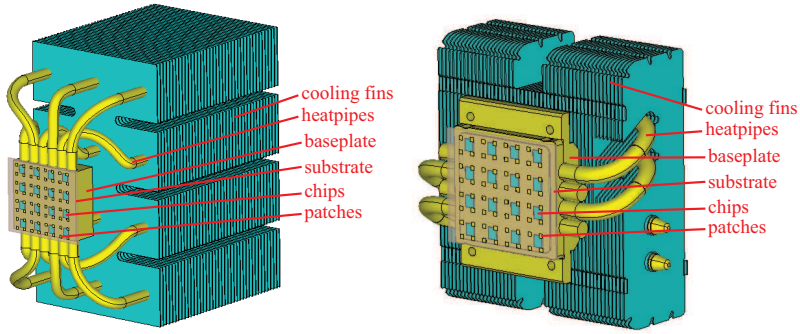
eters (provided by NXP's thermal engineering division) commonly used in this section are provided in Table 8.2. The fabricated single-polarized 8 x 8 integrated antenna array of NXP is modeled with the properties summarized in Table 8.3. Each beamformer IC (16 in total) is placed at the center of a 2 x 2 subarray. The array is integrated with two different CPU heatsinks (Mugen MAX and NH-L9x65) that are shown in Fig. 8.8 and have the properties listed in Table 8.4. The complete simulation model for the array with the two coolers are provided in Fig. 8.9 for better visualization. The damping in the heatpipes

(a) (b)

Figure 8.8: Commercial CPU heatsinks used in the paper, (a) Mugen MAX from Scythe [166], (b) NH-L9x65 from Noctua [167].

Table 8.4: CPU cooler properties

Cooler type	Mugen MAX	NH-L9x65
Overall dimensions	145 x 86 x 145 mm	95 x 95 x 50 mm
Heatsink weight	720 g	340 g
Heatpipe material	Copper (damped, nickel-plated)	Copper (damped, nickel-plated)
Number of heatpipes	6	4
Baseplate material	Copper (nickel-plated)	Copper (nickel-plated)
Baseplate dimensions	38 mm x 43 mm	42 mm x 45 mm
Fin material	Aluminum	Aluminum
Fin plate dimensions	86 x 145 mm	34 x 95 mm
Number of fin plates	40	43
Fin spacing	2.6 mm	1.65 mm
Fin plate thickness	0.5 mm	0.4 mm



(a) (b)

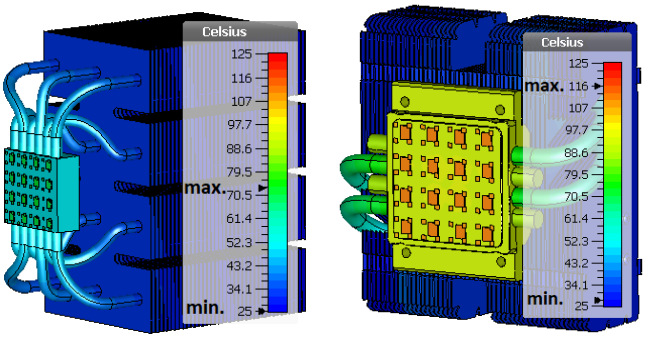
Figure 8.9: CST model of the NXP's integrated antenna array using (a) Mugen MAX, (b) NH-L9x65.

and the nickel-plating are not considered in the simulations for simplicity. Note that the chips (placed on the baseplate of the cooler) and patches are located on the opposite sides of the substrate. Therefore, the radiation pattern is not significantly affected by the heat sink.

A summary of the CST MPS simulation results for the maximum chip junction temperature ($T_{j,\max}$) in different cases (depending on the type of the cooler and the work mode of the chips) is given in Table 8.5. Note that in the idle and transmit modes, there are 16 idle and transmitting chips, respectively. Moreover, the temperatures across the

Table 8.5: Simulation results of the maximum IC junction temperatures with different CPU heatsinks in various work modes

Cooler type	Mugen MAX		NH-L9x65	
Work mode	idle	Tx	idle	Tx
$T_{j,\max}$ (^{o}C)	33	73	39	116



(a) (b)

Figure 8.10: Simulated array temperature distributions for (a) with Mugen MAX in Tx mode, (b) with NH-L9x65 in Tx mode (The substrates are hidden for a clear view of the chip temperatures).

array do not vary much (only by a few degrees) as compared to the maximum. This is because we assume a uniform thermal connection with the same thermal resistance from each IC to the heat sink, at every point of the array. For better visualization of the results, Fig 8.10 shows a few sample temperature distributions in the simulated antennas for the array in transmit (Tx) mode using the Mugen MAX and NH-L9x65 coolers. It is seen that due to its larger natural convection/thermal radiation surface area and better heat conduction with larger number of heatpipes, the Mugen MAX has much better cooling capability compared to the NH-L9x65. This advantage comes at the expense of having a relatively bulkier/heavier cooling structure (see Table 8.4).

8.4. EXPERIMENTAL VALIDATION AND DISCUSSION

The experiments are conducted using the fabricated array of NXP (see Table 8.3 for the details) which is integrated with the Mugen MAX cooler as shown in Fig 8.11. The junc-

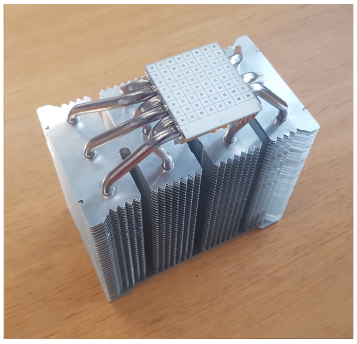


Figure 8.11: Fabricated array of NXP with the Mugen MAX cooling module.

86°C	82°C	78°C	81°C
77°C	73°C	71°C	77°C
76°C	73°C	73°C	74°C
76°C	75°C	73°C	75°C

Figure 8.12: Temperature sensor readings for the fabricated array in Tx mode.

tion temperatures are read from the temperature sensors integrated within the ICs. Fig. 8.12 shows the measured temperatures for the Tx mode. It is seen that despite the approximations used in the model, the predicted 73^oC temperature in Table 8.5 is, on average, in very good agreement with the measured results. It is also realized that at some portions of the array the thermal connection between the chip and the heatsink becomes worse due to the substrate bending during the fabrication process (caused by the thermal expansion coefficient mismatch between the layers) and the unevenness of the thermal interfacing (pad or glue), which result in higher junction temperatures, especially at the edge elements.

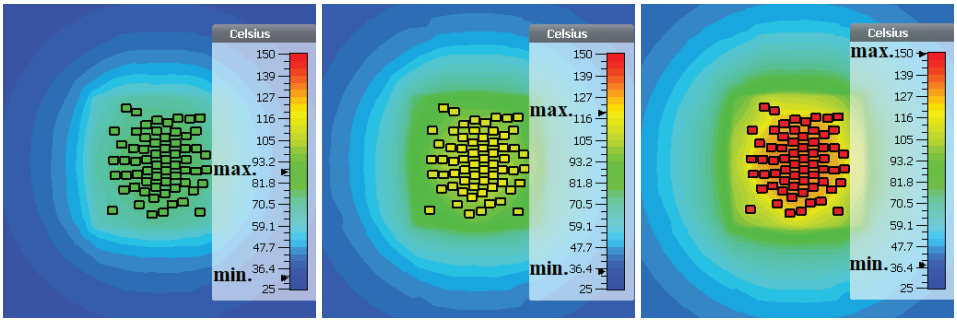
8.5. SYSTEM LIMITATIONS DUE TO PASSIVE-ONLY THERMAL CONTROL

8.5.1. EFFECT OF THE NUMBER OF SDMA USERS

Assuming fully-digital beamforming with one transceiver chip per antenna element and using the thermal modeling and design parameters indicated in [J-6] with a passive CPU cooler heatsink having a heat transfer coefficient of 3000 W/m^{2}Kelvin, we performed the thermal simulations in CST MWS for $K = 4$, $K = 6$ and $K = 8$ simultaneous, co-frequency users. Taking the dissipated heat per chip per user as 0.5 W leads to 2 W, 3 W and 4 W dissipated heat per chip for $K = 4$, $K = 6$ and $K = 8$, respectively. The chip junction temperature results for the optimized array layout shown previously in Fig. 6.28 are given here in Fig. 8.13, which shows that the maximum temperature in the array increases from 85°C to 150°C when K is increased from 4 to 8. This clearly indicates the necessity to use more complex active cooling strategies (with fans or water pumps) for larger number of simultaneous users.

8.5.2. EFFECT OF ANTENNA ELEMENT NUMBER REDUCTION

Due to the limitations in system cost and power consumption, there is a need to find out the lowest number of elements (transceiver chips or RF chains in the case of digital



(a) (b) (c)

Figure 8.13: Maximum chip junction temperature in the passively cooled optimized array for (a) 4 users, (b) 6 users, (c) 8 users.

Table 8.6: Thermal model parameters

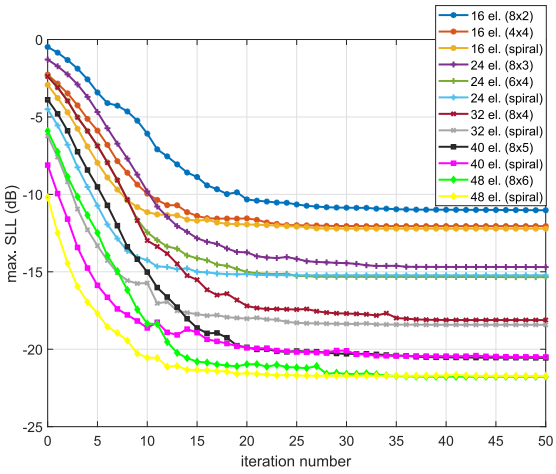
IC junction-to-case resistance	10 W/K
IC junction-to-board resistance	14 W/K
Heat transfer coefficient at the air interfaces	10 W/m^{2}K
Total board dimensions (length x width x height)	45 x 45 x 1.4 mm
Effective board thermal conductivity (length x width x height)	12 x 12 x 3 W/mK
Surface emissivity	0.9
Ambient temperature	25^{o}C

beamforming) at the BS antenna which can sufficiently suppress the inter-user interference and work at a reliable temperature. Yet, to the best of the author's knowledge, there is no prior work that jointly investigates the effect of the number of BS antenna elements on the array temperature and inter-beam interferences in a typical 5G cell deployment scenario. Therefore, in this sub-section, by taking into account the thermal integration with the CPU cooler used in Section 8.4, we use the technique in Section 6.2 to synthesize optimal multibeam antenna layouts with minimized SLLs within the communication sector for different number of total array elements and different initial array topologies. We also perform temperature simulations and search for the "best" array layout with the smallest number of elements[1] that can achieve sufficiently low SLLs and chip temperatures.

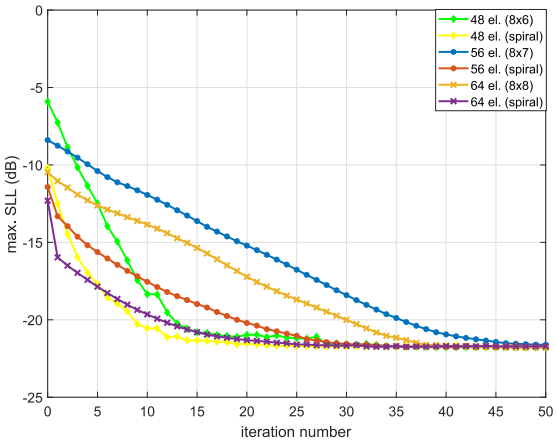
We consider the 5G cell sector introduced as the Use Case B in Section 2.4.2 ,which is defined by a ±15 degree range in elevation and a ±60 degree range in azimuth. The number of elements in the BS antenna is varied from 16 to 64. The minimum element spacing is set to 0.5λ. As explained before, the positions of the elements are to be optimized using the Convex Optimization algorithm proposed in [J-4] with the aim of minimizing the maximum SLL for a beam that is scanned freely inside the sector. Here, differently from the case study in Section 6.2.3.1, we minimize the SLL within the cell sector only and also investigate the effect of the choice of different initial array layouts on the outcome. Assuming a center operating frequency of 38 GHz (λ = 7.9 mm), the aperture size is made restricted to $\pm 2.5\lambda$, which defines the approximate baseplate dimensions of the commercially available CPU coolers. It is worthy of note that in practice, the baseplate of the cooler in Section 8.4 can be made larger with a post-fabrication treatment. However, in that case, as can be seen in [C-7] and in Fig. 8.12, the elements (chips) that are further from the heat pipes will be cooled less and will create a problem. Therefore, while designing the integrated antennas, it is better to stick to the dimensions dictated by the optimally designed CPU coolers. Otherwise, more expensive and customer-specific cooler designs will be needed which are not directly available in the market. In this part, we use the limited aperture size, but we also discuss the impact of unrestricted aperture size on the maximum SLL for different number of elements in order to show the trade-off between the SLLs and complexity of thermal management.

[1] A contrary approach would be to increase the number of elements to achieve more gain, lower side lobes; thus much less power and less heat per chip with more surface for cooling. Although this has a big potential for addressing the thermal issue, due to increased design & processing complexity with more elements, synthesis of large scale aperiodic arrays is given as a futuristic view in Appendix B.

We further consider that the chip integrated antennas are passively cooled using the fanless CPU heatsink Mugen MAX [166]. On the double-sided PCB, the chips are glued to the baseplate of the cooler on one side, while the radiators (patches) are facing the opposite side. Two-resistor thermal model validated in Section 8.4 is used in CST MPS solver to compute the chip temperatures. The common parameters of the thermal model are summarized in Table 8.6. We further assume that the heat generated by a chip is directly proportional to the RF power produced by the chip. Therefore, the heat dissipated per



(a)



(b)

Figure 8.14: Algorithm convergence in peak SLL for different number of total array elements and initial layouts in a 5G cell sector defined by ±15 degree range in elevation and ±60 degree range in azimuth: (a) from 16 to 48 elements, (b) from 48 to 64 elements.

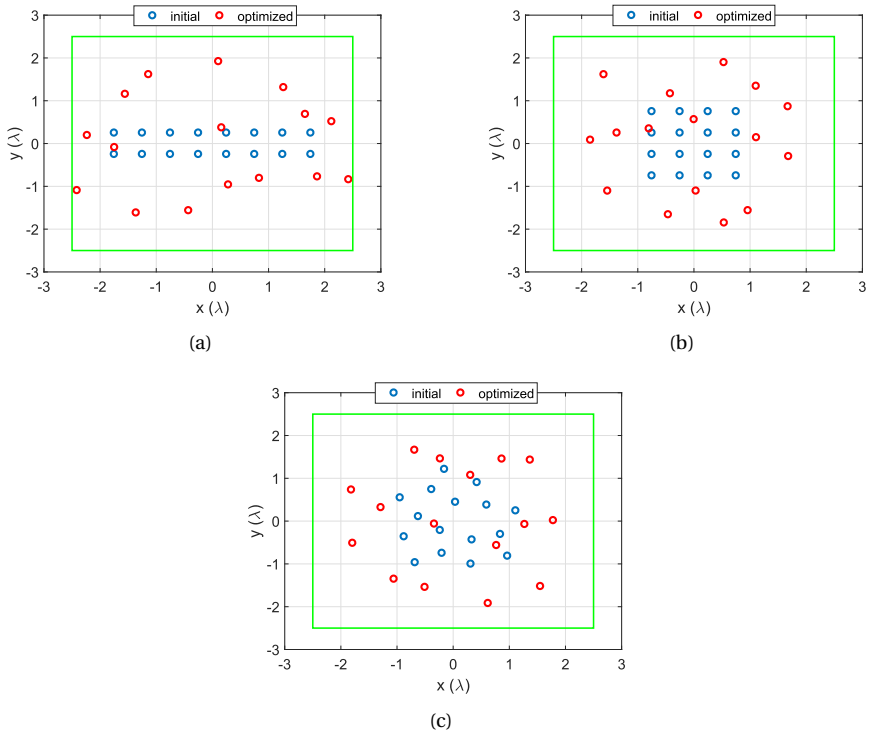


Figure 8.15: Optimized element positions for the 16-element array using the initial layout of (a) 8 x 2 - rectangular, (b) 4 x 4 - square, (c) 16-element sunflower.

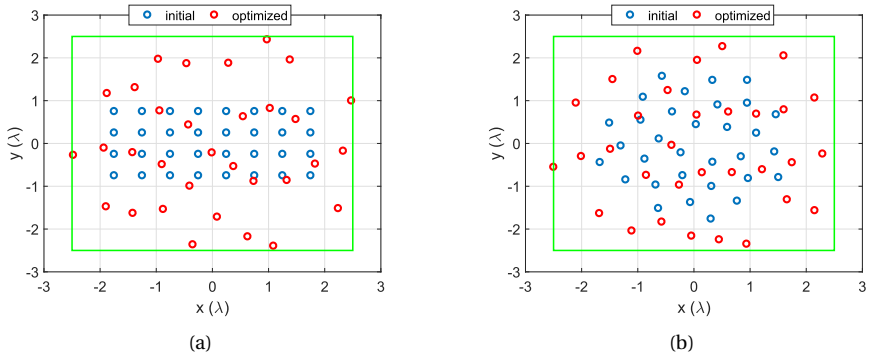


Figure 8.16: Optimized element positions for the 32-element array using the initial layout of (a) 8 x 4 - rectangular, (b) 32-element sunflower.

chip (P_h) is computed depending on the number of elements such that the EIRP is kept the same both for sparse and densely populated arrays. In other words, P_h in the sparser

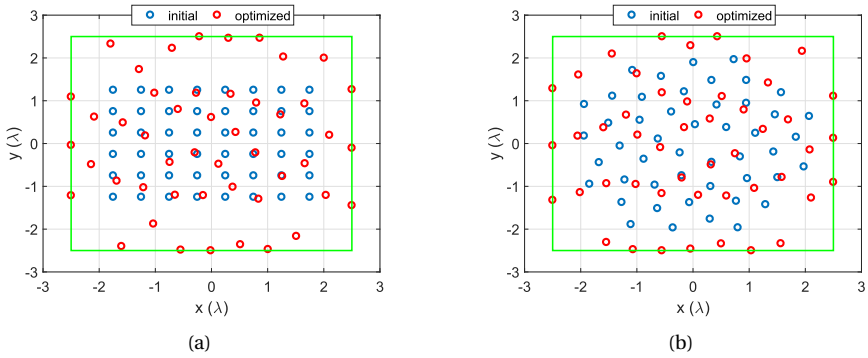


Figure 8.17: Optimized element positions for the 48-element array using the initial layout of (a) 8 x 6 - rectangular, (b) 48-element sunflower.

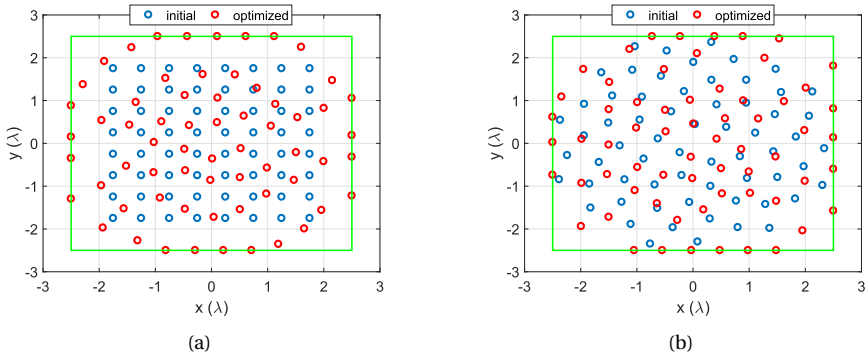


Figure 8.18: Optimized element positions for the 64-element array using the initial layout of (a) 8 x 8 - square, (b) 64-element sunflower.

arrays will be higher than that of the relatively dense ones due to their lower gain, which should be compensated with a larger output power.

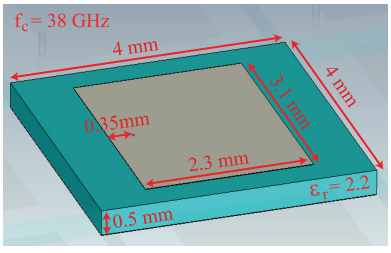


Figure 8.19: Pin-fed patch antenna dimensions used in the full-wave simulations.

First, we present the maximum SLL and radiation pattern results for different total BS antenna element numbers. Fig. 8.14 shows the convergence of the maximum SLL within the sector in Fig. 2.19(a) for different initial layouts (square, rectangular, spiral) and varying element number. From Fig. 8.14(a), it can be seen that using a spiral (sunflower) array layout at the algorithm input generally provides the best result in terms of the SLL and for the maximum aperture edge length of 5λ, the maximum SLL can be reduced from -12.2 dB for 16 elements to -18.4 dB for 32 elements and to -21.7 dB for 48 elements. The effect of increasing the element number more (until a maximum of 64 elements) is given in Fig. 8.14(b). One can observe that after 48 elements, increasing the element number does not help suppressing the SLL further. This is due to the limitation in the aperture size that is enforced by the characteristics of the cooler. Yet, for completeness, we also performed the simulations with no limitation on the aperture size. In that case, for 32, 48 and 64 elements, maximum SLL become -18.4 dB, -22 dB and -25.2 dB, respectively. It can be seen that although there is no change for 32 and 48 element arrays, maximum SLL in the 64 element array can be reduced by 3.5 dB more. This comes at the expense of increased aperture size with an edge length of 6.1λ, which will require a

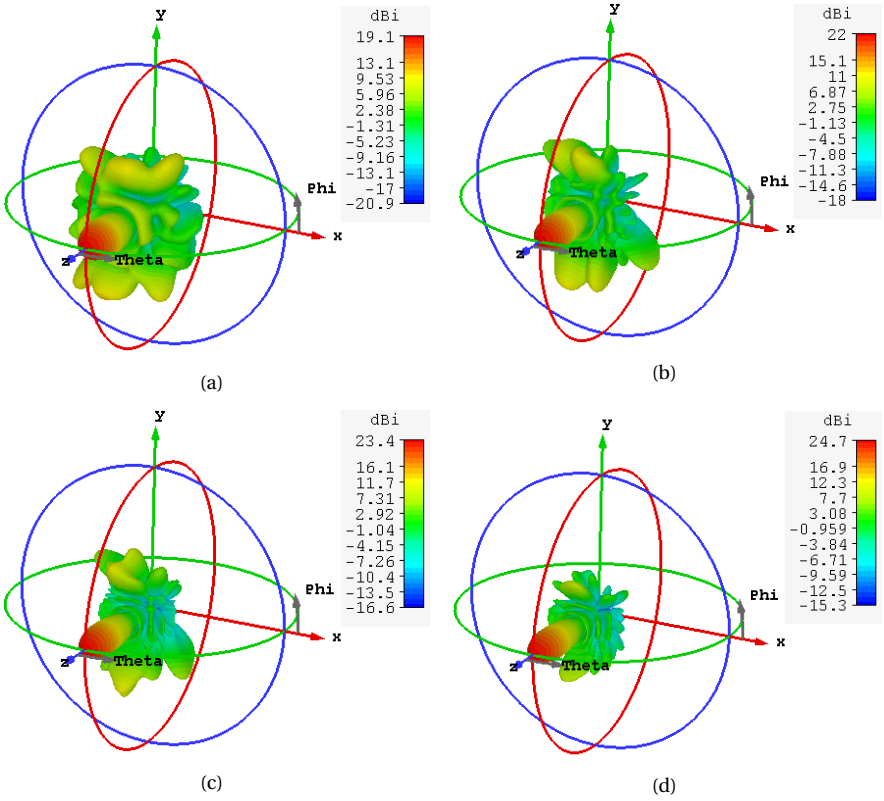


Figure 8.20: Array directivity (in dBi) for the broadside beam using the optimized array layouts with (a) 16, (b) 32, (c) 48, (d) 64 elements.

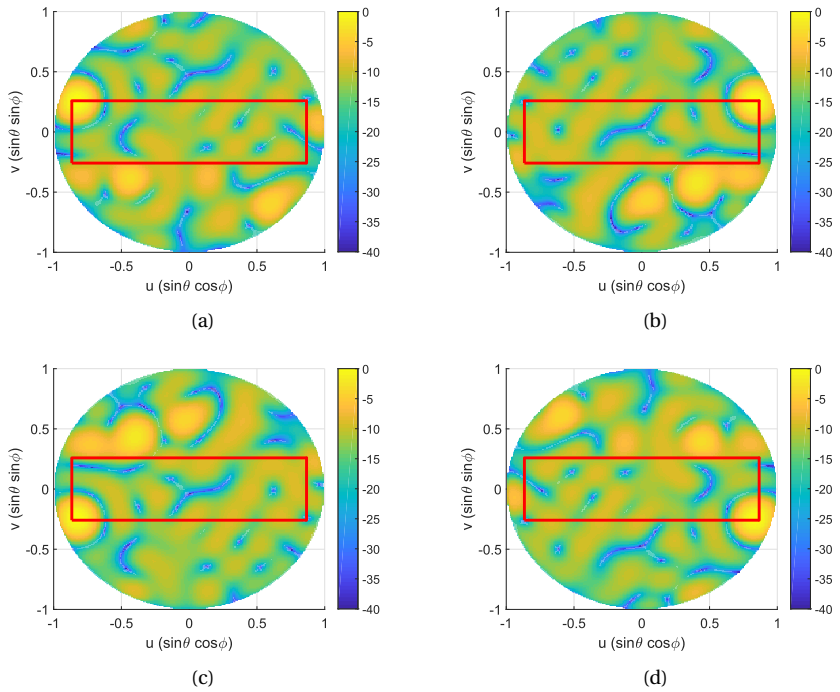


Figure 8.21: Multi-beam radiation patterns (in dB, normalized) of the optimized 16-element array for a beam scanned towards (a) $u = -\sin(\pi/3)$, $v = \sin(\pi/12)$, (b) $u = \sin(\pi/3)$, $v = \sin(\pi/12)$, (c) $u = -\sin(\pi/3)$, $v = -\sin(\pi/12)$, (d) $u = \sin(\pi/3)$, $v = -\sin(\pi/12)$.

costly, design-specific cooler for reliable operation. Therefore, as previously mentioned, here we present the results using a restricted aperture.

The initial and optimized layouts for 16 to 64 element BS arrays are plotted in Fig. 8.15 to Fig. 8.18, respectively. Next, the multibeam radiation patterns (scanned towards the four sector corners) of the optimized 16, 32, 48 and 64 element arrays are given in Fig. 8.21 to Fig 8.24, respectively. These radiation patterns help visualizing the maximum SLL values reported in Fig. 8.14. It is again seen that 48 and 64 element arrays are able to effectively and equally suppress the side lobes within the sector, while the out-of-sector pollution is smaller in the 64-element array.

Next, the optimized (using a sunflower initial layout) 16, 32, 48 and 64 element arrays are simulated in CST MWS to compute the array directivities. The patch element shown in Fig. 8.19 is used as the unit cell. The simulated broadside radiation patterns are shown in Fig. 8.20. It is seen that the gain of 32, 48 and 64 element arrays are 2, 2.7 and 3.6 times more than the gain of the 16 element array. This means that, if we assume per chip output power of P_o and heat generation P_h in the 64-element array, to achieve the same EIRP, P_o (and thus P_h) for the 48, 32 and 16 element arrays will be approximately 1.8, 3.6 and 14.4 times more, respectively. For thermal simulations, we assume that P_h for 64 element array is 0.5 W (which results in P_h equal to 0.9 W, 1.8 W and 7.2 W for 48, 32 and

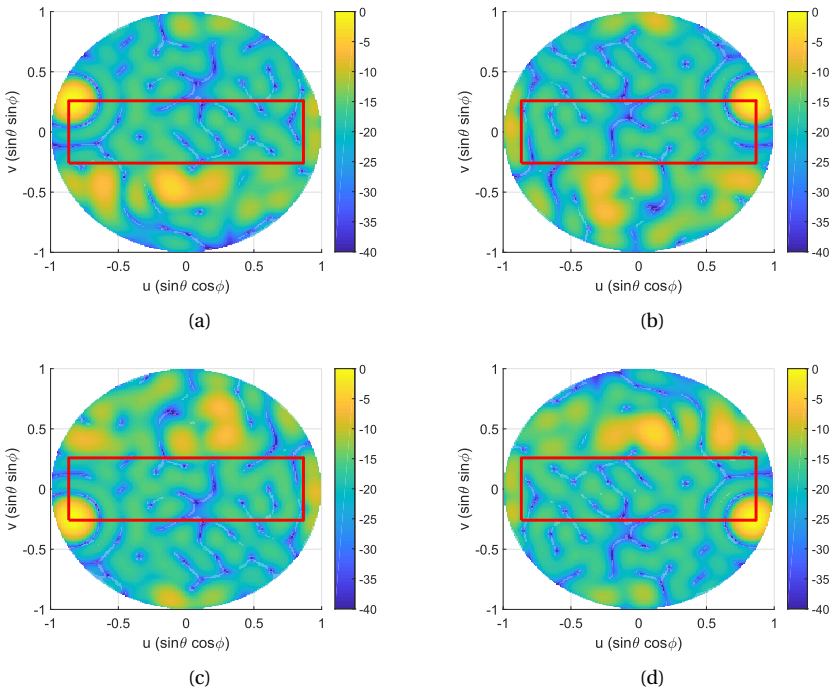


Figure 8.22: Multi-beam radiation patterns (in dB, normalized) of the optimized 32-element array for a beam scanned towards (a) $u = -\sin(\pi/3)$, $v = \sin(\pi/12)$, (b) $u = \sin(\pi/3)$, $v = \sin(\pi/12)$, (c) $u = -\sin(\pi/3)$, $v = -\sin(\pi/12)$, (d) $u = \sin(\pi/3)$, $v = -\sin(\pi/12)$.

16 element arrays, respectively) and the chip dimensions are 0.7 x 0.7 x 0.5 mm. Taking into account the output power needed, we also roughly modify the chip size accordingly such that the 48, 32 and 16 element arrays have chips (per element) with dimensions of 1.4 x 0.7 x 0.5 mm, 1.4 x 1.4 x 0.5 mm and 2.8 x 2.8 x 0.5 mm.

The thermal simulation results are given in Fig. 8.25. It can be seen that the temperature is the highest and not allowable in the 16 element array due to high per-chip heat generation, i.e. P_h. When the element number is increased, P_h goes down as well as the chip temperatures. For the 32, 48 and 64 element arrays, the maximum temperature become 85^oC, 66^oC and 54^oC, respectively. Here, we see that the thermal performances of the 48 and 64 element arrays (which provide the same EIRP) are close to each other and both can guarantee a safe and reliable device operation for a long time. Taking into account the radiation pattern results too, we can conclude that for our specific case with the commercially available passive CPU coolers, using the 48 element array with the optimized layout is the best design solution at the 5G BS.

8.6. CONCLUSIONS

In this chapter, first, the three basic heat transfer mechanisms (conduction, convection and thermal radiation) have been explained. It has been shown that for 5G mm-wave

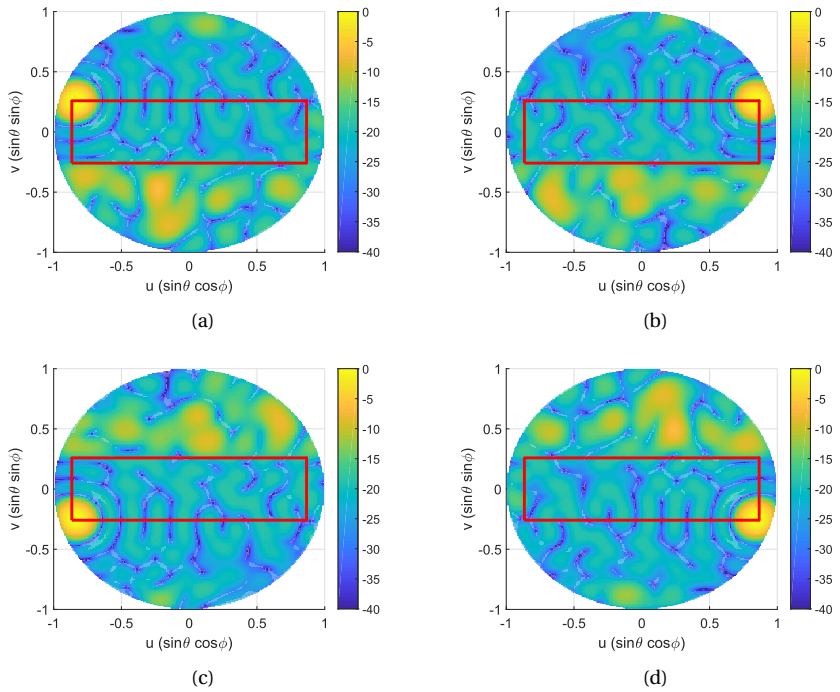


Figure 8.23: Multi-beam radiation patterns (in dB, normalized) of the optimized 48-element array for a beam scanned towards (a) $u = -\sin(\pi/3)$, $v = \sin(\pi/12)$, (b) $u = \sin(\pi/3)$, $v = \sin(\pi/12)$, (c) $u = -\sin(\pi/3)$, $v = -\sin(\pi/12)$, (d) $u = \sin(\pi/3)$, $v = -\sin(\pi/12)$.

antennas, **convection becomes the dominating way of heat dissipation**. Besides, it has been stated that because of the strict cost and energy requirements of high volume 5G systems, thermal management based purely on natural convection/radiation is desired, rather than the active methods that make use of forced air or liquids.

Next, several techniques for thermal system modeling of electronics have been discussed. It has been explained that CTMs are more suitable for system-level studies than DTMs thanks to their computational efficiency. **The two-resistor CTM** has been proposed to be used in the simulations to quickly observe different what-if scenarios with a **good trade-off between the accuracy and model simplicity**.

Later, the use of commercially available, low-cost and **fully-passive CPU heat sinks in the cooling of mm-wave chip-integrated 5G base station antennas** has been shown and validated. Two different CPU heat sinks (Mugen MAX and NH-L9x65) have been modeled and integrated with the NXP's integrated antenna prototype. A good agreement between the experimental junction temperature readings and the predicted values through the simulations has been observed, especially at the chips located in the center of the array with the best thermal connection to the heatsink. The main cause of the discrepancies between the measured and the simulated temperature results for the surrounding chips has been found to be the poorer thermal connection that occurs

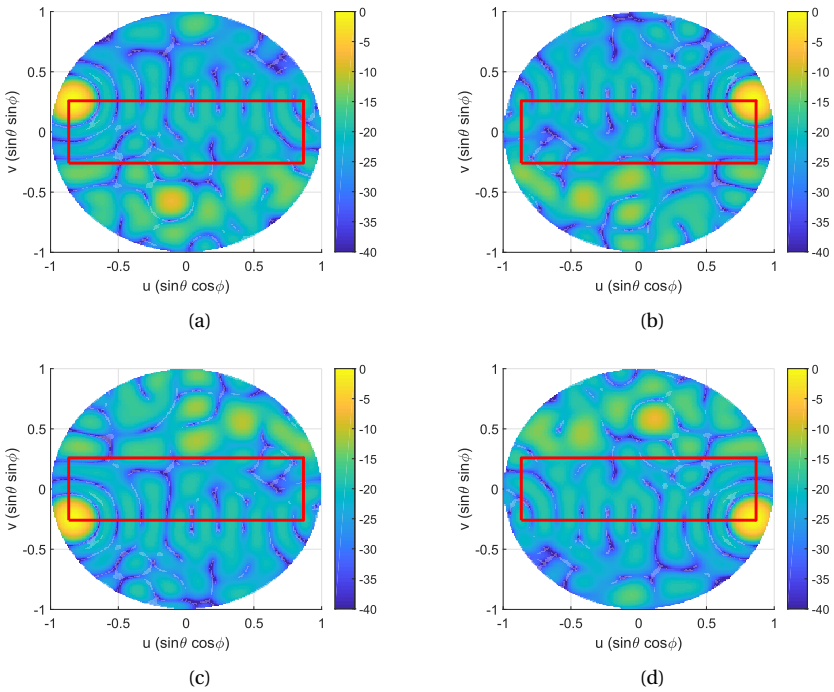


Figure 8.24: Multi-beam radiation patterns (in dB, normalized) of the optimized 64-element array for a beam scanned towards (a) $u = -\sin(\pi/3)$, $v = \sin(\pi/12)$, (b) $u = \sin(\pi/3)$, $v = \sin(\pi/12)$, (c) $u = -\sin(\pi/3)$, $v = -\sin(\pi/12)$, (d) $u = \sin(\pi/3)$, $v = -\sin(\pi/12)$.

due to the substrate bending and unevenness in the thermal pad. From the simulation results, it has been observed that the Mugen MAX has better cooling capacity than the NH-L9x65 due to its larger air-interfacing surface area and better heat conduction with larger number of heatpipes. The experimental results verify that the Mugen MAX cooler can guarantee a safe and reliable operation for both arrays in the work modes used in this study.

The presented research provides **valuable and useful guidelines for the selection of a proper cooler in the initial array design process**.

Furthermore, it has been shown that increased number of simultanous users and consequently, increased heat dissipation in the array raises the temperature above 100°C when using fully-passive cooling. This result shows **a thermal-based limitation in the number of frequency re-use with a passively cooled system**. To increase number of simultaneous users above this limit, it is necessary to have a complex and high-cost active cooling system that requires energy and maintenance.

Last, the effect of varying the total element number in passively-cooled, integrated and space-tapered antenna arrays for 5G on inter-user interference and chip temperatures has been studied. The element number is specified at the input of the layout optimization algorithm and is gradually reduced to create the sparser arrays. The array aper-

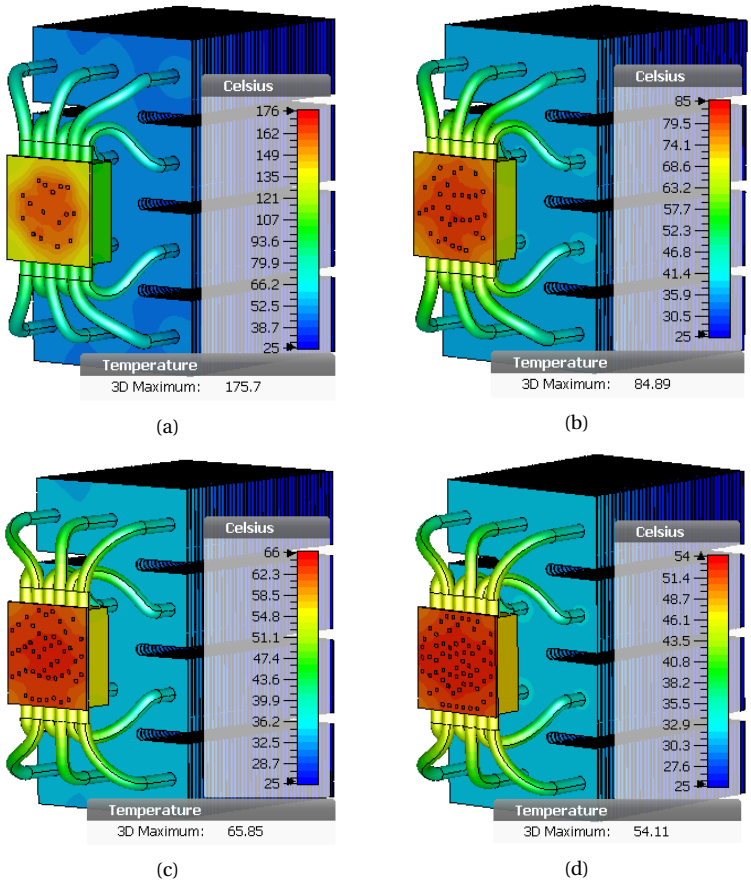


Figure 8.25: Chip temperature distributions across the passively-cooled active integrated antenna arrays with (a) 16 elements - P_h = 7.2 W, (b) 32 elements - P_h = 1.8 W, (c) 48 elements - P_h = 0.9 W, (d) 64 elements - P_h = 0.5 W.

ture has been restricted to the baseplate dimensions of the commercially available CPU heatsinks for direct and low-cost cooling. It has been observed that **the maximum in-sector SLL reduces as the element number increases**. This trend **continues until an optimal element number**, after which adding more elements does not help reduce the SLL further **because of the limited aperture size**. Thermal simulations have also been performed by taking into account the EIRP equalization and it has been seen that increasing the element number decreases the maximum temperature in the array and brings it to a safe level since the heat generated per-element (or per-chip) is reduced. Overall, this study has pointed out the **trade-offs between the electromagnetic/thermal performance and the number of array elements (thus cost, design/fabrication/processing complexity)**. Note that the idea presented can also be applied straightforwardly to different study cases employing different sector definitions, cooling strategies, chip/board characteristics and so on.

9

ALTERNATIVE WAYS FOR PASSIVE COOLING ENHANCEMENT

Parts of this chapter have been published as:
[J-7] Y. Aslan, J. Puskely and A. Yarovoy, "Heat source layout optimization for two-dimensional heat conduction using iterative reweighted L1-norm convex minimization," *Int. J. Heat Mass Transf.*, vol. 122, pp. 432-441, Jul. 2018.
[J-6] Y. Aslan, J. Puskely, J. H. J. Janssen, M. Geurts, A. Roederer and A. Yarovoy, "Thermal-aware synthesis of 5G base station antenna arrays: an overview and a sparsity-based approach," *IEEE Access*, vol. 6, pp. 58868-58882, Oct. 2018.
[C-9] Y. Aslan, J. Puskely, A. Roederer and A. Yarovoy, "Heat transfer enhancement in passively cooled 5G base station antennas using thick ground planes," in *Proc. 13th EuCAP*, Krakow, Poland, Apr. 2019.

9.1. LAYOUT OPTIMIZATION BASED APPROACH

This section is divided into two parts depending on the predominant heat transfer mechanism of the system: the 2-D heat conduction dominated scenarios and the 3-D heat convection dominated scenarios.

9.1.1. 2-D HEAT CONDUCTION DOMINATED SCENARIOS

INTRODUCTION

One of the fundamental interests of electronic cooling engineers is the conduction cooling in which a heat generating volume (such as a computer board) is cooled by a heat sink on the boundary. Conduction is seen to be the only alternative cooling method for the problem due to the impracticality of large-sized ducts for convective cooling and incapability of radiative cooling with a small dissipation area. Such a situation is in fact similar to the thermal management problem of the traditional AESAs that was studied in the thesis previously in Section 7.2 and illustrated in Fig. 7.3.

One effective way to enhance the heat conduction performance of the system and reduce the maximum temperature of the domain is to insert highly thermal conductive materials such as diamond or carbon fiber [168; 169] that are able to reduce the local thermal resistance. In this case, the distribution of such materials is to be optimized with the aim of minimizing the maximum temperature. A variety of approaches have been discussed in the literature including the constructal theory [170–172], entransy theory [173–175], bionic optimization [176; 177] and combinatorial algorithms [178; 179]. The results of these algorithms have shown that the optimized conduction path formed by the inserted materials has a shape similar to a tree with several branches varying in number and size.

Although conduction cooling using high thermal conductivity materials has been shown to be an effective method, the optimal distribution of such materials are difficult to realize in practice and the design costs may increase. An alternative approach is to provide passive cooling by optimizing only the positions of the electronics (or the heat sources). For layout optimization problems with large numbers of degrees of freedom, combinatorial algorithms such as genetic algorithm with artificial neural network [180–183] and particle swarm optimization [184] have been used in the literature. In their recent studies, Chen et al. applied bionic optimization [185; 186] and simulated annealing [187] methods to find out the optimal source distribution for varying number (up to several tens) of heat sources, which provided significant reduction in the maximum temperature of the domain when compared to the randomly or uniformly distributed heat sources.

In this section, the layout optimization is addressed using a new method that has not been investigated in the literature. The solution algorithm is based on the sequential reweighted l_1-norm minimization technique. This technique was first introduced by Candes et al. [188] with the aim of minimizing a convex measure of the l_0-norm which is equal to the cardinality of the sources. Recently, the sequential convex optimization method has been effectively used in the domain of sparse antenna array synthesis [189–191] in order to optimize the positions and excitations of the antenna elements. Considering the amplifiers powering the antennas as heat sources, a direct analogy can be made with the heat conduction problem.

Motivated by this analogy, the mathematical modeling of the two-dimensional heat conduction problem is performed and the positions of the heat sources are optimized using the convex optimization algorithm iteratively. The performance is tested using three typical cases with various boundary conditions that were previously applied by Chen et al. [185–187] and an additional case which takes into account the heat emission from the domain surface (by convection or radiation), while solving the conduction problem.

MATHEMATICAL MODELING OF THE PROBLEM

In this section, the mathematical model for the heat conduction optimization problem in a two-dimensional rectangular domain is revised. Since the notations used in thermal

Table 9.1: Nomenclature for Section 9.1.1

x, y	horizontal and vertical coordinates of the domain, m
T	temperature field, K
k	thermal conductivity, W/(m.K)
ϕ	heat source distribution function, W/m^2
T_0	environmental temperature, K
h	convective heat transfer coefficient, W/(m^2.K)
ϕ_0	intensity of a single heat source, W/m^2
Γ	area of a single heat source, m^2
T_b	background temperature field before adding the heat sources, K
ϕ_b	background heat source distribution before adding the heat sources, W/m^2
T_i	temperature rise due to adding the i^{th} source, K
w_j	selection weight of the source on the j^{th} grid cell, 1
N_g	number of grid cells for placing the heat sources, 1
L	side length of the square domain, m
R_m	normalized maximum temperature rise, 1
σ_m	normalized standard deviation of the temperature field in the domain, 1
σ	standard deviation of the temperature field in the domain, 1
$T_{\max}$	maximum temperature of the domain, K
T_{avg}	average temperature of the domain, K
N_T	number of temperature field samples in the domain, 1
T_{ref}	reference temperature field with uniform heat generation in the domain, K
$\mathbf{w}$	vector (N_g x 1) of selection weights (within [0, 1]) of grid cells, 1
N_s	number of heat sources placed in the domain, 1
$Z^{(j)}$	diagonal weighting matrix (N_g x N_g) in the j^{th} iteration of the optimization
S	separation matrix (N_g x N_g) defining the minimum inter-source spacing, 1
d_g	guaranteed minimum inter-source spacing defined at the algorithm input, m
$d_{\min}$	resulting minimum inter-source spacing at the algorithm output, m
ϵ	parameter for the algorithm stability, 1
l	side length of a square-shaped heat source, m
t_{ref}	computation time for the reference temperature field calculation, s
t_{opt}	total computation time for the optimization with the minimum required number of iterations, s
t_T	computation time for the resulting temperature field after the optimization, s
δ	domain extension length for the heat sink realization in the 3D model, m
Δ	thickness of the domain and the heat sources in the 3D model, m
ε	surface emissivity coefficient, 1
Φ_0	volume heat density of a single heat source in the 3D model, W/(m^3.K)

engineering field have different meanings than the ones in the electromagnetics, the nomenclature for Section 9.1.1 is given in Table 9.1.

As previously stated by Chen et al. [185–187], the temperature field (T) satisfies the following equation at steady state

$$\begin{gathered}\frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y}\right)+\phi(x,y)=0\\ \text{Boundary: } T=T_0 \;\; or \;\; k\frac{\partial T}{\partial \mathbf{n}}=0 \;\; or \;\; k\frac{\partial T}{\partial \mathbf{n}}=h(T-T_0)\end{gathered} \tag{9.1}$$

where k is the thermal conductivity of the domain, ϕ is the heat source distribution function and T_0 is the temperature value at the isothermal boundary or the fluid temperature at the convective boundary. h represents the convective heat transfer coefficient. In fact, (9.1) represents a Poisson problem with Dirichlet (isothermal), Neumann (adiabatic) or Robin (convective) boundary conditions that can be solved with a MATLAB-based finite-difference solver [192].

The heat sources in this study are modeled similar to [185] as follows

$$\phi(x,y)=\begin{cases}\phi_0, & (x,y)\in\Gamma\\ 0, & (x,y)\notin\Gamma\end{cases} \tag{9.2}$$

where ϕ_0 is the (constant) intensity of a single heat source and Γ represents the area of that heat source.

If a background temperature T_b and a source distribution ϕ_b are assumed, (9.1) can be expressed as

$$\begin{gathered}\frac{\partial}{\partial x}\left(k\frac{\partial T_b}{\partial x}\right)+\frac{\partial}{\partial y}\left(k\frac{\partial T_b}{\partial y}\right)+\phi_b(x,y)=0\\ \text{Boundary: } T_b=T_0 \;\; or \;\; k\frac{\partial T_b}{\partial \mathbf{n}}=0 \;\; or \;\; k\frac{\partial T_b}{\partial \mathbf{n}}=h(T_b-T_0)\end{gathered} \tag{9.3}$$

When the i^{th} source is added into the domain, the temperature rise T_i is given by

$$\begin{gathered}\frac{\partial}{\partial x}\left(k\frac{\partial T_i}{\partial x}\right)+\frac{\partial}{\partial y}\left(k\frac{\partial T_i}{\partial y}\right)+\phi_i(x,y)=0\\ \text{Boundary: } T_i=0 \;\; or \;\; k\frac{\partial T_i}{\partial \mathbf{n}}=0 \;\; or \;\; k\frac{\partial T_i}{\partial \mathbf{n}}=hT_i\end{gathered} \tag{9.4}$$

(9.4) shows that the temperature rise field (T_i) for the source i is independent of T_b, ϕ_b and T_0. Thus, the combinatorial optimization problem of the heat source distribution can be expressed as follows

$$\begin{gathered}\text{minimize } (\max(T)) \;\; s.t.\\ T=T_b+\sum_{j=1}^{N_g} w_j T_j\\ w_j=0 \;\; or \;\; 1, \;\; j=1,...,N_g\end{gathered} \tag{9.5}$$

where w_j is the (binary) selection weight of the source at the j^{th} grid cell, T_j is the respective temperature rise and N_g is the number of grid cells suitable for the heat sources.

Normalization of the variables was also performed by Chen et al. [185] in order to investigate the effect of ϕ_0 and k on the optimized layout. For a square computational domain with side lengths L, the normalized variables are given by

$$\overline{R} = \frac{T - T_0}{\phi_0 L^2/k}, \quad \overline{x} = \frac{x}{L}, \quad \overline{y} = \frac{y}{L}, \quad \overline{\phi} = \frac{\phi}{\phi_0} \tag{9.6}$$

Using (9.6), (9.1) and (9.2) are represented as

$$\begin{gathered} \frac{\partial}{\partial \overline{x}}\left(k\frac{\partial \overline{R}}{\partial \overline{x}}\right) + \frac{\partial}{\partial \overline{y}}\left(k\frac{\partial \overline{R}}{\partial \overline{y}}\right) + \overline{\phi}(\overline{x},\overline{y}) = 0, \\ \overline{\phi}(\overline{x},\overline{y}) = \begin{cases} 1, & (\overline{x},\overline{y}) \in \overline{\Gamma}(\overline{l}) \\ 0, & (\overline{x},\overline{y}) \notin \overline{\Gamma}(\overline{l}) \end{cases} \\ \text{Boundary: } \overline{R} = 0 \;\; or \;\; k\frac{\partial \overline{R}}{\partial \mathbf{n}} = 0 \;\; or \;\; k\frac{\partial \overline{R}}{\partial \mathbf{n}} = \frac{hL}{k}\overline{R} \end{gathered} \tag{9.7}$$

The performance is evaluated considering two normalized metrics: normalized maximum temperature rise (R_m) and normalized standard deviation (σ_m) which are defined as

$$R_m = \frac{T_{\max} - T_0}{\phi_0 L^2/k}, \quad \sigma_m = \frac{1}{\phi_0 L^2/k}\sqrt{\frac{1}{N_T}\sum_{i=1}^{N_T}(T_i - T_{\text{avg}})^2} \tag{9.8}$$

where N_T is the total number of temperature field samples in the domain and T_{avg} is the average temperature value.

SOLUTION APPROACH

The combinatorial problem in (9.5) is an NP-hard problem. The main challenge here is the nonconvexity of the objective function resulting from the binary selection weights (i.e. $w_j \in \{0,1\}$). Motivated by this challenge, an alternative method is proposed in this study which is based on convex relaxation of the binary selection weights (i.e. $w_j \in [0,1]$) and an approximation of the objective function in (9.5) through sequential convex minimization of the weighted l_1-norm of a function given by

$$f = T_{\text{ref}}.\mathbf{w} \tag{9.9}$$

where T_{ref} is a vector (of length N_g) showing the temperature distribution obtained from a uniform heat generation in the whole domain, or when the domain is densely filled with equal-intensity heat sources. As previously stated, the vector $\mathbf{w}$ contains the relaxed selection weights that could be equal to any value between 0 and 1. The dot product (".")

is obtained via element-wise multiplication of the two vectors, which is simply multiplying the temperature value at any grid point by its selection weight.

The logic behind the selection of such an objective function can be demonstrated through solving (9.4), which basically shows that the maximum temperature rise takes place at the area of the added heat source. Therefore, T_{ref} in (9.9) shows which regions inside the domain are 'easy-to-cool' and which are potentially inefficient in terms of heat conduction. The idea is to select the sources in 'easy-to-cool' regions with larger weights while assigning smaller weights for the areas with high reference temperature.

Having selected the objective function, the challenge now is to decide on the constraints. Note that in the convex optimization algorithm, the domain is discretized with N_g cells and there is a heat source placed at each cell on the grid. The difference is in the selection weights of the sources. For the ideal case in which the number of sources (N_s) is identified, N_s sources are selected with weight 1 and the remaining ($N_g - N_s$) sources are selected with weight 0. Therefore, the convex algorithm should yield a sparse solution in terms of the selection weights ($\mathbf{w}$) to remove ($N_g - N_s$) heat sources. Besides, the heat sources should not overlap in the final layout, which requires a certain minimum spacing between two selected grid cells depending on the source dimensions.

In this study, in order to obtain a sparse solution, reweighted l_1-norm minimization technique [188] is applied with a defined cardinality ($N_s = \sum w_i$) of $\mathbf{w} \in R^{N_g}$ that represents the desired number of sources in the final layout. A constraint on the inter-element spacing is also defined as proposed by D'Urso et al. [190] to prevent overlapping. Thus, considering the objective function defined in (9.9), the algorithm objective at iteration-j becomes

$$\begin{gathered}
\text{minimize } |Z^{(j)} * (T_{\text{ref}}.\mathbf{w})|_1 \quad s.t. \\
\mathbf{0} \leq \mathbf{w} \leq \mathbf{1} \\
\sum_{i=1}^{N_g} w_i = N_s \\
S^* \mathbf{w} \leq \mathbf{1} \\
Z_i^{(j)} = \frac{1}{w_i^{(j-1)} + \epsilon}, \quad i = 1, 2, ..., N_g
\end{gathered} \tag{9.10}$$

where "." represents the dot product of two vectors and "*" represents the matrix product. Z is a diagonal weighting matrix whose nonzero elements are given by the reciprocal of the sum of the corresponding weights from the previous iteration and ϵ. To have good convergence and stability, ϵ should be set slightly smaller than the expected non-zero values of the entries of $\mathbf{w}$ [188]. It also ensures that a zero value in $\mathbf{w}^j$ does not prevent a nonzero estimate in the next iteration. S is an N_g x N_g matrix defined as the "separation" matrix. For square-shaped heat sources, S centers a square of side lengths equal to the guaranteed minimum inter-source spacing (d_g) whose inside is filled with ones and outside with zeros. The minimum spacing of the sources in the final layout (d_{min}) may be larger than the guaranteed spacing that is given as an input parameter.

The optimization problem in (9.10) is a nonlinear convex problem, namely a second-order cone programming (SOCP) problem [96]. Thus, similarly to the antenna optimiza-

tion problems derived in Chapter 6, it can be efficiently solved by many commercially available solvers (such as CVX [102]) using interior point methods. The solution time of such a process is approximately the same with a linear problem of equivalent size.

RESULTS AND DISCUSSION

A. Parameters of the cases

In this section, the parameters for the layout optimization problem are provided. Since the convex optimization algorithm is proposed as an alternative to the existing methods, the scenarios and design parameters applied by Chen et al. [185–187] are taken as a reference in the first three cases. In the last case, a Volume-to-Volume (VV) problem is introduced, which has not been investigated before in the existing layout optimization techniques.

Throughout the study of the first three cases, optimization of 20 heat sources (N_s = 20) in a square domain with the side length (L) of 0.1 m is aimed. The side length of a single square-shaped heat source (l) is chosen as 0.01 m with an intensity (ϕ_0) of 10,000 W/m^2. The thermal conductivity (k) equals 1 W/(m.K) for both the domain and the heat source. The first three cases studied in this paper are explained as follows.

- Case 1: Volume-to-Point (VP) problem. The cooling is via a 0.001 m wide heat sink located in the middle of the south boundary with uniform temperature at 298 K (Dirichlet BC). The other boundaries are adiabatic (Neumann BC).
- Case 2: Volume-to-Boundary (VB) problem. The cooling is from all the boundaries that are isothermal with uniform temperature at 298 K (Dirichlet BC).
- Case 3: Volume-to-Boundary (VB) problem. The cooling is from the south and west boundaries. The south boundary is isothermal and uniform at 298 K (Dirichlet BC) while the west is convective with heat transfer coefficient 5 W/(m^2.K) and the fluid temperature is 298 K (Robin BC). The other boundaries are adiabatic (Neumann BC).

The fourth case is described as follows.

- Case 4: Volume-to-Volume (VV) problem. Similar to Case 1, the cooling is via a 0.001 m wide heat sink located in the middle of the south boundary with uniform temperature at 298 K. East, west and north boundaries are adiabatic. The difference now is that the domain surface has an emissivity coefficient equal to 0.1.

In the fourth case, a 3-D model is constructed and simulated in CST MultiPhysics (CST MPS) Studio using the thermal stationary solver. The idea is to obtain the reference temperature distribution (T_{ref}) from CST and use it at the input of the convex optimization algorithm. With the resulting source locations, the optimized temperature field is computed again in CST. The values of N_s, L, l and k are selected the same as the ones in the first three cases. To create the 3-D model, a thickness of Δ = 0.001 m is applied for the domain and the heat sources. Since it is not possible to define mixed boundary conditions in CST, the heat sink is realized by extending the length of the middle piece of the south boundary (which has 0.001 m width for the heat sink) by $\delta = 10^{-6}$ m. All the other boundaries are set as adiabatic. The volume density of each source (Φ_0) is set to 14,400

W/m^3 in order to match the temperature distribution result obtained in MATLAB in the case of no surface heat emission (Case 1). The background is vacuum and the ambient temperature is equal to 298 K. The emissivity coefficient is defined at the upper surface of the domain.

In all the cases, the parameter ϵ in (9.10) is set slightly smaller than 1 although the algorithm is seen to be robust in terms of the final layout results for different choices of ϵ. N_g and d_g are varied for a coarse and a fine grid structure for the heat source placement. After several iterations of (9.10), a stable solution is reached where the first 20 elements with the strongest selection weights form the final layout.

All simulations are carried out on an Intel(R) Core(TM) i7-4710HQ 2.5GHz CPU, 16GB RAM computer using CVX and MATLAB. In Case 4, CST MPS is also used for the volume-to-volume simulations.

To have sufficient resolution within a reasonable time, the number of samples for the temperature field calculations is taken as 1000 x 1000 ($N_T = 10^6$). Note that for the convex optimization algorithm, the temperature field is calculated only once using a uniform heat generation in order to obtain the reference temperature distribution (T_{ref}) at the algorithm input. For the selected N_T, the reference temperature field calculation time (t_{ref}) is about 9-10 seconds for all the three cases. After the optimization of 20 heat sources, the resulting temperature field is computed in t_T, which is about 160-180 seconds for Case 1 and Case 3 and 60-80 seconds for Case 2. In Case 4, each temperature field computation in CST MPS takes about 25 seconds.

B. Discussion of the results

Table 9.2 shows the input parameters (N_g and d_g) and convex optimization results (R_m, $T_{\max}$, σ_m, $d_{\min}$ and t_{opt}) in Case 1 for a coarse and a fine grid for the heat source placement. The guaranteed minimum inter-spacing, d_g is equal to the side length of a single heat source.

It is seen that $T_{\max}$ and σ_m are both lower in the coarse grid ($N_g = 37^2 = 1369$) compared to the fine grid ($N_g = 91^2 = 8281$). This implies that the result for the coarse grid is more favorable. However, it is also observed that the resulting minimum inter-source spacing, $d_{\min}$ is larger than 0.01 m (i.e. the smallest value allowable) for the fine grid. This shows the input parameter d_g can be varied in the fine grid to seek for a better result. Table 9.3 provides the results of this analysis. Note that "-" sign in Table 9.3 indicates that there is no feasible solution due to the overlapping heat sources in the final layout.

Table 9.2: Optimized results in Case 1 for a coarse ($N_g = 37^2$) and a fine ($N_g = 91^2$) grid for the heat source placement

Spacing of the square grid (m)	0.0025	0.001
Number of grid cells, N_g	37x37	91x91
Guaranteed min. spacing, d_g (m)	0.01	0.01
Number of iterations	1	3
Normalized max. temperature rise, R_m	0.2983	0.3081
Max. temperature, $T_{\max}$ (K)	327.83	328.81
Normalized std., σ_m (x10^{-2})	1.4263	1.5585
Min. distance between the sources, $d_{\min}$ (m)	0.01	0.012
Optimization time, t_{opt} (s)	1.82	537.31

Table 9.3: Optimized results in Case 1 for $N_g = 91^2$ with varying guaranteed min. spacing

Guaranteed min. spacing, d_g (m)	0.012	0.011	0.009	0.008	0.007
Number of iterations	2	2	2	1	2
Normalized max. temperature rise, R_m	0.3114	0.3050	0.3005	-	-
Max. temperature, $T_{\max}$ (K)	329.14	328.50	328.05	-	-
Normalized std., σ_m (x10^{-2})	1.7034	1.5753	1.3891	-	-
Min. distance between the sources, $d_{\min}$ (m)	0.014	0.013	0.011	-	-
Optimization time, t_{opt} (s)	302.71	329.50	211.41	-	-

When the results in Table 9.2 and Table 9.3 are compared, it can be inferred that the best layout among them is obtained when the fine grid is used with d_g = 0.009 m. Although $T_{\max}$ is lower for the coarse grid, this selection lowers the standard deviation significantly while yielding a comparable maximum temperature value.

Next, a similar study is performed for Case 2 applying the same coarse and fine source placement grids as in Case 1 and using several different d_g values as inputs. The results are summarized in Table 9.4 and Table 9.5 for the coarse and fine grid, respectively.

According to Table 9.4, taking d_g = 0.015 m yields the best solution for the coarse grid. However, using a finer grid gives the opportunity to define a finer step for d_g (through the solution of the inter-element spacing constraint in (9.10)), which could provide a better solution. This is indeed the case since an improved optimization result is observed in Table 9.5 for the fine grid with d_g = 0.014 m.

Having seen the usefulness of applying a fine grid for the heat source placement in Case 1 and Case 2, only the fine grid results with varying d_g values are given for Case 3 in Table 9.6. As previously stated, the "-" sign means overlapping in the final layout, which is not allowed.

According to Table 9.6, the best solution among the tested scenarios in Case 3 is obtained when d_g is equal to 0.009 m.

Table 9.4: Optimized results in Case 2 for $N_g = 37^2$ with varying guaranteed min. spacing

Guaranteed min. spacing, d_g (m)	0.01	0.0125	0.015	0.0175
Number of iterations	1	1	1	2
Normalized max. temperature rise, R_m	0.0037	0.0031	0.0030	0.0067
Max. temperature, $T_{\max}$ (K)	298.37	298.31	298.30	298.67
Normalized std., σ_m (x10^{-2})	0.0790	0.0702	0.0709	0.1582
Min. distance between the sources, $d_{\min}$ (m)	0.01	0.015	0.0175	0.02
Optimization time, t_{opt} (s)	1.58	1.88	2.29	5.32

Table 9.5: Optimized results in Case 2 for $N_g = 91^2$ with varying guaranteed min. spacing

Guaranteed min. spacing, d_g (m)	0.015	0.014	0.013
Number of iterations	1	1	1
Normalized max. temperature rise, R_m	0.0030	0.0030	0.0030
Max. temperature, $T_{\max}$ (K)	298.30	298.30	298.31
Normalized std., σ_m (x10^{-2})	0.0720	0.0694	0.0695
Min. distance between the sources, $d_{\min}$ (m)	0.018	0.016	0.016
Optimization time, t_{opt} (s)	204.39	158.31	150.26

Table 9.6: Optimized results in Case 3 for $N_g = 91^2$ with varying guaranteed min. spacing

Guaranteed min. spacing, d_g (m)	0.011	0.01	0.009	0.008
Number of iterations	2	4	1	2
Normalized max. temperature rise, R_m	0.0328	0.0306	0.0272	-
Max. temperature, $T_{\max}$ (K)	301.28	301.06	300.72	-
Normalized std., σ_m (x10^{-2})	0.6478	0.6976	0.4300	-
Min. distance between the sources, $d_{\min}$ (m)	0.012	0.012	0.011	-
Optimization time, t_{opt} (s)	279.70	685.39	99.93	-

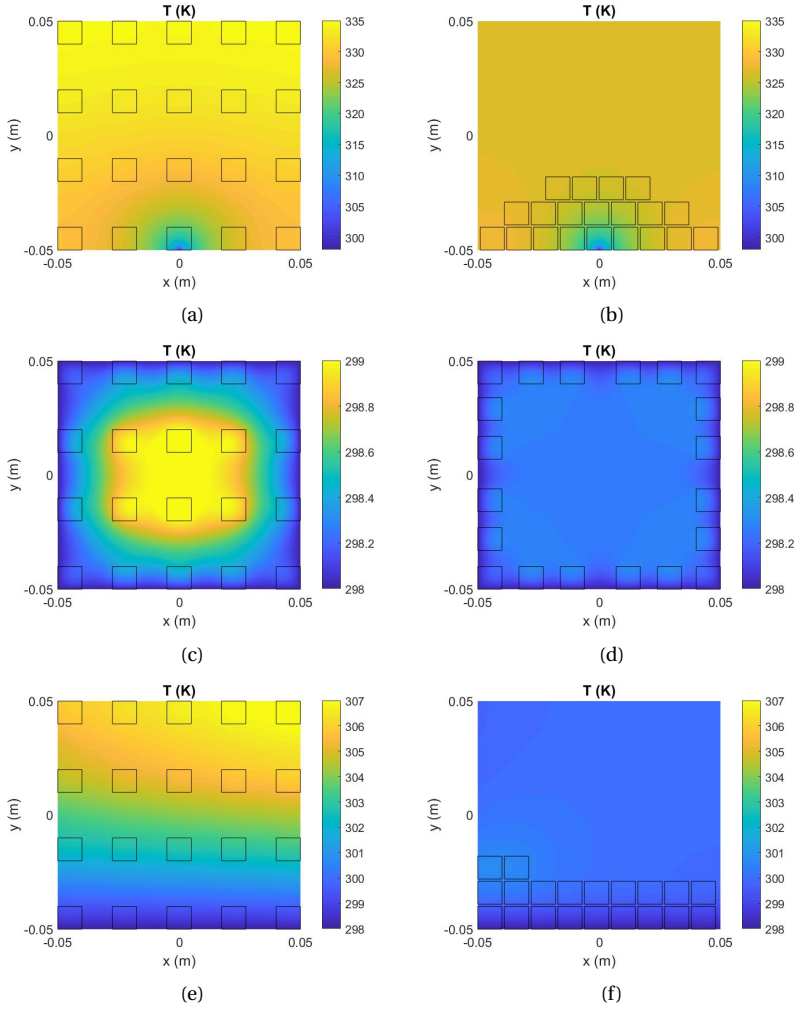


Figure 9.1: Uniform (UD) and convex-optimized (CO) distribution of 20 heat sources in the first three cases: (a) UD, Case 1, (b) CO, Case 1, (c) UD, Case 2, (d) CO, Case 2, (e) UD, Case 3, (f) CO, Case 3.

Table 9.7: Comparison of performance for the uniform and optimized heat source layouts in the first three cases

	UD			CO		
	R_m	$T_{\max}$	σ_m (x10^{-2})	R_m	$T_{\max}$	σ_m (x10^{-2})
Case 1:	0.3694	334.94	3.3365	0.3005	328.05	1.3891
Case 2:	0.0115	299.15	0.3185	0.0030	298.30	0.0694
Case 3:	0.0939	307.39	2.5421	0.0272	300.72	0.4300

The optimized heat source layouts for the first three cases with 20 sources and their respective temperature field distributions are provided in Fig. 9.1 together with the results of uniformly distributed heat sources. The uniform distribution is described as in [185] where the domain is covered by 5 equi-spaced source repetitions in the horizontal and 4 equi-spaced source repetitions in the vertical direction.

The comparison of the convex optimization and uniform distribution results is given in Table 9.7 for completeness. It is seen that via convex optimization, the normalized maximum temperature rises (R_m) in Case 1, Case 2 and Case 3 are reduced by 18.65%, 73.91% and 71.03%, respectively, when compared to the uniform distribution. Besides, the values of the normalized standard deviations (σ_m) are reduced by 58.37%, 78.21% and 83.08% in Case 1, Case 2 and Case 3, respectively.

Fig. 9.2 shows the results of CST MPS simulations for the 'imitated' version of Case 1 with a heat sink in the middle of the south boundary for the uniform source layout and the effect of considering the surface emission in such a scenario, which represents Case 4. If the heat emission is absent (Fig. 9.2(a)), the CST simulation results in the same temperature distribution as shown in Fig. 9.1(a), and thus verifies the CST simulation model. After defining the surface emissivity $\varepsilon = 0.1$, the maximum temperature of the domain is reduced to 303.02 K (see Fig. 9.2(b)). Using the temperature distribution in Fig. 9.2(b) as T_{ref}, the source positions are optimized using the proposed convex optimization (CO) algorithm. The results are provided in Table 9.8 which show that the lowest maximum temperature and standard deviation in Case 4 is obtained when d_g is equal to 0.018 m.

In order to have a comparative study, the built-in genetic algorithm (GA) of CST is

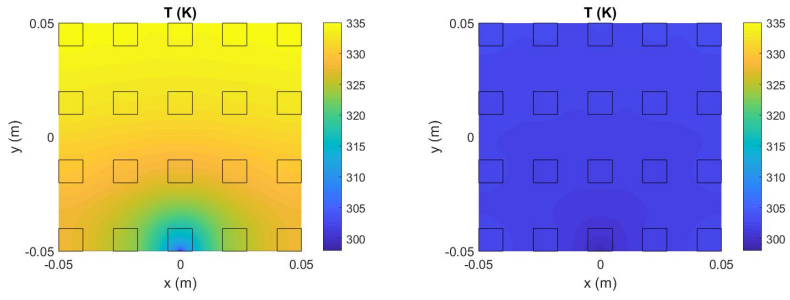


Figure 9.2: Temperature field results from CST MPS for uniform distribution (UD) of 20 heat sources in Case 1 and Case 4: (a) UD, Case 1 (no emission), (b) UD, Case 4 (with $\varepsilon = 0.1$).

Table 9.8: Optimized results in Case 4 for $N_g = 91^2$ with varying guaranteed min. spacing

Guaranteed min. spacing, d_g (m)	0.018	0.016	0.014	0.012
Number of iterations	3	5	2	2
Max. temperature, $T_{\max}$ (K)	302.58	302.83	302.89	302.89
Std., σ	0.2768	0.3202	0.8318	0.9886
Min. distance between the sources, $d_{\min}$ (m)	0.021	0.019	0.016	0.014
Optimization time, t_{opt} (s)	1822.40	2133.38	543.91	410.91

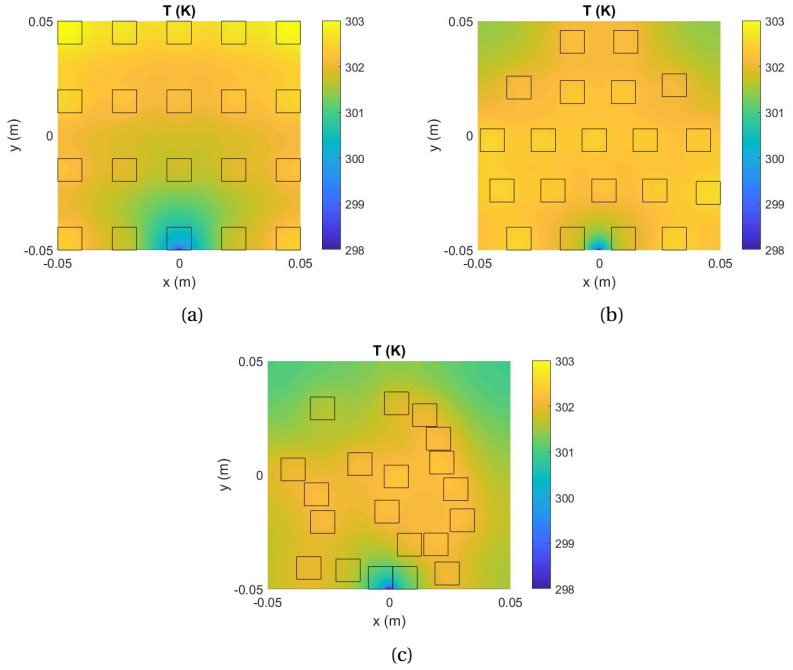


Figure 9.3: The layouts of 20 heat sources and temperature distributions for uniform distribution (UD), convex optimization (CO) and genetic algorithm (GA) in Case 4: (a) UD, (b) CO, (c) GA.

also used for the layout optimization in Case 4 with the aim of minimizing the maximum temperature. In GA, 120 iterations are used to obtain the final result where each iteration takes about 25 seconds. Fig. 9.3 gives the source locations and temperature distributions for the uniform distribution (UD), convex optimization (CO) and genetic algorithm (GA). The comparison results of these three approaches are summarized in Table 9.9. It is seen that CO and GA can provide a slight reduction in the maximum temperature, but the change is not noticeable for the selected parameters. On the other hand, optimized solutions possess much more uniform temperature distributions (with temperature standard deviation reduction by 41.78% and 26.46% for CO and GA, respectively).

Next, the results of the first three cases are compared with the results of the existing layout optimization methods, namely bionic optimization (BO) [185] and simulated

Table 9.9: Comparison of the optimized layouts and the uniform layout in Case 4 for different methods (N_s = 20)

UD		CO		GA	
T_{max} (K)	σ	T_{max} (K)	σ	T_{max} (K)	σ
303.02	0.4754	302.58	0.2768	302.25	0.3496

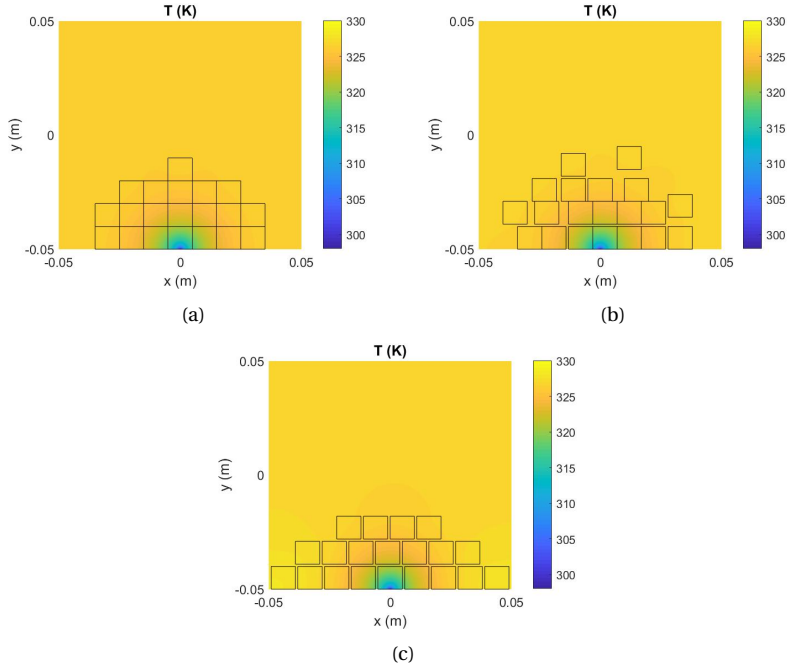


Figure 9.4: Optimized layouts of 20 heat sources for bionic optimization (BO), simulated annealing (SA) and convex optimization (CO) in Case 1: (a) BO, (Chen et al. [185]), (b) SA, (Chen et al. [187]), (c) CO.

annealing (SA) [187]. In Fig. 9.4, Fig. 9.5 and Fig. 9.6, the source location and temperature field distribution comparisons between these methods and the convex optimization (CO) are provided for Case 1, Case 2 and Case 3, respectively. The comparison results of these three methods are summarized in Table 9.10.

In Case 1, the performance of CO is worse than BO and SA. In CO, R_m is increased by 6.56% and 4.85% and σ_m is increased by 8.92% and 1.17% when compared to BO and SA, respectively. In Case 2, CO achieves the same results with BO and it performs better than SA. Compared to SA, R_m and σ_m are reduced by 9.09% and 6.85%, respectively. In Case 3, CO outperforms BO by decreasing R_m and σ_m by 12.82% and 13.59%, respectively. When compared to SA, although R_m increases by 9.24%, σ_m is reduced by 6.62% leading to a more uniform temperature distribution.

The computational costs for BO and SA are given in [187] in terms of the total number of temperature field calculations. It is stated that for 20 heat sources, the total number of temperature field calculations, which is proportional to the number of sources, is 420

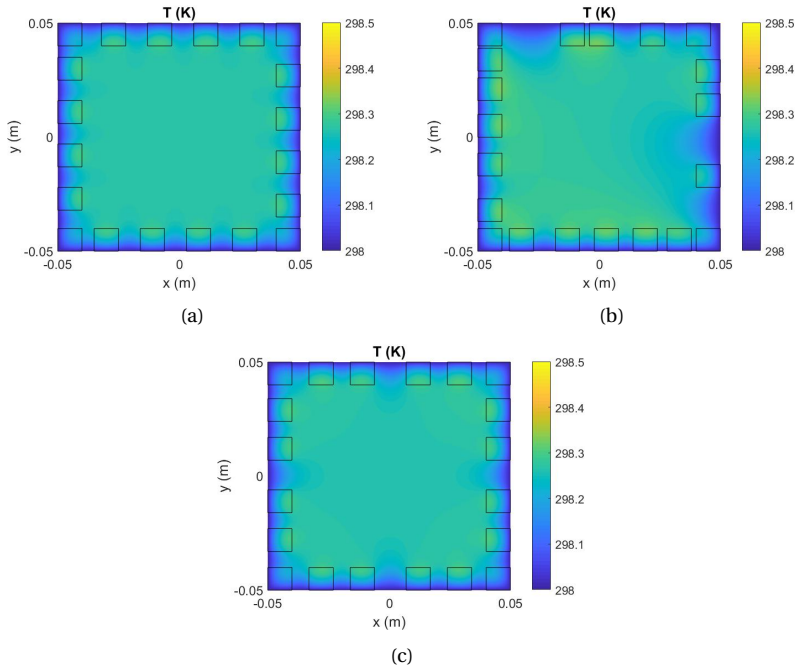


Figure 9.5: Optimized layouts of 20 heat sources for bionic optimization (BO), simulated annealing (SA) and convex optimization (CO) in Case 2: (a) BO, (Chen et al. [185]), (b) SA, (Chen et al. [187]), (c) CO.

Table 9.10: Comparison of the optimization results in the first three cases for different methods ($N_s = 20$)

	BO		SA		CO	
	R_m	σ_m (x10^{-2})	R_m	σ_m (x10^{-2})	R_m	σ_m (x10^{-2})
Case 1:	0.2820	1.2753	0.2866	1.3730	0.3005	1.3891
Case 2:	0.0030	0.0694	0.0033	0.0745	0.0030	0.0694
Case 3:	0.0312	0.4976	0.0249	0.4605	0.0272	0.4300

9

and 1760 for BO and SA, respectively. On the other hand, in CO, the temperature field is calculated only once in order to obtain the reference temperature field with uniform heat generation in the domain. Besides, the optimization time does not depend on the number of sources. In fact, in CO, the main factors affecting the optimization time are the number of grid cells for placing the heat sources (N_g) and the required number of iterations.

Lastly, the convergence of the convex optimization algorithm is shown in Fig. 9.7 for the case that requires the most number of iterations among all the cases studied in the paper (Case 3 for $N_g = 91^2$ and $d_g = 0.01$ m with 4 iterations, as seen in Table 9.6). In Fig. 9.7(a), the histogram plot of the heat source selection weights (that are larger than 0.1) is given up to the fourth iteration. Assuming that a source with a selection weight larger than 0.1 is 'active' (or 'selected'), the number of selected sources at each iteration is plotted in Fig. 9.7(b). It is seen that as iteration number increases, less sources are selected.

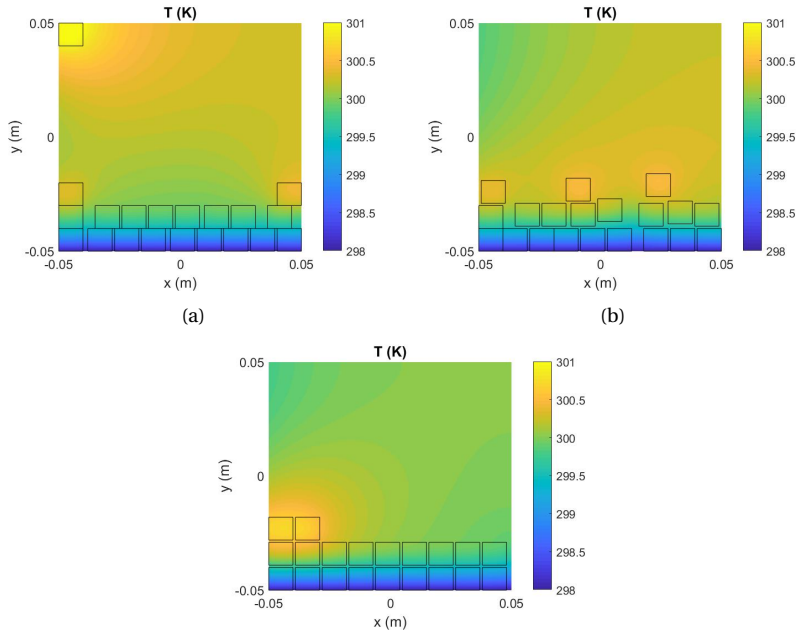


Figure 9.6: Optimized layouts of 20 heat sources for bionic optimization (BO), simulated annealing (SA) and convex optimization (CO) in Case 3: (a) BO, (Chen et al. [185]), (b) SA, (Chen et al. [187]), (c) CO.

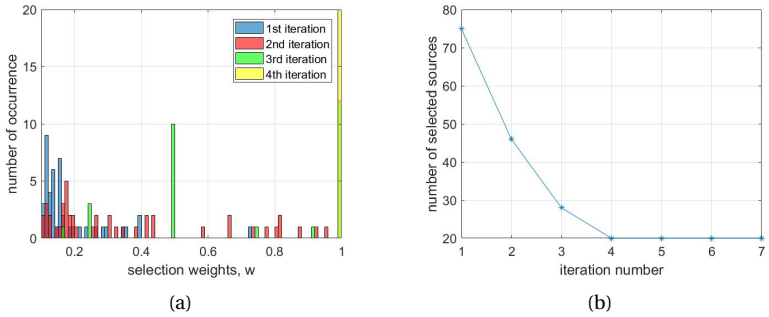


Figure 9.7: Convergence analysis of the convex optimization algorithm in Case 3 with d_g = 0.01 m: (a) Histogram of source selection weights at each iteration, (b) Number of selected sources at each iteration.

Besides, larger weights are assigned to satisfy the cardinality constraint, which provides the increased sparsity. At the fourth iteration, the number of active sources converges to 20 when 20 sources are selected with weight 1 and the result does not change afterwards.

CONCLUSION

In this section, cooling enhancement via heat source distribution optimization for two dimensional heat conduction problem has been studied. Using the convex algorithm, a new approach to decrease the maximum temperature and temperature non-uniformity has been proposed. This solution is based on obtaining a sparse solution from a fully-populated heat source array while keeping the pre-defined number of sources at potentially easy-to-cool regions of the domain. On top of being easy to implement, the proposed convex optimization algorithm has been seen to provide comparable or better results than the existing methods in terms of the maximum temperature and temperature uniformity. Moreover, convex optimization requires only a single temperature field calculation to be used as a reference, while for the existing methods, the number of temperature field calculations are proportional to the number of sources.

Three test cases with different boundary conditions have been revised to evaluate the performance of the proposed layout optimization algorithm, and the results have been compared with the uniform distribution of 20 heat sources. In all cases, it has been seen that using iterative reweighted l_1-norm convex minimization, the maximum temperature of the domain can be effectively reduced.

These scenarios taken from the thermal engineering field are in fact analogous to thermal management strategies in traditional AESAs (see Section 7.2) where the heat must be efficiently transferred to the array sides where the cooling is achieved. Therefore, it is possible to extend the presented convex optimization algorithm to a joint electromagnetics and thermal optimization. The way to do this is to assume that "$\mathbf{w}$" in (9.10) represents the excitation weights of the densely-packed antenna elements and to put additional constraints in (9.10) on the radiation pattern using the array factor formula. If the cooling is achieved from all sides, the reference temperature will be similar to the one shown in Fig. 9.1(c) and the optimizer will try to increase the number of elements at the array edges.

However, as we already explained in Section 7.1, the 5G industry uses planar AESAs for which the convective cooling to the third dimension becomes dominating. Therefore, a fourth case, representing a volume-to-volume situation, has also been introduced which took the heat emission from the domain surface into account. Convex optimization (CO) and genetic algorithm (GA) were used to optimize the positions of the heat sources and the results were compared with the uniformly distributed (UD) layout. The results indicated that for the selected surface emissivity coefficient, heat emission from the surface becomes the dominant cooling mechanism which is further enhanced by the conduction cooling. That is why the highest temperature value is almost the same (slightly improved) for CO and GA when compared to UD, but much better temperature uniformity can be achieved via layout optimization.

The parameters of the fourth case investigated here have been adopted from the other three cases and thus, it does not represent a realistic 5G integrated array scenario. That is why in the next section, the 3-D heat convection dominated 5G antennas are studied and the function of layout optimization in antenna cooling enhancement is explained.

9.1.2. 3-D HEAT CONVECTION DOMINATED SCENARIOS

INTRODUCTION

The problem of cooling of electronics under a conduction dominated cooling scenario was studied in the previous section by solving the heat equation in 2-D and optimizing the positions of the heat sources inside an aperture depending on the position of the heatsink. Combining this approach mathematically with the array factor and thus the radiation pattern of the array (by assuming the antennas are at the same 2-D location with the heat-generating transceiver chips) would result in a joint optimization of both thermal and electromagnetic performances. Such conduction-based thermal management techniques can be applied to the traditional phased arrays where electronic circuit cards are placed orthogonally to the array and the cooling is achieved only from the outer edges of the array (as previously illustrated in Fig. 7.3 and Fig 7.4). For such systems, intuitively, it can be inferred that more elements (chips) should be placed at the edges of the aperture since otherwise, the heat gets trapped in the middle of the array. However, such traditional arrays are not preferred in 5G since they are incompact and they need a large number of RF boards with additional cabling for signal routing (see Section 7.1). What industry prefers today for 5G is a single, multi-layer RF board antenna that is low-cost and low-profile. In such designs, the radiators are placed on one side of the board while the chips are located on the opposite side where the heatsink is attached. The dominating cooling mechanism in this case becomes the natural convection to the third dimension instead of the conduction in 2-D towards the array edges since the heat is not trapped anymore, but is directly transferred to the heatsink and finally to the surrounding air. Thermal modeling of such systems is done via thermal resistors representing the temperature difference among different contact points (board node, chip case, solder bumps etc.) in the design along the third dimension (see Section 8.2), which decouples the thermal and electromagnetic aspects in terms of the element positions. The only practice in this case for an antenna array topology designer is to artificially imply thermal-awareness in the 2-D element optimization procedure by putting smart constraints in the electromagnetic optimization.

In the case of having relatively low thermal conductivity antenna boards, the ICs act as hot-spots in the design. To enhance the cooling in such a case, one must increase the surface area around each IC, and correspondingly around each antenna. Thus, in addition to the side lobe constraints on the radiation pattern, there should be a minimum inter-element spacing constraint, which results in sparse arrays in terms of the average distance between the array elements.

In this section, both electromagnetic radiation and thermal management aspects of 5G base station antennas are jointly studied by examining and comparing various sparse array layouts and element position optimization techniques. The sparsity in the layout is formed by keeping the element number same and increasing the inter-element spacings. Two-resisor thermal model is used in thermal simulations. Several existing techniques are exploited to create irregularity in the layout with the aim of enhancing the cooling performance and, at the same time, to reduce the high side lobes or grating lobes. In this work, we use simple, conventional patch antennas that are connected to the chips at the same 2-D locations to avoid practical challenges on chips integration with the antennas. It is worth to note that by using smartly designed antenna elements and ir-

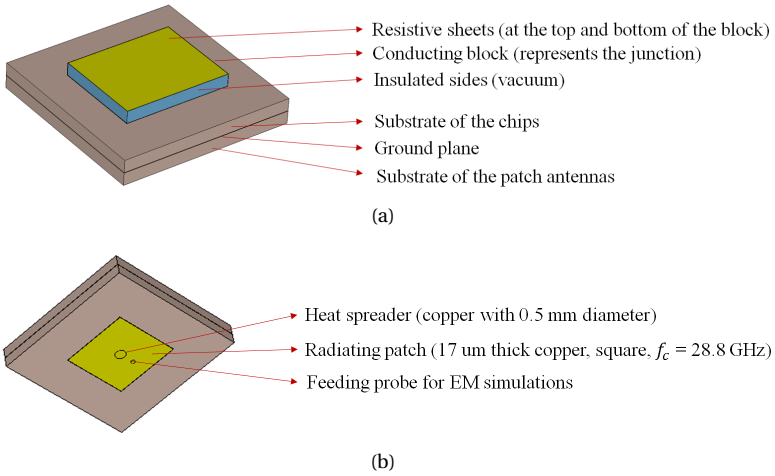


Figure 9.8: Design of a unit cell in CST, (a) back side with a chip, (b) front side with a patch.

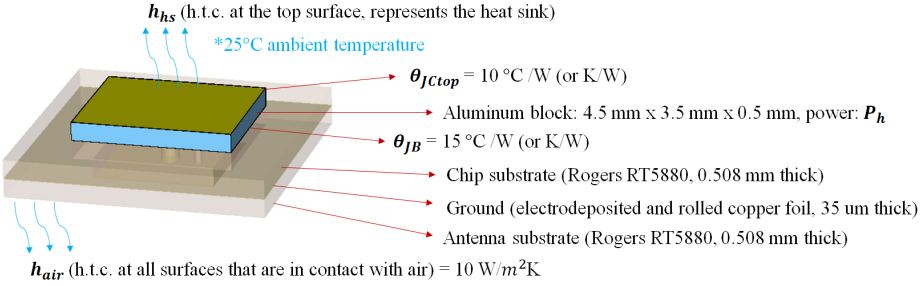


Figure 9.9: Realization of the two-resistor model in CST and selected parameters.

regularly clustered sub-arrays instead of the patches, EM performance might be further improved. Next, first, the simulation settings in CST and realization of the 3-D model are described. The section is continued with the analyses and discussions on several array topologies that are based on arrays of regular and irregular linear arrays, spiral arrays, sparse circular arrays, thinned arrays and heat sink antenna arrays.

9

SIMULATION MODEL

The design of a unit cell structure is visualized in Fig. 9.8. Fig. 9.8(a) shows the back side of the unit cell in which the conducting aluminum block representing the IC junction is visible. The sides of the block are isolated with vacuum. Resistive sheets are used at the top and bottom of the block. The substrates for the chip and the patch antenna are also shown together with the ground plane. Fig. 9.8(b) presents the front side of the cell for which the radiating patch (with tuned dimensions to obtain the center frequency, f_c, at 28.5 GHz for the arrays) is seen. Locations of the feeding probe that is used for EM simulations and the heat spreader (used only later in this Section in the "*F. Heat Sink Antenna Arrays*" part) are also visible.

Realization of the two-resistor model and several parameter values that are common to all simulations presented in this paper are provided in Fig. 9.9. Apart from the fixed patch dimensions (3.13 mm x 3.13 mm without the heat spreader and 3.20 mm x 3.20 mm with the heat spreader), it is assumed that the chip size is 4.5 x 3.5 x 0.5 mm^3, θ_{JCtop} is equal to 10 K/W and θ_{JB} is taken as 15 K/W. Note that these values are taken as a reference which should be modified depending on the particular chip to be used. In addition, both for the antenna and chip, Rogers RT5880 substrate (with 0.508 mm thickness) is used, considering its low cost and sturdiness. The ambient temperature is taken as 25 degree Celsius. Heat transfer coefficient (H.T.C.) for all the surfaces that are in contact with air (h_{air}) and cooled via natural convection is assumed to be equal to 10 W/m^2K. The emissivity coefficient at all the surfaces is assumed to be equal to 0.9 (emissivity represents the energy emitted via thermal radiation, whose effect is nearly negligible compared to the convective cooling). The heating power generated by each chip (P_h) and the equivalent H.T.C. representing the heat sink (h_{hs}) are varied for several cases that are used in the parametric studies throughout the paper. It is worthy to note that the detailed CPU heat sinks used in Section 8.3 is replaced here with equivalent H.T.C.s, which are on the order of thousand Watts per meter-square Kelvin (1000 W/m^2K) [159], in order to perform the parametric studies more quickly.

RESULTS AND DISCUSSION

A. Regular Array Layouts

As a start, a parametric study for 8 x 8 regular square array layouts has been performed by changing the heat generated at each RF channel (P_h), equivalent H.T.C. of the passive CPU cooler (h_{hs}) and the spacing between the elements (d_e). Center frequency of 28.5 GHz has been used in calculation of λ, and thus, the physical locations of the chips. At the output, the maximum junction temperature for the whole array ($T_{\max}$) has been computed which is given by the highest temperature at the aluminum blocks. Note that, although, after several trials, it has been seen that the effect of increased substrate dimension on $T_{\max}$ is very small (less than 1°C), the substrate and ground plane dimen-

Table 9.11: A parametric study on $T_{\max}$ with the 8 x 8 regular square array layout by varying P_h, h_{hs} and d_e

Case #	P_h (W)	h_{hs} (W/m^2K)	d_e (λ)	$T_{\max}$ (°C)
Case-1	2	1000	0.5	164
Case-2	2	1000	1	155
Case-3	2	1000	2	141
Case-4	2	3000	0.5	86
Case-5	2	3000	1	84
Case-6	2	3000	2	81
Case-7	3	1000	0.5	232
Case-8	3	1000	1	218
Case-9	3	1000	2	197
Case-10	3	3000	0.5	117
Case-11	3	3000	1	114
Case-12	3	3000	2	109

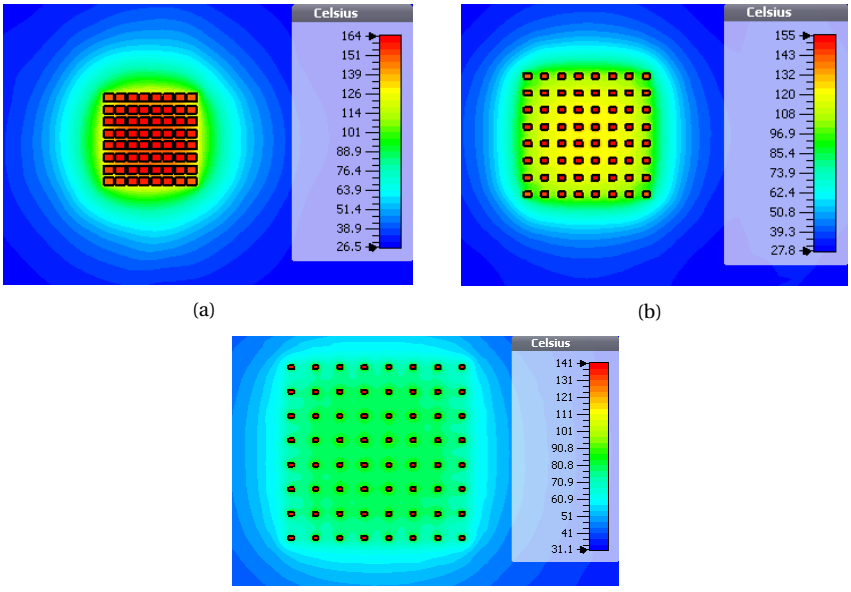


Figure 9.10: Sample temperature distributions with passive CPU cooling for the 8x8 regular square array along the horizontal cut in the middle of the conducting blocks representing the junctions, (a) Case-1, (b) Case-2, (c) Case-3.

sions are equal to 18λ x 18λ in each case for fairer thermal comparison among all the cases with different aperture sizes.

The results of the parametric study[1] with the 8x8 regular array layouts are given in Table 9.11. Fig. 9.10 provides sample temperature distributions at the IC junctions for Case-1, Case-2 and Case-3 for better visualization of the results.

EM simulations have also been performed for the same 8x8 regular patch arrays with $d_e = 0.5\lambda$, $d_e = 1\lambda$ and $d_e = 2\lambda$. As for matching, the active reflection coefficients have been computed to see the center frequency and operational bandwidth. The results for the two cases with the smallest and largest inter-element spacings are given in Fig. 9.11 for the broadside beam, which confirms that the arrays are nicely matched at the center frequency of 28.5 GHz with a bandwidth of more than 1 GHz.

Lastly, the realized gains for such arrays are provided in Fig. 9.12 for broadside beams. The formation of grating lobes is clearly seen in Fig. 9.12(b) and Fig. 9.12(c) for which $d_e = 1\lambda$ and $d_e = 2\lambda$, respectively. On the other hand, Table 9.11 shows that increasing the inter-element spacing helps increase the cooling performance, whose extent depends very much on the selected parameters such as the heating power of the chip and the capability of the attached heat sinks. This trade-off between the EM and cooling perfor-

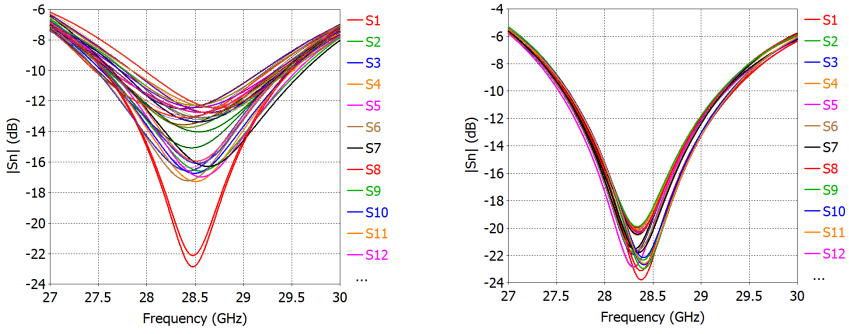
[1]The dissipated heat per channel depends heavily on the EIRP requirements of the system. The values used here are set to relatively higher values than those of a typical small-cell base station serving one user at a time, with the objective of providing sufficient EIRP to more concurrent users per array and/or larger cells, for which the cooling problem becomes significant.

mance is the motivation to search for alternative array layouts that are capable of yielding satisfying outcomes in both EM and thermal aspects. Examples of such layouts are given next employing several techniques that were previously addressed in the literature.

B. Array of Linear Irregular Arrays

From Fig. 9.12(b) and Fig. 9.12(c), it was seen that increasing the inter-element spacing for cooling enhancement leads to grating lobes in the radiation pattern. In the case of having users sharing the same time-frequency resource and located in the directions of grating lobes, the interference towards those users is going to be very high. Zero-forcing or orthogonal beam forming are commonly used digital beamforming algorithms at the base stations which allows killing the inter-user interference in a multi-user network completely by placing nulls in the directions of the interferers [193–195]. However, when two (or more) users are positioned close to the grating lobes (which may occur frequently for randomly distributed users), zero-forcing fails and a null is placed in the direction of each user. In order to prevent this, irregular array layouts can be used which help reduce the level of grating lobes and ensure good performance with zero-forcing [196; 197].

For the layout given in this section, the convex optimization method proposed in [C-5] has been used to synthesize an 8-element linear array with a guaranteed minimum inter-element spacing, $d_{\min}$, of 2λ for the aim of minimizing the maximum side lobe level in the visible region for the broadside beam. Using the optimized linear array, an



(a) (b)

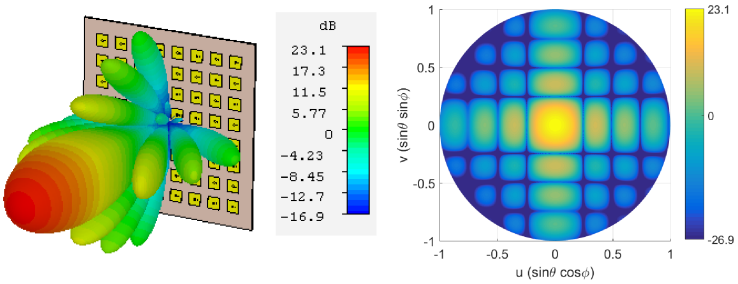
Figure 9.11: Active reflection coefficients at each port of the 8x8 regular square array for the broadside beam, (a) $d_e = 0.5\lambda$, (b) $d_e = 2\lambda$.

Table 9.12: Comparison of maximum temperature values in several cases for 8x8 regular array layout with $d_e = 0.5\lambda$ and irregular optimized array layout with $d_{\min} = 2\lambda$

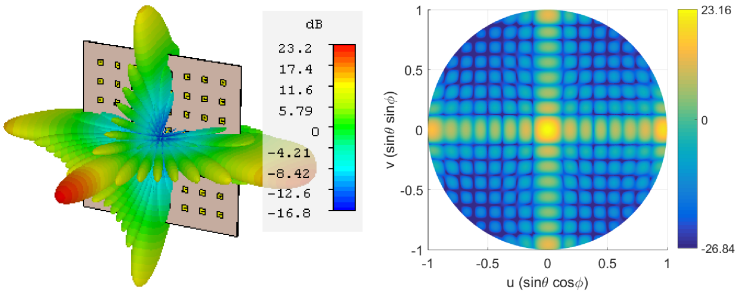
Case #	$T_{\max}$ (°C) regular array (from Table 9.11)	$T_{\max}$ (°C) irregular array
Case-1	164	140
Case-4	86	81
Case-7	232	196
Case-10	117	109

8 x 8 irregular planar array has been created for performance comparison with the 8 x 8 regular array with 2λ inter-element spacing.

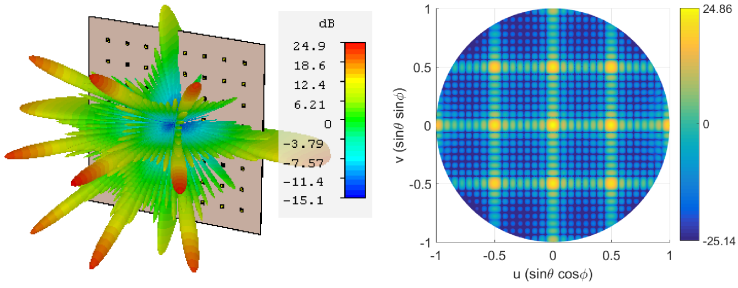
Fig. 9.13 shows the optimized element locations for the synthesized irregular array. The radiation pattern of the array at broadside with the new element locations is given in Fig. 9.14. Compared to the radiation from the regular array given in Fig. 9.12(c), around 5 dB grating lobe suppression is obtained with the new topology. At the same time, since 2λ minimum inter-element spacing is guaranteed, much better cooling performance is obtained as compared to the regular counterparts with $d_e = 0.5\lambda$ (see Table 9.12).



(a)



(b)



(c)

Figure 9.12: Realized gains of 8x8 regular square array layouts for broadside beams at 28.5 GHz, (a) $d_e = 0.5\lambda$, (b) $d_e = 1\lambda$, (c) $d_e = 2\lambda$.

For the multi-user scenario, two simultaneous co-frequency users located at 100 meter distance from the base station with angular positions $\theta_1 = 0$, $\phi_1 = 0$ and $\theta_2 = 30$ degrees, $\phi_2 = 0$ have been considered. Note that both users are located at the grating lobes shown in Fig. 9.12(c). Zero forcing is performed for these users by inverting the channel matrix that is formed by assuming a LoS-only propagation path. The formulation of this technique have already been widely studied [193–196]. Therefore, only the resulting excitation coefficients at each element in the 8 antenna columns (including 8 elements arranged vertically) are given here with a discussion on the outcome.

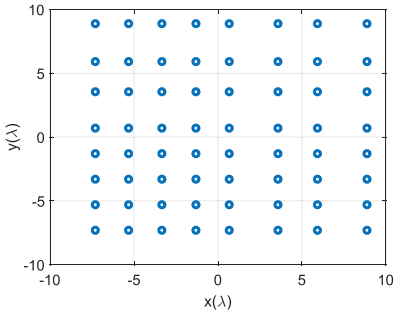


Figure 9.13: Element locations in the 8x8 optimized irregular array layout with $d_{\min} = 2\lambda$ with a QR code including the positions (in λ).

Table 9.13: Zero-forcing excitations at each column of the 8x8 regular square array with $d_e = 2\lambda$ for user-1 ($\theta_1 = 0$, $\phi_1 = 0$) and user-2 ($\theta_2 = 30$ degrees, $\phi_2 = 0$) located at 100 meter distance from the base station

Column #	User-1		User-2	
	Amplitude	Phase (in degrees)	Amplitude	Phase (in degrees)
1	0.5410	-158.1766	0.5410	-158.2201
2	0.0978	171.9686	0.0975	172.1213
3	0.2314	39.3215	0.2313	39.0852
4	0.3812	31.9362	0.3811	31.7519
5	0.3812	28.2898	0.3812	28.1056
6	0.2315	21.2354	0.2316	20.9992
7	0.0969	-111.0076	0.0973	-110.8543
8	0.5391	-140.0903	0.5392	-140.1336

Table 9.14: Zero-forcing excitations at each column of the 8x8 optimized irregular array with $d_{\min} = 2\lambda$ for user-1 ($\theta_1 = 0$, $\phi_1 = 0$) and user-2 ($\theta_2 = 30$ degrees, $\phi_2 = 0$) located at 100 meter distance from the base station

Column #	User-1		User-2	
	Amplitude	Phase (in degrees)	Amplitude	Phase (in degrees)
1	0.1994	118.2254	0.1992	2.1942
2	0.1995	117.6059	0.1993	1.9589
3	0.1995	117.1821	0.1993	1.7976
4	0.1995	116.9539	0.1994	1.7105
5	0.1994	116.9214	0.1995	1.6977
6	0.6185	127.6195	0.6185	-169.7093
7	0.5950	107.3294	0.5951	150.0215
8	0.2541	147.1600	0.2543	-59.1379

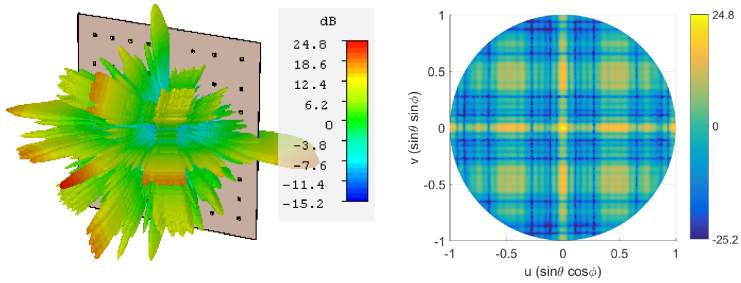
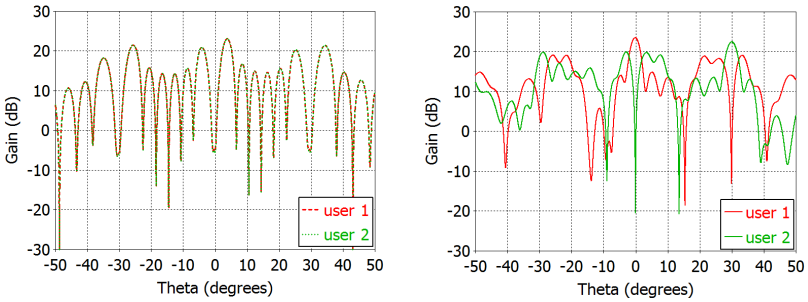


Figure 9.14: Realized gain of the 8x8 optimized irregular array layout with $d_{\min} = 2\lambda$ for the broadside beam at 28.5 GHz.



(a) (b)

Figure 9.15: Zero forcing results (in terms of realized gains) for 8x8 arrays with two users sharing the same time-frequency resource and located at 100 meter distance from the base station at two grating lobe positions of the regular layout ($\theta_1 = 0$, $\phi_1 = 0$, $\theta_2 = 30$ degrees, $\phi_2 = 0$), (a) regular array with $d_e = 2\lambda$, (b) optimized irregular array with $d_{\min} = 2\lambda$.

Table 9.13 and Table 9.14 provide the related excitation coefficients for the regular and optimized irregular array layouts, respectively. The realized gains (at the $\phi = 0$ cut) in this case are given in Fig. 9.15. As mentioned before, due to the grating lobes, zero forcing does not work for the regular array layout (see Fig. 9.15(a)). However, the irregular layout and resulting grating lobe suppression make formation of maxima and nulls possible at the corresponding user locations as seen in Fig. 9.15(b).

C. Spiral Arrays

Spiral arrays[2] have been shown to provide low side lobe levels and grating lobe free scanning capabilities with large inter-element spacings [198]. Ring arrays have been used for space applications since the 80's and spiral arrays for space use were revisited in the last decade [198; 199].

In polar coordinates, the elements in a spiral can be arranged according to the following equations [198],

[2]A generic sunflower array topology is used in this section for demonstration purposes. It is worth to note that the sunflower is naturally optimal for a circular field-of-view. Ideally, the layout optimization approach proposed in Section 6.2 should be used to synthesize an optimal layout for the desired field-of-view.

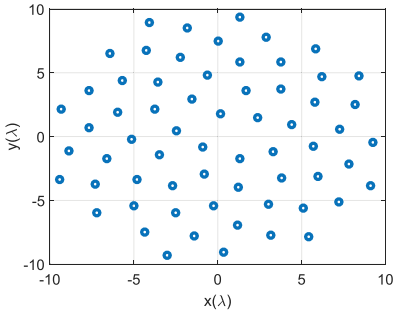


Figure 9.16: Element locations in the 64-element spiral array layout with $d_{\min} = 2\lambda$.

Table 9.15: Comparison of maximum temperature values in several cases for 8x8 regular array layout with $d_e = 0.5\lambda$ and 64-element sunflower array layout with $d_{\min} = 2\lambda$

Case #	$T_{\max}$ (°C) regular array (from Table 9.11)	$T_{\max}$ (°C) sunflower array
Case-1	164	138
Case-4	86	81
Case-7	232	194
Case-10	117	108

$$\rho_n = \alpha\sqrt{\frac{n}{\pi}} \tag{9.11}$$

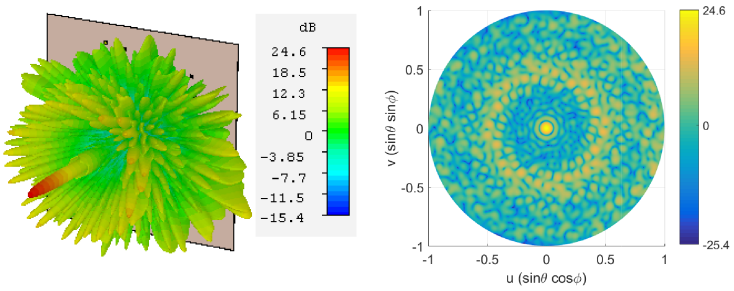
$$\Psi_n = 2\pi n\beta \tag{9.12}$$

where ρ_n is the radial distance of the n^{th} element, the parameter α relates to the mean distance between the neighboring elements and Ψ_n is the angular displacement of each element which depends on the parameter β.

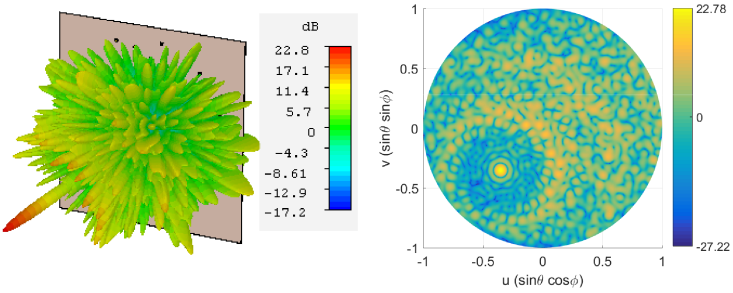
A special case occurs when β is equal to the golden angle ($\beta = \frac{\sqrt{5}+1}{2}$), which resembles to the distribution of sunflower seeds. Being an irrational number, this particular selection of β prevents any rotational periodicity and shadowing of radiators.

Due to these appealing features and its impact on side lobe level suppression even for relatively large inter-element spacings, a uniform sunflower array topology is proposed in this section for passively cooled 5G base stations. Note that density tapering can also be applied on such arrays by optimizing the element positions to further reduce the side lobe levels [199] while keeping the optimum power efficiency without amplitude tapering, which is not addressed here.

In this paper, $\alpha = 2.213\lambda$ is used to have $d_{\min} = 2\lambda$. The element locations for this sunflower array antenna topology for an 64-element array is given in Fig. 9.16. The resulting radiation pattern for the broadside beam is given in Fig. 9.17(a). In order to scan the beam, appropriate phase shifts can also be applied at each element using the following relation



(a)



(b)

Figure 9.17: Realized gain of the 64-element sunflower array with $d_{\min} = 2\lambda$ at 28.5 GHz, (a) broadside beam, (b) scanned beam ($\theta_s = 30$ degrees and $\phi_s = 225$ degrees).

$$\Phi_{n,s} = e^{-jk(\sin\theta_s \cos\phi_s x_n + \sin\theta_s \sin\phi_s y_n)} \tag{9.13}$$

where $\Phi_{n,s}$ is the excitation coefficient at the n^{th} element for the scan angle (θ_s, ϕ_s) and $k = 2\pi/\lambda$ is the wavenumber. As an example, for $\theta_s = 30$ degrees and $\phi_s = 225$ degrees, the realized gain in Fig. 9.17(b) is obtained. In both cases, no grating lobes appear and the maximum side lobe level becomes around -13 dB when compared to the broadside gain. Besides this advantage on the EM performance, the sunflower layout with $d_{\min} = 2\lambda$ is able to achieve much better cooling when compared to regular square layout with $d_e = 0.5\lambda$ (see Table 9.15).

D. Thinned Arrays

In this section, the aim is to investigate the effect of array thinning on the EM and thermal performance of the arrays. The 8 x 8 regular square grid with $d_e = 0.5\lambda$ has been used as the reference and 15 elements have been randomly selected to be switched off. Here, the cooling performance of the thinned array is compared both with the 8 x 8 regular array (which has 15 extra heat sources) and 7 x 7 regular array (which has the same number of sources, but densely populated with $d_e = 0.5\lambda$). EM performance of the thinned array layout is also compared with the regular counterparts.

A realization of random array thinning is given in Fig. 9.18, which is used for the

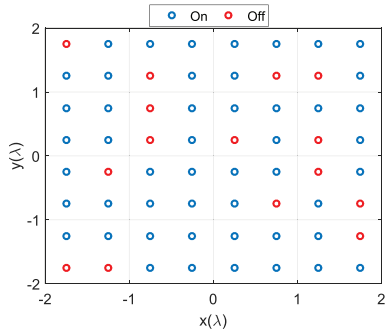


Figure 9.18: Element locations in a 49-element array obtained via random thinning of an 8x8 regular square array layout.

Table 9.16: Comparison of the maximum temperature values in several cases between the 8 x 8 regular array layout with $d_e = 0.5\lambda$, 7 x 7 regular array layout with $d_e = 0.5\lambda$ and the 49-element randomly thinned array

Case #	$T_{\max}$ (°C) 8x8 regular array (from Table 9.11)	$T_{\max}$ (°C) 7x7 regular array	$T_{\max}$ (°C) 49-element thinned array
Case-1	164	164	158
Case-4	86	86	85
Case-7	232	232	224
Case-10	117	117	115

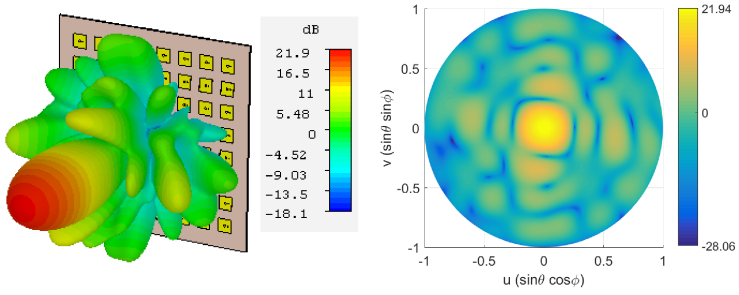


Figure 9.19: Realized gain of the 49-element randomly thinned array (whose layout is given in Fig. 9.18) for the broadside beam at 28.5 GHz.

results given in this section. Table 9.16 provides the comparison of $T_{\max}$ between the 8x8 regular array layout with $d_e = 0.5\lambda$, 7 x 7 regular array layout with $d_e = 0.5\lambda$ and the 49-element randomly thinned array whose layout is shown in Fig. 9.18.

In the thermal analysis given in Table 9.16, it is seen that due to the same inter-spacing of the heat sources, the maximum temperature values are the same in each case analyzed in this section for the 49-element arrays as compared to the 64-element arrays with regular layouts. However, due to the sparsity of the thinned array, enhancement in cooling can be obtained, especially in the cases with relatively low h_{hs} (Case-1 and Case-7). This improvement comes at the expense of reduced EM performance.

As for the EM performance, compared to the 64-element array, the realized gain is reduced by 1.2 dB for the 49-element thinned array and the maximum side lobe level is approximately 3 dB larger compared to the regular square layout (see Fig. 9.19). Besides, since the number of transceiver modules is lower for the thinned arrays, both the transmit power and the gain and consequently, the EIRP at the users are lower than those of the fully-populated arrays. Moreover, unless complex interference suppression techniques are used, the high level of side lobes in Fig. 9.19 could result in high interference levels at the other co-frequency users that are served simultaneously. Therefore, an alternative sparse array topology is introduced next, which helps reduce the level of side lobes while keeping a comparable cooling performance.

E. Sparse Circular Ring Arrays

In sparse circular ring array layout, the sparsity is achieved along the radial direction, which helps for the cooling, and circular arrays are exploited to enhance the EM performance by reducing the level of high side lobes that occur in the case of the randomly thinned arrays. Here, the compressive-sensing method proposed in [200] has been used to synthesize a low side lobe 49-element circular isophoric (i.e. uniform amplitude) sparse array.

Fig. 9.20 shows the optimized element locations for which the inter-element spacings are larger than 0.525λ (i.e. $d_{\text{min}} = 0.525\lambda$). Comparison of the maximum temperature values between the 49-element regular array with $d_e = 0.5\lambda$ and the designed sparse circular ring array is provided in Table 9.17. It is seen that enough sparsity is achieved with the circular ring array to have almost the same cooling performance as with the randomly thinned array in each study case (Case-1, Case-4, Case-7 and Case-10).

In terms of the EM performance, the realized gain of the circular ring array (given in Fig. 9.21) shows that the maximum side lobe level is reduced by nearly 7 dB when compared to the thinned array, which results in around -17 dB maximum side lobe level with respect to the broadside gain. Note that, since the layout is optimized only for the broadside beam, wide-angle scanning could yield higher side lobes in the visible region. This issue might be handled by using the multiple beam optimization technique proposed in Section 6.2 or using additional amplitude tapering at the expense of decreased power efficiency.

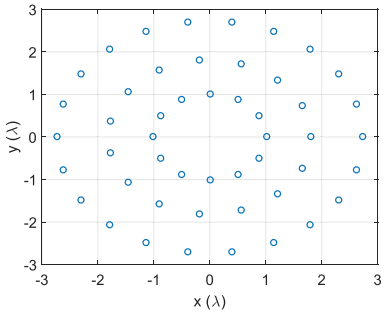


Figure 9.20: Element locations in the 49-element optimized sparse circular ring array with $d_{\text{min}} = 0.525\lambda$.

Table 9.17: Comparison of the maximum temperature values in several cases for the 7 x 7 regular array layout with $d_e = 0.5\lambda$ and 49-element sparse circular ring array layout with $d_{\min} = 0.525\lambda$

Case #	$T_{\max}$ (°C) thinned array (from Table 9.16)	$T_{\max}$ (°C) sparse circular ring array
Case-1	164	159
Case-4	86	85
Case-7	232	225
Case-10	117	115

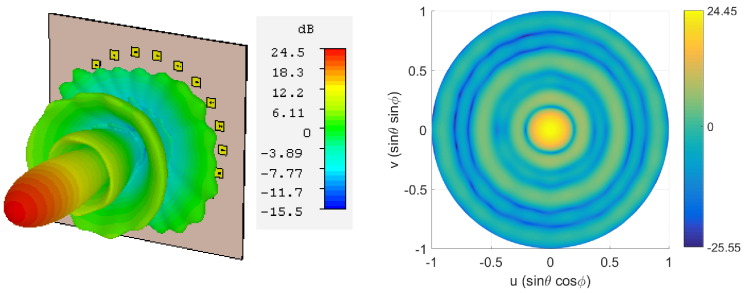


Figure 9.21: Realized gain of the 49-element sparse circular ring array for the broadside beam at 28.5 GHz.

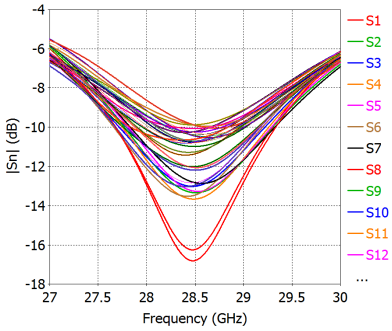
F. Heat Sink Antenna Arrays

Heat sink antenna arrays can be designed by placing a copper heat spreader pin in the middle of each patch (as in [154]), connecting the radiating patch to the chip and the ground plane (revisit Fig. 9.8(b) for visualization). The details of such a design and related design parameters were previously explained for a single element in a unit cell.

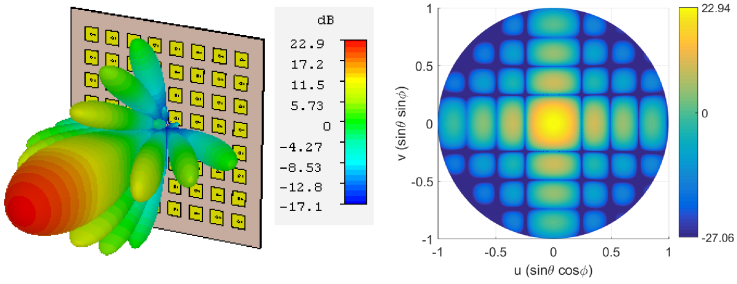
Here, the effect of exploiting heat spreaders in antenna arrays is studied by using the

Table 9.18: A parametric study on $T_{\max}$ with the 8 x 8 regular square array layout by exploiting heat spreaders and varying P_h, h_{hs} and d_e

Case #	$T_{\max}$ (°C) without spreader (from Table 9.11)	$T_{\max}$ (°C) with spreader
Case-1	164	164
Case-2	155	153
Case-3	141	134
Case-4	86	86
Case-5	84	84
Case-6	81	80
Case-7	232	232
Case-8	218	214
Case-9	197	187
Case-10	117	117
Case-11	114	113
Case-12	109	107



(a)



(b)

Figure 9.22: EM simulation results of the 8x8 regular square array layout with $d_e = 0.5\lambda$ for the broadside beam including the heat spreaders, (a) active reflection coefficients, (b) realized gain at 28.5 GHz.

8 x 8 square array layout as the reference. The thermal simulations for the cases used in Table 9.11 with no spreaders have been repeated including the heat spreaders. The comparison of $T_{\max}$ with and without the spreaders for the 8 x 8 regular square layouts is provided in Table 9.18.

9

It is seen that the improvement in heat removal via the heat spreaders is most significant when the heat dissipation per chip is relatively large (equal to 3 W among the cases studied in the paper), H.T.C. of the heat sink is relatively low (equal to 1000 W/m^{2}K among the cases studied in the paper) and the inter-element spacing is relatively large (equal to 2λ in this paper).

Fig. 9.22 presents the EM simulation results (active reflection coefficients and the realized gain) for the 8 x 8 regular square array layout with $d_e = 0.5\lambda$, including the heat spreaders. When compared to the results in Fig. 9.11(a) and Fig. 9.12(a) with no spreaders, it can be inferred that the effect of heat spreaders on the EM performance is not significant since they are located in the middle of the radiating patches, which prevents the parasitic radiation. Thus, heat spreaders between the chips and the patches can be integrated for cooling enhancement in all the layouts studied throughout the paper.

CONCLUSION

In this section, a novel relation between thermal management and layout sparsity has been found. In order to enhance the cooling performance in CPU cooler integrated arrays, sparsity has been introduced in the antenna layouts by increasing the inter-element spacings. Several sparse array topologies have been proposed and evaluated in terms of their thermal and electromagnetic performances.

From the simulation results, the following main observations have been made.

1. In the case of having low thermal conductivity substrates, highly inefficient ICs and a heat sink with lower heat transfer capability, layout sparsity could provide significant reduction in the maximum junction temperature of the array.

2. Although capable to achieve additional cooling, increasing the inter-element distance in regular square layouts to create sparsity leads to the formation of grating lobes, which might result in very high interference among the simultaneous co-frequency users.

3. Alternative sparse array layouts that are based on linear or planar irregular arrays, spiral arrays, thinned arrays or circular ring arrays can be used to obtain better electromagnetic performance (compared to the regular square arrays) while keeping the enhanced cooling capacity.

4. It has been found that for different sparse topologies investigated, the major factor to reduce the temperature is the average distance between the elements. A sunflower-like topology (with further optimization, as performed in Section 6.2, for a few typical 5G angular sectors) has been found to be the best one from this point of view since the inter-element spacing in this topology can be the largest among all investigated ones, while keeping low side lobe levels.

Furthermore, creating a conduction path between the chips, ground plane and the radiators by exploiting heat spreaders located in the middle of the patches provides additional cooling without affecting the radiation performance. Cooling with spreaders increases with increasing power dissipation per chip and decreasing H.T.C. of the heat sink attached to the chips.

9.2. SUBSTRATE THERMAL CONDUCTIVITY & SIZE ENHANCEMENT BASED APPROACH

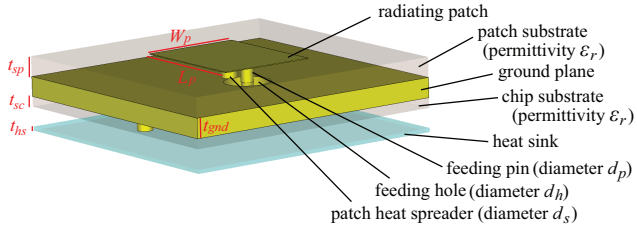
9.2.1. INTRODUCTION

Considering their advantages, fanless CPU coolers were proposed for array cooling in low-cost and low-profile planar 5G AESAs in Section 8.3. Furthermore, following the work on heat source layout optimization in Section 9.1.1, sparse antenna (or chip) array topologies were investigated for cooling extension in 5G antennas in Section 9.1.2. It was seen that fanless CPU coolers are able to remove most of the heat from the chips, while topology optimization can provide additional cooling by increasing the surface area around each antenna element. However, the work in the previous section on layout

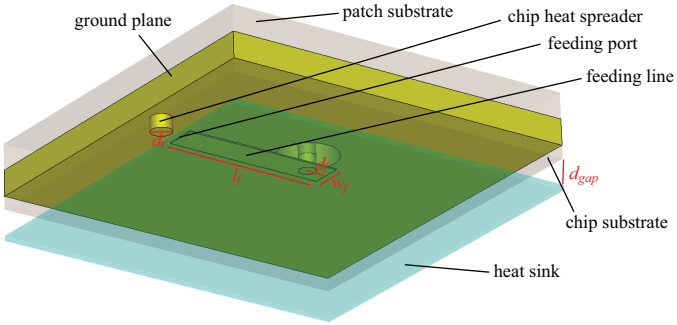
optimization has focused on thin (35 µm copper) ground planes, which mimics scenarios where each source creates a hot-spot at the corresponding chip position. In fact, a thick ground plane can be used to aid in heat transfer by increasing the thermal conductivity of the substrate and spreading the hot-spots [201; 202]. Although qualitatively mentioned in the antenna-related publications, there is no prior work on thermal simulations with thick ground plane antennas which shows the effectiveness of the design in cooling. Therefore, in this section, the aim is to investigate the extent of cooling enhancement in passively cooled (by a fanless CPU cooler) planar AESAs with increased ground plane thickness. The effect of total aperture size on the maximum temperature of the array is also investigated.

9.2.2. SIMULATION MODEL

For the electromagnetic (EM) model, Rogers RT5880 is used as the substrate material. Copper is used for the feeding pin, feed line and ground plane. To have a more realistic feeding model as compared to the one in Section 9.1.2, a 50 Ohm feeding microstrip line is designed on the chip substrate. The losses are included in the simulations as well. The patch is excited by a feeding pin by vertically connecting the patch to the feeding line, as in [203]. Although it may present manufacturing reliability problems, designing a pin-fed patch is easier compared to a non-contact transition in the case of having thick ground planes [204]. The decreasing active input resistance and increasing reactance of pin-fed elements (see Fig. 9.29) can easily be compensated by adjusting the patch



(a)



(b)

Figure 9.23: EM design parameters in a unit cell, (a) top view, (b) bottom view.

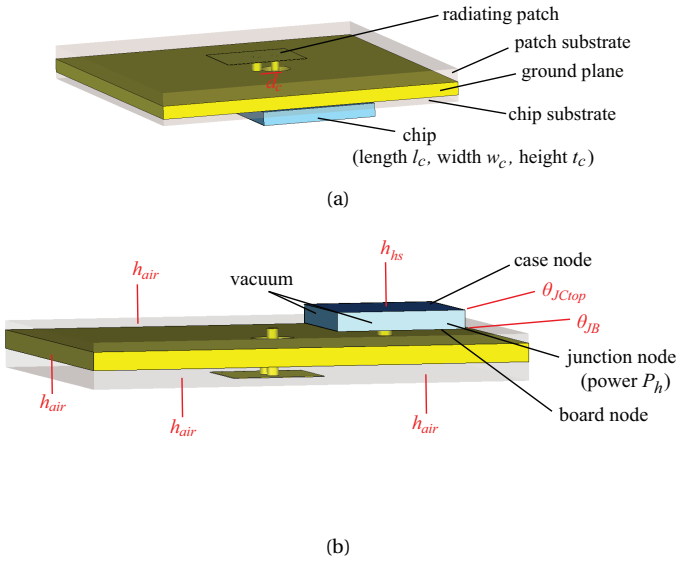


Figure 9.24: Thermal design parameters in a unit cell, (a) top view, (b) bottom view.

dimensions (smaller width and larger length, see Fig. 9.30). For aperture coupling, on the other hand, the coupling decays with ground plane thickness, while increasing the slot length to compensate the decay may lead to high back lobes. An aluminum heat sink plate is also placed at the back of the array, representing the CPU cooler. The EM model for a unit cell is given in Fig. 9.23.

The thermal model used in the paper is the two-resistor model [163]. The chips are modeled as aluminum blocks with insulated sides. Heating power (P_h) is applied to the chip. Heat transfer coefficients (H.T.C's) for the naturally cooled substrate-to-air surfaces (h_{air}) and for the heat sink (h_{hs}) (CPU cooler) at the chip case are defined. Thermal resistances for junction-to-case node and junction-to-board node are also defined to be used in the conduction-based thermal solver CST MPS. Copper heat spreader pins are used to connect the patch and chip to the ground plane. The thermal model for a unit cell structure is given in Fig. 9.24.

The designed 8 x 8 integrated antenna array (formed by the unit cells) is shown in Fig. 9.25. A list of all the EM and thermal design parameters (based on the values in Section 9.1.2) is provided in Table 9.19.

9.2.3. RESULTS AND DISCUSSION

In this part, electromagnetic and thermal simulation results are given in two separate sub-sections.

A. Thermal Simulation Results

First, t_{gnd} is varied between 0.05 mm and 2 mm by taking $L_{\max} = 5\lambda_0$. The maximum temperature in the 8 x 8 array ($T_{\max}$) (maximum junction temperature) is plotted in Fig. 9.26. It is seen that $T_{\max}$ decreases by about 4^oC when t_{gnd} increases from 0.05 mm to 2

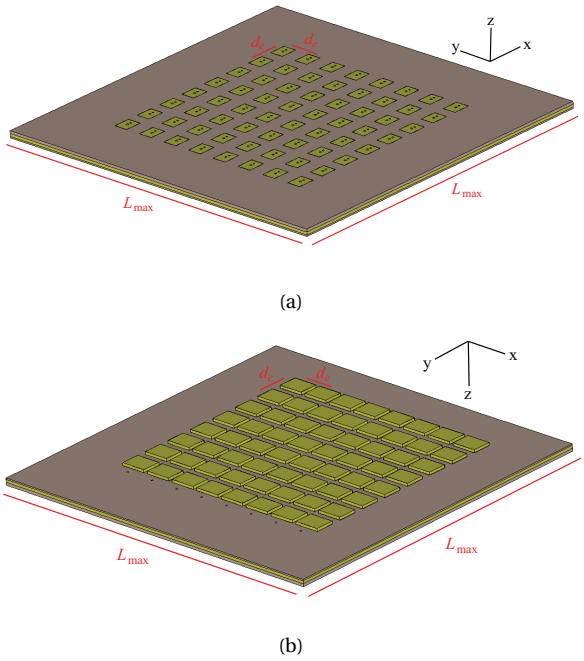


Figure 9.25: 8x8 integrated antenna array model, (a) patch side, (b) chip side.

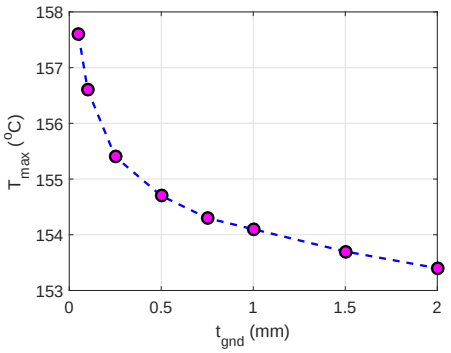


Figure 9.26: $T_{\max}$ for varying t_{gnd} when $L_{\max} = 5\lambda_0$.

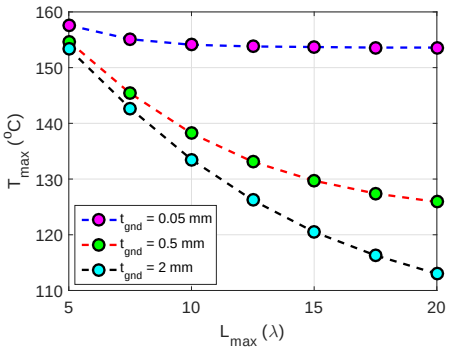


Figure 9.27: $T_{\max}$ at several t_{gnd} values for varying $L_{\max}$.

mm.

Next, $L_{\max}$ is varied between $5\lambda_0$ and $20\lambda_0$ and $T_{\max}$ is plotted in Fig. 9.27 for several t_{gnd} values. It is seen that for the thinnest ground plane (t_{gnd} = 0.05 mm), increasing $L_{\max}$ does not help decrease $T_{\max}$ since heat gets trapped around the chip and creates a hot-spot. On the other hand, for thicker ground planes, the heat generated at the chip can be spread to the sides of the array, which is then removed from the array surface.

Table 9.19: List of design parameters used in CST MWS & MPS

Parameter definition	Symbol	Value
Center frequency	f_0	28.5 GHz
Relative permittivity of the substrates	ϵ_r	2.2
Patch width	W_p	2.50 mm
Patch length	L_p	3.20 mm
Feed line width	w_f	0.6 mm
Feed line length	l_f	3.1 mm
Chip block width	w_c	3.5 mm
Chip block length	l_c	4.5 mm
Patch substrate thickness	t_{sp}	0.508 mm
Chip substrate thickness	t_{sc}	0.254 mm
Heat sink plate thickness	t_{hs}	0.5 mm
Chip block height	t_c	0.5 mm
Ground plane thickness	t_{gnd}	0.05 mm to 2 mm
Feeding pin diameter	d_p	0.32 mm
Feeding hole diameter	d_h	1.2 mm
Heat spreader diameter	d_s	0.4 mm
Feeding pin offset	d_f	0.4 mm
Chip offset relative to the patch	d_c	0.75 mm
Inter-element spacing	d_e	$0.5\lambda_0$
Heat sink plate offset	d_{gap}	0.5 mm
Dissipated power per chip	P_h	2 W
H.T.C of the heat sink	h_{hs}	1000 W/m^{2}K
H.T.C of air	h_{air}	10 W/m^{2}K
Junction-to-case thermal resistance	θ_{JCtop}	10 W/K
Junction-to-board thermal resistance	θ_{JB}	15 W/K
Maximum aperture edge length	$L_{\max}$	$5\lambda_0$ to $20\lambda_0$
Ambient temperature	T_{amb}	25^{o}C

Table 9.20: Effective substrate thermal conductivities (k_t) in lateral (x, y) and vertical (z) directions with varying t_{gnd}

t_{gnd} (mm)	k_{tx} (W/mK)	k_{ty} (W/mK)	k_{tz} (W/mK)
0.05	24.88	24.88	0.21
0.1	46.70	46.70	0.23
0.25	99.21	99.21	0.27
0.5	159.00	159.00	0.33
0.75	199.01	199.01	0.40
1	227.67	227.67	0.46
1.5	265.98	265.98	0.59
2	290.42	290.42	0.72

Since surface area increases with $L_{\max}$, the best cooling performance is achieved when $t_{gnd} = 2$ mm and $L_{\max} = 20\lambda_0$. The increase in lateral (k_{tx}, k_{ty}) and vertical (k_{tz}) effective thermal conductivities with t_{gnd} is given in Table 9.20. The values are obtained using the formulas given in [205]. Note that the conductivities will be slightly larger in the model compared to the theoretical values due to the heat spreader vias.

Sample temperature distributions (zoomed-in to center) for $t_{gnd} = 0.05$ mm & $L_{\max} = 5\lambda_0$, $t_{gnd} = 2$ mm & $L_{\max} = 5\lambda_0$ and $t_{gnd} = 2$ mm & $L_{\max} = 20\lambda_0$ are shown in Fig. 9.28(a),

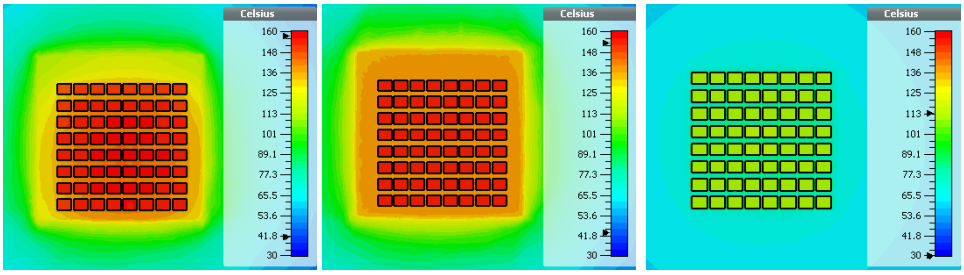
Fig. 9.28(b) and Fig. 9.28(c), respectively. It can be seen that in Fig. 9.28(b), the heat is more spread along the aperture which decreases $T_{\max}$ about 4^oC compared to Fig. 9.28(a). In Fig. 9.28(c), $T_{\max}$ becomes about 45^oC less since the aperture size is much larger, where the heat can be naturally removed.

B. Electromagnetic Simulation Results

In this part, first, the EM performances of the two arrays with t_{gnd} = 0.05 mm & $L_{\max}$ = $5\lambda_0$ and t_{gnd} = 2 mm & $L_{\max}$ = $20\lambda_0$ are compared in terms of active S-parameters to see if it is possible to achieve antenna matching with the design parameters given in Table 9.19. The results are given in Fig. 9.30.

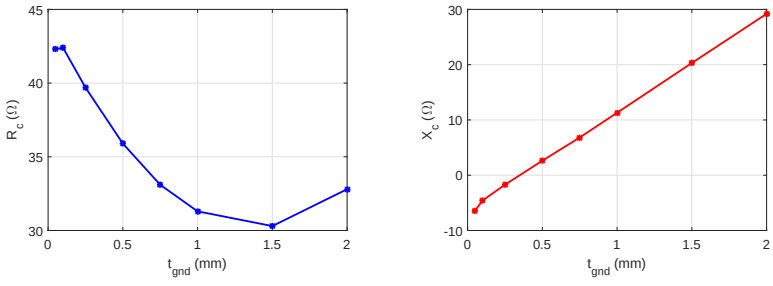
From Fig. 9.30, it can be inferred that due to the change in mutual coupling, the matching of the array becomes much worse when the thick ground plane is used. This effect can be compensated by changing the dimensions of the patch. Modifying the patch dimensions to L_p = 3.26 mm, W_p = 1.80 mm allows us to achieve good matching with no significant impact on the cooling performance (see Table 9.21).

Next, the full EM performance of the designed 8 x 8 arrays with the worst and the



(a) (b) (c)

Figure 9.28: Sample temperature distributions (zoomed-in to center) with passive CPU cooling for the 8x8 regular square array along the horizontal cut in the middle of the conducting blocks representing the junctions, (a) t_{gnd} = 0.05 mm & $L_{\max}$ = $5\lambda_0$, (b) t_{gnd} = 2 mm & $L_{\max}$ = $5\lambda_0$, (c) t_{gnd} = 2 mm & $L_{\max}$ = $20\lambda_0$.



(a) (b)

Figure 9.29: Active impedance of a center element at f_0 with varying t_{gnd} for $L_{\max}$ = $5\lambda_0$, L_p = 3.20 mm, W_p = 2.50 mm, (a) resistance, (b) reactance.

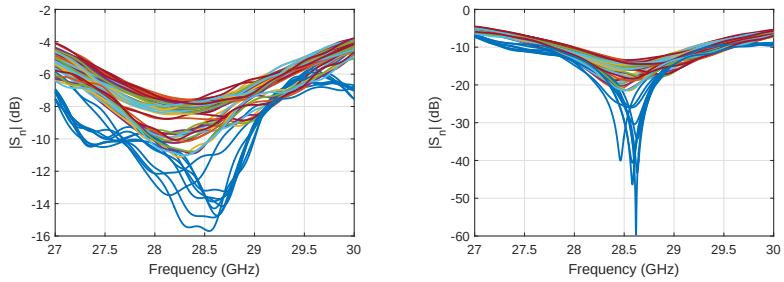


Figure 9.30: Active S-parameters with t_{gnd} = 2 mm & $L_{\max} = 20\lambda_0$, (a) L_p = 3.20 mm, W_p = 2.50 mm, (b) L_p = 3.26 mm, W_p = 1.80 mm.

Table 9.21: Patch dimension effect on $T_{\max}$ with t_{gnd} = 2 mm for varying $L_{\max}$

$L_{\max}$ (λ_0)	$T_{\max}$ (^{o}C) (L_p = 3.20 mm, W_p = 2.50 mm)	$T_{\max}$ (^{o}C) (L_p = 3.26 mm, W_p = 1.80 mm)
5	153.4	153.7
7.5	142.7	142.9
10	133.5	133.6
12.5	126.2	126.2
15	120.5	120.6
17.5	116.3	116.4
20	113.0	113.0

best cooling performance are given in terms of the active reflection coefficients, total efficiency, mutual couplings for a center element and radiation pattern. The broadside beam and scanned (towards θ_s = -π/9, $\phi_s = \pi/4$) beam results of the array with the lowest t_{gnd} and $L_{\max}$ (thus the worst thermal performance) are given in Fig. 9.31 and Fig. 9.33, respectively. Fig. 9.32 and Fig. 9.33 present the EM performance of the array with the largest t_{gnd} and $L_{\max}$ (thus the best thermal performance) for the broadside and scanned (towards θ_s = -π/9, $\phi_s = \pi/4$) beam, respectively.

Overall, it can be seen that similar EM performance can be achieved with the proposed thick ground plane and extended aperture size antenna as compared to the thin ground plane, standard aperture size counterpart. The advantage of increasing t_{gnd} and $L_{\max}$ appears in the cooling performance due to increased effective thermal conductivity of the substrate and substrate-to-air surface area.

Note that due to selecting an electrically thin, low-permittivity substrate in the array designs, the surface wave losses are not significant and radiation efficiency is very high as seen in Figs. 9.31-9.33. This remains valid for all the preliminary antenna array designs used for demonstration purposes throughout the thesis. Also, in the case of the SIW-based arrays given in Section 5.2, the problem of surface waves is efficiently mitigated. However, in other designs for which the substrate is electrically thicker, multiple surface wave modes may exist, which reduces the radiation efficiency and causes per-

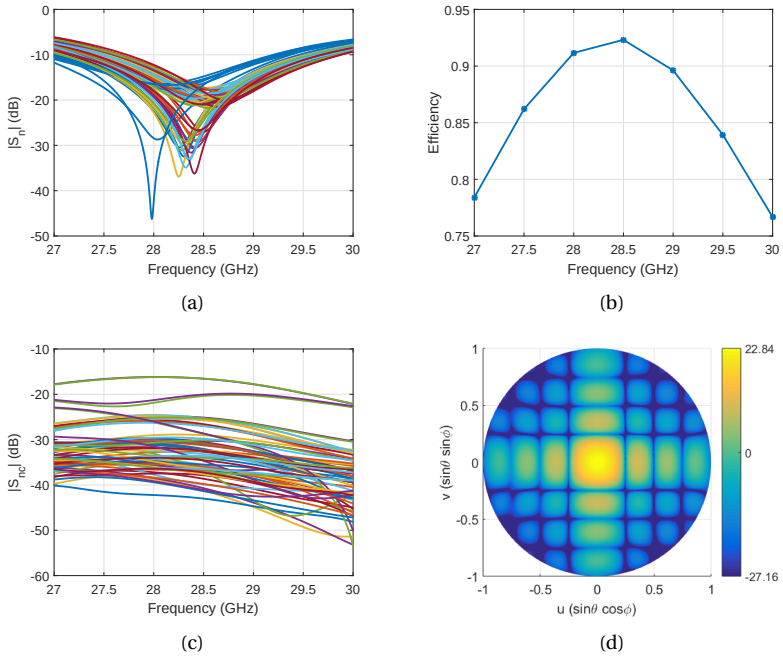


Figure 9.31: EM performance results for the 8x8 broadside array with t_{gnd} = 0.05 mm & $L_{\max} = 5\lambda_0$, L_p = 3.20 mm, W_p = 2.50 mm, (a) active S-parameters, (b) total efficiency, (c) mutual couplings of a center element, (d) realized gain.

turbation in the pattern due to the uncontrolled diffraction of the surface waves at the edges. In that case, using cavity-backed patches or applying typical methods such as using a metamaterial fence or a defected ground plane might be needed. The sparse array topologies introduced in Section 6.2 are also helpful in such scenarios as they allow extra spacing between the elements for inserting additional structures to suppress the surface wave propagation.

9

9.2.4. CONCLUSION

In this section, an active integrated, passively cooled patch antenna array with a thick ground plane and an extended aperture size has been proposed for 5G base stations with fully-digital beamforming capabilities. A double sided PCB has been designed with patch antennas on one side and transceiver chips on the other side. For the EM simulations, a microstrip feeding line with a vertical pin transition to the patch has been realized. The thermal simulations have been performed following the two-resistor model and assuming that there is a fanless CPU cooler sinking the heat from the chips.

An 8 x 8 array of patches and chips has been simulated using an electromagnetic and a thermal solver. From the thermal aspect, it has been found out that increasing the ground plane thickness leads to an increase in the effective thermal conductivity of the substrate, which helps to spread the hot-spots that are formed in the vicinity of the chip

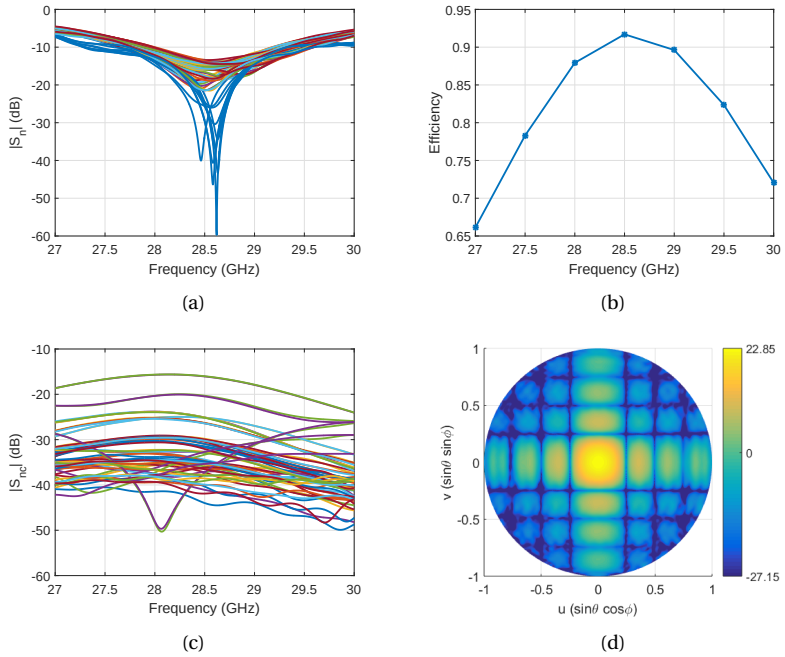


Figure 9.32: EM performance results for the 8x8 broadside beam array with t_{gnd} = 2 mm & $L_{\max} = 20\lambda_0$, L_p = 3.26 mm, W_p = 1.80 mm, (a) active S-parameters, (b) total efficiency, (c) mutual couplings of a center element, (d) realized gain.

to a larger aperture area. Besides, increasing this aperture size contributes to the cooling since the surface area that is in contact with air gets larger. In conclusion, as compared to the standard counterparts, much better cooling performance can be achieved using relatively thick ground planes with extended aperture sizes, with no significant effect on the electromagnetic properties.

9.3. CONCLUSIONS

In Chapter 8, it was seen that fully passive cooling of 5G base station antennas need bulky and heavy CPU heat sinks which might not be applicable in certain deployment scenarios. ***To relax the requirements on the CPU coolers and miniaturize the base station units***, it is necessary to form extra cooling paths in the design. It was seen previously in Chapter 7 that the antenna itself can be used as a heat sink to provide an additional cooling facility. However, such an approach results in complex 3-D antenna designs that are impractical for the first phase of the 5G systems. Therefore, in this chapter, we have proposed ***two novel ways for passive cooling enhancement that are compatible with the planar arrays: a layout optimization based approach and a substrate thermal conductivity and size increase based approach***.

For the layout optimization based approach, considering the main interest of the researchers from the field of thermal engineering, first, optimization of heat source dis-

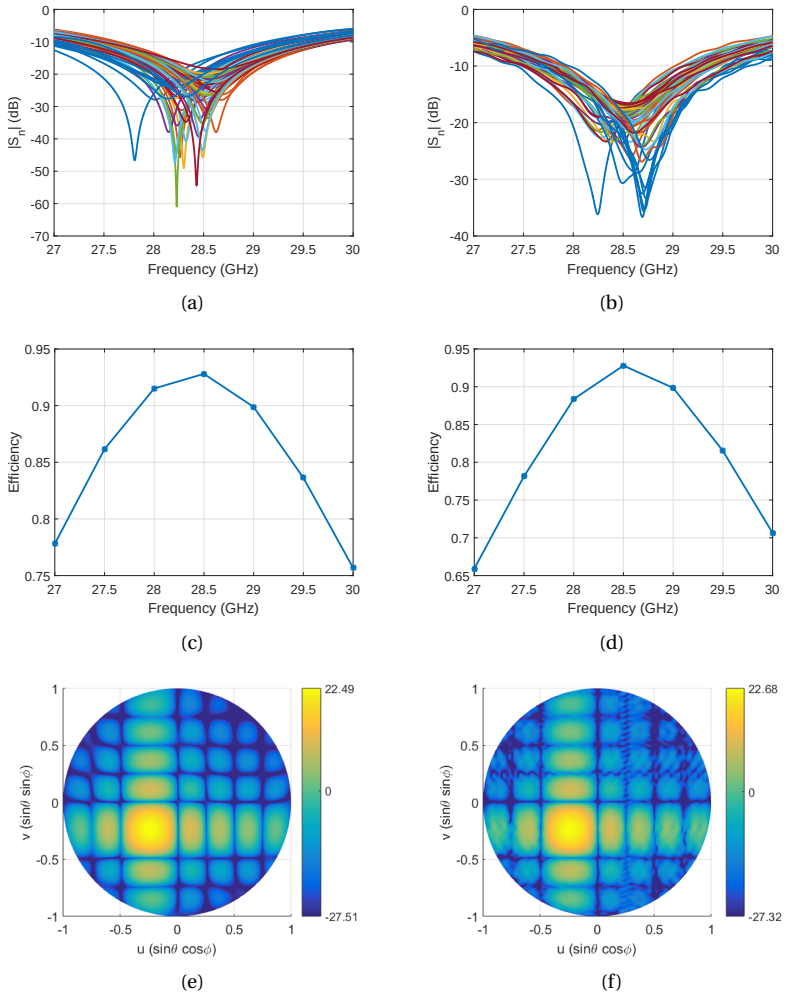


Figure 9.33: EM performance results for the 8x8 scanned beam (θ_s = -π/9, ϕ_s = π/4) array with t_{gnd} = 0.05 mm & $L_{\max} = 5\lambda_0$, L_p = 3.20 mm, W_p = 2.50 mm, (a) active S-parameters, (c) total efficiency, (e) realized gain; and with t_{gnd} = 2 mm & $L_{\max} = 20\lambda_0$, L_p = 3.26 mm, W_p = 1.80 mm, (b) active S-parameters, (d) total efficiency, (f) realized gain.

tribution in two dimensional heat conduction for electronic cooling problem has been considered. Then, the analogy between the conduction dominated heat transfer in thermal engineering and cooling of phased arrays in antenna system engineering has been investigated. It has been inferred that the problem in the thermal-domain is similar to the thermal management challenge in the traditional AESAs and the optimization algorithm can be straightforwardly modified to address the dual (electromagnetic and thermal) performance criteria in the traditional phased arrays. However, as previously discussed in Section 7.1, the focus in 5G is not on the traditional AESAs, but on the planar counterparts. Since in planar AESAs the dominating heat transfer mechanism becomes

convection, the use of layout optimization in 3-D convection dominated scenarios has been discussed next.

It has been explained that **layout sparsity for planar phased array cooling is most useful in the case of having relatively low thermal conductivity antenna and chip boards**, such that the ICs act as hot-spots in the design. In such a situation, it is necessary to increase the surface area around each IC (or antenna element). Therefore, heat removal capabilities and radiation performances of several sparse antenna array topologies have been studied in Section 9.1.2 where **both electromagnetic and thermal aspects have been jointly considered for the first time in array layout optimization** and a **novel connection between the extent of layout sparsity and thermal management** has been presented. The **aperiodic sparse optimized (with respect to the required field-of-view) topologies have been proposed as a good candidates** among all the sparse topologies investigated, which has provided **up to 15% reduction in the maximal array temperature**.

Finally, in Section 9.2, the condition of the low thermal conductivity board has been repealed by using thick ground planes in the planar AESA designs. The thermal and electromagnetic effects of **varying the ground plane thickness and aperture size of the densely- (0.5λ-) spaced planar phased arrays have been investigated for the first time**. The simulation results have shown that as compared to the standard counterparts, **much better cooling performance (up to 30% less maximal temperature) can be achieved using relatively thick ground planes with extended aperture sizes**, with **no significant effect on the electromagnetic properties**.

The selection between the two proposed techniques depends on the requirements and limitations in terms of the radiation pattern (beamwidth, side lobes etc.), board thermal conductivity, minimum inter-element spacing, maximum array size and fabrication cost.

- - - - - - - - - - - - - - - - - - -

The thermal studies in Part III has resulted in a collaboration with the Celestia Technologies Group BV for a proposal [P-1] for the H2020 call for SPACE related to phased array technologies matching with the SPACE-29-TEC-2020 Satellite Communication Technologies topic. The planned contribution is to analyze the cooling technology options and thermal aspects for the multibeam ground station prototype to be developed that can provide dual Earth Observation and Communications services with several satellites simultaneously.

10

CONCLUSIONS AND FUTURE WORK

The last chapter of this thesis is a concluding chapter which summarizes the research findings that have been discussed throughout the previous chapters and proposes possible future research directions.

10.1. CONCLUSIONS

10.1.1. NOVELTIES, OVERALL IMPORTANCE AND ADDED VALUES

In this thesis an innovative approach to 5G system has been proposed and structurally investigated: aperiodic arrays. The focus of the thesis is on getting generic system solutions for 5G which will be valid regardless of the antenna element type, rather than the design of a particular array. For this aim, novel methods/ideas for irregular/sparse array topology optimization, beamforming and heat management have been proposed. In particular,

- For the first time, the array topology problem has been formulated in an advance way, which is a system-based optimization (that considers different 5G base station deployment scenarios and mm-wave link budgets) to achieve a certain system performance within the angular sector statistically;
- A new iterative element position perturbation algorithm has been introduced, which is superior to all the existing array synthesis methods in terms of the 5G compatibility (better SLL suppression in wide angle scanning with uniform excitation amplitudes for optimum power efficiency);
- An original minimum element spacing control feature has been proposed in irregular array synthesis and the layout sparsity (i.e. large inter-element distance criterion) is used as a passive heat removal mechanism for the first time;
- The layout irregularity has been combined with the sequential rotation technique, for the first time, to achieve modularity and ease the design, fabrication and calibration requirements of integrated 5G arrays.

The impact of these novel ideas can be summarized as following. By keeping a similar system complexity as a referenced fully populated equally spaced array in a rectangular topology, an aperiodic sparse array can (simultaneously)

- improve total SINR by preserving the gain while suppressing the side lobes within the same scanning area, whose extent depends on the number of elements and the available aperture size;
- decrease the maximal temperature of the beamforming chips, whose extent depends on the sparsity of the array and the thermal properties of the ICs/boards;
- provide more space for placing the electronics and for circuit routing.

Consequently, the quasi-modular aperiodic array synthesis method has been demonstrated to be the best design approach for the near-future mm-wave phased arrays (at the current technology level) in terms of the SLL suppression capability, design and fabrication complexity, power efficiency and thermal handling capacity.

Furthermore, the system studies, which have been performed to formulate the demands to the 5G front-end beamforming performance, have resulted in the following findings:

- Forming a single-lobe, directive and low side lobe beam towards the strongest multipath of a user yields the most optimal system performance in terms of the frequency re-use and beamforming complexity.

- With the current level of PA efficiency and computational capabilities, fully-digital beamforming is a too costly tool for the mm-wave multibeam base station phased arrays. The proposed novel hybrid multiple beamforming techniques in this thesis provide a much better performance/cost trade-off.

- Using the proposed phase-only synthesis technique, the phase shifters used for beam steering in the currently proposed conventional 5G phased arrays can be exploited to obtain low sidelobes and narrow or wide nulls while scanning the beam inside a wide angular sector.

- The thermal problem in the active mm-wave arrays occurs at the chip level, which makes the integration of a large heatsink at the chip side necessary in the case of a passive-only cooling. However, the size (weight, cost) of the cooler can be reduced by integrating a heat dissipation functionality into the antenna array design (via layout sparsity or thick ground planes).

As 5G front-end and communication system design is a very complex multidimensional optimization problem, there is no single system solution which will be optimal from all points of view and in all scenarios. That is why the diversity of the proposed solutions has an additional value from the system design point of view. Due to the same reasons, the applicability of each technical solution proposed has been carefully discussed in the thesis and placed in the system perspective. The suggested array topologies and beamforming strategies have been verified via high-precision full-wave simulations using a software tool (CST Studio Suite) which is considered as the industrial standard. For a limited amount of designs (cosecant antennas), these simulations have been verified with experiments as well. A good agreement has been achieved. The work on the analysis of the system impact of the measured patterns of the fabricated cosecant array with respect to the ideal/simulated ones is in progress.

Following the structure of the thesis, the major results are listed next under the proposed three main aspects in 5G antenna design: communication system-, electromagnetic- and thermal-perspectives. Moreover, using the complementary connections among the three aspects established throughout the thesis, novel interdisciplinary conclusions are given under the “joint” perspectives.

10.1.2. COMMUNICATION SYSTEM PERSPECTIVE

The communication system aspects have been covered in Chapter 2 and Chapter 3, forming Part I. In Chapter 2, a multiuser LoS SDMA system model has been developed and integrated with a meaningful 5G mm-wave link budget to assess the statistical QoS in terms of the user SINRs. The study has been extended to a NLoS urban outdoor scenario to see the impacts of multipath and investigate the possibilities to make use of it. In both chapters, the system trade-offs have been stated to highlight the strong coupling between the implementation complexity of beamforming, the QoS and the computa-

tional cost of precoding. The research performed in Part I has led to the following major observations and results:

- *To obtain the best attainable SINR performance, full flexibility in the formation of the multiple beams is needed, which comes at the expense of complex & costly beamforming architectures (with large heat generation!) that require computationally heavy precoding algorithms.* To handle multiple data streams and generate multiple beams within a wide-angle sector from one array with ideally zero interference between them, fully flexible multibeam architectures have to be used. Such topologies (either digital or analog) are too difficult to realize due to their extremely complex design and heavy processing burden. The system discussions in Section 2.3 have shown that it is crucial in 5G to simplify the beamforming architecture, reduce the signal processing burden and find effective ways for the thermal management at the base stations, which has provided the basis of this thesis.

- *A certain minimum angular spacing among the simultaneous co-frequency users has to be satisfied in coordination with the MAC layer user scheduling.* The system simulations in Section 2.4 has shown that the currently proposed random user selection technique significantly deteriorates the QoS in 5G SDMA. Therefore, **a new simultaneous, co-frequency user selection technique has been proposed for which no other co-channel users are allowed to enter into a user's main lobe**. The simulation results has demonstrated that the proposed scheduling scheme guarantees high SNR values in ZF and low interference levels in CB precoding.

- *ZF precoding yields the best statistical SINR performance in the 5G's interference-dominated scenarios. However, it is not robust against the errors in the estimated channel matrix or quantization in the excitation weights.* ZF precoding can completely cancel-out the inter-user interference with careful user selection under ideal system conditions. However, the system non-idealities (inaccurate channel feedback, quantization errors etc.) might lead to filled nulls and eventually, reduced QoS. In Section 2.4, the negative impact of system imperfections on the performance of ZF precoding has been quantified and **it has been demonstrated that even CB with a smart array layout (that suppresses the in-sector SLLs) could yield competitive QoS with the ZF precoding**.

- *Under ideal system conditions, reduced complexity ZF precoding algorithms provide a good compromise between the digital signal processing complexity and the QoS.* Full ZF with $(K-1)$ nulls for K users causes a large processing burden due to the large size of the channel matrix that is to be inverted. CB, on the other hand, has the least processing complexity, but it generally comes at the expense of reduced QoS. Considering this, two different reduced complexity ZF algorithms (based on the number of zeroes, static or adaptive) have been investigated in Section 2.4 as a trade-off between the computational complexity and QoS. Through the simulations, it has been shown that **the reduced complexity ZF can provide an order of magnitude less complexity than the ideal ZF, while achieving better QoS than the CB precoding under ideal system conditions**.

- *There is a single dominant path in most of the cases in a 5G urban outdoor environment at mm-waves.* The ray tracing simulations in Section 3.3 have demonstrated that in NLoS scenarios, receiving a signal from an indirect path (via reflection or diffraction) could lead to a few tens of dB reduction in the ray power. Therefore, **in most of the cases, the main beam direction of a user can be determined from the characteristics of the strongest ray connecting the user to the base station**.

- *Even in the presence of several equally-strong multipaths, a single-lobe beam directed towards the strongest multipath provides the best system performance.* The discussions in Chapter 3 have shown that one could use multiple simultaneous single lobed beams and ports per user and recombine signals coherently after time synchronization of each. That should provide the optimal gain, but it limits greatly same frequency re-use by others and increases the beamforming complexity. Therefore, **if it satisfies the minimum SNR threshold, forming a single beam with one lobe to the strongest ray seems preferable**.

10.1.3. ELECTROMAGNETISM PERSPECTIVE

The system analysis in Part I has revealed the demand in 5G for reduction in the complexity of beamforming in terms of both hardware design and computation effort. Besides, the statistical studies have clearly demonstrated the importance of effective interference suppression on the QoS improvement. Therefore, in Part II (Chapter 4, Chapter 5 and Chapter 6), the electromagnetism driven aspects have been covered. In Chapter 4, the currently proposed multibeam generation architectures have been revisited and their drawbacks have been pointed out. Having seen the potential of hybrid schemes in 5G, two subarray based hybrid beamforming concepts have been introduced in Chapter 5, which are based on (i) an array of active multiport subarrays and (ii) an array of cosecant subarrays. Chapter 6 has focused on the co-channel interference minimization, in which two novel array optimization strategies have been introduced: (i) a phase-only control technique and (ii) a position-only control technique. The research performed in Part II has led to the following major results:

- *The narrow angular sector coverage or high implementation complexity of the existing hybrid beamforming techniques can be improved by using active multiport subarrays.* Hybrid beamforming is currently proposed in 5G so as to reduce the number of RF chains, and thus system cost and complexity. In Section 4.2, two existing hybrid techniques, namely, array of phased subarrays and passive multiple fixed beam arrays, have been discussed. It has been demonstrated that for the array of phased subarrays, the angular coverage is narrow and grating lobes might occur in the sector. For the passive multiple fixed beam arrays, on the other hand, the array design becomes complex and bulky with high insertion loss, power inefficiency and parasitic radiation. Therefore, **a novel hybrid beamforming topology has been proposed** in Section 5.1. **The new scheme is based on relatively small-sized active subarrays (less lossy, more compact and power-efficient) with multiple Butler ports generating multiple fixed beams that are digitally controlled in the baseband**. The major advantage of the active multiport subarrays over the array of phased subarrays has been reported as the increased angular coverage via

the use of multiple Butler beams.

- *Cosecant subarray beamforming is efficient in serving multiple SDMA users in the case of a fairness-motivated LoS communication thanks to its low complexity and power equalization capability.* **An alternative hybrid beamforming architecture has been proposed** in Section 5.2. **The idea is based on forming a single analog beam in the elevation plane with a cosecant-square shape and controlling the beams digitally in the azimuth plane**. Cosecant beamforming has been proven to be useful in clear LoS propagation scenarios that do not require a large amount of vertical scanning range (i.e. when the base station is on a tower at a certain height and transmits signals down towards the receiver units). It has been shown that using the cosecant beam, the power flux at the users is equalized which provides more fairness with predictable minimum SNR and the near-far problem is naturally prevented. **Several novel cosecant subarray design strategies have also been proposed** in Section 5.2.2.

- *Once the users' angular positions are known, effective side lobe suppression and null formation can be achieved via low complexity phase-only excitation control, which comes at the expense of array gain reduction.* The ZF precoding relies on accurate channel information; however, by making use only of users' angular location information, it is also possible to create (ZF-like) orthogonal beams. An efficient way to do this is to make use of the phase shifters used for beam steering in phased arrays also to obtain low SLLs and nulls. In Section 6.1, **a novel phase-only control has been introduced, which yields a relatively simpler feed network and higher power efficiency as compared to the amplitude-tapered arrays**. Through comparative study cases, it has been demonstrated that the proposed approach outperforms the existing methods when radiation pattern nulling in certain sectors is required and achieves at least comparative results in cases without nulling.

- *Aperiodic arrays combined with multibeam optimization suppress the side lobes remarkably, while keeping optimal power efficiency.* Apart from controlling the excitation amplitudes and phases, element positions can be modified to suppress the SLLs. In fact, by keeping uniform amplitudes and linear progressive phase shifts, position-only optimization (i.e. space tapering) provides the optimal power efficiency. In Section 6.2, **a novel position perturbation technique has been proposed** and the method's superior qualifications regarding the 5G SDMA system compatibility has been listed. It has been demonstrated that **the proposed technique is the first algorithm that can efficiently combine uniform amplitude excitation, element spacing constraint and simultaneous multi-beam optimization in a single optimization routine**.

10.1.4. THERMAL PERSPECTIVE

In the sytem studies, it has also been affirmed that one of the strongest challenges in mm-wave 5G base station antennas is to deal with the extreme heat generation which might cause reliability issues and decrease the chip lifetime. Therefore, in Part III (Chapter 7, Chapter 8 and Chapter 9), the thermal driven aspects of 5G antennas have been

examined. In Chapter 7, an overview of the cooling strategies for the active integrated phased arrays has been given. Considering the energy efficiency, cost and maintenance requirements of 5G base stations, in Chapter 8, a new passive cooling strategy, based on CPU heatsinks, has been proposed and its performance has been verified by simulations (using an efficient thermal system model) and by experiments (using temperature sensors). With the aim of reducing the size and weight of the heatsinks, Chapter 9 has focused on novel alternative ways for cooling enhancement that are based on (i) layout optimization & sparsity and (ii) increase in substrate thermal conductivity & dimensions. The research performed in Part III has led to the following major results:

- *Two-resistor thermal model provides a good initial guess of chip junction temperatures in active integrated antenna arrays.* The thermal system modeling in electronics design has been discussed in Section 8.2 and **the two-resistor CTM, standardized by the JEDEC Solid State Association, has been proposed to be used for investigating system-level designs and exploring what-if scenarios because of its simplicity and high computational efficiency**. Through experimental validations of the thermal simulations in Section 8.4, it has been demonstrated that the model is advisable to be applied in the initial integrated array design stage to simply obtain the first predictions on the chip temperatures.

- *Commercially available fanless CPU cooler heatsinks can be efficiently used for fully-passive thermal management in 5G base station antennas.* High-volume and low-cost communication market demands fully-passive thermal management with high efficiency and low maintenance in active planar phased arrays. For this purpose, in Section 8.3, **it has been proposed to make use of the commercial fanless CPU heatsinks whose baseplate is attached to the beamforming ICs on the back side of a double-sided substrate**. The results in Section 8.3 and Section 8.4 have shown that such heatsinks can bring the temperature down to a safe value below the 125^oC limit.

- *Thermal problem of 5G antenna systems can be relaxed by using a sparse array layout in the case of having: (i) highly inefficient transceiver ICs, (ii) substrates with low thermal conductivity, (iii) heat sinks with low heat transfer coefficient.* **The effect of antenna element (or chip) layout sparsity on the array temperature has been investigated** in Section 9.1.2. The parametric thermal simulations in that section has demonstrated that **the increased surface area around each chip contributes to the cooling. The extent of contribution can reach up to 15% reduction in the maximum array temperature**, in the case of the above listed design criteria (i-iii). It has been demonstrated that this is due to the "hot-spot"-like behaviour of the chips for which the heat is trapped around each IC.

- *To ease the thermal problem in the case of densely-packed arrays, both effective thermal conductivity of the substrate and the substrate dimensions should be increased simultaneously.* **An alternative cooling enhancement technique has been proposed** in Section 9.2 **which is based on increasing the thermal conductivity and dimensions of the substrate**. The aim with this method is to spread the heat in the "hot-spots" to a larger surface that is in contact with the surrounding air. Since

the element positions are kept unchanged, the technique is also suitable for the densely packed (0.5λ-spaced) arrays. The thermal simulations in Section 9.2.3 has indicated that **the proposed method can decrease the maximum array temperature up to 30%, which can help reduce the size of the bulky CPU heatsink and miniaturize the wireless unit**.

10.1.5. JOINT PERSPECTIVES

At this point, it is critical to notice that the results from the communication system-, electromagnetism-, and thermal-related perspectives are closely connected to each other. For the overall 5G network optimality, joint perspectives combining these three different areas of expertise should be considered as well. Therefore, this section reiterates the various major interdisciplinary conclusions drawn throughout the thesis. To better highlight the multidisciplinary nature of the outcomes, the concluding statements given below are categorized by the acronyms (CS for Communication System-, EM for ElectroMagnetism-, and TH for THermal-driven aspects) placed in front of the items.

- ***(CS&EM) Irregular, position-optimized arrays with minimized SLL achieve better QoS than the currently proposed regular arrays with square layouts. Using irregular arrays, computational complexity in beamforming can be greatly reduced.*** The system studies of the optimized irregular arrays have been performed in Section 6.2.3. A novel relation between the SLL and the QoS has been established. It has been shown that the suppressed SLLs of the space-tapered arrays with CB could provide similar statistical QoS as compared to the regular arrays using ZF, especially under non-ideal system conditions. Thus, compared to their regular counterparts, the irregular arrays achieve better QoS and potentially help decrease the computational burden of beamforming.

- ***(CS&EM) The complex fabrication, circuit routing and calibration requirements of the fully aperiodic arrays can be relaxed with the help of quasi-modular arrays having a rotational symmetry.*** In Section 6.2.3.3, a rotational symmetry has been enforced in the proposed layout optimization algorithm in order to form quasi-modular arrays that consist of rotationally repeated irregular subarrays of only a few different types. In this way, instead of the fully-aperiodic array, only 3-4 unique subarray clusters have been optimized and repeated rotationally. The synthesized quasi-modular arrays will potentially simplify the design/fabrication complexity and the calibration demands. The advantage of modularity comes at the expense of increased side lobes as compared to SLLs of the fully-aperiodic arrays. However, the quasi-modular topologies still perform better than the periodic arrays in terms of the inter-beam interference suppression and the statistical QoS.

- ***(CS&EM&TH) The number of antenna elements can be reduced to some extent while keeping the same EIRP, same maximal SLL and a similar maximal temperature compared to a fully-populated array.*** In Section 8.5, it has been shown that as the element number increases, the maximum temperature in the array decreases and reaches to a safe level. Moreover, for the optimized space-tapered arrays, the maximum SLL reduces as more elements are added. This trend continues until an optimal element number, after which adding more elements does not

help reduce the SLL further because of the limited aperture size enforced by the baseplate of the CPU heatsink.

- ***(CS&EM&TH) Large-scale arrays have a big potential to address the radiation pattern and thermal requirements of 5G base stations.*** Contrary to decreasing the number of elements as in the previous item, adding more elements into the design allows us to achieve more gain and lower side lobes; thus much less power and less heat per element with more surface for cooling. The price to pay is the much-increased design, integration and processing complexity. Sample aperiodic large-scale arrays have been presented as a futuristic view in Appendix B.

- ***(CS&TH) The number of simultaneous co-frequency users is restricted by the thermal requirements.*** When the number of SDMA users is increased, the power output per chip should be increased as well in order to keep sufficient SNR at the users. This would result in more heat generation within each chip, which brings the system limitation in terms of the mamimum number of users that can be served using passive-only thermal control. In Section 8.5, the quantitative relation between the number of users and the maximum array temperature has been given for the CPU heatsink integrated active arrays.

- ***(CS&EM&TH) Layout sparsity for cooling must be combined with irregularity to prevent grating lobes and possible interference. Optimized sparse aperiodic topologies are good candidates.*** The thermal simulations in Section 9.1.2 have demonstrated that the inter-element (or inter-chip) spacing should be greater than 1λ to have an observable decrease in the array temperature. Regular layouts with such sparsity suffer from the grating lobes due to their symmetrical nature and decrease the frequency reuse & QoS. Therefore, different irregular array layouts have been invesitgated in Section 9.1.2, where the sunflower-like topology has been observed as a good option. The proposed space tapering algorithm in Section 6.2 can also potentially be used to improve the radiation pattern performance further, depending on the desired field-of-view.

- ***(EM&TH) Thick ground plane antennas can be used for cooling extension with no negative impact on the radiation characteristics.*** In Section 9.2, it has been demonstrated that the effective thermal conductivity of the substrate can be increased by using thick ground planes in the antenna design. Changing the ground plane thickness affects the matching performance of the antenna and decreases the radiation efficiency, unless taken care of. However, the electromagnetic simulations in Section 9.2 have revealed that the negative effect can simply be compensated by properly modifying the patch dimensions.

10.2. RECOMMENDATIONS FOR FUTURE RESEARCH

Several research lines that can be thought as a continuation or extension of the work presented in this thesis are listed below:

- ***Design of the feeding, IC routing and calibration for the proposed and further optimized irregular arrays/tiles.*** In order to have the fully or partially aperiodic

arrays operational, the practical fabrication and post-processing challenges have to be addressed. Within the scope of the thesis, we have only focused on the high-level synthesis of the optimal layouts and their system performance. Therefore, the realization and fabrication of the chip-integrated antenna arrays with proper calibration is an encouraging future research direction.

- ***Thinning-based layout optimization for adaptive dynamic activation of optimal site-specific subsets of an oversized array.*** The proposed layout optimization algorithm presented in Section 6.2 can be used to obtain unique and site-specific array topologies with minimal SLLs, depending on the angular cell sector definition. A more practical alternative would be to have a single-type oversized aperiodic array which could be used at every base station and whose optimal subsets are dynamically activated depending on the deployment site. Such an approach would require the inclusion of array thinning algorithms in the irregular array synthesis. The SLL suppression performance degradation due to the fixed array grid, fabrication cost and thermal aspects are to be investigated and compared with the currently proposed technique.

- ***Investigation of reconfigurable beamforming matrices and different (sub)array topologies in the array of active multiport subarrays.*** In Section 5.1, an array of active multiport subarrays has been proposed as an alternative hybrid beamforming technique which is superior to the existing hybrid methods in terms of the implementation complexity and angular coverage. To prevent the formation of the grating lobes within the communication sector, an intuitive, diamond-like array layout has been proposed. A possible future research line here could be to aim to further improve the radiation pattern performance of such arrays by introducing reconfigurability in the beamforming matrix and (optimized) aperiodicity within/among the linear/planar subarrays.

- ***Proof of simultaneous co-frequency multi-beam generation concepts.*** The fabricated cosecant subarray based integrated antennas presented in Section 5.2.2 and in Appendix A have realized a single beam at a time. However, for true multibeam systems, it is crucial to form several beams in a single time-frequency slot using a single array. As a future research agenda, it is worth validating the SDMA communication concept by exploiting the fabricated array sub-modules (antenna panel, chip boards, combiners/dividers). To achieve this, the user data streams should be generated, several chip boards per user stream should be used to feed the whole antenna panel, and different user signals at each antenna port must be added.

- ***More efficient implementation of the layout synthesis technique.*** In the position optimization technique given in Section 6.2, the initial topology plays an important role since the number of iterations (thus the computational burden) is reduced if the starting layout is chosen close to optimal. Therefore, it would be very useful to obtain (intuitively or via simple optimization techniques) a suitable initial array topology, possibly with some space tapering and coverage-type array shape optimization by proper element selection. It would also be possible to further optimize the code itselft to speed up the computation.

- ***Position and phase control for SLL suppression.*** The position and phase optimization methods proposed in Chapter 6 can be combined to have a phase distribution added to the position control, which would help suppress the side lobes further. One way to achieve this is to perform the phase optimization after the position optimization. A better alternative would be to jointly optimize the positions and phases, but such an approach requires a different optimization strategy than the proposed convex technique.

- ***Improved thermal modeling.*** Because of its low-complexity, the two-resistor CTM of JEDEC industry standards committee has been used in this thesis for the study of various what-if scenarios. Although the model provides a good initial guess of the array temperature at the steady state, the results are not very accurate and the transient phenomena are not considered. As a future work, the model can be extended to the DELPHI CTM [164] by adding more nodes into the network and even further to a dynamic CTM in which a capacitance value is associated with each node of the network. Unlike the static models, there is no universal methodology for the dynamic thermal modeling, but several techniques have been proposed in the literature [206; 207].

- ***Coupled multiphysics (electro-thermal) modeling and simulations.*** The presented thermal and electromagnetic simulations in this thesis are not coupled. In other words, the impacts of the temperature changes on the chip outputs, antenna matching and radiation characteristics are not taken into account. However, to increase the performance predictability, the 5G community has recently started to examine coupled electro-thermal multiphysics models [208–210]. Therefore, it is undoubtedly a promising future research direction to develop a coupled solver that can accurately show the temperature effects on the antenna performance.

- ***Extension of the system model towards a full network simulation.*** To assess and compare the statistical SINR performance for various antenna layouts and beamforming strategies, a primitive multiuser LoS communication system model has been developed in Section 2.2, which then has been extended to NLoS scenarios via intelligent ray tracing tools in Section 3.2. If a detailed network planning is demanded, a more enthusiastic approach is to perform complete end-to-end simulations for a particular antenna system. As a future work, the presented system model can be extended to a more comprehensive one that takes into account all types of losses, mutual coupling and edge effects, the modulation scheme, access strategy, beamforming in receivers, receiver mobility, multiple sectors/cells, inter-sector/inter-cell interference, cell load/traffic and so on.

APPENDIX

A. COSECANT SUBARRAYS: SIMULATED AND MEASURED RESULTS

This chapter presents the design parameters, simulations and measurements for the two different kinds of array of cosecant subarrays that were discussed in Section 5.2.2. The first part shows the performance of the vertically (along the waveguide axis) polarized AC-MPA array fed by SIW, while the second part focuses on the vertically polarized AC-MPA array fed by microstrip line.

Both antenna design projects were led by Dr. Jan Puskely within the project framework. The arrays were fabricated by CIBEL, France and populated/soldered in NXP Semiconductors. The measurements were performed in the TU Delft anechoic chamber DUCAT using available RF measurement equipment with the help of the TU Delft technician, ir. Pascal Aubry.

Parts of this chapter will be published as:
[J-2] J. Puskely, T. Mikulasek, Y. Aslan, A. Roederer and A. Yarovoy, "5G SIW Based Phased Antenna Array with Cosec Shape Pattern," *IEEE Antennas Wirel. Propag. Lett*, under review.
[J-8] J. Puskely, Y. Aslan, A. Roederer and A. Yarovoy, "Analog beamforming cosecant squared pattern active phased array antennas for 5G cellular systems," *IEEE Antennas Propag. Mag.*, 2020, in preparation.

SIW FED VERTICALLY POLARIZED ARRAY

Fig. 10.1(a) and Fig. 10.1(b) show the antenna array structure and the design parameters for the single element and the 12-element subarray, respectively. The corresponding (optimized) dimensions are listed in Tables 10.1 and 10.2.

A photo of the fabricated antenna panel consisting of 16 half-wavelength spaced subarrays is shown in Fig. 10.2.

The simulated and measured reflection coefficients of the SIW-fed array are provided

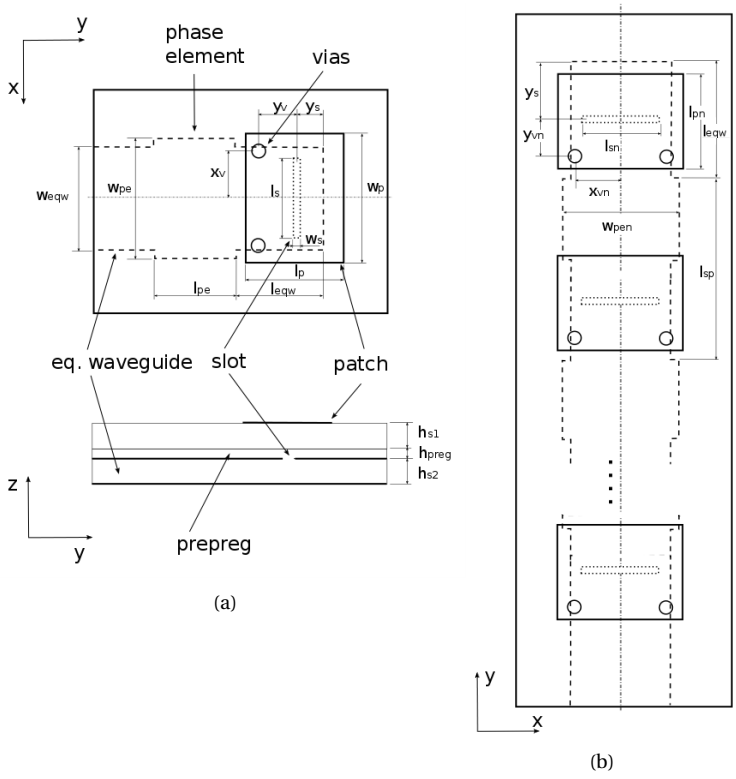


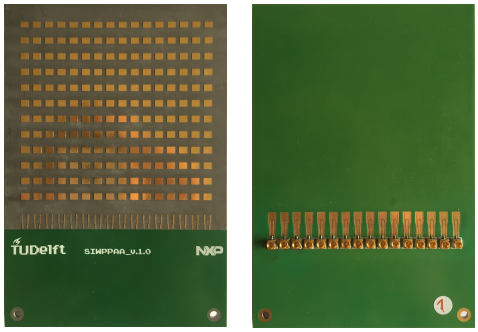
Figure 10.1: Geometry of the SIW-fed AC-MPA subarray.

Table 10.1: Dimensions of the SIW-fed antenna

Parameter	Value (in mm)	Parameter	Value (in mm)
l_{pe}	3.000	l_{sp}	6.799
l_{eqw}	5.299	d_l	1.660
w_{eqw}	2.866	d_{via}	0.300
y_s	3.399	h_{s1}	0.508
w_p	3.500	h_{s2}	0.635
w_s	0.200	h_{preg}	0.096

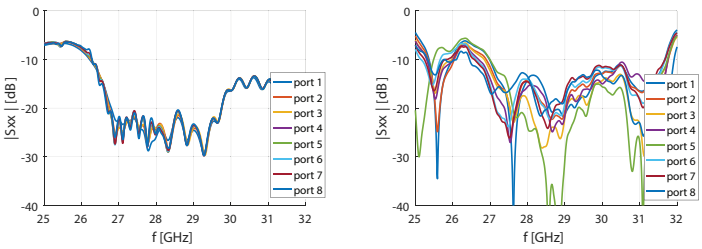
Table 10.2: Design parameters of the SIW-fed subarray

Element index	l_p	l_s	w_{pe}	x_v	y_v	Element index	l_p	l_s	w_{pe}	x_v	y_v
1	2.906	2.135	2.866	1.287	0.859	7	2.931	2.211	2.547	1.278	1.104
2	2.915	2.431	3.295	1.285	0.749	8	2.857	2.259	2.561	1.258	0.869
3	2.879	2.724	2.571	1.316	0.596	9	2.857	2.199	2.517	1.249	1.001
4	2.994	2.336	2.547	1.277	1.238	10	2.809	2.099	2.597	1.216	0.895
5	2.926	2.257	2.574	1.295	0.997	11	2.657	2.170	2.687	1.096	0.922
6	2.933	2.401	2.497	1.265	1.149	12	2.574	2.472	2.352	no via	no via



(a) (b)

Figure 10.2: Fabricated SIW-fed AC-MPA array panel: (a) top view, (b) bottom view.



(a) (b)

Figure 10.3: Reflection coefficient of the SIW-fed AC-MPA array: (a) simulated, (b) measured.

in Fig. 10.3. Note that the results are given only for half of the ports due to symmetry. Fig. 10.4 shows the simulated and measured active reflection coefficient at the central port for several scan angles in azimuth. The simulated and measured mutual coupling among the subarrays is given in Fig. 10.5.

As for the radiation patterns, Fig. 10.6 provides the excitation coefficients of the 12 elements in the subarray and the resulting approximated cosecant-squared pattern in elevation at 28.5 GHz. The measured vertical radiation patterns of the SIW-fed antenna array at varying frequencies are given in Fig. 10.7. It is worth to mention that due to the

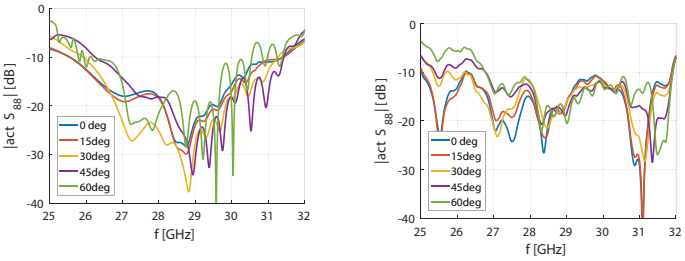


Figure 10.4: Active reflection coefficient of the SIW-fed AC-MPA array at the middle port (Port #8): (a) simulated (from an infinite array of subarrays), (b) measured.

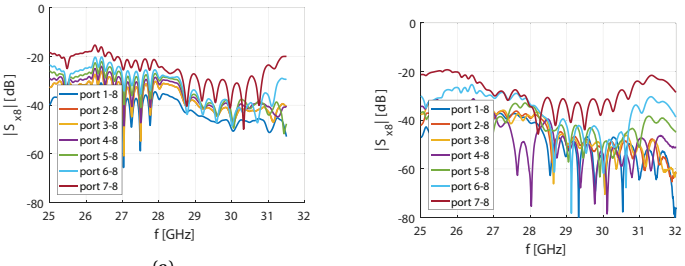


Figure 10.5: Mutual coupling between the SIW-fed subarrays: (a) simulated, (b) measured.

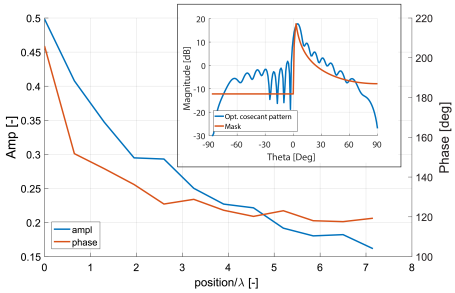


Figure 10.6: Excitation coefficients for the 12-element sub-array with the cosecant shaped beam and the coresponding radiation pattern in elevation with the defined mask.

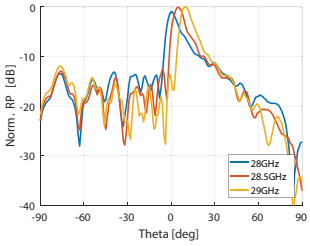


Figure 10.7: Measured vertical raditation pattern of the SIW-fed AC-MPA array panel at varying frequencies.

frequency-dependent nature of the design, beam squinting occurs, which unwantedly changes the main beam direction in elevation.

The simulated and measured embedded vertical radiation patterns (with no scannning in azimuth) at different frequencies are also provided in Fig. 10.8.

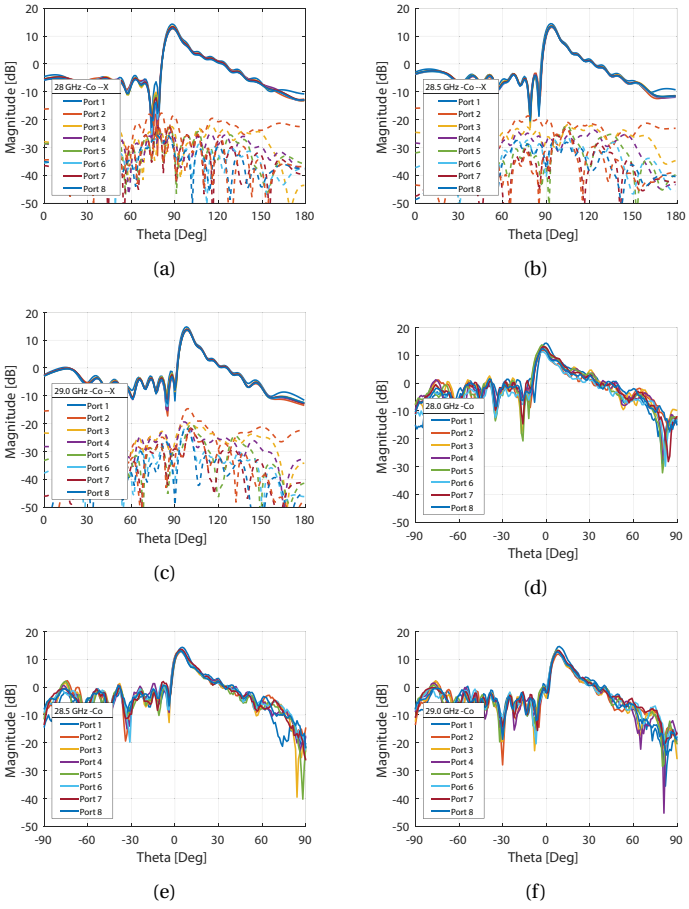
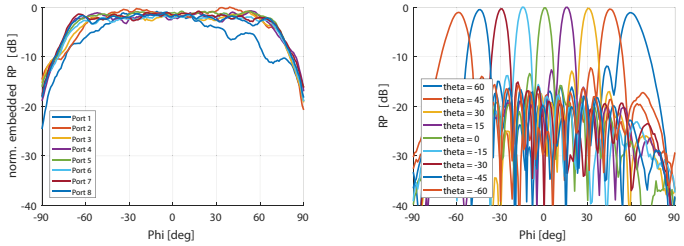


Figure 10.8: Vertical cuts of the embedded radiation patterns: (a) simulated at 28 GHz, (b) simulated at 28.5 GHz, (c) simulated at 29 GHz, (d) measured at 28 GHz, (e) measured at 28.5 GHz, (f) measured at 29 GHz.

Finally, the measured horizontal scan performance at 28 GHz is given in Fig. 10.9, which shows the embedded patterns of the subarrays and beam steering in azimuth.

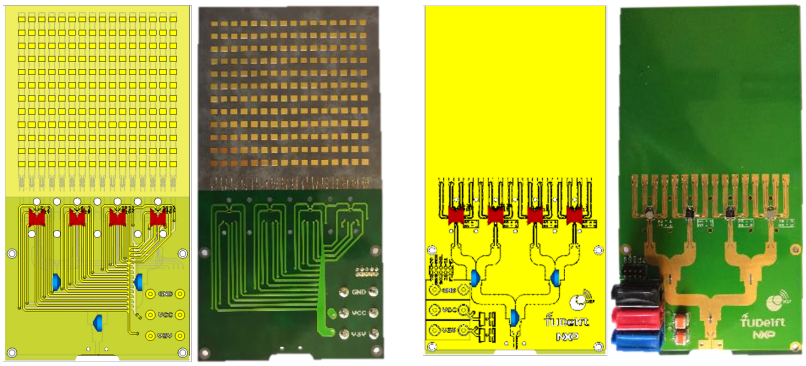
Next, the antenna panel was integrated with the NXP's analog beamforming chips in a complete antenna system including the array panel, chip board and power divider. Fig. 10.10 shows the fabricated complete SIW-fed AC-MPA antenna system. 4 chips are used to adjust the amplitude and phase at each of the 16 subarrays. The chips are controlled from a computer via USB connection using the National Instrument's NI-845x I2C/SPI hardware. The measurement setup in the anechoic chamber is shown in Fig. 10.11.

Fig. 10.12 provides the measured reflection coefficient of the SIW-based complete antenna system for several scan angles. The measured vertical radiation patterns at varying frequencies are given in Fig. 10.13. In the end, the azimuthal scanning performance of the array is shown at different frequencies in Fig. 10.14.



(a) (b)

Figure 10.9: Measured horizontal radiation pattern of the phased array at 28GHz: (a) embedded radiation patterns, (b) scanning performance.



(a) (b)

Figure 10.10: Fabricated complete SIW-fed AC-MPA antenna system (array panel + chip board + power divider): (a) top view, (b) bottom view.

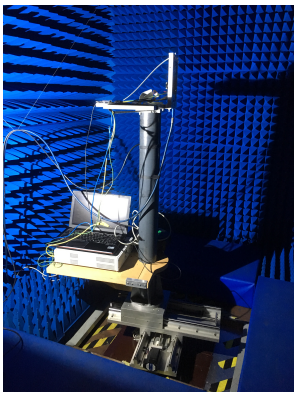


Figure 10.11: Antenna system measurement setup in DUCAT.

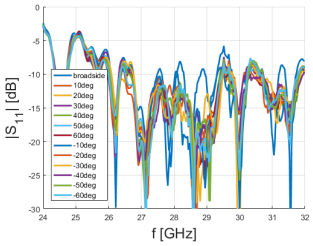


Figure 10.12: Measured reflection coefficient of the SIW based full antenna system.

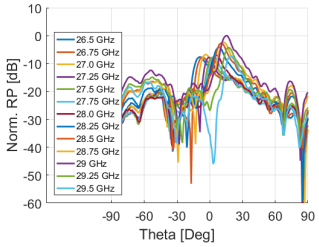
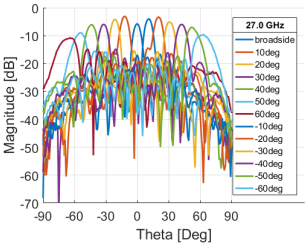
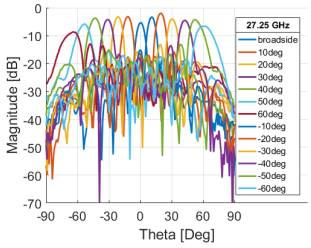


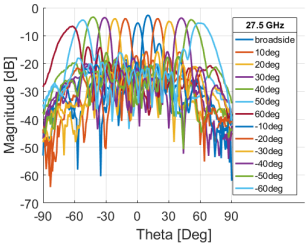
Figure 10.13: Measured vertical radiation pattern of the SIW based complete antenna system at varying frequencies.



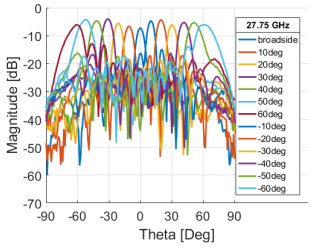
(a)



(b)



(c)



(d)

Figure 10.14: Measured horizontal scanning performance of the SIW based complete antenna system at various frequencies: (a) 27 GHz, (b) 27.25 GHz, (c) 27.5 GHz, (d) 27.75 GHz.

MICROSTRIP LINE FED VERTICALLY POLARIZED ARRAY

This section summarizes the performance of the microstrip line fed alternative of the cosecant subarray based 5G antennas.

The S-parameters were measured using a Vector Network Analyzer in a setup shown in Fig. 10.15 with the fabricated antenna panel. The measured reflection coefficient at each port for half of the array (the rest is not shown due to symmetry) is shown in Fig. 10.16. Fig. 10.17 provides the measured mutual coupling among the subarrays. The measured active reflection coefficient at the central subarray for different horizontal scan angles is given in Fig. 10.18.

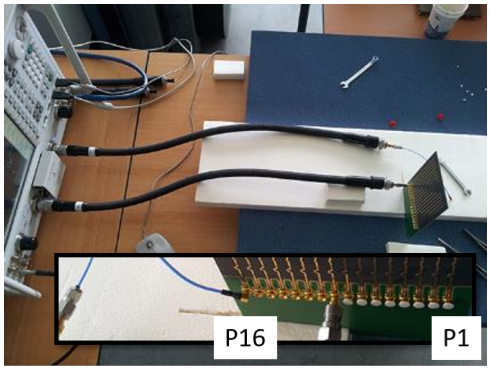


Figure 10.15: Reflection coefficient measurement setup.

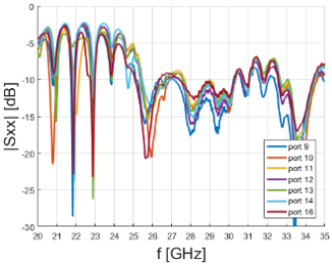


Figure 10.16: Measured reflection coefficient of the microstrip line fed AC-MPA array.

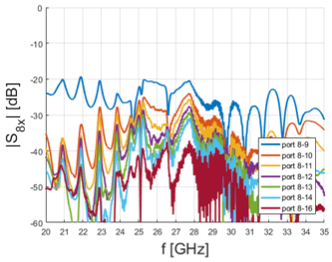


Figure 10.17: Measured mutual coupling of the microstrip line fed AC-MPA array.

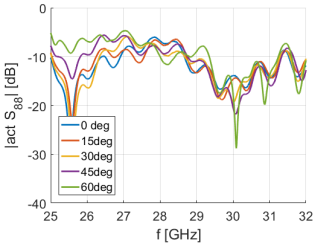


Figure 10.18: Measured active reflection coefficient of the middle subarray (port #8) in the microstrip line fed AC-MPA array.

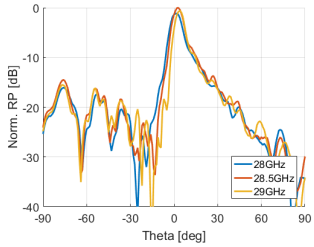


Figure 10.19: Measured vertical raditation pattern of the microstrip line fed AC-MPA array.

Regarding the radiation characteristics, the measured vertical raditation pattern of the microstrip line fed array of subarrays for varying frequency of operation is given in Fig. 10.19. The corresponding embedded vertical radiation patterns measured at varying frequencies for different subarrays are shown in Fig. 10.20. Finally, Fig. 10.21 provides the measured horizontal scanning performance of the microsrip line fed AC-MPA

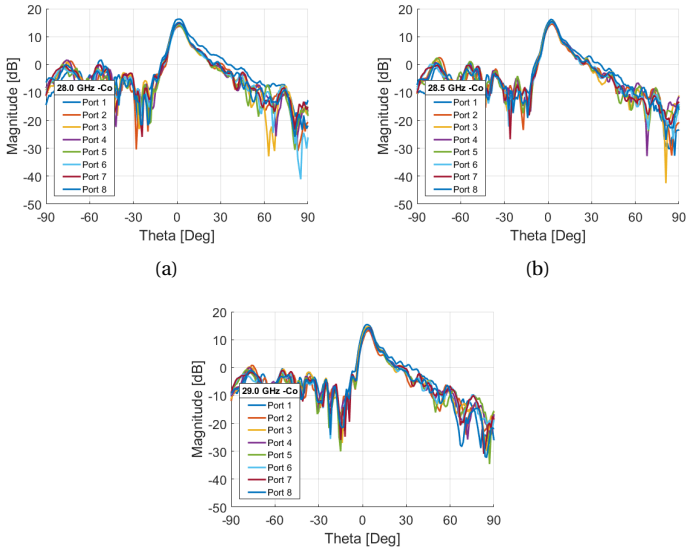


Figure 10.20: Measured vertical cuts of the embedded radiation patterns at varying frequencies: (a) 28 GHz, (b) 28.5 GHz, (c) 29 GHz.

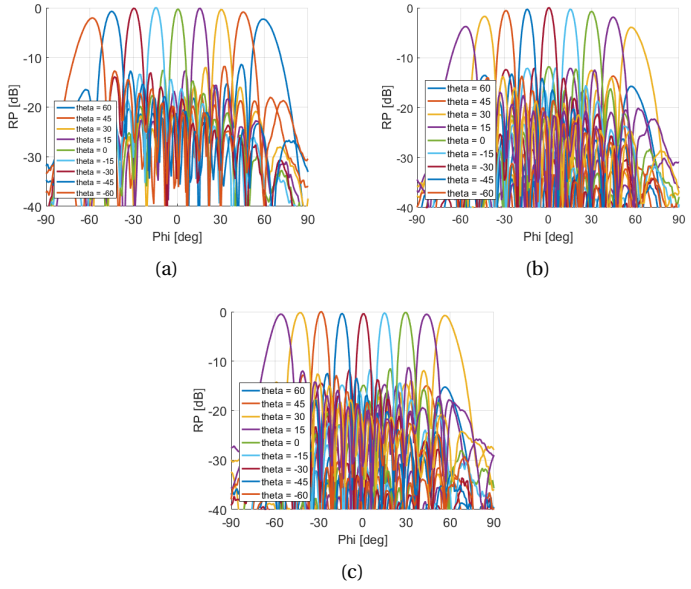
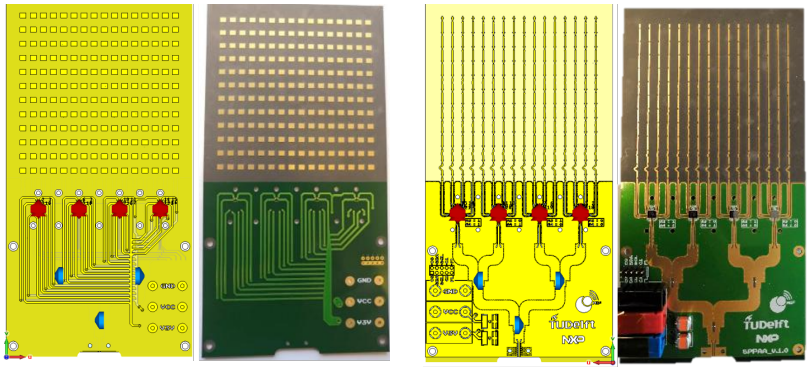


Figure 10.21: Measured horizontal scanning performance of the microsrip line fed AC-MPA antenna panel at various frequencies: (a) 28 GHz, (b) 28.5 GHz, (c) 29 GHz.



(a) (b)

Figure 10.22: Fabricated complete microstrip line fed AC-MPA antenna system (array panel + chip board + power divider): (a) top view, (b) bottom view.

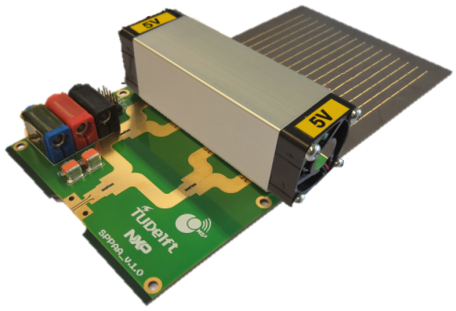


Figure 10.23: Complete microstrip line fed AC-MPA antenna system with active cooling.

antenna panel at various frequencies.

Next, the antenna panel was integrated with the NXP's analog beamforming chips in a complete antenna system including the array panel, chip board and power divider. Fig. 10.22 shows the fabricated complete microstrip line fed AC-MPA antenna system. Similarly to the SIW-fed array, 4 chips are used to adjust the amplitude and phase at each of the 16 subarrays. The measurements for the complete antenna system is still in progress.

It is also worth to mention here that active cooling (2 fans and a rectangular heatsink covering all the chips) is used in both the SIW and microstrip line fed integrated arrays as visualized in Fig. 10.23, which help us maintain the maximum IC temperature around 80^{o}C in the transmit mode.

B. SYNTHESIS OF LARGE-SCALE APERIODIC ARRAYS: A FUTURISTIC VIEW

The layout optimizations throughout the thesis have been performed using a maximum of 64 elements, considering the baseline of the state-of-the-art mm-wave 5G antenna arrays from the leading companies, such as Ericsson, Nokia, IBM, NXP and NEC. Although adding more elements into the design increases the processing & hardware, larger arrays can provide multiple system advantages. For example, when compared to the 64-element aperiodic array, an optimized 256-element array (with a similar average inter-element spacing) has:

- 4 times the gain and better angular resolution,
- 1/4 of the radiated power (thus with 1/16 of the power per PA) to serve the same number of users,
- 6 dB (or more) additional SLL suppression (similar to or better than the random arrays),
- 4 times the area to dissipate 1/4 of the heat.

It is also very important to note that the large-scale arrays have a key advantage on reducing the electricity consumption, with the help of the high antenna gain and low side lobes.

Few companies (such as Anokiwave) have already released 256-element active antenna arrays for mm-wave 5G. However, the array designs are based on regular square-grid dense topologies with high side lobes, whereas, as already shown in Chapters 6 and 9, the aperiodic arrays presented in this thesis have key advantages on interference suppression and thermal handling. This also has a big potential for space applications.

In this section, we present novel large-scale (256-element) irregular array layouts that are optimized for the (statistically) highest in-sector interference suppression and that can potentially achieve fully-passive cooling without an external heatsink. Moreover, the system advantages and performance trade-offs of the proposed arrays are discussed in detail from an interdisciplinary (electromagnetic, thermal and communication system) perspective.

Although the proposed concept has a big potential for addressing the thermal issue, due to the increased hardware and design complexity with more elements, synthesis of large scale aperiodic arrays can be considered as a futuristic view which would be suitable for the next phase of the 5G or 6G radios.

Next, the approaches and settings used in the electromagnetic, communication system and thermal simulations are described and the simulation results are shown.

Parts of this chapter will be published as:
[J-9] Y. Aslan, A. Roederer and A. Yarovoy, "System advantages of using large-scale aperiodic array architectures in mm-wave 5G base stations: an interdisciplinary look," *IEEE Syst. J.*, 2020, under review.

SIMULATION APPROACH AND SETTINGS

ELECTROMAGNETIC ASPECTS

The optimization approach and settings used for the electromagnetic simulations (in MATLAB) are listed as follows:

- Synthesis algorithm: Iterative convex element position perturbation (uniform amplitudes, linear phases) [J-4]
- Optimization strategy: Multibeam peak SLL minimization within the sector
- Sector definition: ±60 / ±15 degrees in azimuth / elevation
- Number of elements: 64 or 256
- Element pattern: $\cos\theta$ (common embedded pattern)
- Operation frequency, f_0: 28 GHz
- Minimum element spacing, $d_{\min}$: 0.5λ or 1λ
- Enforced symmetry: Rotational, if applied.

Due to the increased computational demands with the larger arrays, the optimizations in this part have been carried out in TU Delft's 32-processor Intel(R) Xeon(R) CPU E5-2650 v2 @2.60GHz 128GB RAM Red Hat Enterprise Linux Server.

COMMUNICATION SYSTEM ASPECTS

This part considers the use-case scenario and user selection & beamforming strategies for the evaluation of the Quality-of-Service (QoS) performance and processing complexity of the proposed optimal array topologies.

Similarly to the study in Chapter 2, we consider an isolated cell in which a base station with N antenna elements is serving K single antenna users in pure Line-of-Sight simultaneously in the same narrow frequency sub-band using Space Division Multiple Access (SDMA). The users' locations are randomly selected from uniformly distributed points in the uv-plane (within the ±60 by ±15 degree window), using the interference-aware (i.e., with well-separated user beams) scheduling algorithm proposed in [J-1]. For comparison, we use two different minimum angular separation conditions (= 1/(Array Length in λ) in the uv-plane) between the users, which are suitable for arrays larger in size than a $\lambda/2$-spaced 8-by-8 and a 16-by-16 reference antennas. Furthermore, we apply adaptive power transmission with equalized Signal-to-Noise Ratios (SNRs) and for comparison, we consider two different maximal per-user SNR values (achieved only with a uniformly fed array with progressive phase shifts) that are in line with the currently proposed 5G link budgets. It is worthy of note that we assume the user-dependent path losses, impedance matching performance of the antennas and scanning losses are taken into account in the transmission power control so that the maximal SNR of each user is kept fixed and the main focus of the study remains as the investigation of the antenna array topology impact on the interference. For the precoding, we exploit the two commonly applied techniques: Conjugate Beamforming (CB) and Zero Forcing (ZF).

Table 10.3: Communication system simulation parameters

Center frequency (GHz)	28
Number of array elements at the base station	64, 256
Number of simultaneous co-frequency users	4, 8
Maximal communication range (m)	200
Maximal per-user SNR (dB)	20, 25
Minimum angular spacing between users (in uv-plane)	0.28, 0.13
Number of random user location realizations	10,000

Later, in the results, the statistical system QoS performance evaluation (in MATLAB) is given in terms of the cumulative distribution function (CDF) of Signal-to-Interference-plus-Noise Ratio (SINR) at the user ends, while the computational cost is given in terms of Floating Point Operations Per Second (FLOPS) (see Section 2.4.2).

The simulation parameters used in the system studies are summarized in Table 10.3.

THERMAL ASPECTS

The modeling approach and settings for the thermal simulations (in CST MPS) are listed as follows:

- Simulation model: Two-Resistor Compact Thermal Model [J-6].
- Cooling strategy: Passive-only cooling (natural convection & radiation) with a non-finned flat plate
- Design strategy: A double-sided design (two substrates supporting patches and ICs on opposite sides, with a ground plane in between), one IC per element
- Heat per element is directly proportional (with a constant 1) to the RF power required for multiple users
- Board properties: Same material and size in all topologies for fair comparison. Two design concepts are used: (i) conventional antenna (i.e., with a thin ground plane), (ii) planar heatsink antenna (i.e. with a thick ground plane and extended board edge length) [C-9].

A list of thermal simulation parameters is given in Table 10.4.

Table 10.4: Thermal model parameters

Heat produced per element (W)	2 (for N = 64) 0.125 (for N = 256)
Patch / IC board material	Rogers RT5880
Patch / IC board thickness (mm)	0.508
Ground plane thickness, t_g (mm)	0.05 (thin) / 2 (thick)
Board edge length, L_b (λ)	12 (for t_g = 0.05 mm) 20 (for t_g = 2 mm)
Chip dimensions (mm)	3 x 3 x 0.5
IC junction-to-case resistance (W/K)	10
IC junction-to-board resistance (W/K)	15
Heat transfer coefficient at the air interfaces (W/m^{2}K)	10
Surface emissivity	0.9
Ambient temperature (^{o}C)	25

SIMULATION RESULTS

ELECTROMAGNETIC ASPECTS

Using the above-mentioned electromagnetic settings, four different array topologies have been synthesized:

(i) 64-element, fully aperiodic, $d_{\min} = 0.5\lambda$ (layout is provided in Fig. 10.24(a), radiation pattern is shown in Fig. 10.25)

(ii) 256-element, fully aperiodic, $d_{\min} = 0.5\lambda$ (layout is provided in Fig. 10.24(b), radiation pattern is shown in Fig. 10.26)

(iii) 256-element, quasi-modular, $d_{\min} = 0.5\lambda$ (layout is provided in Fig. 10.24(c), radiation pattern is shown in Fig. 10.27)

(iv) 256-element, fully aperiodic, $d_{\min} = 1\lambda$ (layout is provided in Fig. 10.24(d), radiation pattern is shown in Fig. 10.28)

The maximal SLL at f_0 (with respect to the array's broadside field strength) for each topology is given in Table 10.5. From the results, the following main observations have been made:

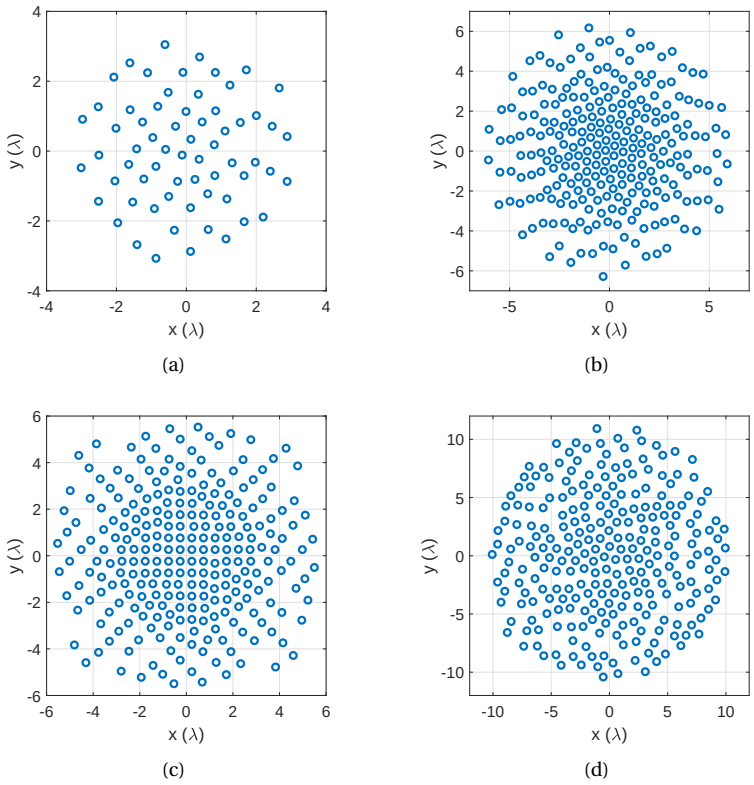


Figure 10.24: Optimal aperiodic array layouts: (a) 64-element, fully aperiodic, $d_{\min} = 0.5\lambda$, (b) 256-element, fully aperiodic, $d_{\min} = 0.5\lambda$, (c) 256-element, quasi-modular, $d_{\min} = 0.5\lambda$, (d) 256-element, fully aperiodic, $d_{\min} = 1\lambda$.

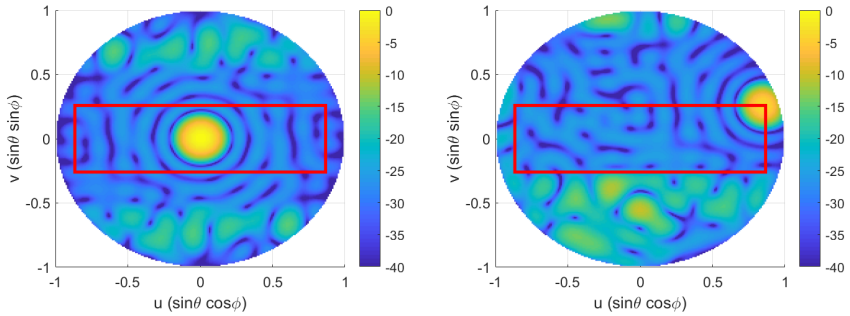


Figure 10.25: Radiation pattern at f_0 (normalized w.r.t. broadside gain, in dB) of the 64-element, fully aperiodic array with $d_{\min} = 0.5\lambda$: (a) broadside beam, (b) corner beam.

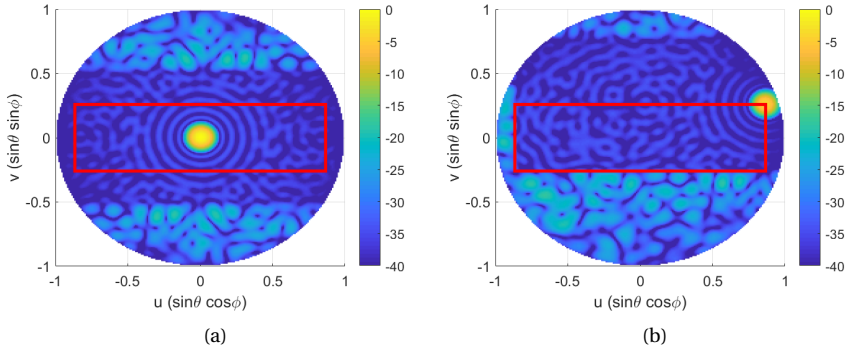


Figure 10.26: Radiation pattern at f_0 (normalized w.r.t. broadside gain, in dB) of the 256-element, fully aperiodic array with $d_{\min} = 0.5\lambda$: (a) broadside beam, (b) corner beam.

(1) Aperiodic arrays provide much better maximal SLL suppression than the periodic counterparts (reduction from -13.2 dB to -25.2 dB for 64-elements).

(2) As compared to the 64-element counterpart, the 256-element aperiodic array has much better SLL suppression capability (reduction from -25.2 dB to -33.4 dB).

(3) Applying layout modularity (to have design & calibration simplification) causes an increase in the peak SLL (from -33.4 dB for no symmetry to -28.1 dB for 4-th order rotational symmetry with 256 elements).

(4) Increasing the minimum inter-element distance (to have extra physical space for the ICs & feeding lines, to provide extra cooling) has a significant impact on the maximal SLL (increase from -33.4 dB for $d_{\min} = 0.5\lambda$ to -22.3 dB for $d_{\min} = 1\lambda$ with 256 elements).

Note that the maximal SLL values given in Table 10.5 can only be maintained over a narrow bandwidth since the topology optimization is performed at a single frequency. The effect of operation bandwidth on the irregular array pattern results is visualized in Fig. 10.29, for which the corner beam radiation pattern of the 256-element, fully aperiodic array with $d_{\min} = 0.5\lambda$ at the design frequency f_0 is taken as the reference (see Fig.

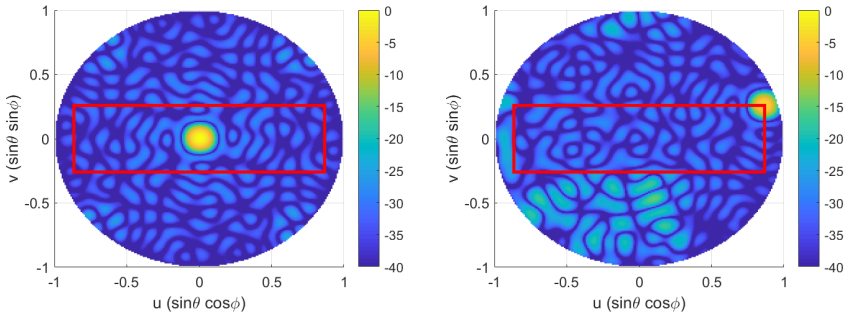


Figure 10.27: Radiation pattern at f_0 (normalized w.r.t. broadside gain, in dB) of the 256-element, quasi-modular array with $d_{\min} = 0.5\lambda$: (a) broadside beam, (b) corner beam.

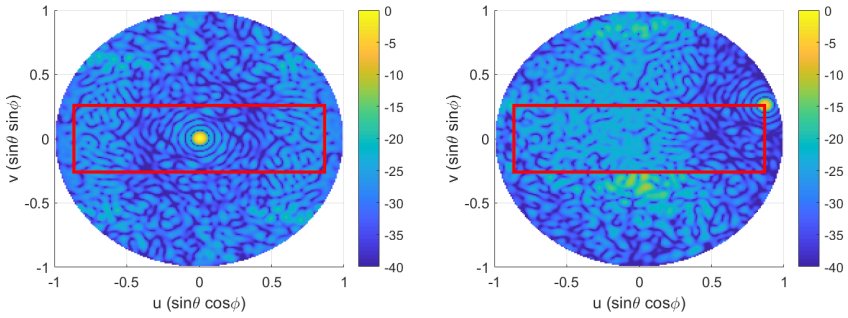


Figure 10.28: Radiation pattern at f_0 (normalized w.r.t. broadside gain, in dB) of the 256-element, fully aperiodic array with $d_{\min} = 1\lambda$: (a) broadside beam, (b) corner beam.

10.26(b)). It is seen that, due to the scaling of the array (in terms of λ):

(i) decreasing the operating frequency increases the beamwidth. Therefore, it will be necessary to increase the minimum angular spacing between the simultaneous co-frequency users to maintain a high QoS.

Table 10.5: Maximal in-sector SLLs (with respect to the field strength at broadside) and array directivities for multiple steerable beams within the sector

Array topology	Maximal SLL (dB)	Broadside beam directivity (dBi)	Corner beam directivity (dBi)
64-element, periodic, 0.5λ-spaced	-12.8	23.0	20.4
64-element, fully aperiodic, $d_{\min} = 0.5\lambda$	-25.2	24.4	20.1
256-element, periodic, 0.5λ-spaced	-13.2	29.1	25.7
256-element, fully aperiodic, $d_{\min} = 0.5\lambda$	-33.4	29.7	25.7
256-element, quasi-modular, $d_{\min} = 0.5\lambda$	-28.1	30.2	25.5
256-element, fully aperiodic, $d_{\min} = 1\lambda$	-22.3	30.7	26.0

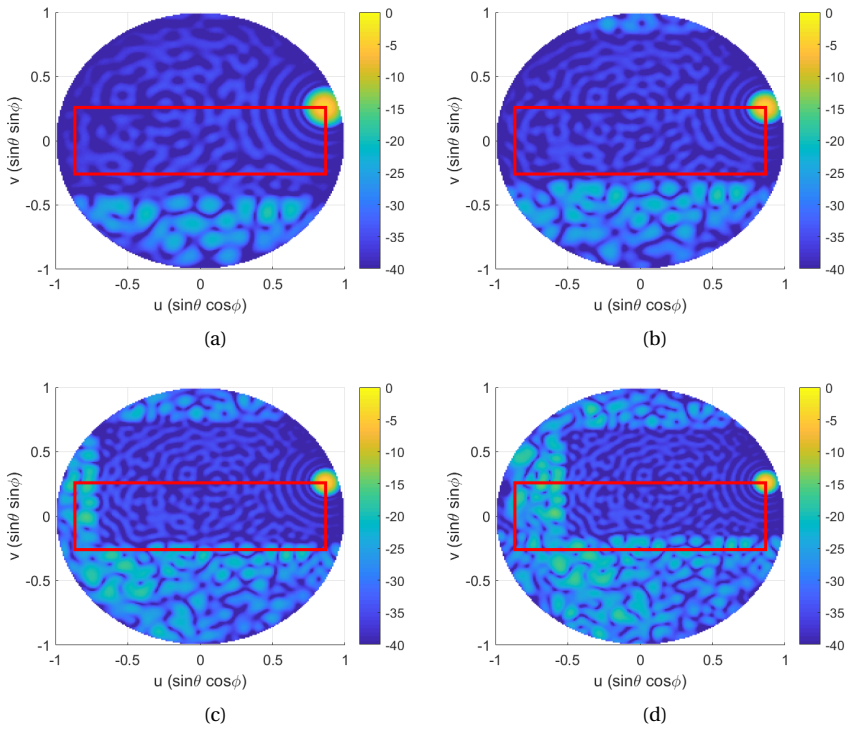


Figure 10.29: Corner beam radiation pattern (normalized w.r.t. broadside gain, in dB) of the 256-element, fully aperiodic array with $d_{\min} = 0.5\lambda$ at different operating frequencies: (a) $0.75f_0$, (b) $0.9f_0$, (c) $1.1f_0$, (d) $1.25f_0$.

(ii) increasing the frequency leads to the appearance of high side lobes in the field-of-view, which will deteriorate the statistical QoS.

In mm-wave 5G applications, the relative bandwidth is narrow and the impact of the array scaling on the pattern is not so significant. For example, the f_0 = 28-GHz *n257* 5G NR frequency band extends from 26.5 GHz ($0.95f_0$) to 29.5 GHz ($1.05f_0$). Similar figures are obtained in other 5G mm-wave bands (such as *n258, n260, n261*) as well.

For broadband applications, on the other hand, the algorithm in [J-4] (which is what we used in this section) can be straightforwardly extended to a multi-frequency multi-beam peak SLL minimization tool. Depending on the desired bandwidth, such an optimization would yield higher SLL than the ones reported in Table 10.5. However, the achieved level would be preserved over a wider range of frequencies. Moreover, considering its practical interest and further SLL suppression capability, a phase-only optimization [J-3] can be added to the already optimized positioning (at the expense of some power efficiency loss), which is not considered here. With increased computational demands, it is also possible to include mutual coupling in the topology optimization procedure via embedded pattern simulations as demonstrated in [C-6]. In this work, for simplicity, we used a cosine-shaped common embedded pattern for all the elements as a reasonable approximation.

COMMUNICATION SYSTEM ASPECTS

This part relates the radiation pattern results of different array topologies and precoding strategies to the system's QoS statistically. Using the pre-given settings, the communication system simulations have been performed.

Fig. 10.30 and Fig. 10.31 show the CDF of SINR for 4/8 simultaneous co-frequency users with 20/25 dB maximal SNR per user, when the minimum distance between the users in the uv-plane is equal to 0.28 and 0.13 units, respectively. For more insight, note that at the maximal communication range of 200 meters, for users located at a similar height with the base station near its broadside, the uv-plane user separation of 0.28 and 0.13 units corresponds to approximately 56 and 26 meters, respectively. The formulation

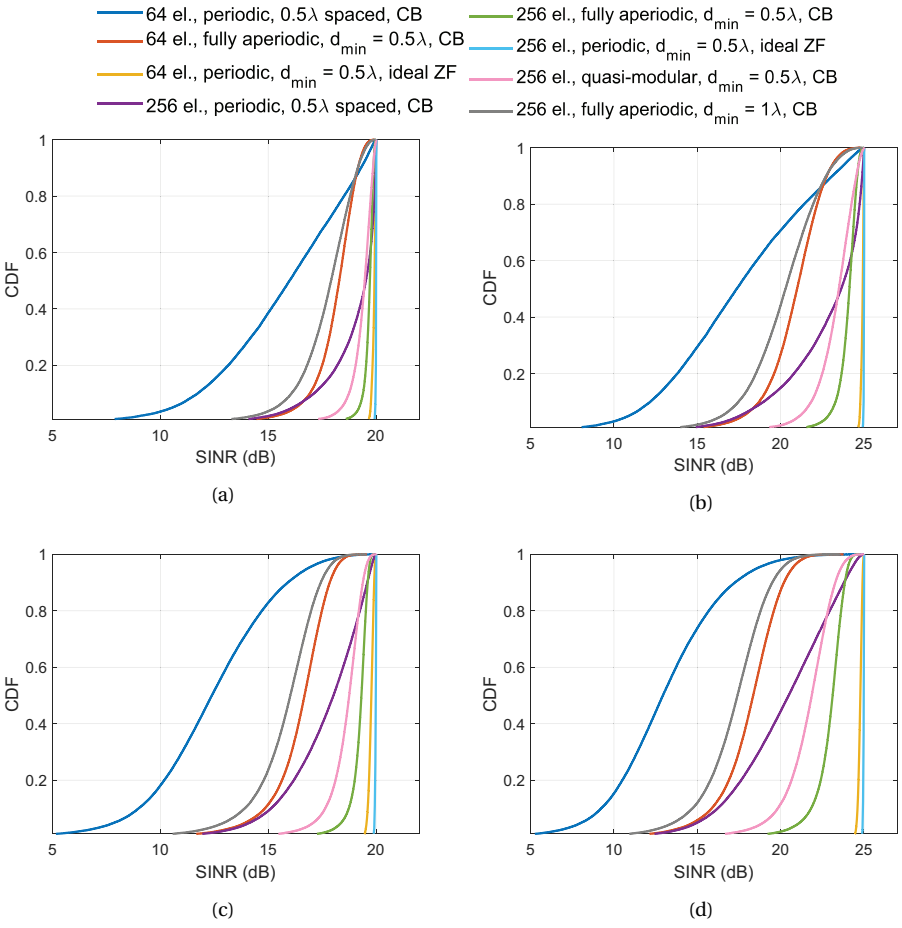


Figure 10.30: CDF of SINR for different array topologies & precoding strategies for the minimum inter-user uv-plane distance of 0.28 units, with: (a) 4 SDMA users & 20 dB maximal SNR per user, (b) 4 SDMA users & 25 dB maximal SNR per user, (c) 8 SDMA users & 20 dB maximal SNR per user, (d) 8 SDMA users & 25 dB maximal SNR per user.

of conversion between the Cartesian and uv-plane coordinates can be found in Section 2.3.

Corresponding to Fig. 10.30(c) and Fig. 10.31(c), Table 10.6 and Table 10.7 exemplifies the minimal SINR attained in 95% of the total occurrences, with 8 users and 20 dB maximal per-user SNR, for the minimum inter-user uv-plane distance of 0.28 and 0.13 units, respectively.

The performance of a 64-element (8-by-8) periodic (square-grid) array with $d_{\min} = 0.5\lambda$ using ZF (i.e., ideally with no inter-beam interference) has been taken as the benchmark here.

From the simulation results (with the statistical criterion of 95%), the following main observations have been made:

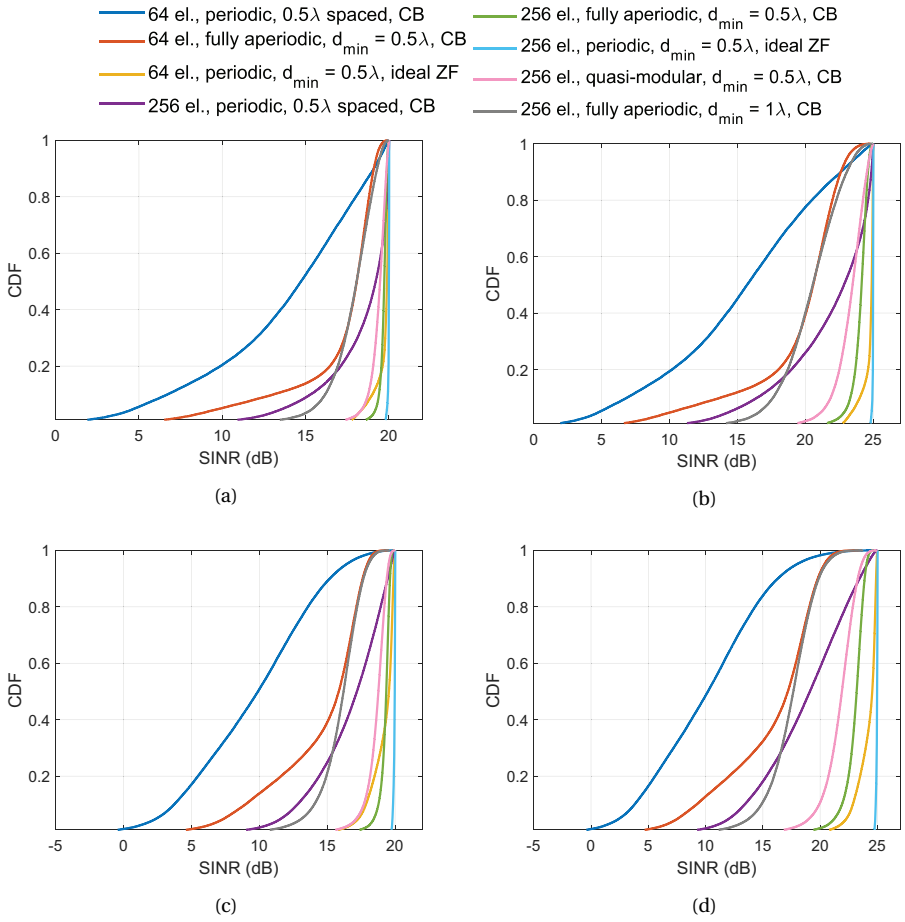


Figure 10.31: CDF of SINR for different array topologies & precoding strategies for the minimum inter-user uv-plane distance of 0.13 units, with: (a) 4 SDMA users & 20 dB maximal SNR per user, (b) 4 SDMA users & 25 dB maximal SNR per user, (c) 8 SDMA users & 20 dB maximal SNR per user, (d) 8 SDMA users & 25 dB maximal SNR per user.

Table 10.6: Minimal SINR (statistically achieved for more than 95% of total occurrences) for 8 simultaneous co-frequency users with minimum uv-plane angular spacing of 0.28 units and maximal per user SNR of 20 dB, corresponding to Fig. 10.30(c)

Array topology & precoding strategy	Minimal SINR (dB)
64-element, periodic, 0.5λ-spaced, CB	7.9
64-element, fully aperiodic, $d_{\min} = 0.5\lambda$, CB	14.0
64-element, periodic, 0.5λ-spaced, ideal ZF	19.6
256-element, periodic, 0.5λ-spaced, CB	14.2
256-element, fully aperiodic, $d_{\min} = 0.5\lambda$, CB	18.4
256-element, periodic, 0.5λ-spaced, ZF	19.9
256-element, quasi-modular, $d_{\min} = 0.5\lambda$, CB	17.1
256-element, fully aperiodic, $d_{\min} = 1\lambda$, CB	13.0

Table 10.7: Minimal SINR (statistically achieved for more than 95% of total occurrences) for 8 simultaneous co-frequency users with minimum uv-plane angular spacing of 0.13 units and maximal per user SNR of 20 dB, corresponding to Fig. 10.31(c)

Array topology & precoding strategy	Minimal SINR (dB)
64-element, periodic, 0.5λ-spaced, CB	2.4
64-element, fully aperiodic, $d_{\min} = 0.5\lambda$, CB	7.4
64-element, periodic, 0.5λ-spaced, ideal ZF	17.3
256-element, periodic, 0.5λ-spaced, CB	11.6
256-element, fully aperiodic, $d_{\min} = 0.5\lambda$, CB	18.5
256-element, periodic, 0.5λ-spaced, ZF	19.8
256-element, quasi-modular, $d_{\min} = 0.5\lambda$, CB	17.2
256-element, fully aperiodic, $d_{\min} = 1\lambda$, CB	13.1

(1) As the exact information of the channel is practically not known and CB can be applied more reliably in most of the cases instead of ZF, fully aperiodic arrays provide significant SINR improvements as compared to their periodic counterparts. For example, Table 10.6 and Table 10.7 show around 6 dB (for 64 elements) and 4 dB (for 256 elements) increase in the minimal SINR in the case of minimal angular spacing of 0.28 units and 5 dB and 7 dB for 64 and 256 elements correspondingly) in the case of minimal angular spacing of 0.13 units.

(2) The 256-element periodic/aperiodic arrays with CB/ZF provides much better statistical QoS as compared to their 64-element counterparts, thanks to their higher beam resolution. In the periodic topology, for the minimal user spacing of 0.28 units, the far side lobes (which are much lower than the first side lobe) of the 256-element array become the major source of interference. In line with this reasoning, Fig. 10.30 shows that the 256-element periodic array can achieve even a slightly better statistical QoS performance than the 64-element fully irregular array. For the minimal user spacing of 0.13 units, the 64-element array cannot resolve two users with an angular separation less than 0.28 units, which causes a significant reduction in the QoS as compared to that of a 256-element array (see Fig. 10.31).

(3) The 256-element fully aperiodic array with CB achieves a close statistical QoS to the one of the reference case (i.e. 64-element periodic array with ideal ZF). The minimal SINR is even higher (by 1.2 dB) for the irregular 256-element array with CB as compared to the reference in one study case given in Fig. 10.31(c) and Table 10.7. Since this performance is achieved with CB instead of ZF (thanks to the large-scale aperiodic layout),

the QoS becomes more robust against channel impurities. In addition, by inserting K = 4 (and 8) and N = 256 (for CB) and 64 (for ZF), it can be computed that $[\#\mathcal{F}_{CB}/\#\mathcal{F}_{ZF}]$ = 0.01 (and 0.004), which clearly highlights the potential of large-scale arrays in decreasing the computational burden of precoding.

(4) Introducing layout modularity or increasing the sparsity in the 256-element arrays causes a decrease in the QoS with CB. The effect is relatively small when changing from the fully-aperiodic to the quasi-modular array with $d_{\min} = 0.5\lambda$ (1.3 dB reduction is seen in Table 10.6-10.7). The impact on statistical SINR is more serious for the array with $d_{\min} = 1\lambda$ (5.4 dB reduction is seen in Table 10.6-10.7), which shifts the SINR curve just behind / ahead of that of the 256-element periodic CB-precoded array with $d_{\min} = 0.5\lambda$ in Fig 10.30 / Fig. 10.31.

(5) Since we assume that the maximal SNR per user is fixed, increasing the number of users results in increased power demand. Besides, the total interference level increases (with the QoS decreasing accordingly) with the number of users. The effect is more visible for the periodic arrays when compared to the aperiodic ones due to the relatively higher side lobes in the regular topologies. For example, by comparing Fig. 10.31(a) and Fig. 10.31(c), it is seen that the minimal SINR statistically achieved for more than 95% of total occurrences is 2.1 dB and 0.8 dB higher for 4 users as compared to the 8 user case in the case of using the 256-element, periodic, 0.5λ-spaced array with CB and the 256-element, fully aperiodic, $d_{\min} = 0.5\lambda$ array with CB, respectively.

Next, the impact of bandwidth on the statistical SINR results is studied. Fig. 10.32 shows the CDF of SINR curves (zoomed in to the region-of-interest) at different frequency sub-bands for the 256-element, fully aperiodic, $d_{\min} = 0.5\lambda$ array under CB precoding in the case of inter-user uv-plane separation of 0.13 units and maximal per-user

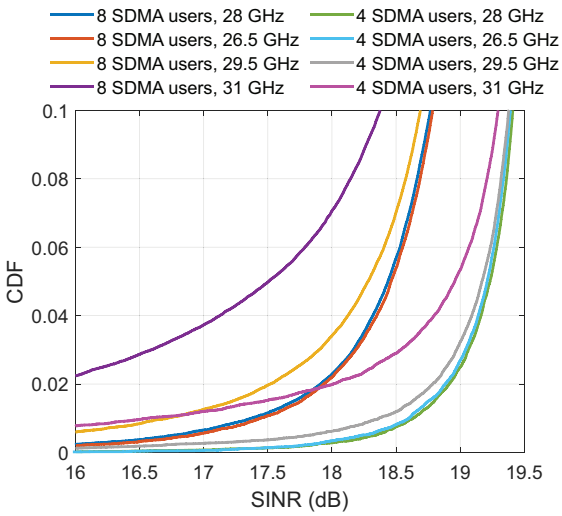


Figure 10.32: CDF of SINR at different frequency sub-bands for the minimum inter-user uv-plane distance of 0.13 units and maximal per-user SNR of 20 dB with the 256-element, fully aperiodic, $d_{\min} = 0.5\lambda$ array under CB precoding.

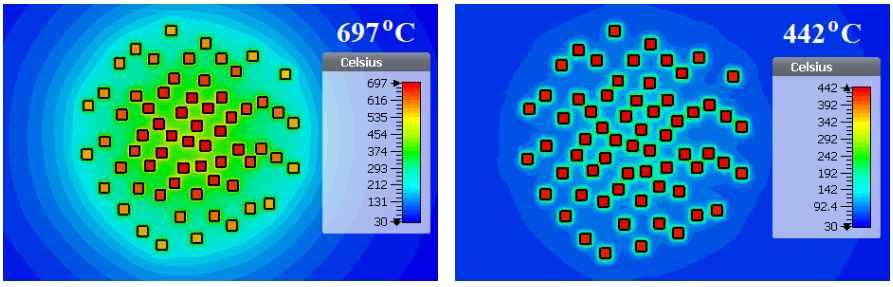
SNR of 20 dB. In line with the results in Fig. 10.29, it is seen that the statistical SINR performance is very stable within the 5G NR band (*n257* mm-wave band: 26.5 - 29.5 GHz is taken as an example). The statistical QoS degrades if the optimized array is used beyond that range. The SINR performance at the 31 GHz sub-band is shown on Fig. 10.29 as an example. As already mentioned, such broadband applications would require a multi-beam array which is jointly optimized for multiple frequencies in a wide range.

THERMAL ASPECTS

Lastly, thermal simulations have been performed for the three selected topologies:

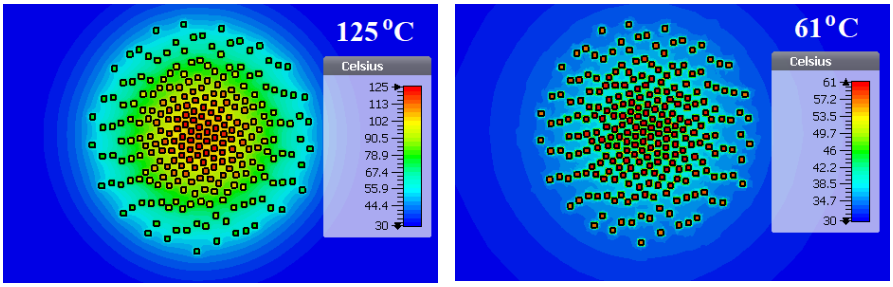
(i) 64-element, fully aperiodic, $d_{\min} = 0.5\lambda$ (see Fig. 10.24(a))
(ii) 256-element, fully aperiodic, $d_{\min} = 0.5\lambda$ (see Fig. 10.24(b))
(iii) 256-element, fully aperiodic, $d_{\min} = 1\lambda$ (see Fig. 10.24(d))

for the two design approaches (namely the conventional antenna, with $t_g = 0.05$ mm and the heatsink antenna, with $t_g = 2$ mm, as described previously in the thermal settings listed in Table 10.4. Note that the 25^oC ambient temperature setting is an optimistic assumption that will not apply to some countries in summertime. Therefore, if any, the relative difference in the ambient temperature should be added to the given simulation



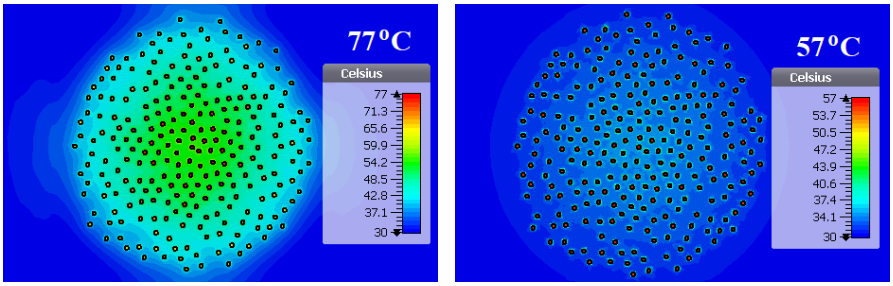
(a) (b)

Figure 10.33: IC temperature distribution in the 64-element, fully aperiodic array with $d_{\min} = 0.5\lambda$: (a) conventional antenna: $L_b = 12\lambda$ & $t_g = 0.05$ mm, (b) heatsink antenna: $L_b = 20\lambda$ & $t_g = 2$ mm.



(a) (b)

Figure 10.34: IC temperature distribution in the 256-element, fully aperiodic array with $d_{\min} = 0.5\lambda$: (a) conventional antenna: $L_b = 12\lambda$ & $t_g = 0.05$ mm, (b) heatsink antenna: $L_b = 20\lambda$ & $t_g = 2$ mm.



(a) (b)

Figure 10.35: IC temperature distribution in the 256-element, fully aperiodic array with $d_{\min} = 1\lambda$: (a) conventional antenna: $L_b = 12\lambda$ & $t_g = 0.05$ mm, (b) heatsink antenna: $L_b = 20\lambda$ & $t_g = 2$ mm.

Table 10.8: Maximal IC junction temperatures (with constant EIRP)

Array topology	Maximal temperature (^oC)
64-element, fully aperiodic, $d_{\min} = 0.5\lambda$, conventional antenna	697
64-element, fully aperiodic, $d_{\min} = 0.5\lambda$, heatsink antenna	442
256-element, fully aperiodic, $d_{\min} = 0.5\lambda$, conventional antenna	125
256-element, fully aperiodic, $d_{\min} = 0.5\lambda$, heatsink antenna	64
256-element, fully aperiodic, $d_{\min} = 1\lambda$, conventional antenna	77
256-element, fully aperiodic, $d_{\min} = 1\lambda$, heatsink antenna	57

results for correct evaluation of the thermal performance under different environmental conditions.

The temperature distributions for the three cases are shown in Figs. 10.33-10.34-10.35, respectively. Note that the figures are zoomed-in to the array centers. Table 10.8 summarizes the outcome. From the results, the following main conclusions have been drawn:

(1) Due to the extremely large heat per element, the 64-element layout results in unacceptably high IC temperatures. This clearly shows that using a single non-finned flat plate is not sufficient for cooling of such arrays, and there must be a big finned heatsink behind the radiators for passive-only cooling (see Chapter 8).

(2) For the same EIRP requirement with the 64-element array, the 256-element topology (with $d_{\min} = 0.5\lambda$) can bring the temperature down to the edge of the acceptable limit of 125^oC with the conventional antenna, under the current settings. Using a heatsink antenna design can provide additional passive cooling, which yields a maximal IC temperature of 64^oC.

(3) The sparse 256-element topology (with $d_{\min} = 1\lambda$) can provide a safe temperature (77^oC), even with the conventional design (i.e., with a low equivalent thermal conductivity board). The temperature can be further decreased (up to 57^oC) by using a more complex heatsink antenna design employing a thick ground plane.

CONCLUSION

In this section, novel large-scale (256-element) aperiodic multibeam arrays have been proposed to be used in the future 5G/6G base stations. Adding more elements into the design has allowed to achieve more gain and much lower side lobes; thus much less power and less heat generation per element with more surface for cooling. This brings the following key system advantages of the proposed arrays over the existing/discussed periodic/aperiodic 64-element antennas:

(1) The electricity consumption is reduced remarkably.

(2) A similar statistical QoS performance to the one of the ideal ZF is achieved by using CB precoding, which significantly decreases the computational complexity (to less than 1% of ZF) and increases the robustness against the operation bandwidth and the non-ideal system conditions (such as channel impurities, quantization errors, estimation errors etc.).

(3) Fully passive cooling (via natural convection and thermal radiation) to a safe and reliable maximal IC temperature is achieved by using a single, non-finned, flat antenna/IC board, with no additional finned heatsink behind the array.

(4) The layout sparsity can be much increased (to provide more space for the electronics, or to enhance the passive cooling capacity), with only limited impact on the QoS statistically.

Apart from the apparent increase in the (RF) hardware (but without the massive thermal hardware with the fins of existing designs), the price to pay is the increased design/fabrication complexity (with irregular IC feeding & routing), and calibration requirements. This can be partially compensated by enforcing (quasi-) modularity in the synthesis procedure, which comes with some compromise against the SLL suppression capability.

BIBLIOGRAPHY

[1] D. Warren and C. Dewar, "Understanding 5g: Perspectives on future technological advancements in mobile," *GSMA Intelligence*, Dec. 2014.

[2] "5G network architecture: A high-level perspective," White Paper, Huawei, Jul. 2016.

[3] K. Kusume and M. Fallgren, "Updated scenarios, requirements and KPIs for 5G mobile and wireless systems with recommendations for future investigations," ICT-317669-METIS/D1.5, METIS, Apr. 2015.

[4] "Massive-element antenna systems technology for 5G base stations," R&D Highlights, Mitsubishi Electric, Mar. 2017.

[5] J. Thompson, X. Ge, H.-C. Wu, R. Irmer, H. Jiang, G. Fettweis, and S. Alamouti, "5g wireless communication systems: Prospects and challenges part 1," *IEEE Comm. Mag.*, vol. 52, pp. 62–64, Feb. 2014.

[6] ——, "5g wireless communication systems: Prospects and challenges part 2," *IEEE Comm. Mag.*, vol. 52, pp. 24–26, May 2014.

[7] W. Hong, Z. H. Jiang, C. Yu, J. Zhou, P. Chen, Z. Yu, H. Zhang, B. Yang, X. Pang, M. Jiang, Y. Cheng, M. K. T. Al-Nuaimi, Y. Zhang, J. Chen, and S. He, "Multibeam antenna technologies for 5G wireless communications," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6231–6249, Dec. 2017.

[8] "A straight path towards 5G," White Paper, Straight Path Communications Inc., Aug. 2015.

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

[10] R. Jos, "Managing power dissipation in 5G antenna design," *EE-Times Europe*, Jun. 2016.

[11] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive MIMO: A survey," *IEEE Comm. Mag.*, vol. 55, no. 9, pp. 134–141, Sep. 2017.

[12] J. Aasted and A. G. Roederer, "A multi-beam array system for communications with mobiles," in *Proc. AIAA 7th Communications Satellite Conference*, Apr. 1978.

[13] A. G. Roederer, "Multiple beam satellite antenna concepts for mobile communications," in *Proc. 3rd International Conference on Satellite Systems for Mobile Communications and Navigation*, 1983.

[14] A. G. Roederer and C. van't Klooster, "Apparatus for electronically controlling the radiation pattern of an antenna having one or more beams of variable width and/or direction," Patent EU Patent EP0 497 652, Nov., 1994.

[15] D. Whitefield, R. Gopal, and S. Arnold, "Spaceway now and in the future: on-board IP packet switching satellite communication network," in *Proc. IEEE MILCOM*, Washington, DC, USA, Oct. 2006.

[16] T. Cameron, "Bits to beams: Rf technology evolution for 5gmillimeter wave radios," in *Tech. Art.*, Analog Devices Inc., 2018.

[17] A. Yarovoy, "Antenna topologies and front-end configurations for multiple beam generation," NWO-NXP Partnership on Advanced 5G Solutions, Research Proposal, Delft University of Technology, Sep. 2016.

[18] "Mobile telephony - wide area coverage - case 20564," Bell Telephone Laboratories, USA, 1947.

[19] H. Magri, N. Abghour, and M. Ouzzif, "Key concepts of 5th generation mobile technology," *WASET IJECE*, vol. 9, no. 4, pp. 471–474, 2015.

[20] N. Odeh, S. Khatun, B. MohdAli, and A. Ismail, "Combined CDMA-SDMA performance analysis and code assignment algorithm," in *Proc. IEEE ICICT*, Dhaka, Bangladesh, Mar. 2007.

[21] C.-S. Park, Y.-S. Byun, A. M. Bokiye, and Y.-H. Lee, "Complexity reduced zero-forcing beamforming in massive MIMO systems," in *Proc. IEEE ITA Workshop*, San Diego, CA, USA, Feb. 2014.

[22] A. A. Glazunov, "Impact of deficient array antenna elements on downlink massive MIMO performance in RIMP and random-LoS channels," in *Proc. 12th EuCAP*, London, UK, Apr. 2018.

[23] E. Degirmenci, "Emf test report: Ericsson AIR 5121," Ericsson AB, Stockholm, Sweden, Tech. Rep. GFTB-17:001589 Uen Rev B, Jan. 2018.

[24] G. Oliveri, G. Gottardi, and A. Massa, "A new meta-paradigm for the synthesis of antenna arrays for future wireless communications," *IEEE Trans. Antennas Propag.*, vol. 67, no. 6, pp. 3774–3788, Jun. 2019.

[25] A. F. Molisch, A. Karttunen, R. Wang, C. U. Bas, S. Hur, J. Park, and J. Zhang, "Millimeter-wave channels in urban environments," in *Proc. 10th EuCAP*, Davos, Switzerland, Apr. 2016.

[26] L. C. Godara, *"Handbook of Antennas in Wireless Communications.* CRC Press, 2002.

[27] V. Degli-Esposti, F. Fuschini, E. M. Vitucci, M. Barbiroli, M. Zoli, L. Tian, X. Yin, D. A. Dupleich, R. Müller, C. Schneider, and R. S. Thoma, "Ray-tracing-based mm-wave beamforming assessment," *IEEE Access*, vol. 2, pp. 1314–1325, 2014.

[28] F. Fuschini, M. Zoli, E. M. Vitucci, M. Barbiroli, and V. Degli-Esposti, "A study on millimeter-wave multiuser directional beamforming based on measurements and ray tracing simulations," *IEEE Trans. Antennas Propag.*, vol. 67, no. 4, pp. 2633–2644, Apr. 2019.

[29] R. Hoppe, G. Wölfle, and U. Jakobus, "Wave propagation and radio network planning software winprop added to the electromagnetic solver package FEKO," in *Proc. IEEE ACES*, Florence, Italy, May 2017.

[30] "Effects of building materials and structures on radiowave propagation above about 100 mhz," in *Recommendation ITU-R P.2040-1*, International Telecommunication Union, Jul. 2015.

[31] A. M. Bokiye, "Ray-based propagation modeling for ofdm-based mobile networks," in *MSc. Thesis*, Delft University of Technology, The Netherlands, Jun. 2009.

[32] M. H. Vogel, "MIMO system simulation in WinProp," in *White Paper*, Altair Engineering, Inc.

[33] R. Deng, Z. Jiang, S. Zhou, and Z. Niu, "How often should CSI be updated for massive MIMO systems with massive connectivity?" in *Proc. IEEE GLOBECOM.*, Singapore, Singapore, Dec. 2017.

[34] R. Valkonen, "Compact 28-GHz phased array antenna for 5G access," in *Proc. IEEE/MTT-S IMS*, Philadelphia, Pennsylvania, Jun. 2018.

[35] B. Sadhu, Y. Tousi, J. Hallin, S. Sahl, S. Reynolds, O. Renstrom, K. Sjogren, O. Haapalahti, N. Mazor, B. Bokinge, G. Weibull, H. Bengtsson, A. Carlinger, E. Westesson, J.-E. Thillberg, L. Rexberg, M. Yeck, X. Gu, D. Friedman, and A. Valdes-Garcia, "A 28 GHz 32-element phased-array transceiver IC with concurrent dual polarized beams and 1.4 degree beam-steering resolution for 5G communication," in *Proc. IEEE ISSCC*, San Francisco, CA, USA, Feb. 2017.

[36] *IEEE Standard for Definitions of Terms for Antennas.* IEEE Standard 145-2013, Dec. 2013.

[37] S. Rao, L. Shafai, and S. Sharma, *Handbook of reflector antennas and feed systems.* Norwood, MA, USA: Artech House, 2013.

[38] O. Quevedo-Teruel, M. Ebrahimpouri, and F. Ghasemifard, "Lens antennas for 5G communications systems," *IEEE Commun. Mag.*, vol. 56, no. 7, pp. 36–41, Jul. 2018.

[39] J. Butler and R. Lowe, "Beamforming matrix simplifies design of electronically scanned antennas," *Electron. Des.*, vol. 9, pp. 170–173, Apr. 1961.

[40] A. Jacomb-Hood and E. Lier, "Multibeam active phased arrays for communications satellites," *IEEE Microw. Mag.*, vol. 1, no. 4, pp. 40–47, Dec. 2000.

[41] I. Ahmed, H. Khammari, A. Shahid, A. Musa, K. S. Kim, E. D. Poorter, and I. Moerman, "A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives," *IEEE Commun. Surv. Tut.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.

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

[43] T. Kim, J. Park, J.-Y. Seol, S. Jeong, J. Cho, and W. Roh, "Tens of gbps support with mmwave beamforming systems for next generation communications," in *Proc. IEEE GLOBECOM*, Atlanta, GA, USA, Dec. 2013.

[44] G. W. Kant, P. D. Patel, S. J. Wijnholds, M. Ruiter, and E. van der Wal, "EMBRACE: A multi-beam 20,000-element radio astronomical phased array antenna demonstrator," *IEEE Trans. Antennas Propag.*, vol. 59, no. 6, pp. 1990–2003, Jun. 2011.

[45] P. L. Metzen, "Globalstar satellite phased array antennas," in *Proc. IEEE Int. Conf. Phased Array Syst. Technol.*, Dana Point, CA, USA, May 2000.

[46] D.-W. Kang, K.-J. Koh, and G. M. Rebeiz, "A Ku-band two-antenna four-simultaneous beams SiGe BiCMOS phased array receiver," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 4, pp. 771–780, Apr. 2010.

[47] K. J. Maalouf and E. Lier, "Theoretical and experimental study of interference in multibeam active phased array transmit antenna for satellite communications," *IEEE Trans. Antennas Propag.*, vol. 52, no. 2, pp. 587–592, Feb. 2004.

[48] H. Steyskal and J. F. Rose, "Digital beamforming for radar systems," *Microw. J.*, vol. 32, no. 1, pp. 121–136, Jan. 1989.

[49] R. Miura, T. Tanaka, I. Chiba, A. Horie, and Y. Karasawa, "Beamforming experiment with a dbf multibeam antenna in a mobile satellite environment," *IEEE Trans. Antennas Propag.*, vol. 45, no. 4, pp. 707–714, Apr. 1997.

[50] S. S. Ahmed, A. Genghammer, A. Schiessl, and L.-P. Schmidt, "Fully electronic E-band personnel imager of 2 m2 aperture based on a multistatic architecture," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 1, pp. 651–657, Jan. 2013.

[51] W. Menzel and A. Moebius, "Antenna concepts for millimeter-wave automotive radar sensors," *Proc. IEEE*, vol. 100, no. 7, pp. 2372–2379, Jul. 2012.

[52] B. Yang, Z. Yu, J. Lan, R. Zhang, J. Zhou, and W. Hong, "Digital beamforming-based massive MIMO transceiver for 5G millimeter-wave communications," *IEEE Trans. Microw. Theory Tech.*, vol. 66, no. 7, pp. 3403–3418, Jul. 2018.

[53] M. A. B. Abbasi, H. Tataria, V. Fusco, and M. Matthaiou, "Performance of a 28 GHz two-stage rotman lens beamformer for millimeter wave cellular systems," in *Proc. 13th EuCAP*, Krakow, Poland, Apr. 2019.

[54] J. Wang, H. Zhu, N. J. Gomes, and J. Wang, "Frequency reuse of beam allocation for multiuser massive MIMO systems," *IEEE Trans. Wirel. Commun.*, vol. 17, no. 4, pp. 2346–2359, Apr. 2018.

[55] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, Feb. 2014.

[56] P. Xingdong, H. Wei, Y. Tianyang, and L. Linsheng, "Design and implementation of an active multibeam antenna system with 64 RF channels and 256 antenna elements for massive MIMO application in 5G wireless communications," *China Commun.*, vol. 11, no. 11, pp. 16–23, Nov. 2014.

[57] G. Adamidis and I. Vardiambasis, "Smart antenna design and implementation: A simple switched-beam antenna array based on a 8x8 butler-matrix network," in *Proc. 10th WSEAS ICCOM*, Athens, Greece, Jul. 2006.

[58] T. Azar, "Overlapped subarrays: review and update," *IEEE Antennas Propag. Mag.*, vol. 55, no. 2, pp. 228–234, Apr. 2013.

[59] C. K. Ng, M. Ismail, B. Ali, S. Khatun, and S. Jamuar, "Impact of inter-cell interference on capacity in the joint multiple access (cdma and sdma) system," in *Proc. IEEE APCC*, Perth, WA, Australia, Oct. 2005.

[60] T. Kopacz and D. Heberling, "Impact of the elevation scanning angle on the vertical compliance distance of 5G massive MIMO antennas," in *Proc. 13th EuCAP*, Krakow, Poland, Apr. 2019.

[61] W. Stutzman, "Synthesis of shaped-beam radiation patterns using the iterative sampling method," *IEEE Trans. Antennas Propag.*, vol. 19, no. 1, pp. 36–41, Jan. 1971.

[62] O. M. Bucci, G. Franceschetti, G. Mazzarella, and G. Panariello, "A general projection approach to array synthesis," in *Proc. IEEE AP-S Int. Symp.*, San Jose, CA, USA, Jun. 1989.

[63] B. Fuchs, "Synthesis of sparse arrays with focused or shaped beampattern via sequential convex optimizations," *IEEE Trans. Antennas Propag.*, vol. 60, no. 7, pp. 3499–3503, May 2012.

[64] N. V. S. N. Sarma and R. Chandrasekharam, "Shaped beam radiation pattern synthesis using genetic algorithm," in *Proc. IEEE ICEMIC*, Hyderabad, India, Dec. 1999.

[65] A. K. Behera, A. Ahmad, S. K. Mandal, G. K. Mahanti, and R. Ghatak, "Synthesis of cosecant squared pattern in linear antenna arrays using differential evolution," in *Proc. IEEE CICT*, Thuckalay, Tamil Nadu, India, Apr. 2013.

[66] X. M. Zhang, K. M. Luk, Q. F. Wu, T. Ying, X. Bai, and L. Pu, "Cosecant-square pattern synthesis with particle swarm optimization for nonuniformly spaced linear array antennas," in *Proc. IEEE ISAPE*, Kunming, China, Nov. 2008.

[67] A. Akdagli and K. Guney, "Shaped-beam pattern synthesis of equally and unequally spaced linear antenna arrays using a modified tabu search algorithm," *Microw. Opt. Technol. Lett.*, vol. 36, no. 1, pp. 16–20, Jan. 2003.

[68] A. Dastranj, H. Abiri, and A. Mallahzadeh, "Design of a broadband cosecant squared pattern reflector antenna using iwo algorithm," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3895–3900, Mar. 2013.

[69] M. Milijic, A. D. Nesic, and B. Milovanovic, "Design, realization, and measurements of a corner reflector printed antenna array with cosecant squared-shaped beam pattern," *IEEE Antennas Wirel. Propag. Lett.*, vol. 15, pp. 421–424, Jun. 2015.

[70] E. Carrasco, M. Barba, J. A. Encinar, M. Arrebola, F. Rossi, and A. Freni, "Design, manufacture and test of a low-cost shaped-beam reflectarray using a single layer of varying-sized printed dipoles," *IEEE Trans. Antennas Propag.*, vol. 61, no. 6, pp. 3077–3085, Jun. 2013.

[71] M. Arrebola, J. A. Encinar, , and M. Barba, "Multifed printed reflectarray with three simultaneous shaped beams for lmds central station antenna," *IEEE Trans. Antennas Propag.*, vol. 56, no. 6, pp. 1518–1527, Jun. 2008.

[72] J. Lei, G. Fu, L. Yang, and D. M. Fu, "Wide band linear printed antenna array with low sidelobe cosecant square-shaped beam pattern," *Prog. Electromagn. Res. C*, vol. 15, pp. 233–241, 2010.

[73] A. Kedar, P. N. S. Kutiyal, M. Garg, and U. K. Revankar, "Wide band low profile linear microstrip antenna array with cosecant square-shaped beam pattern," *Microw. Opt. Technol. Lett.*, vol. 49, no. 4, pp. 963–965, Apr. 2007.

[74] M. Koubeissi, L. Freytag, C. Decroze, and T. Monediere, "Design of a cosecant-squared pattern antenna fed by a new butler matrix topology for base station at 42 ghz," *IEEE Antennas Wirel. Propag. Lett.*, vol. 7, pp. 354–357, Mar. 2008.

[75] Z. C. Hao and M. He, "Developing millimeter-wave planar antenna with a cosecant squared pattern," *IEEE Trans. Antennas Propag.*, vol. 65, no. 10, pp. 5565–5570, Oct. 2017.

[76] H. Chu, P. Li, and Y. X. Guo, "A beam-shaping feeding network in series configuration for antenna array with cosecant-square pattern and low sidelobes," *IEEE Antennas Wirel. Propag. Lett.*, vol. 18, no. 4, pp. 742–746, Apr. 2019.

[77] J. Hirokawa, C. Yamazaki, and M. Ando, "Postwall waveguide slot array with cosecant radiation pattern andnull filling for base station antennas in local multidistributed systems," *Radio Sci.*, vol. 38, no. 2, Dec. 2002.

[78] S. Yamamoto, J. Hirokawa, and M. Ando, "A beam switching slot array with a 4-way butler matrix installed in a single layer post-wall waveguide," in *IEEE AP-S Int. Symp.*, San Antonio, TX, USA, Jun. 2002.

[79] C. Sacchi, T. F. Rahman, N. Bartolomei, S. Morosi, A. Mazzinghi, and F. Ciabini, "Design and assessment of a ce-ofdm-based mm-wave 5g communication system," in *Proc. IEEE GLOBECOM.*, Washington, DC, USA, Dec. 2016.

[80] L. Yan, W. Hong, K. Wu, and T. J. Cui, "Investigations on the propagation characteristics of the substrate integrated waveguide based on the method of lines," *IEE Proc. Microw. Antennas Propag.*, vol. 152, no. 1, pp. 35–42, Feb. 2005.

[81] R. S. Elliott, *Antenna theory and design, Revised edn., Chapter VIII.* (John Wiley & Sons, Inc., Hoboken (NJ), 2003.

[82] J. F. DeFord and O. P. Gandhi, "Phase-only synthesis of minimum peak sidelobe patterns for linear and planar arrays," *IEEE Trans. Antennas Propag.*, vol. 36, no. 2, pp. 191–201, Feb. 1988.

[83] A. F. Morabito, T. Isernia, and L. D. Donato, "Optimal synthesis of phase-only reconfigurable linear sparse arrays having uniform-amplitude excitations," *Prog. Electromagn. Res.*, vol. 124, pp. 405–423, 2012.

[84] E. Ercil, "An alternative method for phase only array pattern synthesis," in *Proc. IEEE APS-URSI*, Chicago, IL, USA, Jul. 2012.

[85] J. Liang, X. Fan, W. Fan, D. Zhou, and J. Li, "Phase-only pattern synthesis for linear antenna arrays," *IEEE Antennas Wirel. Propag. Lett.*, vol. 16, pp. 3232–3235, Nov. 2017.

[86] D. W. Boeringer and D. H. Werner, "Adaptive mutation parameter toggling genetic algorithm for phase-only array synthesis," *Electron. Lett.*, vol. 38, no. 25, pp. 1618–1619, Dec. 2002.

[87] O. Bucci, G. Mazzarella, and G. Panariello, "Reconfigurable arrays by phase-only control," *IEEE Trans. Antennas Propag.*, vol. 39, no. 7, pp. 919–925, Jul. 1991.

[88] B. Fuchs, "Application of convex relaxation to array synthesis problems," *IEEE Trans. Antennas Propag.*, vol. 62, no. 2, pp. 634–640, Feb. 2014.

[89] P. J. Kajenski, "Phase only antenna pattern notching via a semidefinite programming relaxation," *IEEE Trans. Antennas Propag.*, vol. 60, no. 5, pp. 2562–2565, Mar. 2012.

[90] G. Buttazzoni and R. Vescovo, "Reconfigurable antenna arrays with phase-only control in the presence of near-field nulls," *Journ. Inf. Comm. Tech.*, vol. 3, pp. 88–93, Jan. 2017.

[91] P. Harikumar, V. B. Bikkani, G. Mahanti, and B. Mahato, "Phase-only side lobe level reduction of uniformly excited linear array antenna using iterative fast fourier transform," in *Proc. IEEE INDICON*, Hyderabad, India, Dec. 2011.

[92] P. Rocca, R. L. Haupt, and A. Massa, "Sidelobe reduction through element phase control in uniform subarrayed array antennas," *IEEE Antennas Wirel. Propag. Lett.*, vol. 8, pp. 437–440, Feb. 2009.

[93] T. H. Ismail, D. I. Abu-Al-Nadi, and M. J. Mismar, "Phase-only control for antenna pattern synthesis of linear arrays using the levenberg-marquardt algorithm," *Electromagnetic.*, vol. 24, no. 7, pp. 555–564, Jun. 2010.

[94] M. J. Mismar and T. H. Ismail, "Pattern nulling by iterative phase perturbation," *Prog. Electromagn. Res.*, vol. 22, pp. 181–195, 1999.

[95] G. Buttazzoni, M. Comisso, F. Ruzzier, and R. Vescovo, "Phase-only antenna array reconfigurability with gaussian-shaped nulls for 5G applications," *Int. J. Antennas Propag.*, vol. 2019, Feb. 2019.

[96] M. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra and its Applications*, vol. 284, no. 1-3, pp. 193–228, 1998.

[97] F. Alizadeh and D. Goldfarb, "Second-order cone programming," *Math. Program*, vol. 95, no. 1, pp. 3–51, 2003.

[98] Y.-J. Kuo and H. Mittelmann, "Interior point methods for second-order cone programming and OR applications," *Computat. Optim. Appl.*, vol. 28, no. 3, pp. 255–285, 2004.

[99] Z. Cai and K. Toh, "Solving second order cone programming via a reduced augmented system approach," *SIAM J. Optim.*, vol. 17, no. 3, pp. 711–737, 2006.

[100] M. Muramatsu, "A pivoting procedure for a class of second-order cone programming," *Optim. Methods Softw.*, vol. 21, pp. 295–314, 2006.

[101] T. Hasuike, "On an exact optimal solution for a second-order cone programming problem," in *Proc. IMECS*, Hong Kong, Mar. 2011.

[102] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, 2014.

[103] J. F. Sturm, "Using SeDuMi1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods Softw.*, vol. 11–12, pp. 625–653, 1999.

[104] H. B. Van, S. N. Jha, and C. Craeye, "Fast full-wave synthesis of printed antenna arrays including mutual coupling," *IEEE Trans. Antennas Propag.*, vol. 64, no. 12, pp. 5163–5171, Dec. 2016.

[105] S. E. Nai, W. Ser, Z. L. Yu, and H. Chen, "Beampattern synthesis for linear and planar arrays with antenna selection by convex optimization," *IEEE Trans. Antennas Propag.*, vol. 58, no. 12, pp. 3923–3930, Dec. 2010.

[106] B. Fuchs, A. Skrivervik, and J. R. Mosig, "Synthesis of uniform amplitude focused beam arrays," *IEEE Antennas Wireless Propag. Lett.*, vol. 11, pp. 1178–1181, Oct. 2012.

[107] M. D'Urso, G. Prisco, and R. M. Tumolo, "Maximally sparse, steerable, and non-superdirective array antennas via convex optimizations," *IEEE Trans. Antennas Propag.*, vol. 64, no. 9, pp. 3840–3849, Sep. 2016.

[108] G. Toso, P. Angeletti, and C. Mangenot, "A comparison of density and amplitude tapering for transmit active arrays," in *Proc. 3rd EuCAP*, Berlin, Germany, Mar. 2009.

[109] R. L. Haupt, "Thinned arrays using genetic algorithms," *IEEE Trans. Antennas Propag.*, vol. 42, no. 7, pp. 993–999, Jul. 1994.

[110] D. G. Kurup, M. Himdi, and A. Rydberg, "Synthesis of uniform amplitude unequally spaced antenna arrays using the differential evolution algorithm," *IEEE Trans. Antennas Propag.*, vol. 51, no. 9, pp. 2210–2217, Sep. 2003.

[111] L. Cen, Z. L. Yu, W. Ser, and W. Cen, "Linear aperiodic array synthesis using an improved genetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 60, no. 2, pp. 895–902, Feb. 2012.

[112] O. M. Bucci, M. D'Urso, T. Isernia, P. Angeletti, and G. Toso, "Deterministic synthesis of uniform amplitude sparse arrays via new density taper techniques," *IEEE Trans. Antennas Propag.*, vol. 58, no. 6, pp. 1949–1958, Jun. 2010.

[113] J. L. A. Quijano, M. Righero, and G. Vecchi, "Sparse 2-d array placement for arbitrary pattern mask and with excitation constraints: A simple deterministic approach," *IEEE Trans. Antennas Propag.*, vol. 62, no. 4, pp. 1652–1662, Apr. 2014.

[114] C. Bencivenni, M. V. Ivashina, R. Maaskant, and J. Wettergren, "Synthesis of maximally sparse arrays using compressive sensing and full-wave analysis for global earth coverage applications," *IEEE Trans. Antennas Propag.*, vol. 64, no. 11, pp. 4872–4877, Jul. 2016.

[115] D. Pinchera, M. D. Migliore, and G. Panariello, "Synthesis of large sparse arrays using IDEA (inflating-deflating exploration algorithm)," *IEEE Trans. Antennas Propag.*, vol. 66, no. 9, pp. 4658–4668, Sep. 2018.

[116] C. Yan, P. Yang, and S. Y. Huang, "Synthesis of planar sparse arrays with minimum spacing constraint," *IEEE Antennas Wireless Propag. Lett.*, vol. 17, pp. 1095–1098, Jun. 2018.

[117] X. Li, B. Duan, J. Zhou, L. Song, and Y. Zhang, "Planar array synthesis for optimal microwave power transmission with multiple constraints," *IEEE Antennas Wireless Propag. Lett.*, vol. 16, pp. 70–73, Apr. 2016.

[118] C. Bencivenni, M. V. Ivashina, R. Maaskant, and J. Wettergren, "Design of maximally sparse antenna arrays in the presence of mutual coupling," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 159–162, Sep. 2014.

[119] H. B. Van, S. N. Jha, and C. Craeye, "Efficient array synthesis of printed arrays including mutual coupling," in *Proc. 10th EuCAP*, Davos, Switzerland, Apr. 2016.

[120] J. I. Echeveste, M. A. G. de Aza, J. Rubio, and C. Craeye, "Gradient-based aperiodic array synthesis of real arrays with uniform amplitude excitation including mutual coupling," *IEEE Trans. Antennas Propag.*, vol. 65, no. 2, pp. 541–551, Feb. 2017.

[121] C. Y. Chiu, C. H. Cheng, R. D. Murch, and C. R. Rowell, "Reduction of mutual coupling between closely-packed antenna elements," *IEEE Trans. Antennas Propag.*, vol. 55, no. 6, pp. 1732–1738, Jun. 2007.

[122] E. Rajo-Iglesias, O. Quevedo-Teruel, and L. Inclan-Sanchez, "Mutual coupling reduction in patch antenna arrays by using a planar EBG structure and a multilayer dielectric substrate," *IEEE Trans. Antennas Propag.*, vol. 56, no. 6, pp. 1648–1655, Jun. 2008.

[123] Y. Aslan and A. Yarovoy, "Reduction of mutual coupling between closely spaced patch antennas using dielectric stratification technique," in *Proc. 47th EuMC*, Nuremberg, Germany, Oct. 2017.

[124] G. Oliveri, L. Poli, and A. Massa, "Maximum efficiency beam synthesis of radiating planar arrays for wireless power transmission," *IEEE Trans. Antennas Propag.*, vol. 61, no. 5, pp. 2490–2499, May 2013.

[125] N. Anselmi, P. Rocca, M. Salucci, and A. Massa, "Irregular phased array tiling by means of analytic schemata-driven optimization," *IEEE Trans. Antennas Propag.*, vol. 65, no. 9, pp. 4495–4510, Sep. 2017.

[126] Z.-Y. Xiong, Z.-H. Xu, S.-W. Chen, and S.-P. Xiao, "Subarray partition in array antenna based on the algorithm X," *IEEE Antennas Wireless Propag. Lett.*, vol. 12, pp. 906–909, Jul. 2013.

[127] P. Rocca, R. J. Mailloux, and G. Toso, "Ga-based optimization of irregular subarray layouts for wideband phased arrays design," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 131–134, Sep. 2015.

[128] Q. Luo, S. Gao, M. Sobhy, J. T. S. Sumantyo, J. Li, G. Wei, J. Xu, and C. Wu, "Dual circularly polarized equilateral triangular patch array," *IEEE Trans. Antennas Propag.*, vol. 64, no. 6, pp. 2255–2262, Jun. 2016.

[129] A. B. Smolders and U. Johannsen, "Axial ratio enhancement for circularly-polarized millimeter-wave phased-arrays using a sequential rotation technique," *IEEE Trans. Antennas Propag.*, vol. 59, no. 9, pp. 3465–3469, Sep. 2011.

[130] T. Teshirogi, M. Tanaka, and W. Chujo, "Wideband circularly polarised array antenna with sequential rotations and phase shift of elements," in *Proc. ISAP*, Tokyo, Japan, 1985, pp. 117–120.

[131] J. Huang, "A technique for an array to generate circular polarization with linearly polarized elements," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 9, pp. 1113–1124, Sep. 1986.

[132] A. G. Roederer, "Some comments on the gain of large planar arrays," *IEEE Antennas Propag. Soc. Newslet.*, vol. 31, no. 5, pp. 63–64, Oct. 1989.

[133] Y. J. Hu, W. P. Ding, and W. Q. Cao, "Broadband circularly polarized microstrip antenna array using sequentially rotated technique," *IEEE Antennas Wirel. Propag. Lett.*, vol. 10, pp. 1358–1361, Nov. 2011.

[134] A. B. Smolders, "Random sequential rotation: A new technique for sidelobe control of circularly-polarized phased arrays," in *Proc. IEEE APS-URSI*, Memphis, TN, USA, Jul. 2014.

[135] E. Schubert and A. Zimek, "ELKI: A large open-source library for data analysis - ELKI release 0.7.5 "heidelberg"," *CoRR*, 1902.03616 2019.

[136] N. Hussain, M. J. Jeong, J. Park, and N. Kim, "A broadband circularly polarized fabry-perot resonant antenna using a single-layered PRS for 5G MIMO applications," *IEEE Access*, vol. 7, pp. 42 897–42 907, Apr. 2019.

[137] J. S. Herd and M. D. Conway, "The evolution to modern phased array architectures," *Proc. IEEE*, vol. 104, no. 3, pp. 519–529, Mar. 2016.

[138] "An AESA revolution utilizing the disruptive technology of highly-integrated silicon ICs," White Paper, Anokiwave, Nov. 2015.

[139] E. McCune, "Energy efficiency maxima for wireless communications: 5G, IoT, and massive MIMO," in *Proc. IEEE CICC*, Austin, TX, USA, May 2017.

[140] "SKY66184-11 LTE power amplifier data sheet," Skyworks Inc., Massachusetts, USA.

[141] C. Fager, K. Hausmair, T. Eriksson, and K. Buisman, "Analysis of thermal effects in active antenna array transmitters using a combined EM/circuit/thermal simulation technique," in *Proc. IEEE INMMiC*, Taormina, Italy, Oct. 2015.

[142] D. G. Kam, D. Liu, A. Natarajan, S. K. Reynolds, and B. A. Floyd, "Organic packages with embedded phased-array antennas for 60-GHz wireless chipsets," *IEEE Trans. Compon. Packaging Manuf. Technol.*, vol. 1, no. 11, pp. 1806–1814, Nov. 2011.

[143] D. Price, "A review of selected thermal management solutions for military electronic systems," *IEEE Trans. Compon. Packag. Technol.*, vol. 26, no. 1, pp. 26–39, Mar. 2003.

[144] M. Parlak and R. J. McGlen, "Cooling of high power active phased array antenna using axially grooved heat pipe for a space application," in *Proc. IEEE RAST*, Istanbul, Turkey, Jun. 2015.

[145] M. Nakagawa, E. Morikawa, Y. Koyama, R. Suzuki, and Y. Yasuda, "Development of thermal control for phased array antenna," in *Proc. AIAA ICSSC*, Yokohama, Japan, Apr. 2003.

[146] J. Wilson, "Challenges in thermal control of military electronics systems," *Electron. Cooling Mag.*, vol. 9, pp. 18–24, Feb. 2003.

[147] B. J. Döring, "Cooling system for a Ka band transmit antenna array," German Aerospace Center (DLR), Köln, Germany, Tech. Rep. IB554-06/02, Dec. 2005.

[148] C. J. M. Lasance and R. E. Simons, "Advances in high-performance cooling for electronics," *Electron. Cooling Mag.*, vol. 11, pp. 22–39, Nov. 2005.

[149] B. Hanafi, "Design of silicon power ampliers and arrays for millimeter wave applications," Ph.D. dissertation, Dept. Elect. Eng., Univ. California, San Diego, CA, USA, 2014.

[150] S. Shakib, H.-C. Park, J. Dunworth, V. Aparin, and K. Entesari, "A 28 GHz efficient linear power amplifier for 5G phased arrays in 28nm bulk CMOS," in *Proc. ISSCC 2016*, San Francisco, CA, USA, Feb. 2016.

[151] B. Rohrdantz, K. Kuhlmann, A. Stark, A. Geise, and A. F. Jacob, "Digital beamforming antenna array with polarisation multiplexing for mobile highspeed satellite terminals at Ka-band," *IET J. Eng.*, vol. 2016, no. 6, pp. 180–188, Feb. 2016.

[152] "TGA4509-EPU, 27-31 GHz 1 W power amplifier data sheet," TriQuint Semiconductor Texas, USA, Jan. 2005.

[153] "NEC develops antenna heat dissipation technology enabling compact 5G-compatible radio units," *NEC Corporation, Press Release*, Sept. 2018.

[154] A. Alnukari, P. Guillemet, Y. Scudeller, and S. Toutain, "Active heatsink antenna for radio-frequency transmitter," *IEEE Trans. Adv. Packag.*, vol. 33, no. 1, pp. 139–146, Feb. 2010.

[155] L. Covert and J. Lin, "Simulation and measurement of a heatsink antenna: A dual-function structure," *IEEE Trans. Antennas Propag.*, vol. 54, no. 4, pp. 1342–1345, Apr. 2006.

[156] J. J. Casanova, J. A. Taylor, and J. Lin, "Design of a 3-D fractal heatsink antenna," *IEEE Antennas and Wireless Propagation Letters*, vol. 9, pp. 1061–1064, Nov. 2010.

[157] V. Lakshminarayanan and N. Sriraam, "The effect of temperature on the reliability of electronic components," in *Proc. IEEE CONECCT*, Bangalore, India, Jan. 2014.

[158] D. Nicholls, *"System Reliability Toolkit.* Reliability Information Analysis Center / Data and Analysis Center for Software, Dec. 2005.

[159] J. Wei, "Challenges in cooling design of CPU packages for high-performance servers," *Heat Transfer Eng.*, vol. 29, no. 2, pp. 178–187, 2008.

[160] "Power semiconductor cooling solutions," Thermacore Inc., Lancaster, USA.

[161] "AC401 CPU air cooler data sheet," FSP Technology Inc., Taiwan.

[162] S. Shidore, "Compact thermal modeling in electronics design," *Electronics Cooling*, vol. 13, no. 2, May 2007.

[163] "Two-resistor compact thermal model guideline," JESD-1, JEDEC, Jul. 2008.

[164] "DELPHI compact thermal model guideline," JESD-4, JEDEC, Oct. 2008.

[165] A. Chen, "How to properly evaluate junction temperature with thermal metrics," Texas Instruments, Dallas, Texas, Tech. Rep. SLUA844A, Mar. 2018.

[166] "Mugen MAX CPU cooler SCMGD-1000," Scythe EU GmbH, Germany, Jul. 2014.

[167] "NH-L9x65 65-mm low profile cooler for HTPC and small form factor," Noctua, Austria, Mar. 2015.

[168] P. Mankowski, A. Dominiak, R. Domanski, M. J. Kruszewski, and L. Ciupinski, "Thermal conductivity enhancement of copper–diamond composites by sintering with chromium additive," *Journal of Thermal Analysis and Calorimetry*, vol. 116, no. 2, pp. 881–885, 2014.

[169] J. Fukai, M. Kanou, Y. Kodama, and O. Miyatake, "Thermal conductivity enhancement of energy storage media using carbon fibers," *Energy Conversion and Management*, vol. 41, no. 14, pp. 1543–1556, 2000.

[170] A. Bejan, "Constructal-theory network of conducting paths for cooling a heat generating volume," *International Journal of Heat and Mass Transfer*, vol. 40, no. 4, pp. 799–811, 1997.

[171] L. Ghodoossi, "Conceptual study on constructal theory," *Energy Conversion and Management*, vol. 45, no. 9, pp. 1379–1395, 2004.

[172] L. Chen, "Progress in study on constructal theory and its applications," *Science China Technological Sciences*, vol. 55, no. 3, pp. 802–820, 2012.

[173] Z. Guo, H. Zhu, and X. Liang, "Entransy a physical quantity describing heat transfer ability," *International Journal of Heat and Mass Transfer*, vol. 50, no. 13, pp. 2545–2556, 2007.

[174] Q. Chen, H. Zhu, N. Pan, and Z. Guo, "An alternative criterion in heat transfer optimization," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 467, no. 2128, pp. 1012–1028, 2011.

[175] Q. Chen, X. Liang, and Z. Guo, "Entransy theory for the optimization of heat transfer–a review and update," *International Journal of Heat and Mass Transfer*, vol. 63, pp. 65–81, 2013.

[176] X. Cheng, Z. Li, and Z. Guo, "Constructs of highly effective heat transport paths by bionic optimization," *Science in China Series E: Technological Sciences*, vol. 46, no. 3, pp. 296–302, 2003.

[177] Z. Xia, X. Cheng, and Z. Li, "Bionic optimization of heat transport paths for heat conduction problems," *Journal of Enhanced Heat Transfer*, vol. 11, no. 2, pp. 119–132, 2004.

[178] X. Xu, X. Liang, and J. Ren, "Optimization of heat conduction using combinatorial optimization algorithms," *International Journal of Heat and Mass Transfer*, vol. 50, no. 9, pp. 1675–1682, 2007.

[179] X. Cheng, X. Xu, and X. Liang, "Homogenization of temperature field and temperature gradient field," *Science in China Series E: Technological Sciences*, vol. 52, no. 10, pp. 2937–2942, 2009.

[180] R. R. Madadi and C. Balaji, "Optimization of the location of multiple discrete heat sources in a ventilated cavity using artificial neural networks and micro genetic algorithm," *International Journal of Heat and Mass Transfer*, vol. 51, no. 9, pp. 2299–2312, 2008.

[181] T. V. V. Sudhakar, C. Balaji, and S. P. Venkateshan, "Optimal configuration of discrete heat sources in a vertical duct under conjugate mixed convection using artificial neural networks," *International Journal of Thermal Sciences*, vol. 48, no. 5, pp. 881–890, 2009.

[182] T. K. Hotta, C. Balaji, and S. P. Venkateshan, "Optimal distribution of discrete heat sources under mixed convection a heuristic approach," *Journal of Heat Transfer*, vol. 136, no. 10, p. 104503, 2014.

[183] ——, "Experiment driven ann-ga based technique for optimal distribution of discrete heat sources under mixed convection," *Experimental Heat Transfer*, vol. 28, no. 3, pp. 298–315, 2015.

[184] S. Soleimani, D. D. Ganji, M. Gorji, H. Bararnia, and E. Ghasemi, "Optimal location of a pair heat source-sink in an enclosed square cavity with natural convection through pso algorithm," *International Communications in Heat and Mass Transfer*, vol. 38, no. 5, pp. 652–658, 2011.

[185] K. Chen, S. Wang, and M. Song, "Optimization of heat source distribution for two-dimensional heat conduction using bionic method," *International Journal of Heat and Mass Transfer*, vol. 93, pp. 108–117, 2016.

[186] ——, "Temperature-gradient-aware bionic optimization method for heat source distribution in heat conduction," *International Journal of Heat and Mass Transfer*, vol. 100, pp. 737–746, 2016.

[187] K. Chen, J. Xing, S. Wang, and M. Song, "Heat source layout optimization in two-dimensional heat conduction using simulated annealing method," *International Journal of Heat and Mass Transfer*, vol. 108, pp. 210–219, 2017.

[188] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted l1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 877–905, 2008.

[189] M. D'Urso and G. Prisco, "Maximally sparse arrays via sequential convex optimizations," *IEEE Antennas and Wireless Propagation Letters*, vol. 11, pp. 192–195, 2012.

[190] M. D'Urso, G. Prisco, and R. M. Tumolo, "Maximally sparse, steerable, and non-superdirective array antennas via convex optimizations," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 9, pp. 3840–3849, 2016.

[191] S. E. Nai, W. Ser, Z. L. Yu, and H. Chen, "Beampattern synthesis for linear and planar arrays with antenna selection by convex optimization," *IEEE Transactions on Antennas and Propagation*, vol. 58, no. 12, pp. 3923–3930, 2010.

[192] A. S. Reimer and A. F. Cheviakov, "A matlab-based finite-difference solver for the poisson problem with mixed dirichlet–neumann boundary conditions," *Computer Physics Communications*, vol. 184, no. 3, pp. 783–798, 2013.

[193] H. Q. Ngo, *Massive MIMO: Fundamentals and System Designs*. Nottingham, U.K: Linkoping Univ. Electronic Press, 2015.

[194] P. Wongchampa and M. Uthansakul, "Orthogonal beamforming for multiuser wireless communications: Achieving higher received signal strength and throughput than with conventional beamforming," *IEEE Antennas Propag. Mag.*, vol. 59, no. 4, pp. 38–49, Jun. 2017.

[195] E. Ali, M. Ismail, R. Nordin, and N. F. Abdulah, "Beamforming techniques for massive MIMO systems in 5G: overview, classification, and trends for future research," *Front. Inform. Technol. Electron. Eng.*, vol. 18, no. 6, pp. 753–772, Jun. 2017.

[196] C. Bencivenni, "Aperiodic array synthesis for telecommunications," Ph.D. dissertation, Dept. Elect. Eng., Chalmers Univ. of Tech., Gothenburg, Sweden, 2017.

[197] C. Bencivenni, A. A. Glazunov, R. Maaskant, and M. V. Ivashina, "Effects of regular and aperiodic array layout in multi-user MIMO applications," in *Proc. IEEE USNC/URSI NRSM*, San Diego, CA, USA, Jul. 2017.

[198] M. C. Vigano, "Sunflower array antenna for multi-beam satellite applications," Ph.D. dissertation, Dept. Elect. Eng., Math. and Comp. Sci., Delft Univ. of Tech., Delft, The Netherlands, Oct. 2011.

[199] M. C. Vigano, G. Toso, G. Caille, C. Mangenot, and I. E. Lager, "Sunflower array antenna with adjustable density taper," *Int. J. Antennas Propag.*, vol. 2009, Jan. 2009.

[200] C. Bencivenni, M. V. Ivashina, and R. Maaskant, "Synthesis of circular isophoric sparse arrays by using compressive-sensing," in *Proc. IEEE APS URSI*, Fajardo, Puerto Rico, Jul. 2016.

[201] Z. N. Chen and K. Hirasawa, "On the effect of ground-plane thickness on an aperture-coupled dielectric resonator antenna," *Int. J. RF. Microw. C. E.*, pp. 271–277, Apr. 2000.

[202] J. M. Lee, Y. H. Cho, C. S. Pyo, and I. G. Choi, "A 42-GHz wideband cavity-backed slot antenna with thick ground plane," *ETRI J.*, vol. 26, no. 3, pp. 262–264, Jun. 2004.

[203] S. Masihi, P. Rezaei, and M. Panahi, "Compact chip-resistor loaded active integrated patch antenna for ISM band applications," *Wireless Pers. Commun.*, vol. 97, no. 4, pp. 5733–5746, Dec. 2017.

[204] P. R. Haddad and D. M. Pozar, "Analysis of an aperture coupled microstrip patch antenna with a thick ground plane," in *Proc. APS/URSI*, Seattle, USA, Jun. 1994.

[205] J. Cuntala, P. Spanik, M. Frivaldsky, and A. Kondelova, "Replacement of multilayer PCB for equivalent composite board at temperature simulation in COMSOL environment," in *Tech. Comp. Prague*, Prague, Czech Republic, 2013.

[206] E. Monier-Vinard, C. T. Dia, V. Bissuel, O. Daniel, and N. Laraqi, "Extension of the DELPHI methodology to dynamic compact thermal model of electronic component," in *Proc. IEEE THERMINIC*, Paris, France, Sep. 2011.

[207] E. Monier-Vinard, C. T. Dia, V. Bissuel, N. Laraqi, and O. Daniel, "Dynamic compact thermal model for stacked-die components," in *Proc. 28th IEEE SEMI-THERM*, San Jose, CA, USA, Mar. 2012.

[208] C. Blair, S. L. Ruiz, and M. Morales, "5G, a multiphysics simulation vision from antenna element design to systems link analysis," in *Proc. IEEE ICEAA*, Granada, Spain, Sep. 2019.

[209] C. Fager, K. Hausmair, T. Eriksson, and K. Buisman, "Analysis of thermal effects in active antenna array transmitters using a combined em/circuit/thermal simulation technique," in *Proc. IEEE INMMiC Workshop*, Taormina, Italy, Oct. 2015.

[210] E. Baptista, K. Buisman, J. C. Vaz, and C. Fager, "Analysis of thermal coupling effects in integrated MIMO transmitters," in *Proc. IEEE MTT-S IMS*, Honolulu, HI, USA, Jun. 2017.

List of Acronyms

5G	Fifth Generation
3-D	Three Dimensional
2-D	Two Dimensional
ABF	Analog Beam Forming
AC-MPA	Aperture Coupled Microstrip Patch Antenna
AD/DA	Analog-to-Digital/Digital-to-Analog
AESA	Active Electronically Scanned Array
AF	Array Factor
AiP	Antenna in Package
AMSPAA	Active Multiport Subarray Phased Array Antenna
AWGN	Additive White Gaussian Noise
BF	Beam Forming
BM	Butler Matrix
BO	Bionic Optimization
BPF	Band Pass Filter
BS	Base Station
BW	Bandwidth
CB	Conjugate Beamforming
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CFD	Computational Fluid Dynamics
CO	Convex Optimization
CSI	Channel State Information
DBF	Digital Beam Forming

DPX/S	Duplexer/Switch
EEP	Embedded Element Pattern
EIRP	Effective Isotropically Radiated Power
FDMA	Frequency Division Multiple Access
FLOP	Floating Point Operations
FNBW	First Null Beam Width
FSS	Frequency Selective Surface
GA	Genetic Algorithm
HBF	Hybrid Beam Forming
HPBW	Half Power Beam Width
HTC	Heat Transfer Coefficient
IC	Integrated Circuit
INR	Interference to Noise Ratio
IPM	Interior Point Method
LNA	Low Noise Amplifier
LoS	Line of Sight
MAC	Medium Access Control
MBA	Multi Beam Array
MBO	Multiple Beam Optimization
MBPAA	Multiple Beam Phased Array Antenna
MC	Mutual Coupling
MFBA	Multiple Fixed Beam Antenna
MIMO	Multiple Input Multiple Output
NF	Noise Figure
NLoS	Non Line of Sight
PA	Power Amplifier
PCB	Printed Circuit Board
QoS	Quality of Service

Rx	Receiver
SA	Simulated Annealing
SDMA	Space Division Multiple Access
SINR	Signal to Interference plus Noise Ratio
SIW	Substrate Integrated Waveguide
SLL	Side Lobe Level
SNR	Signal to Noise Ratio
SOCP	Second Order Cone Programming
TDMA	Time Division Multiple Access
TIM	Thermal Interface Material
Tx	Transmit
UE	User Equipment
ZF	Zero Forcing

ACKNOWLEDGEMENTS

To everyone who has helped me and stood with me throughout my journey - THANK YOU! Thank you for your time, your knowledge, your patience, your investment and so much more.

First of all, I would like to express my sincere gratitude to my promotor Prof. Alexander Yarovoy for his confidence in me and his continuous support during my Ph.D. study. His progressive guidance, encouragement and advice helped me a lot in the past 4 years. He has always been very supportive and has given me the freedom to pursue various research directions and find my own way. It is thanks to him that I have become a motivated and independent researcher.

Second, I am deeply grateful to my advisor Dr. Antoine Roederer for his endless guidance. His scientific advice and immense knowledge on antennas has led to many insightful discussions and suggestions, even on Whatsapp! He played a major role in helping me define my research topics, obtain novel results and present them in the best possible way. It was both fun and inspiring to have Dr. Roederer in the team. I feel very lucky to get to know him. His enthusiasm and love for scientific work is contagious like Covid-19. I hope that I could be as enthusiastic and energetic as him at his age.

Third, I would like to thank my daily supervisor and my office roommate Dr. Jan Puskely. Not every Ph.D. student has the privilege to have a person like Jan who is always willing to help. His wide knowledge and experience on antenna design and simulation has been invaluable for the progress of my work. Besides the technical part, I have enjoyed the time sharing the office with Jan. His friendship and positive attitude has made the hard work easier.

I would also like to thank the Netherlands Organisation for Scientific Research (NWO) and NXP Semiconductors for their financial support during my research. Special thanks to the project leader Marcel Geurts and the company engineers Paul Mattheijssen and John Janssen for their helpful comments on my research. I also give thanks to my fifth-generation Ph.D. colleagues at TU Delft, TU Eindhoven, KU Leuven and Chalmers for the collaboration and the fruitful exchange of ideas we had during the regular meetings and conferences.

I also wish to acknowledge the help of Pascal Aubry for the antenna measurements, Dr. Bert-Jan Kooij for the Dutch translations of my summary and propositions, and Minke van der Put, Minaksie Ramsoekh and Esther de Klerk for their administrative assistance.

I would like to thank my Ph.D. defense committee members as well, Prof. Dr. Nuria Llombart Juan, Prof. Dr. Leo de Vreede, Prof. Dr. Geert Leus, Prof. Dr. Christian Fager, Prof. Dr. Christophe Craeye and Dr. Marcel Geurts for accepting to be a part of the committee, for their time and comments.

To all the former and present colleagues in the MS3 group, thank you for the insightful discussions and the fun time we spent together (even when we got stuck in the el-

evator for an hour!). Of the many, I want to thank: Nikita Petrov, Jianping Wang, Max Schöpe, Utku Kumbul, Shengzhi Xu, Ozan Dogan, Ronny Guendel, Matteo Unterhorst, Bert-Jan Kooij, Hans Driessen, Francesco Fioranelli, Oleg Krasnov, Faruk Uysal, Pascal Aubry, Peter Swart, Dinh Tran, Etienne Goossens, Fred van der Zwan, Yun Lu, Nannan Chen, Rossiza Gourova, Tworit Dash, Merel Verhoef, Nick Cancrinus, Salman, Saravanan Nagesh, Nikola Bogdanovic, Inna Ivashko, Stefano Medagli, Arun Muraleedharan, Sharef Neemat, Albert OudeNijhuis and Shilong Sun. It was unfortunate to go through the last phase of my Ph.D. without you (or I must say with you, but virtually). I hope we will celebrate my graduation soon, and not on Zoom!

Special thanks go to Prof. Luca Perregrini and my Italian colleagues from the Microwave Lab at the University of Pavia for the very nice hospitality and collaboration during my short internship.

To my former and present housemates in Delft, the lovely people of Group 8: Annalena, Suzanne, Helmer, Eva, Konstanze, Ruben, Julia, Floortje, Anathe and Evert, thank you for being such nice neighbors and caring friends during these years. They say "get the right neighbors rather than the right house". I think I have got both thanks to you guys!

To my precious and lifelong brothers/sisters from Turkey: Dogancan, Ekin, Selardis, Lacin, Cosku, Cemre, Zeynep, Busra, Taylan, Ipek, Ali Can, thank you for your support and friendship. It has always been so relaxing and fun to see you guys in Turkey after a long period of hard work. I feel very lucky to be surrounded by wonderful people like you!

I am infinitely grateful to my fiancée, my love, Buse for her constant encouragement and unwavering support. Thank you for your faith in me and for sticking by me, even when I was depressed and unbearable. Thank you for being so positive and encouraging during the chaotic times. I know that our commitment and love to each other will remain immortal and will help us get through difficult situations. I am looking forward to being in lockdown with you for the rest of our lives.♡

Last and the most important, I would like to thank to my family, my mother, my father and my sister. I dedicate this thesis to you for your unconditional love, distance-independent support, encouragement and sacrifice. I know you always wanted me to be a doctor, but you did not say a medical doctor. So, here we are. I hope I made you proud! I love you with all my heart.

Dear reader, I also would like to thank to you for reading this thesis to the end. I sincerely hope that the provided information will be useful and inspiring to you!

CURRICULUM VITÆ

Yankı ASLAN

BIOGRAPHY

Mr. Aslan was born on May 29, 1991 in Ankara, Turkey. He received his BSc degree in 2014 from Middle East Technical University, Department of Electrical and Electronic Engineering, with double specialization in Communications and Microwaves & Antennas. With the Justus & Louise van Effen Scholarship from Delft University of Technology, he completed his MSc in Electrical Engineering, Telecommunications and Sensing Systems track with Cum Laude Honor in 2016. In October 2016, he joined the Microwave Sensing, Signals and Systems group at TU Delft as a PhD candidate. He has worked on the project "Antenna Topologies and Front-end Configurations for Multiple Beam Generation" which is a part of the NWO & NXP Partnership Program on "Advanced 5G Solutions". His work has resulted in 10 journal publications and 10 conference papers. With his achievements during his PhD project, he received the IEEE APS Doctoral Research Grant in 2018 and the first EuMA Internship Award in 2019. With the latter award, he carried out his internship in the University of Pavia, Italy where he worked on Substrate Integrated Waveguide based leaky wave antenna array designs for 5G. He is a member of IEEE Antennas and Propagation Society and he serves as a reviewer for IEEE Transactions on Antennas and Propagation, IEEE Antennas and Wireless Propagation Letters, International Journal of Antennas and Propagation, IEEE Access and Microwave and Optical Technology Letters. His main research interests include multibeam antennas, array optimization methods, antenna front-end architectures, beamforming algorithms, communication system modeling and antenna cooling.

EDUCATION

2016–2020	PhD. Electrical Engineering Delft University of Technology, The Netherlands *Thesis:* Antenna Array Synthesis and Beamforming for 5G Applications: A Multidisciplinary Approach *Promotor:* Prof. dr. A.G. Yarovoy
2014–2016	MSc. Electrical Engineering Delft University of Technology, The Netherlands Telecommunications and Sensing Systems Track C.GPA: 9.11/10 *Thesis:* Reduction of Mutual Coupling Between Closely Spaced Patch Antennas Using Dielectric Contrast Techniques *Promotor:* Prof. dr. A.G. Yarovoy
2009–2014	BSc. Electrical Engineering Middle East Technical University, Turkey Communications, Microwaves and Antennas Area C.GPA: 4.00/4.00

AWARDS

2019	European Microwave Association (EuMA) Internship Award
2018	IEEE Antennas and Propagation Society Doctoral Research Grant
2016	Master of Science - Cum Laude, TU Delft, The Netherlands
2014–2016	Justus & Louise Van Effen Scholarship, TU Delft, The Netherlands
2014	1st rank graduate from the Middle East Technical University, Turkey

ACADEMIC CERTIFICATES

2018	Certificate of Completion ESoA Course on Short Range Propagation Technische Universitat Braunschweig, Germany
2018	Certificate of Completion ESoA Course on Arrays and Reflectarrays Université catholique de Louvain, Belgium
2018	Certificate of Acknowledgement ESoA Course on 5G Phased Arrays Delft University of Technology, The Netherlands

LIST OF PUBLICATIONS

Journal Papers:

[J-1] **Y. Aslan**, J. Puskely, A. Roederer and A. Yarovoy, "Trade-offs between the quality of service, computational cost and cooling complexity in interference-dominated multi-user SDMA systems," *IET Comm.*, vol. 14, no. 1, pp. 144-151, Jan. 2020.

[J-2] J. Puskely, T. Mikulasek, **Y. Aslan**, A. Roederer and A. Yarovoy, "5G SIW based phased antenna array with cosecant-squared shaped pattern," *IEEE Trans. Antennas Propag.*, 2020, under review.

[J-3] **Y. Aslan**, J. Puskely, A. Roederer and A. Yarovoy, "Phase-only control of peak sidelobe level and pattern nulls using iterative phase perturbations," *IEEE Antennas Wirel. Propag. Lett.*, vol. 18, no. 10, pp. 2081–2085, Oct. 2019.

[J-4] **Y. Aslan**, J. Puskely, A. Roederer and A. Yarovoy, "Multiple beam synthesis of passively cooled 5G planar arrays using convex optimization," *IEEE Trans. Antennas Propag.*, vol. 68, no. 5, pp. 3557-3566, May 2020.

[J-5] **Y. Aslan**, J. Puskely, A. Roederer and A. Yarovoy, "Synthesis of quasi-modular circularly polarized 5G base station antenna arrays based on irregular clustering and sequential rotation," *Microw. Opt. Technol. Lett.*, 2020, under review.

[J-6] **Y. Aslan**, J. Puskely, J. H. J. Janssen, M. Geurts, A. Roederer and A. Yarovoy, "Thermal-aware synthesis of 5G base station antenna arrays: an overview and a sparsity-based approach," *IEEE Access*, vol. 6, pp. 58868-58882, Oct. 2018.

[J-7] **Y. Aslan**, J. Puskely and A. Yarovoy, "Heat source layout optimization for two-dimensional heat conduction using iterative reweighted L1-norm convex minimization," *Int. J. Heat Mass Transf.*, vol. 122, pp. 432-441, Jul. 2018.

[J-8] J. Puskely, **Y. Aslan**, A. Roederer and A. Yarovoy, "Analog beamforming cosecant squared pattern active phased array antennas for 5G cellular systems," *IEEE Antennas Propag. Mag.*, 2020, in preparation.

[J-9] **Y. Aslan**, A. Roederer and A. Yarovoy, "System advantages of using large-scale aperiodic array architectures in mm-wave 5G base stations: an interdisciplinary look," *IEEE Syst. J.*, 2020, under review.

[J-10] **Y. Aslan**, A. Roederer and A. Yarovoy, "Synthesis of optimal 5G array layouts with wide-angle scanning and zooming ability for efficient link setup and high-QoS communication," *IEEE Antennas Wirel. Propag. Lett.*, vol. 19, no.9, Sep. 2020.

Conference Proceedings:

[C-1] **Y. Aslan**, S. Salman, J. Puskely, A. Roederer and A. Yarovoy, "5G multi-user system simulations in line-of-sight with space-tapered cellular base station phased arrays," in *Proc. 13th EuCAP*, Krakow, Poland, Apr. 2019.

[C-2] **Y. Aslan**, J. Puskely, A. Roederer and A. Yarovoy, "Performance comparison of single- and multi-lobe antenna arrays in 5G urban outdoor environments at mm-waves via intelligent ray tracing," in *Proc. 14th EuCAP*, Copenhagen, Denmark, Mar. 2020.

[C-3] **Y. Aslan**, J. Puskely, A. Roederer and A. Yarovoy, "Active multiport subarrays for 5G communications," in *Proc. IEEE APWC*, Granada, Spain, pp. 298-303, Sept. 2019.

[C-4] J. Puskely, **Y. Aslan**, A. Roederer and A. Yarovoy, "SIW based antenna array with power equalization in elevation plane for 5G base stations," in *Proc. 12th EuCAP*, London, UK, Apr. 2018.

[C-5] **Y. Aslan**, J. Puskely, A. Roederer and A. Yarovoy, "Synthesis of multiple beam linear arrays with uniform amplitudes," in *Proc. 12th EuCAP*, London, UK, Apr. 2018.

[C-6] **Y. Aslan**, M. Candotti and A. Yarovoy, "Synthesis of multi-beam space-tapered linear arrays with side lobe level minimization in the presence of mutual coupling," in *Proc. 13th EuCAP*, Krakow, Poland, Apr. 2019.

[C-7] **Y. Aslan**, C. E. Kiper, A. Biggelaar, U. Johannsen and A. Yarovoy, "Passive cooling of mm-wave active integrated 5G base station antennas using CPU heatsinks," in *Proc. 16th EuRAD*, Paris, France, pp. 121-124, Oct. 2019.

[C-8] **Y. Aslan**, J. Puskely, A. Roederer and A. Yarovoy, "Effect of element number reduction on inter-user interference and chip temperatures in passively-cooled integrated antenna arrays for 5G," in *Proc. 14th EuCAP*, Copenhagen, Denmark, Mar. 2020.

[C-9] **Y. Aslan**, J. Puskely, A. Roederer and A. Yarovoy, "Heat transfer enhancement in passively cooled 5G base station antennas using thick ground planes," in *Proc. 13th EuCAP*, Krakow, Poland, Apr. 2019.

[C-10] S. Salman, **Y. Aslan**, J. Puskely, A. Roederer and A. Yarovoy, "System modeling and simulation in 5G: A hybrid beamforming approach with power flux equalization in the elevation plane," in *Proc. 49th EuMC*, Paris, France, pp. 746-749, Oct. 2019.

Thesis Co-Supervised:

[T-1] S. Salman, "Analysis of antenna and RF front-end topologies for multi-beam systems," *MSc. Thesis*, Delft University of Technology, Aug. 2018.

Awards:

[A-1] European Microwave Association (EuMA) Internship Award (2019).

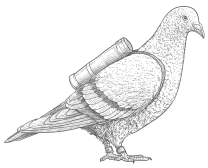
[A-2] IEEE Antennas and Propagation Society Doctoral Research Grant (2018-2019).

Invited Talks:

[W-1] **Y. Aslan**, "Thermal-aware multi-beam antenna synthesis in 5G," *Workshop on integration and multi-physics challenges in 5G mm-wave system design*, European Microwave Week (EuMW), Paris, France, Oct. 2019.

Invited Projects:

[P-1] Celestia Technologies Group UK, "Multipurpose and multibeam hybrid COM/EO ground station for satellite megaconstellations - MULTISCAN," *H2020 - Research and Innovation Action*, SPACE-29-TEC-2020: Satellite Communication Technologies.



"When wireless is perfectly applied,
the whole earth will be converted into a huge brain."
Nikola Tesla, 1926

This doctoral thesis presents the first ever irregular/sparse and subarray based reduced-complexity 5G base station antennas that are integrated with low temperature - high efficiency power amplifiers and have wide-angle scan multibeam capability, while providing low level of side lobes and radiation nulls. The developed antenna arrays and beam generation concepts could also have an impact over a broad range of applications where they should help overcome the capacity problem by use of multiple adaptive antennas, improve reliability and reduce interference.

Yankı Aslan
Delft, 2020