



AFRL-RI-RS-TR-2023-149

PROTOCOL STACK DESIGN FOR FREQUENCY-AGILE ULTRA-BROADBAND AIRBORNE NETWORKS IN THE THZ BAND

NORTHEASTERN UNIVERSITY

JULY 2023

FINAL TECHNICAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

STINFO COPY

**AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE**

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09. This report is available to the general public, including foreign nations. Copies may be obtained from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>).

AFRL-RI-RS-TR-2023-149 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE CHIEF ENGINEER:

/ S /

NGWE THAWDAR
Work Unit Manager

/ S /

GREGORY J. HADYNSKI
Assistant Technical Advisor
Computing and Communications Division
Information Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

1. REPORT DATE		2. REPORT TYPE		3. DATES COVERED	
JULY 2023		FINAL TECHNICAL REPORT		START DATE APRIL 2020	END DATE JANUARY 2023
4. TITLE AND SUBTITLE PROTOCOL STACK DESIGN FOR FREQUENCY-AGILE ULTRA-BROADBAND AIRBORNE NETWORKS IN THE THZ BAND					
5a. CONTRACT NUMBER N/A		5b. GRANT NUMBER FA8750-20-1-0200		5c. PROGRAM ELEMENT NUMBER 62788F	
5d. PROJECT NUMBER		5e. TASK NUMBER		5f. WORK UNIT NUMBER R2XW	
6. AUTHOR(S) Josep Miquel Jornet					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Northeastern University 360 Huntington Ave Boston MA 02115-5005				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory/RITGA 525 Brooks Road Rome NY 13441-4505			10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/ RI		11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-RI-RS-TR-2023-149
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited. This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The objective of this research is to develop the first network architecture and protocol stack to support ultrabroadband airborne networks in the terahertz (THz) band (0.1-10 THz), by following a bottom-up approach and through a closed-loop combination of mathematical modeling, numerical simulation and experimental evaluation. During this effort, contributions have been made along four thrusts. First, in relation to the development of experimental platforms for terahertz networking research, a software-defined-radio platform able to process 8 GHz of bandwidth in real-time at 120-140 GHz has been designed, implemented and demonstrated. Second, relating to the physical layer, a frequency-domain terahertz channel has been developed specifically to capture the non-homogeneous absorption across the airborne links. In addition, the impact of adverse weather on terahertz links has been experimentally studied by building and leveraging a state-of-the-art correlator-based ultrabroadband channel sounder. Tailored to the channel, low-complexity single-carrier modulations have been developed, including innovative hierarchical bandwidth modulations that can simultaneously transmit different streams to users within the same transmission beam. Third, receiver-initiated and dual-band medium access control protocols that overcome the challenges introduced by the need for high gain directional antennas simultaneously at the transmitter and the receiver at all times as well as the need to share the terahertz spectrum with passive users, have been designed, simulated and experimentally tested. Finally, a cross-layer framework to design optimal multi-hop relaying strategies have been developed, by building on the peculiarities learned in each one of the previous tasks, i.e., from hardware and channel to physical and link layers designs.					
15. SUBJECT TERMS Terahertz communications, ultrabroadband communications, hierarchical bandwidth modulation					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT SAR		18. NUMBER OF PAGES 131
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			
19a. NAME OF RESPONSIBLE PERSON NGWE THAWDAR				19b. PHONE NUMBER (Include area code) N/A	

Contents

1	SUMMARY	1
2	INTRODUCTION	2
3	ULTRABROADBAND NETWORKING TESTBED DEVELOPMENT	5
3.1	Motivation	5
3.2	Methods, Assumptions and Procedures	6
3.2.1	Overview of the Real-time Ultrabroadband Testbed	6
3.2.2	Transmitter Design and Implementation	7
3.2.3	Receiver Design and Implementation	9
3.3	Results and Discussion	10
3.3.1	Experimental Setup	10
3.3.2	Results	11
3.3.3	Future Enhancements to the Testbed	14
4	AIRBORNE TERAHERTZ CHANNEL MODELING	15
4.1	Motivation	15
4.2	Methods, Assumptions and Procedures	16
4.2.1	Scenario and Fundamentals	16
4.2.2	Atmospheric Path Loss Modeling in the Frequency Domain	17
4.2.3	Time-domain Channel Sounder	23
4.3	Results and Discussion	25
4.3.1	Numerical Results	25
4.3.2	Experimental Results	32
5	ULTRABROADBAND MODULATIONS FOR TERAHERTZ COMMUNICATIONS	39
5.1	Motivation	39
5.2	Methods, Assumptions and Procedures	40
5.2.1	Design of M-ary Amplitude Phase Shift Keying Modulations	40
5.2.2	Design of Hierarchical Bandwidth Modulations	45
5.3	Results and Discussion	61
5.3.1	Experimental Validation of Optimized APSK Modulations	61
5.3.2	Experimental Validation of HBM	62
6	MEDIUM ACCESS CONTROL FOR TERAHERTZ NETWORKS	67
6.1	Motivation	67
6.2	Methods, Assumptions and Procedures	68
6.2.1	Design and Performance Analysis of ADAPT	68
6.2.2	Dual-band MAC Protocol for Dynamic Spectrum Sharing above 100 GHz	76
6.3	Results and Discussion	79
6.3.1	Simulation-based Study of ADAPT	79
6.3.2	Experimental Validation of the Dynamic Dual-band Terahertz MAC Protocol	84
7	RELAYING AND ROUTING IN TERAHERTZ NETWORKS	88
7.1	Motivation	88
7.2	Methods, Assumptions and Procedures	89
7.2.1	System Preliminaries	89
7.2.2	Overview of the Quasi Birth Death Markov Process	92
7.2.3	Intra-level states and transition model based on THz MAC Protocol	93

7.2.4	Inter-level analysis of the QBD Markov process	100
7.3	Results and Discussion	103
7.3.1	Validation of the QBD Markov process Model	103
7.3.2	Performance Metrics	104
7.3.3	Numerical Analysis	105
8	CONCLUSIONS	110
9	REFERENCES	111
10	LIST OF ACRONYMS	120

List of Figures

1	Architecture of the developed real-time terahertz communication testbed.	7
2	Block diagram of the digital components in the proposed architecture.	8
3	Baseband in-phase component of Channel 1.	8
4	Block diagram of the analog components in the proposed architecture, supporting multiple channels: a) transmitter, b) receiver.	9
5	Capture of the intermediate frequency signal: a) injected at the transmitting frontend and b) collected at the receiver frontend.	9
6	Experimental setup for the real-time multi-channel terahertz communications testbed.	11
7	IQ constellation for each data sub-carrier in Channel 1, when transmitting a 16 QAM.	12
8	IQ constellations for a selection of data sub-carriers in: a) Channel 1, b) Channel 2, c) Channel 3, and d) Channel 4.	13
9	An illustration of the system model considered in this paper and the several wireless links in it. Picture of Earth by NASA/NOAA's GOES Project.	16
10	The system geometry for the angles and distances in the airplane to satellite system. Picture of Earth by NASA/NOAA's GOES Project.	20
11	System geometry on the angles of the transmission path through the atmosphere. Picture of Earth by NASA/NOAA's GOES Project.	20
12	Comparison of the layer thickness of the plane parallel atmosphere and a round atmosphere as a function of altitude.	22
13	A comparison of the total distance through the atmosphere according to the plane parallel assumption and the model herein as a function of the elevation angle. . . .	23
14	The NU Channel Sounder system: a) transmitter hardware, b) outdoor link, c) receiver hardware, and d) signal processing backend.	24
15	Atmospheric properties as a function of altitude.	25
16	The surface volume mixing ratio of water vapor (left) and the surface height from the sea level (right).	26
17	Rain and cloud attenuation for slant paths as well as for 45-degree angled paths from zenith.	27
18	Path loss from airplane to satellite for LEO and GEO orbits as a function of frequency. The airplane altitude is 11 km.	29
19	Path loss from airplane to satellite for 500 km LEO satellite as a function of frequency and airplane altitude.	29
20	Path loss from Earth to an airplane as a function of airplane altitude and frequency.	30
21	Path loss from Earth to satellite as a function of frequency.	30
22	Path loss from airplane to airplane (common altitude) for a 100-meter link as a function of airplane altitude and frequency.	31
23	Comparison of path losses per kilometer for FSPL, molecular absorption loss at sea level and 11 km altitude, and antenna gains of 0.5 and 1 meter ideal dish antennas. .	31
24	Comparison of A2G and A2S path losses at 660 and 940 GHz center frequencies as a function of the airplane altitude for zenith paths and inclined paths with crossover altitudes where of the minimum loss path.	32
25	Experimental and theoretical received power. Table: mean experimental and theoretical received power received and RMSE.	35

26	Multipath profiles for clear, rain, and snow weather.	36
27	The CDF of K-factor [dB] in different weather conditions for experimental (EXP) results and curve fitting (CF) values.	37
28	The CDF of RMS delay spread in different weather conditions for experimental (EXP) result and curve fitting (CF) value with the Stable and the Rician distributions.	37
29	Constellation design with design criteria: a) 16-APSK with two rings; (b) 32-APSK with three rings to depict the different scenarios.	41
30	SER comparison between M-QAM, M-PSK, and M-APSK with increasing SNR: a) M=16; (b) M=32.	44
31	Hierarchical 4/16-QAM Constellation	46
32	Amplitude and Time Distribution of (a) an example 4/16-QAM HBM System and (b) its frequency spectrum compared with (c) the amplitude and time distribution of a traditional 4/16-QAM HM System and (d) its frequency spectrum.	47
33	Block diagram for hierarchical bandwidth modulator	47
34	Constellation diagram for the non-uniform 4PAM modulation.	49
35	Constellation diagram for the non-uniform 6-PAM modulation observed by a low-resolution receiver.	50
36	4/M-QAM Constellation Diagram with $\lambda = \frac{d_2}{d_1}$	51
37	Constellation diagram for the non-uniform PAM modulation for the 4/M-QAM constellation.	53
38	HBM receiver's $\frac{E_s}{N_0}$ and SER dependence on λ and transmission distance.	55
39	A comparison of the high resolution error rates of HM and HBM for (a) $\lambda = 0.2$, (b) $\lambda = 0.4$, (c) $\lambda = 0.6$, (d) $\lambda = 0.8$	55
40	Achievable data rates for HM and HBM systems.	57
41	The Functional Region for the 4/16 QAM HBM System	58
42	Functional Region for the 4/64-QAM HBM System	59
43	Functional Region for the 4/16-QAM HBM System for various distances	60
44	Comparison of the SER for simulated HBM systems and the theoretical equations.	63
45	TeraNova Testbed	64
46	Received constellations for high- and low-resolution receivers at 3m and 12m respectively for a 4/16-QAM HBM system.	64
47	Received constellations for high and low-resolution receivers at 3m and 12m respectively for a 4/16-QAM HBM system.	65
48	Received Constellation by the far and near receivers in the 4/64-QAM system with $\lambda = 0.4$	65
49	Network architecture: A central turning AP divides space in sectors, and nodes are randomly distributed in a 2-D plane.	69
50	Packet flow in the case of unanswered CTA.	71
51	Packet flow of a 3-way handshake, data transmission and acknowledgement between AP and one client node.	71
52	Packet flow of a 3-way handshake, data transmission and acknowledgement for multiple active nodes in a sector.	72
53	Example of highest possible MCS depending on distance to the AP when using the primary sector (left), and MCS in the primary sector and adjacent sectors for a node close to the AP (right).	73

54	Prototype schematics and dual-band backhaul link deployment.	78
55	Throughput simulation and analytical results for a range of 18 m.	81
56	Packet discard rate simulation results for a range of 18 m.	81
57	Throughput simulation and analytical results for a range of 7.5 m. Fixed node density at 0.28 nodes/m ²	82
58	Cycle time of 1-way-2019 and ADAPT-3 for a range of 18 m.	83
59	Analytical throughput for ADAPT-3 as a function of packet size. Parameters: 50 nodes and 10 ms of T_{ia}	83
60	Experimental result for the dual-band self-adaptive backhaul link.	84
61	Link budget analysis, Aura satellite orbits, and spectrum.	86
62	Capacity in Gbps as a function of the bandwidth B	87
63	Network architecture.	89
64	Absorption loss defined 3 dB bandwidth.	91
65	QBD Markov process model of the system.	93
66	MAC protocol.	94
67	Entire probability vector of the QBD Markov process model in (126) ($D = 100$ m, $R = 100$ Gbps, $M = 50$ packets, $N = 3$).	103
68	Single hop throughput v.s. throughput of E2E multi-hop link ($D = 100$ m, $R =$ 100 Gbps, $M = 500$ packets).	105
69	Inspection of single node buffer status ($D = 100$ m, $R = 100$ Gbps, $M = 500$ packets).106	
70	The maximum buffer occupancy rate and the throughput of E2E multi-hop link with varying R ($D = 100$ m, $N = 5$, $M = 500$ packets).	107
71	The maximum buffer occupancy rate and the throughput of E2E multi-hop link with varying D ($R = 100$ Gbps, $N = 5$, $M = 500$ packets).	107
72	The throughput of E2E multi-hop link with varying R ($D = 1000$ m, $N = 5$, $M =$ 500 packets) and with varying D ($R = 100$ Gbps, $N = 5$, $M = 500$ packets). . . .	108

1 SUMMARY

The objective of this research is to develop the first network architecture and protocol stack to support ultrabroadband airborne networks in the terahertz (THz) band (0.1–10 THz), by following a bottom-up approach and through a closed-loop combination of mathematical modeling, numerical simulation and experimental evaluation. During this effort, contributions have been made along four thrusts. First, in relation to the development of experimental platforms for terahertz networking research, a software-defined-radio platform able to process 8 GHz of bandwidth in real-time at 120-140 GHz has been designed, implemented and demonstrated. Second, relating to the physical layer, a frequency-domain terahertz channel has been developed specifically to capture the non-homogeneous absorption across the airborne links. In addition, the impact of adverse weather on terahertz links has been experimentally studied by building and leveraging a state-of-the-art correlator-based ultrabroadband channel sounder. Tailored to the channel, low-complexity single-carrier modulations have been developed, including innovative hierarchical bandwidth modulations that can simultaneously transmit different streams to users within the same transmission beam. Third, receiver-initiated and dual-band medium access control protocols that overcome the challenges introduced by the need for high gain directional antennas simultaneously at the transmitter and the receiver at all times as well as the need to share the terahertz spectrum with passive users, have been designed, simulated and experimentally tested. Finally, a cross-layer framework to design optimal multi-hop relaying strategies have been developed, by building on the peculiarities learned in each one of the previous tasks, i.e., from hardware and channel to physical and link layers designs.

2 INTRODUCTION

Terahertz communication [1–5] is an upcoming wireless technology that can address three relevant challenges for the U.S. Air Force. First, sub-terahertz and terahertz frequencies (broadly, from 0.1 to 10 THz) remain almost completely unutilized for communication; hence, being the first to exploit this band provides an unmatched technological advantage. Second, the very high propagation loss at terahertz frequencies can be leveraged to create intrinsically secure air-to-air communication channels. On the one hand, the high spreading loss requires the utilization of ultra-directional systems, which enable ultra-low probability of intercept & detection (PI/PD) communications. On the other hand, atmospheric absorption at terahertz frequencies drastically limits the ability to intercept or disrupt air-to-air communications from ground systems, thus, further enhancing the security of the air-to-air links. Third, the terahertz band supports huge transmission bandwidths (from hundreds of gigahertz up to a few terahertz, depending on the transmission distance and medium composition). This enables wireless Terabit-per-second (Tbps) links and opens the door to transformative applications, such as in-air big-data sharing among unmanned airborne networks.

For decades, the terahertz band was referred to as the terahertz gap because of the lack of device technologies able to support communication and sensing applications at these frequencies. However, today, there are multiple solutions to realize the critical hardware building blocks of a terahertz system, including analog frontends, antenna systems, and digital signal processing (DSP) backends. In terms of *analog frontends*, which are in charge of generating, modulating, filtering, and amplifying signals at terahertz frequencies with multi-gigahertz bandwidths [6], high-power electronic devices [7, 8] able to emit a few hundred milliWatts at sub-terahertz frequencies have been demonstrated; very low-phase noise photonic devices built on optical downconversion have been built [9, 10]; and plasmonic nano-devices that intrinsically operate at terahertz frequencies have been proposed [11, 12]. When it comes to *antennas*, for the time being, experimental work has largely relied on fixed directional antennas, many times accompanied by lenses, able to provide gains all the way up to 55 dBi. However, there have recently been experimental demonstrations of on-chip antenna arrays [13, 14], that open the door to dynamic communication networks at terahertz frequencies. Finally, the *DSP engine* and, particularly, the high-speed data converters required to exploit the bandwidths available at terahertz frequencies, pose a critical challenge. For the time being, most experimental terahertz communication platforms rely on high-performance arbitrary waveform generators (AWG) and digital storage oscilloscopes (DSO) with data converters able to sample at a few hundred Gigasamples-per-second (GSps) [15, 16]. However, their size and power consumption (approaching 1 kW) limit their application to laboratory setups.

In parallel with the development of new device technologies, there has been significant progress when it comes to the study of the terahertz channel and the design of the physical layer of ground-based terahertz communication systems. In terms of *terahertz propagation*, mostly aided by the availability of terahertz channel sounding platforms primarily in the sub-terahertz band [17], channel studies have evolved from physics-based models [18, 19] to data-driven empirical models [20–22]. Similarly, there have been many works aimed at defining the modulation for terahertz systems, ranging from conventional single-carrier modulations to orthogonal frequency division modulation (OFDM) [23], single-carrier OFDM or DFT-Spread-OFDM [24], orthogonal time frequency space (OTFS) [25], DFT-Spread-OTFS [26], and different forms of spread spectrum [27, 28]. Similarly, beyond single-input single-output (SISO) systems, multiple-input multiple-output (MIMO) systems that can help reduce the bandwidth requirements (and, thus, relax the constraints on the DSP engine),

are also being considered, currently still with a large gap between what theory predicts [29, 30] and what experimentally we can show today [31, 32]. Still, the majority of these works have been focused on ground applications, and not much beyond the physical layer.

Overall, thanks to parallel major advancements in hardware technologies and communication solutions enabled by a better understanding of the terahertz channel, applications that once were thought of as impossible, have been demonstrated. For example, while terahertz communications were often thought for short-distance applications only, our team integrated by Northeastern University with the US Air Force Research Laboratory (AFRL) and the NASA Jet Propulsion Laboratory (JPL) has experimentally demonstrated multi-kilometer and multi-Gigabit-per-second (Gbps) links [33]. These results motivate the exploration of terahertz communications for non-terrestrial network applications, ranging from drones [34, 35] and planes to satellite communication networks [36, 37].

The objective of this research is **to develop the first network architecture and protocol stack to support ultrabroadband airborne networks in the terahertz band**, by following a bottom-up approach and through a closed-loop combination of mathematical modeling, numerical simulation and experimental evaluation. More specifically, contributions have been made along the following four intertwined thrusts:

- **Thrust 1: Ultrabroadband Networking Testbed Development:** We have designed, built and experimentally demonstrated the fastest software-defined-radio (SDR) platform existing to date, able to process 8 GHz of bandwidth in real-time. This platform leverages the state-of-the-art in Radio Frequency Systems on Chip (RFSoc) running a custom-developed multi-channel multi-carrier-based physical layer and a custom-designed analog frequency multiplexing network, all glued by tailored software. We have demonstrated its over-the-air performance when interfaced with analog front ends at 120-140 GHz, demonstrating effective data rates approaching 16 Gbps (physical-layer bit-rate of 32 Gbps). The platform is described in Sec. 3. A non-provisional application patent has been submitted to protect the generated intellectual property and potentially support a spin-out of this project.
- **Thrust 2: Ultrabroadband Physical Layer Design:** We have developed a mathematical model to characterize the airborne terahertz channel in the frequency domain, which takes into account the non-homogeneous properties of the atmosphere that impact molecular absorption. In addition, we have experimentally studied the impact of adverse weather phenomena on critical channel metrics, beyond power loss. More details are provided in Sec. 4. In addition, we have developed innovative low-complexity single-carrier modulations tailored to the hardware and the channel, including an optimized amplitude and phase shift keying (APSK) modulation that minimizes the peak-to-average power ratio (PAPR), and hierarchical bandwidth modulations, which leverage the distance-dependent channel bandwidth resulting from molecular absorption to effectively multiplex users even when they are within the same antenna transmission beam. These works, including their experimental validation, are described in Sec. 5. Moreover, our optimized APSK constellation is now included in the revised IEEE 802.15.3d standard [38].
- **Thrust 3: Ultrabroadband Link Layer Design:** We have developed new receiver-initiated Medium Access Control (MAC) protocols able to support user throughputs in excess of 100 Gbps when operating with the physical layer defined in the IEEE 802.15.3d standard. We

have evaluated the performance of the protocols both analytically and with simulations using TeraSim [39], our team's developed open-source extension for *ns-3* officially available in App Store. In addition, we have designed and experimentally demonstrated the first dual-band MAC protocol able to dynamically switch between two transmission windows within the terahertz band while minimally impacting on the user experience. Our work has been presented in the context of dynamic spectrum sharing between on-the-ground communication infrastructure passive sensing satellites orbiting the Earth, which have exclusive access to narrow bands that break up the otherwise large contiguous bandwidth at sub-terahertz frequencies. This work signifies the first experimental demonstration of a MAC protocol above 100 GHz. The details of the two solutions are provided in Sec. 6

- **Thrust 4: Ultrabroadband Network Layer Design:** Moving beyond (dynamic) point-to-point links and as a way to increase the coverage of terahertz networks, we have studied new cross-layer relaying strategies for multi-hop communications in the terahertz band. More specifically, we have derived a mathematical framework to smartly allocate the resources across layers by optimizing the distribution of relays and ultimately achieving the smallest end-to-end(E2E) delay. The developed framework builds on the outcomes of each of the previous thrusts and takes into account the effects from not only the physics of the hardware and the channel, e.g., transmission power, antenna beamwidth, transmission distance and operating bandwidth in the case of line-of-sight propagation, but also the MAC layer protocol design, e.g., interference and re-transmission, and the queuing delay when transmitting at multi-Gbps rates with buffer/memory limitations. The details on the framework, as well as numerical results, are presented in Sec. 7.

3 ULTRABROADBAND NETWORKING TESTBED DEVELOPMENT

3.1 Motivation

As mentioned in the introduction, developing a DSP engine able to process signals with multi-gigahertz bandwidth in real-time is a critical challenge. According to the Nyquist sampling theorem, the digital processing of the signals has to happen at a frequency f_s , that is greater than twice the information-bearing bandwidth B of the signal. When considering traditional signal processing backends, this necessitates on the one hand operating the data converters at rates beyond $2B$ samples per second and on the other hand having the digital processing circuits be clocked at the same speeds to perform signal processing in real-time.

On the one hand, the use of interleaved *data converter architectures* along with modern process technologies have led to digital-to-analog and analog-to-digital converters (DACs and ADCs, respectively) designs exceeding 100 GSps. For example, a 240 GSps DAC that supports 100 GHz of bandwidth is reported in [40]. A few commercial devices also exist that feature DACs that support up to 70 GHz of bandwidth [41]. Similarly, ADCs with 110 GHz of analog bandwidth sampling at 256 GSps are commercially available [42]. However, the size and power consumption of such data converters limit the utilization of such data converters in practical applications.

On the other hand, with Moore's law slowing down significantly, the clock frequencies that *digital hardware* can be operated at have become saturated. Currently, clock frequencies of application-specific integrated circuits (ASICs) using cutting-edge CMOS processes are maxed out at a few gigahertz. Since terahertz technologies and standards are still being figured out, it is still too early to think of ASICs and research communities currently need programmable digital backends to conduct dynamic experimentation. Programmable digital hardware like field programmable gate arrays (FPGAs) have slower fabric and the maximum realizable circuit speeds are much lower than the ASICs due to the programmability involved. The maximum operable frequencies of the state-of-the-art FPGAs for modest designs max out at about 500 MHz and this imposes a major hindrance in realizing real-time wireless links at terahertz frequencies that can support at least a few gigahertz of bandwidth. Thus there is a need of finding innovative solutions for alternative ways to do digital processing of ultrabroadband signals.

To digitally process signals with tens of gigahertz of bandwidth, one feasible approach is to utilize a channelized backend signal processing architecture where multiple low-bandwidth channels are frequency multiplexed to form an ultrabroadband link. Such an architecture can be realized using an RFSoc [43] at a moderate cost. The RFSocs are state-of-the-art RF digital programmable hardware where multiple RF data converters are integrated onto the same chip alongside the programmable logic as a single system-on-chip. In the most recent generation (Gen-3) of RFSocs [44], each individual data converter supports up to 10 GSps sample rates at the DAC side and up to 5 GSps speeds at the ADC side. With multiple (8 or 16) data converters integrated into a single System on Chip (SoC), a single RFSoc can accommodate up to a few tens of GHz in overall bandwidth. Thus, such architectures can be used to implement a channelized real-time signal processing backend [45].

As part of this effort, we have designed, built and experimentally demonstrated the **fastest SDR platform as of today**, and interfaced it with our analog front ends at 120-140 GHz.

3.2 Methods, Assumptions and Procedures

3.2.1 Overview of the Real-time Ultrabroadband Testbed

System Architecture In Fig. 1, we illustrate a high-level overview of our real-time testbed with all its components and interconnections. At the uppermost layer, an external computer executes a high data-rate demanding application. The outgoing/incoming data streams pass between the application-running machine and the digital baseband processor through a high-speed interface such as a Peripheral Component Interconnect Express (PCIe) or Quad Small Form-Factor Pluggable (QSFP). The digital baseband processor is implemented in an RFSoc and performs all physical layer, MAC, and synchronization operations. The RFSoc contains two main components: the programmable logic (PL) and the processing system (PS). We utilize the PL mainly for the physical layer implementation due to its nature for parallelism and high-speed performance when compared to the PS. With that in mind, we program all the necessary wireless communication blocks on the PL alongside all the sub-systems necessary for the system to function. On the PS side, we program all the driver code and user space applications needed to interact with the PL to provide real-time control over the physical layer. We also envision the upper networking layers to be implemented on the PS as well. The RFSoc then utilizes the on-chip data converters to convert the digital data stream to analog or vice versa. At the output of the RFSoc, we connect the data converters via Sub-Miniature Version C (SMC) to Sub-Miniature Version A (SMA) cables to the analog chain. The analog chain acts as the middleman between the baseband RFSoc and the RF frontends. It receives the I and Q streams for each channel and modulates it with a unique frequency and finally performs channel multiplexing at the output to be passed to the front ends. Finally, the multiplexed 4 channels are then passed to our front ends via a SMA connection. At the front ends, we up-convert/down-convert the signal to/from 130 GHz by utilizing the frequency multiplying chains. More details on the analog frontends available to our team are provided in [46].

Waveform and Frame Format Selection As discussed in the introduction of this report, there are many ongoing efforts aimed at determining the waveform for sub-terahertz and terahertz communications. In our current implementation, we have chosen OFDM due to its intrinsic parallelized nature and simplistic channel equalization implementation, which facilitates the DSP tasks to perform. Due to the ultra-wide bandwidth supported by the front ends in our testbed, the stronger contribution to the overall channel frequency selectivity comes from the hardware electronics. Initial measurements using the 120-140 GHz frontends were carried out to determine the sub-carrier spacing such that the sub-channels can be regarded as frequency flat. The measurement showed that this is met for bandwidths ≤ 50 MHz per sub-carrier and thus an FFT size of 64 is used to design the PHY layer. Since the channel bandwidth is at 2.048 GHz, the subcarrier spacing for the implementation is 32 MHz. The digital cores are designed to support up to 64-QAM modulations with the option of adaptive modulations at the sub-carrier level. Accordingly, the physical layer frame structure has been implemented with close resemblance to IEEE 802.11a standard due to the choice of 64-point FFT size. This frame includes a short training field (STF) and a long-training field (LTF) for initial synchronization and channel estimation, followed by the data payload.

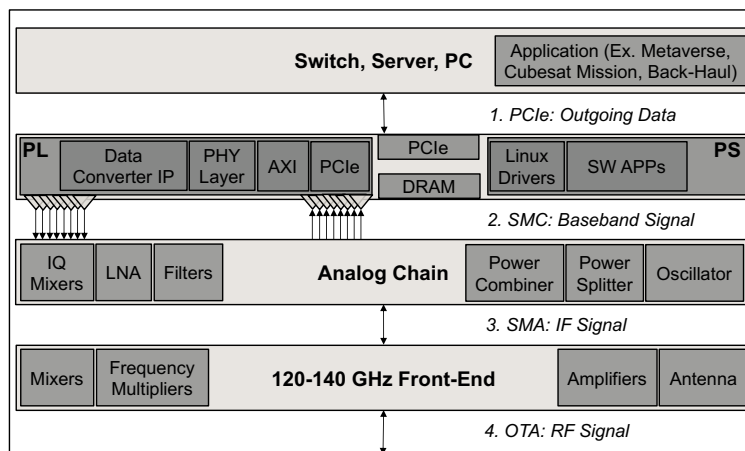


Fig. 1: Architecture of the developed real-time terahertz communication testbed.

3.2.2 Transmitter Design and Implementation

Digital backend The architecture of the transmitter-side digital blocks is shown in Fig. 2a. The transmitter blocks utilize the output clock provided by the on-chip DAC which is configured for 128 MHz, thus allowing the digital to be synchronized with the data-converter block. When parallelizing the digital model by 16 phases we achieve the target sample generation rate of 2.048 GSps. In Fig. 3, a DSO capture of the in-phase component is shown for the first baseband channel. The payload being sent for each packet occupies 52 out of the 64 sub-carriers. Following the IEEE 802.11a standard, 12 sub-carriers are nulled out to provide guard bands and avoid leakage between channels. It is important to note that the header we are utilizing is the same header from IEEE 802.11a and, thus, we can only estimate the channel for 52 sub-carriers and thus we disregard the remaining 12 sub-carriers. Nonetheless, due to the filters utilized post-up-conversion and a 500 MHz guard band between channels, we have seen that those 12 sub-carriers can be utilized for data without any significant change in the error vector magnitude (EVM).

Analog IF Network The on-chip DACs are capable of supporting a max sample rate of 6.554 GSps, however since the ADC can only support up to 4.096 GSps, we proceed with the minimum of the two and with two DACs employed per channel, we generate a 2.048 GHz wide baseband signal. The DAC outputs the baseband signal with 32 mA of output current, thus, roughly -26 dBm of power when measured with a DSO. The technique we utilize to allow us to transmit data over 8 GHz of bandwidth is to utilize all 4 data-converter pairs (2 DACs and 2 ADCs per channel) for simultaneous data transmission/reception via channel multiplexing. In Fig. 4a, we show a detailed transmitter diagram of our analog architecture with all the components used to achieve our results. Currently, we are relying on off-the-shelf analog components, but ultimately the goal is to develop a custom-designed integrated board. With this in mind, the proposed architecture yields the best results considering a multi-channel system. For instance, instead of amplifying the in-phase and

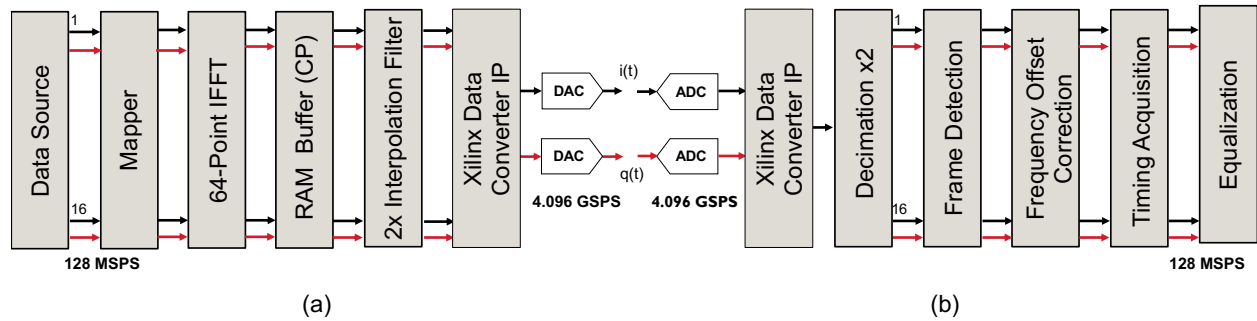


Fig. 2: Block diagram of the digital components in the proposed architecture.

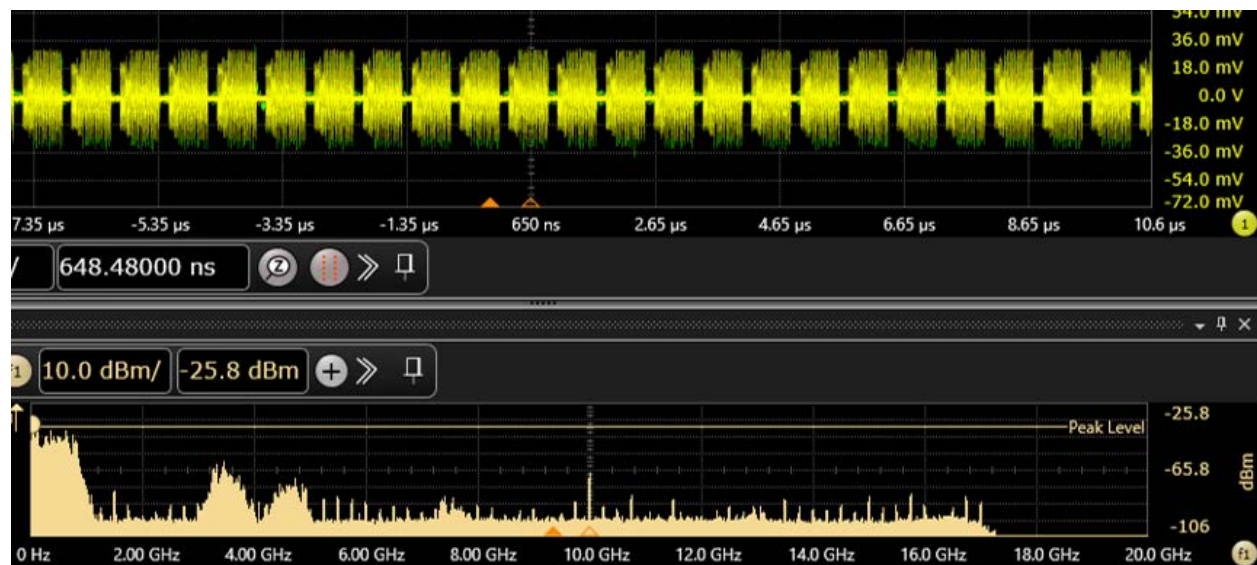


Fig. 3: Baseband in-phase component of Channel 1.

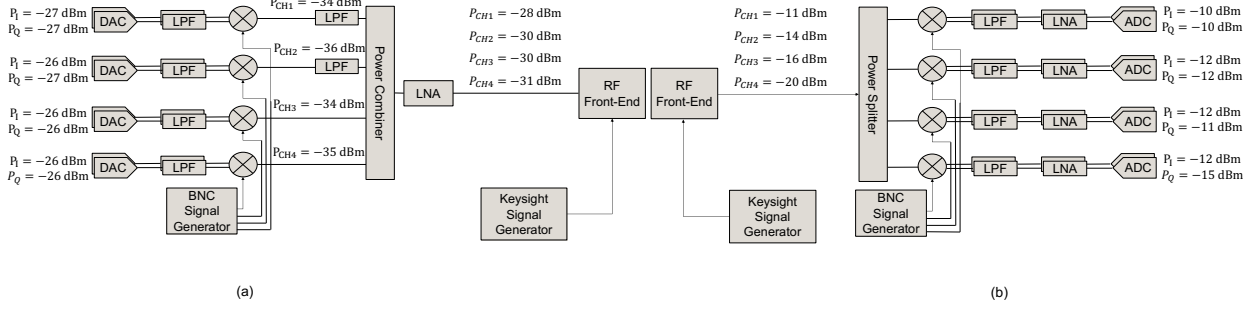


Fig. 4: Block diagram of the analog components in the proposed architecture, supporting multiple channels: a) transmitter, b) receiver.

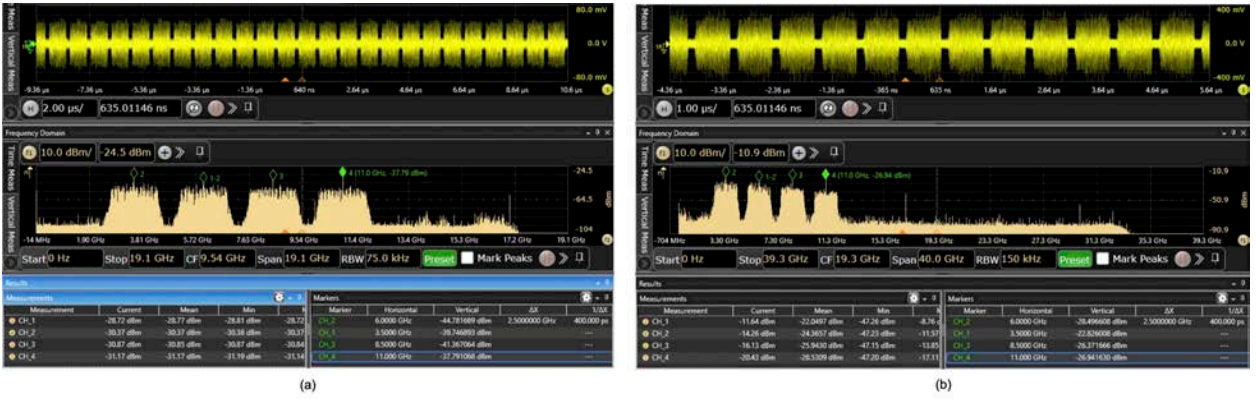


Fig. 5: Capture of the intermediate frequency signal: a) injected at the transmitting frontend and b) collected at the receiver frontend.

quadrature separately, we can amplify the modulated signal post-IQ mixing and reduce the impact of IQ imbalance. However, that would require a custom-made amplifier tailored for the frequency range of each channel. For instance, for Channel 1, an amplifier operating from 2.5 to 4.5 GHz would be needed such that it does not amplify the second copy produced by the IQ mixer which could potentially interfere with the other channels. However, since such amplifiers are not available off the shelf, we instead use a broadband amplifier to amplify all four channels after multiplexing them. Beyond this point, we drive the IQ mixers with a unique LO that is channel-dependent using a Berkeley Nucleonics (BNC) generator with 4 independent LO channels, and we utilize each channel to up-convert the signal during the IQ-mixing stage. In Fig. 5a, we show a DSO capture of the four channels being driven as input to the transmitter side RF frontend.

3.2.3 Receiver Design and Implementation

Digital backend In Fig. 2b, we show the high-level overview of the receiver DSP blocks starting from the decimation to the channel estimation and equalization. Given that the on-chip ADCs run

on a 256 MHz clock, we first convert the clock from 256 to 128 MHz and re-sample the incoming samples at that lower clock to keep the design consistent with the 16-phase architecture being used at both TX and RX digital backend. From there, we pass the samples to the frame detection Intellectual Property (IP) core that performs a 16-time-shifted auto-correlation (16 being the repetition period of the STF in the header). Simultaneously, we are also searching for the LTF peaks by cross-correlating the samples with the known LTF sequence through the timing acquisition IP core. Given that the LTF samples are after the STF samples, we only check the output of the timing acquisition after the frame detection detects a packet. Due to the latency of the frame-detection and timing acquisition blocks, we are required to store the incoming samples in a buffer such that when the timing synchronization is accomplished we still have access to the data symbols.

Analog IF network In Fig. 4b, we show a detailed architecture of the analog system we utilize at the receiver side. As you can see, it is very similar to that of the transmitter side. The main challenge here is to utilize the least amount of analog components to reduce the additional added noise, but at the same time meet the input/output requirements of each analog component. On one hand, if we utilize the high gain 32 dBi antenna horn at the receiver we will need to ensure a distance long enough between the front ends to minimize the impact of multi-path on the signal. On the other hand, if we utilize the 21 dBi gain antenna, the signal is much weaker and below the typical value of which the IQ mixer has its best performance. In Fig. 5b, we show a DSO capture of the four channels being driven out of the receiver RF frontends at a distance of 9 feet utilizing the 32 dBi horn antennas. Considering that the IQ mixer has a conversion loss of almost 10 dB, the signal needs to be further amplified otherwise it is too weak for the ADC to capture it. After down-converting the signal, we use a low pass filter to filter out the out-of-band signals and noise, and then amplify the baseband in-phase and quadrature components.

3.3 Results and Discussion

3.3.1 Experimental Setup

The end-to-end setup of our proposed SDR platform is shown in Fig. 6. The evaluation board used in experiment is a Gen-1 ZU28DR from HTG [47]. To support the 8 GHz target bandwidth, all the data converters on the chip are utilized. In this experiment, we utilize a random data generator running in real-time in the PL while leaving the PCIe interface for application data for future work. Because we are transmitting random data, we rely on IQ constellations and EVM to analyze the system performance. After IQ up-conversion on the TX side and before down-conversion on the RX side, a low-pass filter is utilized for both Channels 1 and 2. This low-pass filter has a cut-off frequency close to the highest frequency of that particular channel. This allows us to filter out any noise and harmonics outside the channel bandwidth that might interfere with the nearby channel. Ideally, a bandpass filter should perform better due to its capability to filter out on both ends outside the bandwidth of interest. However, these are not available off the shelf at our required frequencies (which further motivates a custom board design down the line).

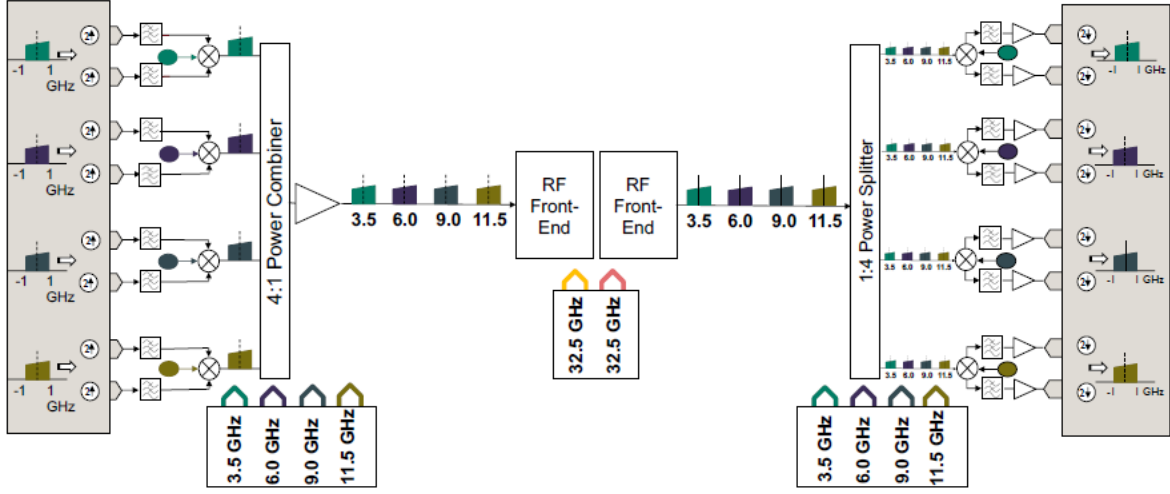


Fig. 6: Experimental setup for the real-time multi-channel terahertz communications tested.

Table 1: Four Channel Real-Time Capability

Channel Number	EVM _{avg}	Bitrate [Gbps]	Data Rate [Gbps]
Channel 1	10.21%	8.2	4
Channel 2	11.24%	8.2	4
Channel 3	11.73%	8.2	4
Channel 4	14.99%	8.2	4
Total		32.8	16

3.3.2 Results

In Fig. 7, we show the IQ constellation of Channel 1 for all 52 sub-carriers. From the results of Channel 1, we notice that this channel provides the best results in terms of EVM. This occurs due to the channel being on the lower end of the frequency band and given that the mixer experiences low leakages and imbalances at this low frequency. In Fig. 8 we show the IQ constellations of 4 randomly selected sub-carriers for Channels 1 to 4. Although the EVM begins to degrade from one channel to the next, we notice that some sub-carriers exhibit much better EVM than others. With that in mind, given the huge bandwidth, it would be more efficient to select different modulation orders for different sub-carriers. In Table 1, we show the average EVM per channel, alongside the bit and data rates. Given our bandwidth and setup, we can reach 32.8 Gbps bit-rates when only considering the bandwidth and modulation order. Unlike the bit rate, the data-rate accounts for meaningful data being transmitted. To calculate the data rate, we take into account, the packet time duration and the amount of data being sent within that packet. The nominal data rate achieved was ≈ 16 Gbps, which is roughly half the bit rate due to the CP length, header length, and the fact that we are only using 52 out of the 64 sub-carriers for data.

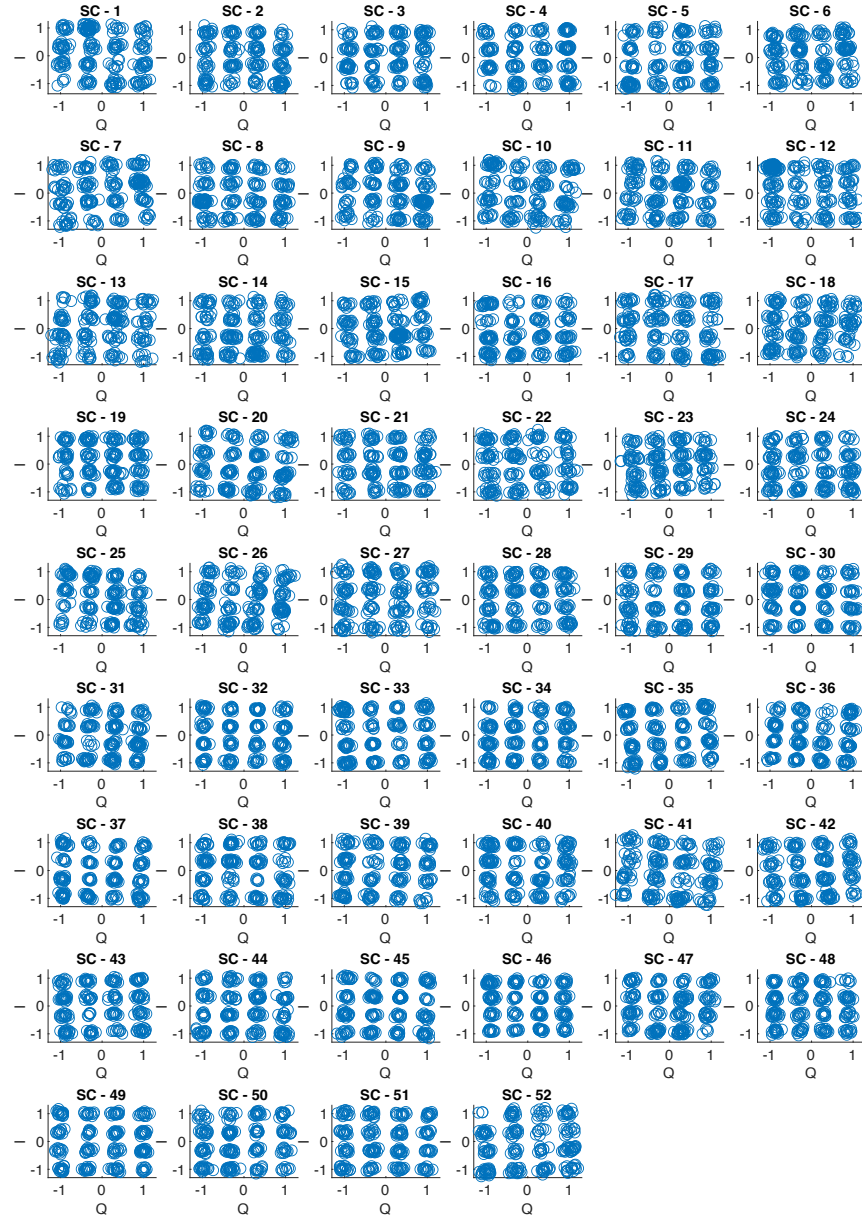
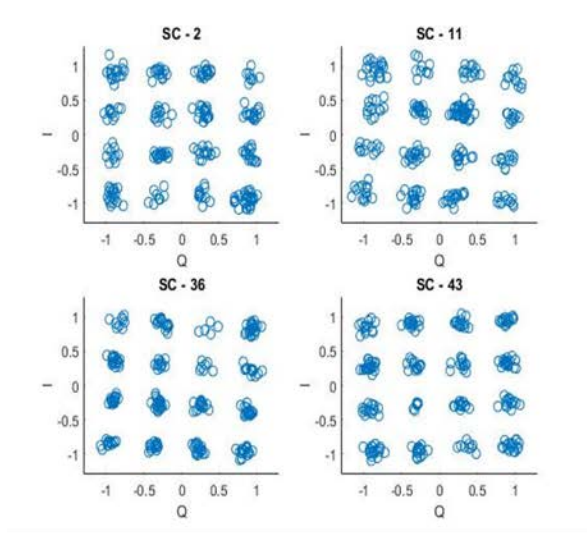
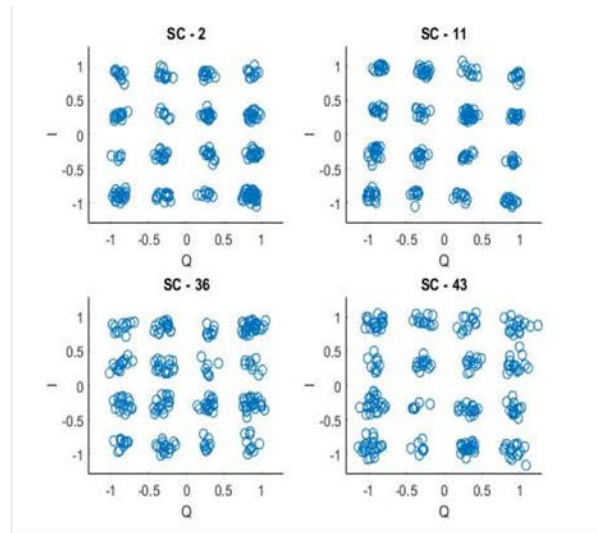


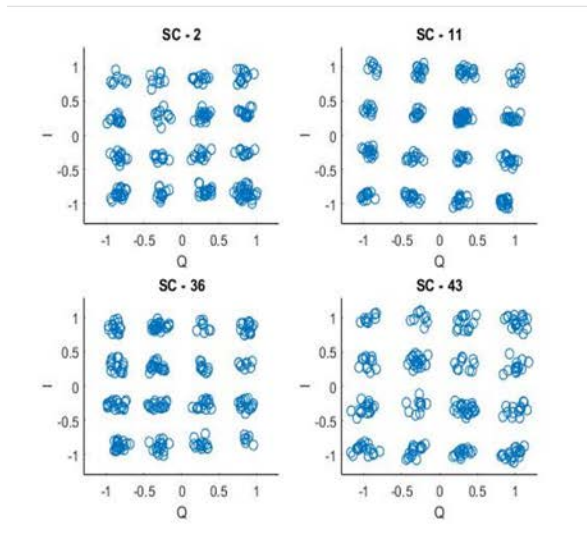
Fig. 7: IQ constellation for each data sub-carrier in Channel 1, when transmitting a 16 QAM.



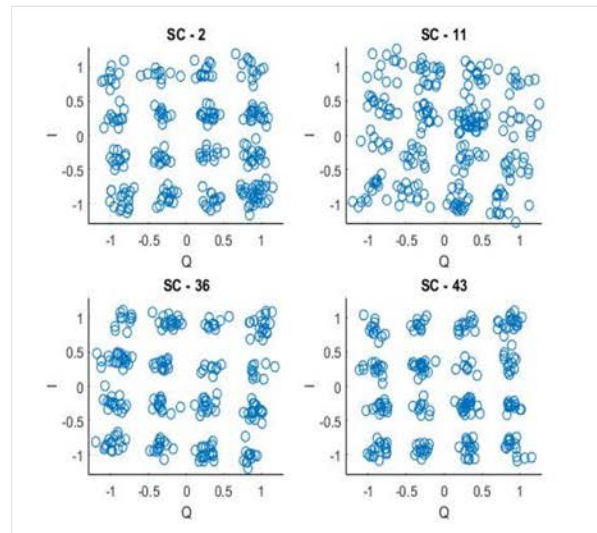
(a)



(b)



(c)



(d)

Fig. 8: IQ constellations for a selection of data sub-carriers in: a) Channel 1, b) Channel 2, c) Channel 3, and d) Channel 4.

3.3.3 Future Enhancements to the Testbed

With this work, the TeraNova platform has evolved from a non-real-time AWG/DSO-based testbed for ultrabroadband communications, to a real-time RFSoc-based system for ultrabroadband networking. This work is the first realization of the system, but not the last. Upcoming enhancements include 1) implementing a single-carrier modulation, potentially DFT-S-OFDM; 2) building a custom-designed printed circuit board including all the components in the analog frequency combining network; 3) finalizing the PCIe and QSFP interfaces; 4) interfacing with digital beamforming antenna arrays at sub-terahertz frequencies; and 5) experimentally testing new MAC, relaying and routing protocols, including those conceptualized in this project.

4 AIRBORNE TERAHERTZ CHANNEL MODELING

4.1 Motivation

As mentioned in the introduction, most of the terahertz propagation studies conducted to date have been focused on terrestrial, on-the-ground scenarios. In order to design wireless communication links between ground, airplanes and satellites, there are several specifics that must be taken into account when performing channel modeling. The topic has been partially covered in prior works. A simulation-based model for terahertz band satellite links has been proposed in [48]. Later, a review of weather impact on outdoor terahertz links has been presented in [49]. The latter study has been complemented in [50], where the effects of the atmosphere's non-homogeneous refractivity on millimeter-wave and terahertz satellite links have been explored. One of the most relevant models for our effort has been presented in [51]. The model first characterizes the absorption losses at different altitudes and then allows the estimation of the average effect when propagating through several atmospheric layers by integrating the obtained data. There are also other models suitable for terahertz band line-by-line absorption loss calculations, such as MODTRAN [52] and ATM [53]. However, as these have not been designed for wireless communications, they are not very agile for dynamic loss calculations as a function of the position of moving network nodes. Still, these models are very accurate but best suited for static link calculations. For dynamic environments and user nodes, it is more efficient to create a model that automatically calculates the path losses and the link budget based on the positions of the network elements. In parallel to academic research, some initial steps towards characterizing the airborne terahertz links have been made by International Telecommunication Union (ITU). Specifically, the approach to estimate the signal attenuation by atmospheric gases has been presented in ITU-R P.676-9 [54] while the noise levels for different frequencies (including the terahertz band) have been estimated in ITU-R P.372-13 [55].

Besides theoretical communication channel models, we have also recently seen measurement-based campaigns through channel sounding techniques, such as those in [20–22]. These models are aimed at studying the temporal characteristics of the channel, which are critical to then define the physical layer of terahertz communication networks. However, these models are either for indoor scenarios or terrestrial outdoor scenarios in the presence of good weather. As step towards characterizing terahertz links through different type of clouds and atmospheric phenomena, it is important to properly characterize the impact of adverse weather conditions (e.g., rain and snow) on the channel metrics and link performance. So far, the only experimental studies with snow focus just on the received power [56, 57] and thus do not provide a full-fledged channel characterization of the impact of adverse weather at these frequency bands.

As part of this effort,

- We have developed a mathematical model to characterize the airborne terahertz channel in the frequency domain, which takes into account (i) the features of the target use case and the deployment geometry, (ii) the peculiarities of the signal propagation through the atmosphere, as well as (iii) the prospective characteristics of terahertz radio equipment to operate in the airborne scenarios. By utilizing this model, we have numerically studied the theoretical performance of airborne networks.
- We have developed a custom sliding-correlator-type channel sounder with 20 GHz of bandwidth (the largest of any existing time-domain channel sounder). This is capable of capturing

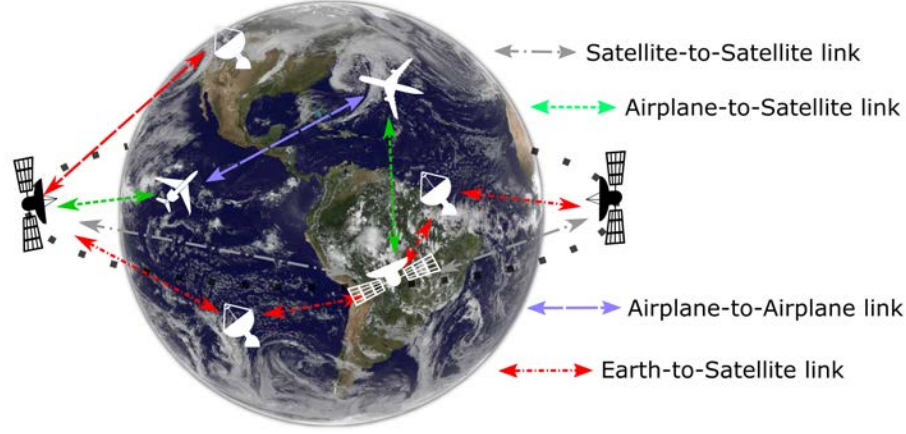


Fig. 9: An illustration of the system model considered in this paper and the several wireless links in it. Picture of Earth by NASA/NOAA's GOES Project.

multipath profiles with high resolution (0.1 ns) and dynamic range (60 dB). With this platform, we have collected extensive channel-sounding data and provided an extensive evaluation of the weather impact on several parameters including path loss, K-factor, and delay spread. We validate the measurements by comparing the experimental results with physics-driven models.

4.2 Methods, Assumptions and Procedures

4.2.1 Scenario and Fundamentals

The considered application scenarios for the channel models in this paper include: (i) airplane-to-satellite (A2S) or satellite-to-airplane (S2A), (ii) Earth-to-airplane (E2A) or airplane-to-Earth (A2E), and (iii) Earth-to-Satellite (E2S) or satellite-to-Earth (S2E). The satellite-to-satellite (S2S) case is not considered in this paper due to strong focus on modeling the molecular absorption and the atmosphere in general. Even the low Earth orbit (LEO) satellites lie in nearly empty space where any atmospheric effects will be small. Therefore, the channel mainly comprises FSPL and antenna systems. Furthermore, although not deeply analyzed, the E2S/S2E use case is equivalent to models required for possible THz band remote area communications.

The target scenario for the A2S and S2A channel modeling is shown in Fig. 9. The target scenario for channel modeling consists of an airplane at an altitude h_a and a satellite at an altitude h_s from the Earth and distance r_{as} from the airplane. The boundary of the atmosphere is r_{atm} away from the airplane. As the separation distance between the considered airplane and satellite is significant, the relative speed of the airplane is considered negligible for the channel modeling. The most important parameters and variables are given in Table 2.

The main phenomena affecting the propagation of THz signals in the target scenarios include:

Spreading Loss This measures the fraction of the power radiated by an isotropic transmitter at a frequency f that an isotropic receiver at a distance r can detect. Under the assumption of spherical

propagation, the spreading loss is given by

$$\text{PL}_{\text{spr}}(f, r) = 4\pi r^2 \frac{4\pi}{(c/f)^2} = \left(\frac{4\pi f r}{c^2} \right), \quad (1)$$

where c is the speed of light in the medium, f is the frequency, and r is the distance from Tx to Rx. In our system, directional antennas (e.g., horn lens antennas, Cassegrain parabolic antennas or beamforming antenna arrays) will expectedly be utilized at the transmitter and the receiver to counter the high channel losses. In that case, their directivity gain $D(\theta, \gamma)$ with respect to an ideal isotropic emitter and detector (commonly given in dBi) needs to be accounted for.

Molecular Absorption Loss The molecular absorption loss measures the fraction of electromagnetic energy that is converted into kinetic energy internal to vibrating molecules [18, 51]. For a homogeneous medium with thickness r , this is given by the Beer-Lambert law and can be written as:

$$\text{PL}_{\text{abs}}(f, r) = e^{\sum_i \kappa_a^i(f)r}, \quad (2)$$

where $\kappa_a^i(f)$ is the absorption coefficient of i th absorbing species (molecule or its isotopologue) at frequency f . We denote the total summed absorption coefficient with $\kappa_a(f)$, i.e., $\kappa_a(f) \equiv \sum_i \kappa_a^i(f)$.

In our system, the medium is not homogeneous, as the density of different absorbing molecules (particularly water vapor molecules) changes drastically with altitude. In the case of vertical paths, the absorption coefficient κ_a depends on distance $\kappa_a(f, r)$ and its variations need to be accurately computed along the signal propagation path.

Adverse Weather Additional attenuation may be caused by adverse weather conditions in A2A, A2S, and S2A scenarios. The ITU-R provides rain [58] and cloud/fog [59] attenuation models. It should be noted that the cloud attenuation is formally valid only up to 200 GHz frequencies [59]. Furthermore, the cloud densities, cloud thicknesses and rain rates vary greatly depending on the type of the cloud, and the probability of different types of clouds vary geographically. However, in any case the Earth to airplane or satellite paths experience added path losses ranging from few decibels up to tens of decibels in the presence of clouds and depending on the weather in general. In

the further, we note the rain and cloud losses with the distance-dependant variables $\delta_{\text{Rain}}(r_r)$ and $\delta_{\text{Cloud}}(r_c)$, respectively. Moreover, in Sec. 4.2.3, we experimentally characterize such phenomena experimentally with our platform.

4.2.2 Atmospheric Path Loss Modeling in the Frequency Domain

As highlighted in the previous section, the main challenge arises from the non-uniform molecular absorption coefficient through the atmosphere. The density, temperature, and molecular composition of the atmosphere depend on the altitude. This section gives the dynamic molecular absorption loss coefficient valid at any altitude by taking into account the impact of atmospheric parameters on the absorption line calculations. After that, we go through the transmission path geometry of slant paths through the atmosphere. Table 2 gives the most important constants, variables, and parameters utilized in this paper.

Non-homogeneous Molecular Absorption Loss Traditionally the channel modeling for the THz to higher altitudes, frequencies has been considered for terrestrial communications. When moving

Table 2: The most important constants, variables, and parameters used in the paper.

Constants		
Constant	Symbol	Value
Radius of Earth	R	6,371,000 m
Speed of light	c	299,792,458 m/s
Standard pressure	p_0	101325 Pa
Standard temperature	T_0	296 K
Variables		
Variable	Symbol	Unit
Absorption coefficient	$\kappa_a(f)$	1/m
Center frequency	f_c	GHz
Parameters		
Parameter	Symbol	Unit
Altitude of airplane	h_a	m
Altitude of satellite	h_s	m
Angle to zenith	θ	rad
Depth of the atmosphere	r_{atm}	m
Distance from plane to satellite	r_{as}	m
Elevation of airplane	ϕ_a	rad
Elevation of satellite	ϕ_s	rad
Frequency	f	Hz
Pressure	p	Pa
Temperature	T	K

the homogeneous assumption for the line shape functions is no longer valid. We can also utilize meteorological data to take into account the global variations in water vapor content as a function of latitude, longitude, and altitude. Those will be discussed briefly in conjunction with the numerical results. Similarly, when moving to higher altitudes, mixing ratios of the molecular species change and this has to be taken into account for accurate molecular absorption modeling.

The line-by-line absorption model and the parameters required for different altitudes are given in [18,60]. Putting the absorption loss model therein together, the altitude-dependent absorption coefficient becomes

$$\kappa_a^i(f, p(h_{\text{atm}}), T(h_{\text{atm}})) = \frac{p(h_{\text{atm}})\mu_i N_A}{R_g T(h_{\text{atm}})} S^i(T(h_{\text{atm}})) F^i(f, p(h_{\text{atm}}), T(h_{\text{atm}})), \quad (3)$$

where $p(h_{\text{atm}})$ is the pressure, N_A is the Avogadro constant, R_g is the gas constant, $S^i(T(h_{\text{atm}}))$ is the line intensity, $T(h_{\text{atm}})$ is the temperature, $F^i(f, p(h_{\text{atm}}), T(h_{\text{atm}}))$ is the absorption line width. The altitude h_{atm} is directly having an impact on the pressure and temperature and it is indirectly affecting on the line-by-line parameters depending on the altitude. The distance through the atmosphere is handled by the integration over the total molecular loss in the atmosphere and is given in the next

section. The line width in the above equation is

$$F^i(f, p(h_{\text{atm}}), T(h_{\text{atm}})) = \begin{cases} F_{VVH}^i(f, p(h_{\text{atm}}), T(h_{\text{atm}})), & \text{if } \alpha_L^i \gg \alpha_D^i, \\ F_D^i(f, T(h_{\text{atm}})), & \text{if } \alpha_L^i \ll \alpha_D^i, \\ F_V^i(f, p(h_{\text{atm}}), T(h_{\text{atm}})), & \text{if } \alpha_L^i \approx \alpha_D^i, \end{cases} \quad (4)$$

where $F_{VVH}^i(f, p(h_{\text{atm}}), T(h_{\text{atm}}))$, $F_D^i(f, T(h_{\text{atm}}))$, and $F_V^i(f, p(h_{\text{atm}}), T(h_{\text{atm}}))$ are the Van Vleck-Weisskopf, Doppler, and Voigt line shape functions, respectively, and α_L^i and α_D^i are Lorentz and Doppler line half widths (half width at half maximum). The Lorentz line shapes are utilized in high-pressure environments and the Doppler line shapes in low-pressure environments. Notice that the altitude defines the pressure, temperature, and other parameters utilized in the calculation of the absorption coefficient. As a consequence, the line shape function requires a dynamic algorithm that defines the correct line shape for every possible set of pressure, temperature, and absorption line center frequency. This can easily be done in the line-by-line algorithm by calculating and comparing the Doppler and Lorentz absorption line half-widths.

Transmission Path Geometry The general system geometry is shown in Fig. 10. The usual way to model the atmosphere is to assume a plane-parallel atmosphere. By assuming a flat surface, the geometry of the system is significantly simplified. However, when talking about applications where the entire atmosphere is penetrated, the plane parallel assumption overestimates the thickness of the atmosphere. The impact is higher the lower the angle to the surface is. In order to rectify the problems with the plane parallel model, in this work we model a realistic curving atmosphere. In such a setting, the problem becomes with the ever-changing angle to the zenith (and to the elevation angle ψ) as shown in Fig. 11. This problem can be modeled geometrically by utilizing the geometry shown in Fig. 10. In this figure, R is the Earth's radius, h_a is the altitude of the airplane, b is the Earth's radius plus the total thickness of the atmosphere, h_s is the altitude of the satellite, ϕ_a is the elevation of the airplane, ϕ_s is the elevation of the satellite, r_{as} is the distance from the airplane to satellite, and r_{atm} is the most interesting parameter, i.e., the distance through the atmosphere from the airplane to the boundary of the atmosphere. This latter distance is important because of the molecular absorption that depends on the altitude and the constantly evolving line parameters as the altitude changes. It should be noted that the elevation herein means an arbitrary angle that is used to define the geometry between the Earth, airplane, and the satellite. Herein, it is bound to the latitude to illustrate the angles easily. However, in the case of moving airplanes, the angular differences have to be bound to both latitude and longitude based on the real locations of the network elements.

The main difference between the curving atmosphere and the plane parallel atmosphere comes from the dynamically changing atmospheric parameters depending on the altitude. This is shown in Fig. 11. Distance from the airplane to the satellite is given by the law of cosines

$$r_{as} = \sqrt{(R + h_a)^2 + (R + h_s)^2 - 2(R + h_a)(R + h_s)\cos(\rho)}, \quad (5)$$

where R is the radius of Earth, h_a is the altitude of the airplane, h_s is the altitude of the satellite, and ρ is the angle between the airplane and satellite looked from the center of the Earth. This angle is obtained as the subtraction of the elevation angles, i.e., $\rho = |\phi_a - \phi_s|$, where ϕ_a is the elevation of the airplane and ϕ_s is the same for the satellite. See details on these variables in Fig. 10. The distance from the airplane to an imaginary line from Earth's core to the satellite is calculated as

$$a = (R + h_a) \sin(\rho). \quad (6)$$

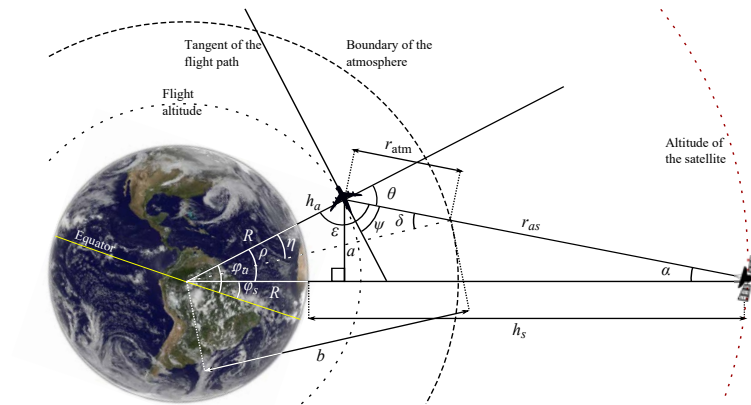


Fig. 10: The system geometry for the angles and distances in the airplane to satellite system. Picture of Earth by NASA/NOAA's GOES Project.

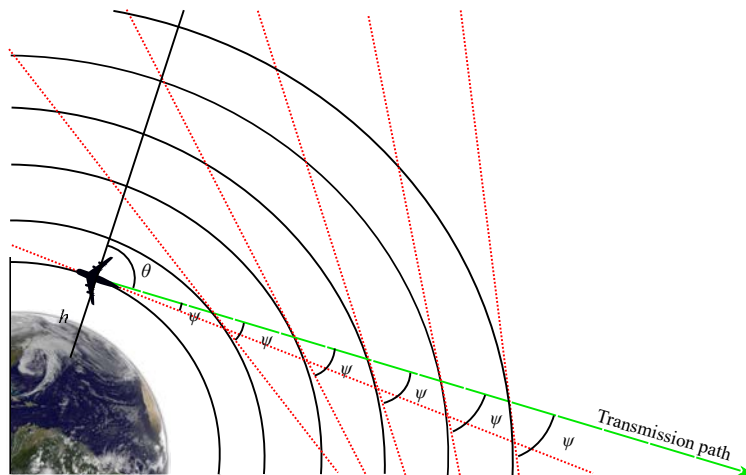


Fig. 11: System geometry on the angles of the transmission path through the atmosphere. Picture of Earth by NASA/NOAA's GOES Project.

Then the angle between the satellite and the airplane, looked from the satellite, is obtained from the two as

$$\alpha = \sin^{-1} \left(\frac{a}{r_{as}} \right). \quad (7)$$

The angle between line from Earth to the airplane and the line from airplane to satellite is given simply by

$$\epsilon = 180^\circ - \alpha - \rho, \quad (8)$$

which is also equivalent to $\epsilon = \psi + 90^\circ$ where ψ is the elevation angle of transmission path from airplane to satellite. These angles also directly give the transmission angle with respect to the zenith as $\theta = 180^\circ - \epsilon = 90^\circ - \psi$. Now we have everything we need in order to calculate the distance from an airplane through the atmosphere r_{atm} . This can be calculated from the triangle suspended by $R + h_a$, distance from the core of the Earth to the boundary of the atmosphere b , and r_{atm} . Let us first calculate the corner angles of this triangle. We can obtain angle δ from the law of sines

$$\delta = \sin^{-1} \left(\frac{(R + h_a) \cos(\epsilon)}{b} \right). \quad (9)$$

Then we can obtain the last angle as $\eta = 180^\circ - \epsilon - \delta$. Finally we can obtain the distance through the atmosphere r_{atm} with the law sines as

$$r_{\text{atm}} = \frac{b \sin(\eta)}{\sin(\epsilon)}. \quad (10)$$

On the other hand, we could also solve r_{atm} with the law of cosines and the quadratic formula as

$$r_{\text{atm}} = (R + h_a) \cos(\epsilon) + \frac{1}{2} \sqrt{(-2(R + h_a) \cos(\epsilon))^2 - 4((R + h_a)^2 - b^2)}. \quad (11)$$

Both of these expressions for r_{atm} are equal, but the former gives perhaps a bit more compact way to calculate the angles and distances. Finally, in order to fit this new approximation for the length of the atmosphere to the transmittance, we need to take a derivative of it. The reason is that the absorption coefficient changes as a function of the altitude, but does not depend on the total distance. Therefore, we need to manipulate the space over which the total absorption is calculated. That is, we multiply the absorption coefficient by $d/dr(r_{\text{atm}})$ instead of $\sec(\theta)$ as in the case of the plane-parallel atmosphere. Then,

$$\tau(f, r) = e^{-\int_{r_1}^{r_2} \kappa_a(f, r) \frac{d}{dr} r_{\text{atm}} dr} = e^{-\int_{r_1}^{r_2} \kappa_a(f, r_{\text{atm}}) dr_{\text{atm}}}. \quad (12)$$

The difference between the plane-parallel assumption and the proposed model is shown in Figs. 12 and 13. In computer simulations/models, we are forced to discretize the parameters due to utilization of the data bases to obtain the parameters and since the computer inherently handles only discrete data. Then we also have to discretize the atmosphere into equally spaced layers. In the plane parallel atmosphere, the distance through each layer is same at all altitudes due to the simple geometry. In the geometry given above, the discretization leads into non-equal propagation distance

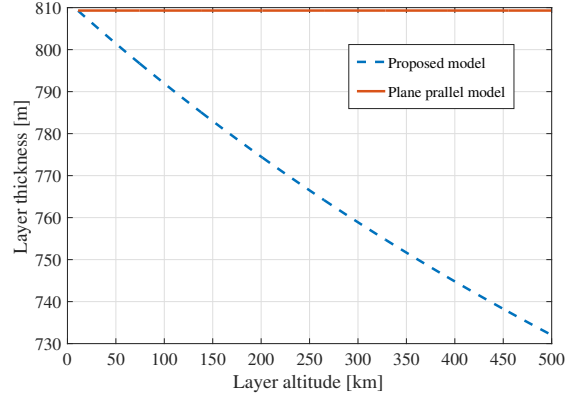


Fig. 12: Comparison of the layer thickness of the plane parallel atmosphere and a round atmosphere as a function of altitude.

through the layer depending on the altitude and the angle of penetration due to the curvature of the atmosphere. It can be seen in Figs. 12 and 13 that the vertical path through the atmosphere is exactly the same for both models as they should be. Notice that the atmosphere has been calculated up to 500 km altitude where it is already very thin and does not contribute much on the propagation anymore. Fig. 12 shows the propagation distance through a single 500 m layer (500 m resolution accuracy) at 38.2 degree angle from the ground. As stated above, the layer altitude has an impact on the dynamic distance through the atmosphere. Thus, the plane parallel assumption tends to overestimate the layer thickness. The problem is the worse the lower is the angle through the atmosphere. This is well visible in Fig. 13, which shows the total distance through the atmosphere. As the molecular absorption is dependent on the propagation through molecular medium, the plane parallel assumption may greatly overestimate the total absorption loss. The next section looks into derivation of the absorption coefficient $\kappa_a(f, r)$ as the geometry of the propagation is one important issue, and the second important aspect are the parameters fed into the model.

Total Path Loss By combining (1) with (12), the total path loss can be written as

$$PL(f, r) = \frac{(4\pi r_{as} f)^2 \delta_{\text{Rain}}(r_r) \delta_{\text{Cloud}}(r_c)}{c^2 G_{\text{Tx}}(\theta_{\text{Tx}}) G_{\text{Rx}}(\theta_{\text{Rx}})} e^{\int_1^{r_2} \kappa_a(f, r_{\text{atm}}) dr_{\text{atm}}}, \quad (13)$$

where $G_{\text{Tx}}(\theta_{\text{Tx}})$ and $G_{\text{Rx}}(\theta_{\text{Rx}})$ are the Tx and Rx antenna gains towards directions θ_{Tx} and θ_{Rx} , respectively, and $\delta_{\text{Rain}}(r_r)$ and $\delta_{\text{Cloud}}(r_c)$ are the channel gains due to rain and cloud losses, where r_r is the distance through the rain and r_c is the distance through the cloud/fog.

The rain loss is given by [58]

$$\delta_{\text{Rain, dB}}(r_r) = k R_r^\alpha r_{\text{eff}}, \quad (14)$$

where k and α are parameters tabulated in [58], R_r is rain rate (mm/hr), and r_{eff} is the effective distance that is a multiplication between the actual distance r_r and a distance scaling factor. The cloud loss is given by [59]

$$\delta_{\text{Cloud, dB}}(r_c) = K_l M r_c, \quad (15)$$

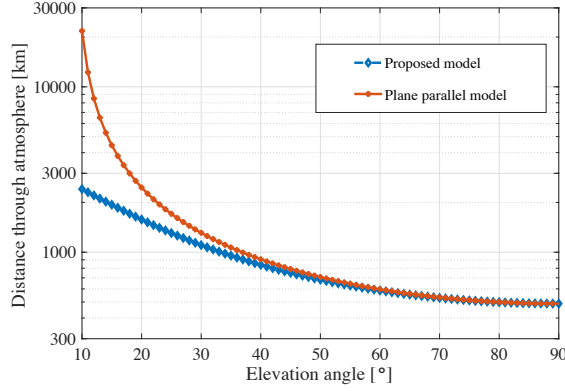


Fig. 13: A comparison of the total distance through the atmosphere according to the plane parallel assumption and the model herein as a function of the elevation angle.

where K_l is the specific attenuation coefficient and M is the liquid water density (g/m^3) in cloud or fog. It should be noticed that the above losses are specified in the ITU-R recommendations in logarithmic scale (hence the dB-subscript in the above equations). More details on the parameters for these models can be found in [58] and [59].

4.2.3 Time-domain Channel Sounder

To experimentally characterize the outdoor channel and, as a major novelty, the impact of adverse weather, we have developed a spread spectrum-based sliding correlator type channel sounder, the “NU Channel Sounder” named after Northeastern University (NU). It is capable of capturing multipath profiles with high resolution and dynamic range and is thus tailored to describe the ultra-broadband nature of the band above 100 GHz. The system is divided into two parts:

Terahertz Sounder Front Ends The key components of the channel sounder front ends are developed based upon the “TeraNova” testbed [46,61]. The transmitter and receiver components are shown in Fig. 14a and c, respectively. The system hardware at the transmitter is based on an AWG, which can generate baseband (BB)/intermediate frequency (IF) channel-sounding signals with up to 32 GHz of bandwidth. The BB/IF signal is fed to our THz RF front end that up-converts the input signal with a mixer, driven by a multiplied LO at 120 GHz, to produce a pass-band signal centered at 140 GHz.

The signal is received by another RF front end that down-converts the pass-band signal in a similar manner. The received BB/IF signal is fed to our digital storage oscilloscope (DSO), which can capture and store signals with 63.2 GHz of operational bandwidth. The saved signals are processed offline by the signal processing backend.

Signal Processing Backend The signal processing backend is divided into transmitter and receiver sections, and its block diagram is shown in Fig. 14d. The backend defines the sounding signal and generates the digital BB/IF signal at the transmitter. At the receiver, the backend processes the received signal to acquire the channel impulse response or multipath profile. We use MATLAB to

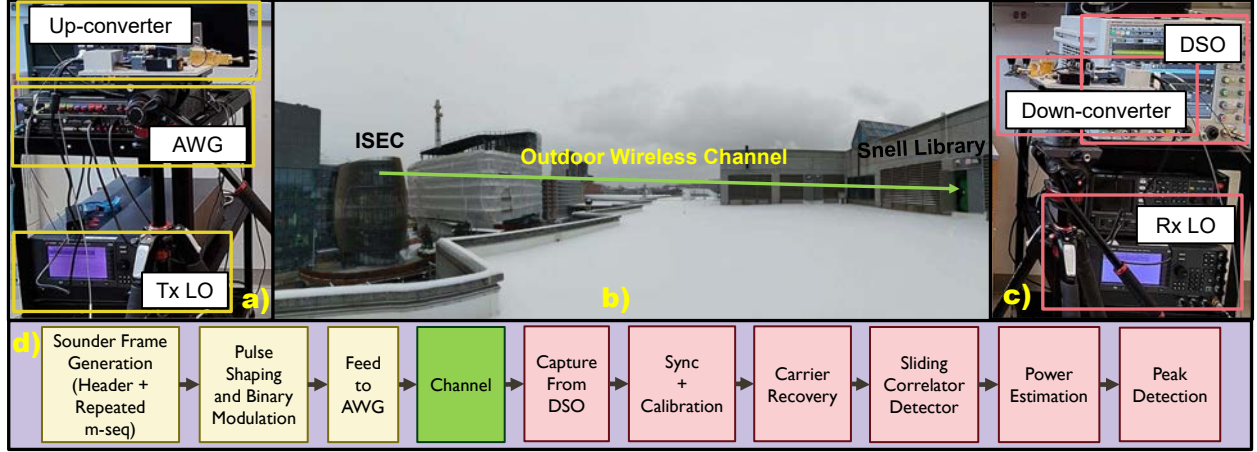


Fig. 14: The NU Channel Sounder system: a) transmitter hardware, b) outdoor link, c) receiver hardware, and d) signal processing backend.

develop the different signal-processing blocks of the backend.

The transmitted frame starts with an 8191-chips-long m-sequence header used for frame synchronization, calibration, and carrier recovery. This is followed by a 4095-chips-long m-sequence repeated 16 times that is used for channel sensing. The full frame is modulated with binary phase-shift keying (BPSK) and pulse shaped with the root-raised-cosine filter to reduce out-of-band leakage. The digital BB/IF signal is actualized by the AWG for further frequency up-conversion and transmission at 130-150 GHz band.

At the receiver, the pass-band signal is down-converted back to the BB/IF band and captured by the DSO for further processing. The header is used to synchronize with the start of a frame. Thereafter, the received signal is calibrated to eliminate the frequency selective response of THz frontend devices and connectors. The measured frequency response of each component is estimated as

$$H(k) = \left(\frac{P_r(k) - P_n}{P_s(k)} \right)^{1/2}, \quad (16)$$

where $P_r(k)$ is the received signal power with noise at the k_{th} frequency, $P_s(k)$ refers to the transmitted signal power, and P_n is noise power for the entire observation bandwidth. The inverse of normalized H calibrates the received signals to mitigate the frequency selectivity of the hardware. However, the utilization of this approach increases the noise floor by approximately 10 dB.

To synchronize with the IF carrier of the received signal, a dual phase-locked loop (PLL) is used. The frequency offset is estimated by the slope of changing/sliding phase of the first PLL. Meanwhile, the phase offset is estimated from the output of the second PLL, which has feedback from the first PLL to eliminate phase sliding. To reduce the phase noise, an average of 16 readings is considered.

The received BB/IF signal is correlated with a locally generated 4095-chips-long m-seq based BB/IF signal to extract the channel impulse response. The impulse response is averaged over 16 such repetitions to reduce the noise floor and falsely detected impulses. The actual received power of the multipath profile is obtained by normalizing the processed channel impulse response with the

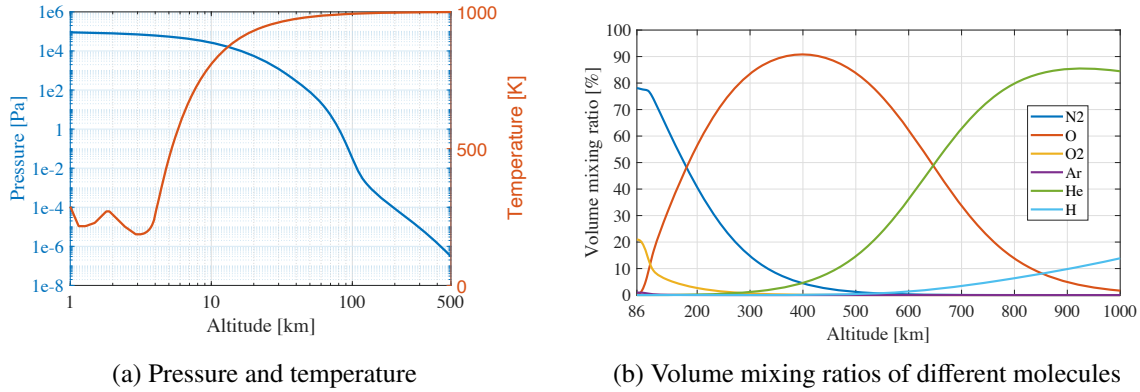


Fig. 15: Atmospheric properties as a function of altitude.

pulse energy and m-seq length (4095). The peaks and corresponding delays are measured through the global convergence method by setting the threshold based on the dynamic range of the channel sounder.

The channel sounder is capable of sounding 20 GHz of RF bandwidth (i.e., the operational bandwidth of the RF frontends' mixers) and can measure multipath components with 0.1 ns time resolution. Further, the sounder system can estimate the multipath channel response up to 130 dB of path loss with a substantial dynamic range of 60 dB (with 38 dBi antenna at both transceiver sides). Also, the sounder can detect the channel response for paths experiencing up to a 160 dB loss for line-of-sight (LoS) point-to-point link scenarios by compromising over dynamic range. It can be used for Doppler measurements less than 135 KHz with $7.38 \mu s$ of measurement time.

4.3 Results and Discussion

4.3.1 Numerical Results

System Parameters One important factor in modeling the atmosphere is to model the altitude-dependent parameters correctly. Those are mainly the temperature and pressure, and volume mixing ratios of different molecules. For these results, we utilize the 1976 US Standard Atmosphere [62]. The corresponding temperature and pressure as a function of altitude are given in Fig. 15a. The volume mixing ratios of common molecules are given in Fig. 15b. Interestingly, the volume mixing ratios of the molecules remain roughly the same to the sea level up until 86 km altitude, after which they start to vary from each other. The pressure, temperature, volume mixing ratio and ground level humidity are shown in Figs. 15a to 15b. In addition, Fig. 16 shows the average humidity around the world (left) and the ground height (right). The data for these were obtained from NASA/GMAO MERRA2 reanalysis data [63]. From [63], one can find large datasets of atmospheric data, including global humidity data at different time scales and much more related to global weather and climate.

The rain and fog attenuations were shortly discussed above. In the real atmosphere, we do not always have clear weather and additional losses by clouds and rain impair the link performance.

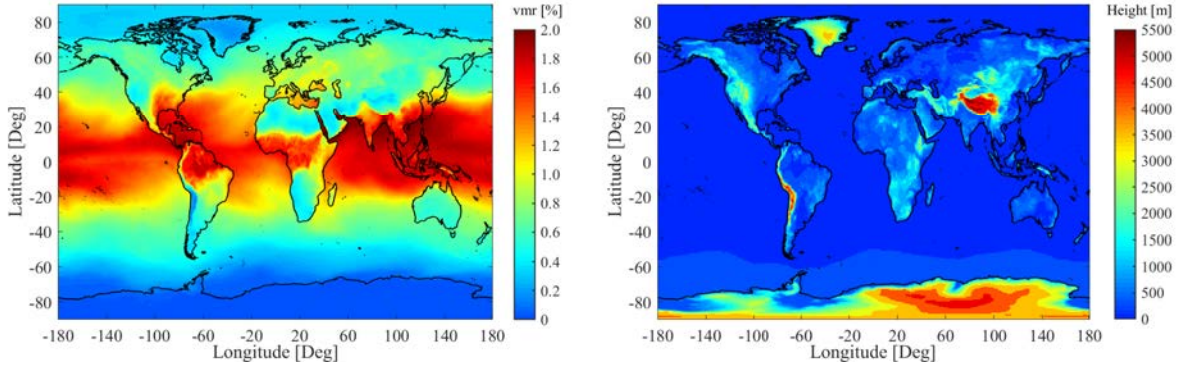


Fig. 16: The surface volume mixing ratio of water vapor (left) and the surface height from the sea level (right).

These are the mostly affecting low altitude communications, below few kilometer heights. As such, they have small impact on A2S and communications while at the typical flight heights. However, some losses by clouds can also exist at very high altitudes. For the take off and landing phases these may have an impact, and also in the G2A use case.

In the example shown in Fig. 17, we assume a typical nimbostratus cloud, which is roughly one kilometer thick and has a cloud bottom height of 0.7 km [64]. The cloud density is assumed to be 0.5 g/m^3 , which is close to a typical density of a nimbostratus cloud [64]. The rain rate is assumed to be 5 mm/hr, which corresponds to a moderate rain. Figure 17 shows the impact of the clouds and rain with the above assumptions for a slant path and path angled at 45 degrees. The angled path, therefore, increases the path lengths inside the rain and clouds by $\sec(\theta = 45^\circ)$, or $\sqrt{2}$. We can see that the losses vary from few decibels to few tens of decibels. Depending on the frequency and other losses, this may cause problems in signal reception and should be taken into account in the link budget considerations. However, as mentioned above, the rain and cloud/fog attenuations are mostly affecting low-altitude communications, such as E2A communication scenario.

The simulation parameters for the rest of the numerical results are given in Table 3 unless otherwise stated. The antenna gains are calculated with a parabolic antenna equation

$$G(f) = A_e \left(\frac{\pi d_A}{\lambda} \right)^2, \quad (17)$$

where A_e is the aperture efficiency, and d_A is the diameter of the antenna. We can see that the antenna gain increases with the square of frequency for fixed size dish when increasing the frequency. Adding dishes at both ends of the link, the ideal gain of the antenna gain can be very large and even overcome the electric size of the antenna visible in the FSPL equation ($\lambda^2/4\pi$).

Parabolic antennas have the advantage of providing a lot of gain in a relatively compact package that can be mounted on the fuselage of the airplane. Such antennas require an aerodynamic cover to protect the antenna and to interfere as little as possible with the overall aerodynamics of the airplane. The downside is the need for physical beamforming that limits the visibility of all the paths and restricts the maximum beamforming angles to at most the horizon with respect to the antenna mount, but most likely less to keep the antenna profile low. If there is a downward link, a separate antenna is required at the bottom of the fuselage.

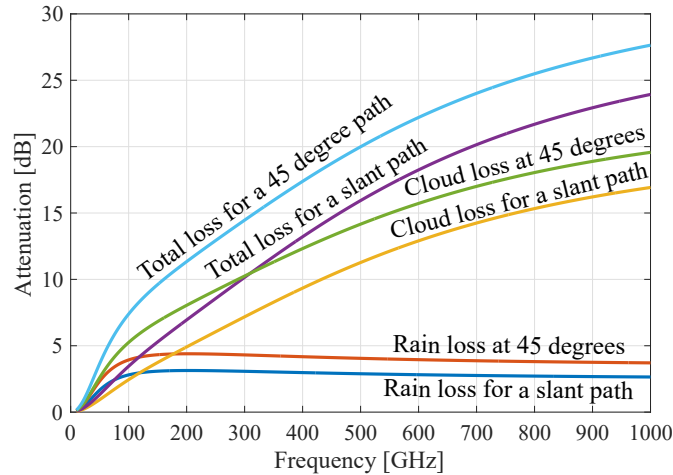


Fig. 17: Rain and cloud attenuation for slant paths as well as for 45-degree angled paths from zenith.

The bandwidth is chosen to be rather low to decrease the noise bandwidth, which is very important design aspect for the prospective long-range THz communications. What is the large enough bandwidth that justifies the THz band, but keeps the noise floor small enough. The antennas are typical parabolic antennas with 70% aperture efficiency. To show the impact of the angular differences between the satellite and the airplane, we utilize 0° , 6.25° , and 12.5° elevation differences between the two. At a Geostationary Orbit (GEO) being at 35,786 km height and LEO orbit at 500 km height, those correspond to 0 – 14.7° angle (14.7° at 0 – 20 km altitude for 12.5° elevation difference) from the zenith at airplane to GEO satellite and 0 – 78° degree angle for the LEO orbits (58.5 – 77.2° zenith angle at ground level and 59.6 – 78.0° at 20 km altitude). These correspond to angles from the horizon of the airplane to satellite varying from 12 – 90° . These elevation differences correspond to link distances in S2A to vary from about 500 to 1,526 km for the LEO orbit at 500 km altitude depending on the airplane altitude (that is rather small compared to the total link distance), and 35,786 – 35,964 km for the GEO orbit.

Path Losses Figure 18 shows the path loss from airplane to satellite for an airplane at 11 km altitude as a function of frequency. The path losses are shown for a geostationary Earth orbit (GEO, ~ 36000 km) and for a LEO at 500 km altitude. The path losses have been calculated for zenith path and for elevation differences 6.25 – 12.5° through the atmosphere for the LEO orbit, and 12.5° elevation difference for the GEO due to smaller differences between the path losses between different angles. The LEO connection suffers more from the angled paths due to the path loss by absorption has more relatively higher impact on the shorter paths. The GEO path loss is dominated by the FSPL and the absorption does not have a that large impact on the total path loss. We can see that the path losses are considerable. Those range from about 190 to 260 dBs disregarding the absorption peaks that are much higher. This means that the required antenna gains need to be very high, but at the same time, the bandwidths need to be low enough to keep the noise level down and to avoid the absorption peaks. Considering fixed aperture antennas, the antenna diameters need to be in the order of 0.5–1 meters to provide enough gain in the ideal case. Furthermore, LEO path loss towards zenith as a function of distance and frequency is given in Fig. 19. This figure shows that the FSPL

Table 3: The simulation parameters used in the numerical results.

Parameter	Symbol	Value
Transmission bandwidth	W	5 GHz
Antenna diameter at airplane	d_A	0.5 m
Antenna diameter at satellite	d_A	1.0 m
Aperture efficiency	A_e	70%
Elevation difference	ρ	0–12.5°
GEO orbit	h_s	35,786 km
LEO orbit	h_s	500 km
Ground level temperature	T	288.15 K
Ground level pressure	p	101325 Pa

dominates vast majority of the high altitude links with molecular absorption loss increasing towards the high frequencies and the lower altitudes. This implies that the A2S/S2A links provide the best performance at the higher flight altitudes as it will be shown later.

The path losses from E2A as a function of airplane altitude are shown in Figs. 20 and 21 for a frequency range from 100 GHz to 1000 GHz. The latter figure assumes a LEO height of 500 km. Any frequencies above 1000 GHz are effectively blocked due to high absorption loss close to Earth. We can see that the below 380–450 GHz frequencies are utilizable for airplane communications from the ground. However, the losses remain high and partially so due to high FSPL at high frequencies. This also means a certain type of bottleneck for feeding the satellites. The airplane-to-satellite links suffer less from atmospheric loss and could theoretically utilize higher frequencies than the E2A or E2S links. This potentially opens a door for frequency division between ground to satellite and air to satellite links. However, this is also dependent on the capacity requirements, and hardware limitations. The latter is a major problem at the THz frequencies and potentially considerable losses have to be expected from inefficiency of them. In this paper we do not consider those, as the development of the THz band hardware is still a major research topic and relatively poorly understood from the viewpoint of realistic system and link performance. It should still be kept in mind that more losses are expected compared to pure atmospheric losses due to rain, clouds/fog, and other possible mechanisms, such as the scattering loss.

The airplane to airplane path loss for a 100-meter link is shown in Fig. 22 as a function of altitude and frequency. As expected, the path loss decreases with altitude and makes the communications more simple. Above 10 km height, the loss is mostly comprised of the FSPL disregarding some spikes in the spectrum. This is also shown in Fig. 23, which shows the path losses per kilometer link on ground level and at 11 km height. As a reference, the antenna gains for 0.5 m diameter dish antenna ('airplane antenna') and a 1-meter diameter dish antenna ('satellite antenna') are also shown. It should be noticed that airplanes in general fly far away from each other and a 100-meter distance would be considered very dangerous. Also, these figures are applicable to on-Earth communications due to varying land heights. On a general note on the high-frequency links, the THz band offers a natural protection against third parties listening to the data traffic as the FSPL and atmospheric losses kill the signals over long distances in the lower atmosphere as seen in

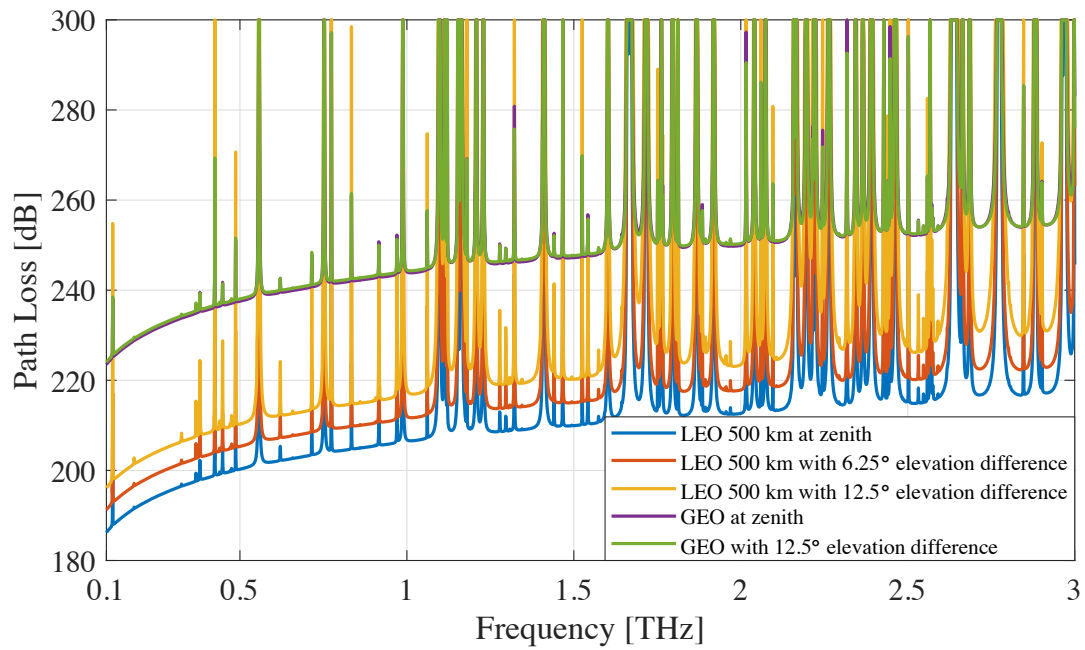


Fig. 18: Path loss from airplane to satellite for LEO and GEO orbits as a function of frequency. The airplane altitude is 11 km.

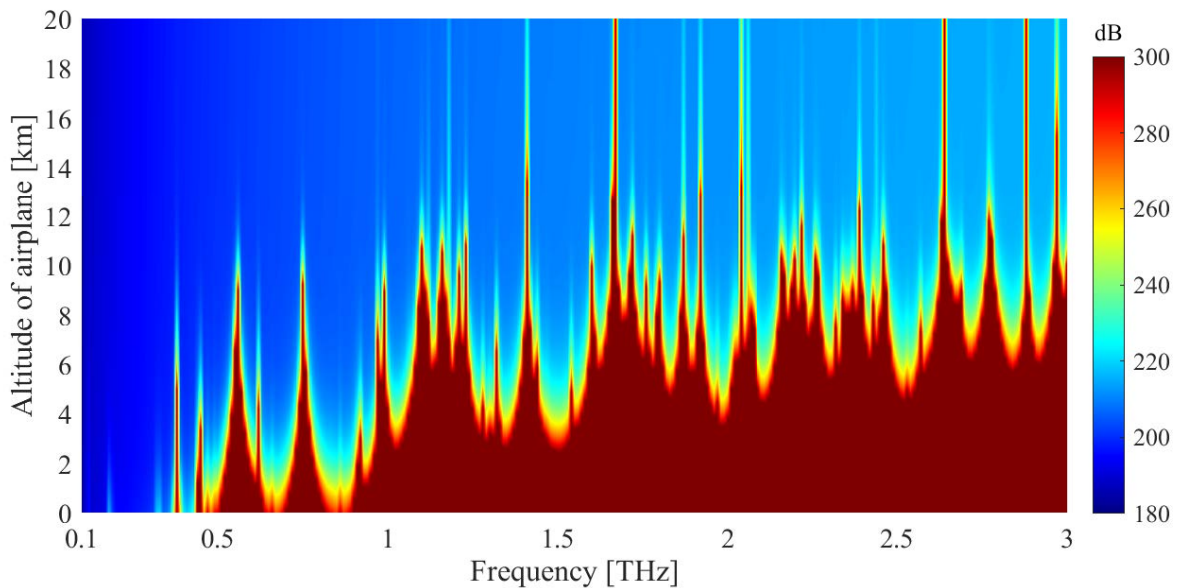


Fig. 19: Path loss from airplane to satellite for 500 km LEO satellite as a function of frequency and airplane altitude.

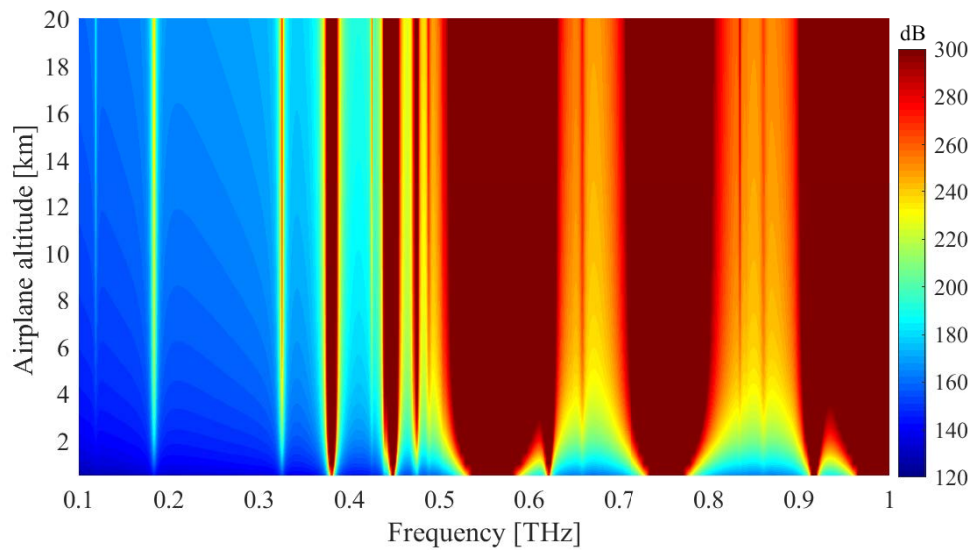


Fig. 20: Path loss from Earth to an airplane as a function of airplane altitude and frequency.

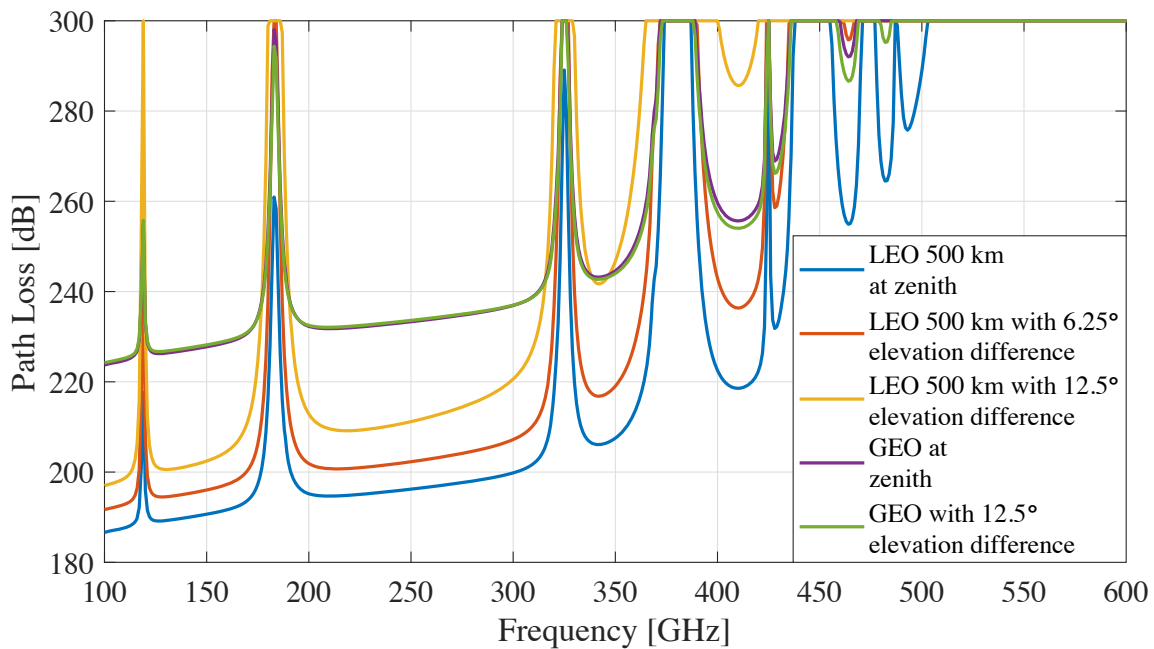


Fig. 21: Path loss from Earth to satellite as a function of frequency.

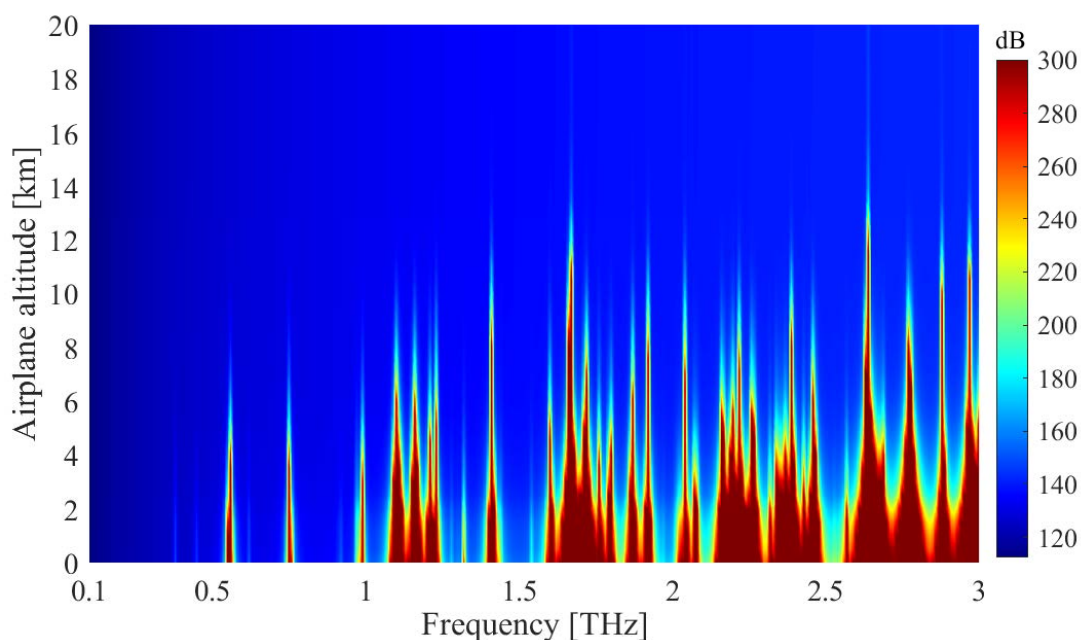


Fig. 22: Path loss from airplane to airplane (common altitude) for a 100-meter link as a function of airplane altitude and frequency.

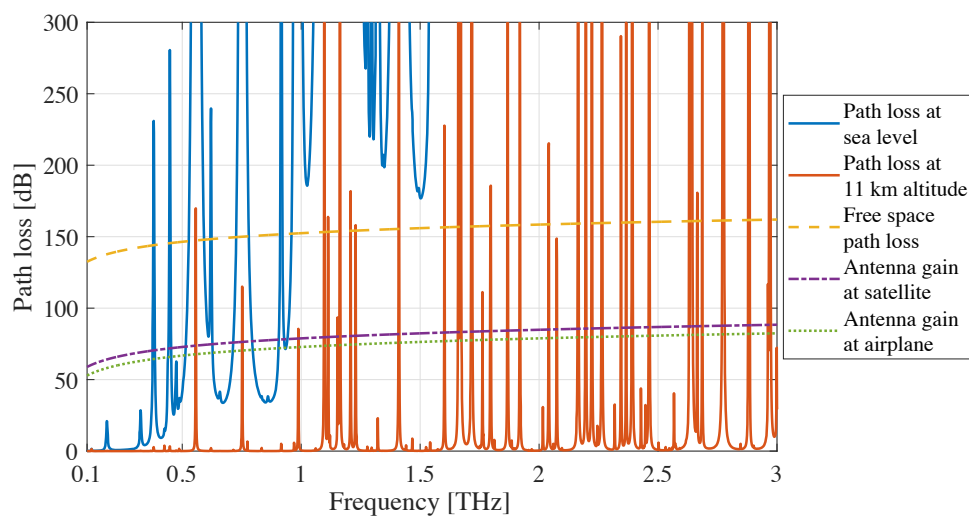


Fig. 23: Comparison of path losses per kilometer for FSPL, molecular absorption loss at sea level and 11 km altitude, and antenna gains of 0.5 and 1 meter ideal dish antennas.

Figs. 20 and 22.

Based on the path losses, we can estimate the performance of the extremely long links.

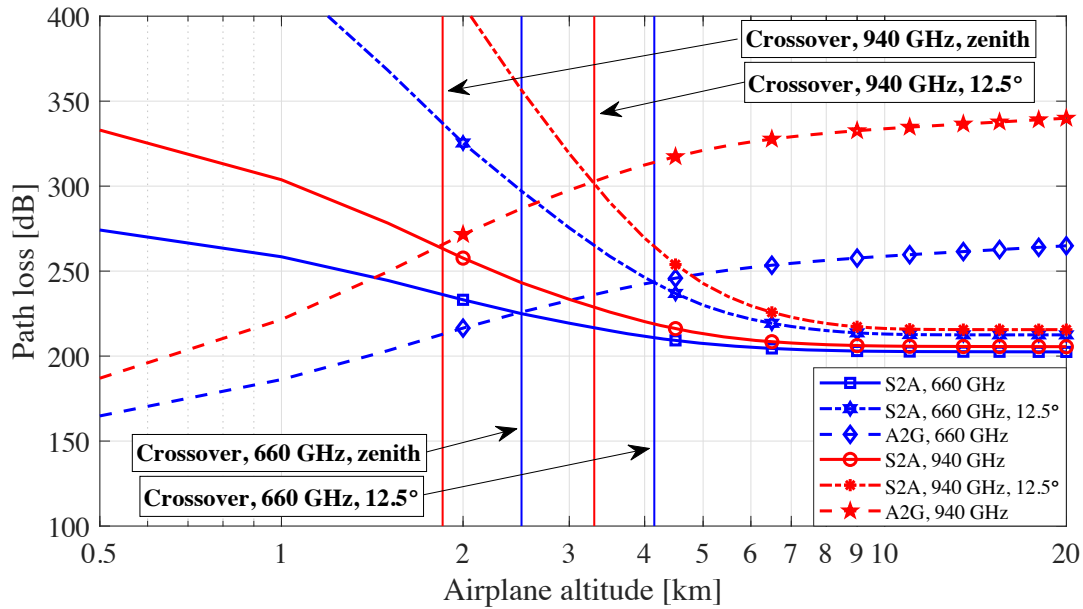


Fig. 24: Comparison of A2G and A2S path losses at 660 and 940 GHz center frequencies as a function of the airplane altitude for zenith paths and inclined paths with crossover altitudes where of the minimum loss path.

Figure 24 shows the path losses from A2G and S2A (or G2A and A2S) as a function of the airplane altitude for 660 GHz and 940 GHz center frequencies. The satellite is in LEO orbit at 500 km height and we show results also for a 12.5-degree elevation difference. As can be expected, there is a crossover altitude for the lower loss connection direction. This crossover point depends on the frequency and the elevations of the airplane and the satellite but suggests that the best way to provide THz connectivity would be to have both links available. During the mid-flight, the satellite link provides better connectivity. Considering that the use of electronic devices is not allowed during the takeoff and landing, the satellite link alone would be enough for passenger use. These crossover altitudes are approximately 1.85 km (at $\rho = 0^\circ$) and 3.3 km (at $\rho = 12.5^\circ$) for 940 GHz frequency and 2.5 km (at $\rho = 0^\circ$) and 4.15 km (at $\rho = 12.5^\circ$) for 660 GHz. It should also be noticed that the calculation herein does not take into account the movement of the airplane that would have a more severe impact toward the ground since the link distance is much lower than towards the satellites. The true path loss towards the ground would therefore also depend on the available ground stations and their locations, which would increase the distance and the path loss, most likely very significantly rendering the low altitude communications challenging. Especially with physical beam steering that is not as agile for handovers as electronic beam steering techniques.

4.3.2 Experimental Results

Sounding Environment and Methodology The measurement campaign is performed in an outdoor urban environment considering a static point-to-point link. The 70 m-long link is set up between two buildings on the Northeastern University campus. We transmit from the 6th floor of

Table 4: Mean and variance (left and right values of each cell) of weather data collected from "Accuweather" weather service.

	Temp. (C)	Water Vapor (g/m ³)	Pressure (hPa)	Precip. (mm/hr)
Clear	12.23, 0.60	5.15, 0.18	1026, 0.51	0, 0
Rain	7.23, 0.39	7.09, 0.05	1014, 0.43	1.84, 1.22
Snow	-2.32, 0.50	3.54, 0.03	1020, 3.30	0.45, 0.08

the Interdisciplinary Science and Engineering Complex (ISEC) to the roof of the Snell Library, shown in Fig. 14b. The channel measurement was performed in three different weather conditions: (i) Clear, (ii) Rain, and (iii) Snow, to characterize link performance in adverse weather. We recorded data in 3-minute intervals over 6 hours, collecting 120 frames of data for each scenario. We use "Accuweather" weather service to record real-time atmospheric conditions that are later used to estimate the theoretical link characteristics.

Link Budget Analysis The link budget incorporates the transmit power, the hardware's gains and losses, the channel losses describe before (FSPL and molecular absorption as well as scattering for the rain and snow scenarios).

The theoretical power received, $P_{r,T}$, is modeled with our link budget

$$P_{r,T} = P_s + G_s + G_r + G_h - L_h - L_{FSPL} - L_{abs} - L_{sc} , \quad (18)$$

where P_s is the transmit power before the antenna, G_s , G_r are the transmitter and receiver antenna gains, G_h is the receiver-side hardware gain, L_h is the receiver-side hardware loss, L_{FSPL} is the FSPL of the center frequency 140 GHz, L_{abs} is the molecular absorption loss, and L_{sc} is the loss caused by scattering.

The RF transmitter outputs $P_s = 16$ dBm, which combined with $G_s = 38$ dBi from the horn antenna produces an equivalent isotropic radiated power (EIRP) of approximately 54 dBm. A similar 38 dBi antenna is used at the receiver. The hardware gain comes from the low noise amplifier (LNA) operating at IF, and the losses come from the mixer conversion loss and the cable insertion loss.

Absorption and Rain Losses Prediction: Utilizing the aforementioned ITU-R P.676-12 and ITU-R P.838-3 models, we predict the loss for frequencies of 1-1000 GHz as functions of temperature, pressure, water vapor, and rainfall rate. The average observed values of the relevant atmospheric readings are given in Table 4.

The scattering loss from rain, L_{rain} , is computed as

$$L_{rain} = k R_{rain}^{\alpha} , \quad (19)$$

where R_{rain} is the rain precipitation rate in mm/hr, while the coefficients k and α are calculated for horizontal polarization [58].

For the 130-150 GHz band and a link distance of 70 m, the losses from molecular absorption and rain scattering were in the order of 0.01 dB and 0.1 dB, respectively, thus ultimately negligible.

Snow Scattering Losses Prediction: Although there is no standard model for snow scattering of terahertz waves, pre-existing literature that can be leveraged [56, 65], which details how to calculate

scattering loss due to snowflakes using Mie scattering theory [66]. Mie scattering theory applies to any spherical dielectric particle which can include snowflakes. The dielectric constant of a snowflake, ϵ , which is a mixture of water and ice, is calculated with Debye's mixture theory [67] using the dielectric constants of water and ice at terahertz frequencies [68]. For propagation through air, the Mie extinction cross-section is defined as [66]

$$\sigma_{ext}(n, \chi) = \frac{2}{\chi^2} \sum_{m=1}^{\infty} (2m+1) \Re\{a_m(n, \chi) + b_m(n, \chi)\}, \quad (20)$$

where $n = \epsilon^{1/2}$ and $\chi = \frac{2\pi r}{\lambda}$, and r is the radius of the snowflake, λ is the center frequency's wavelength, a_m , b_m are the Mie coefficients, and $\Re$ indicates the real part of the coefficients.

To calculate the total attenuation due to snow scattering, L_{snow} , we use a modified form of the equation from [66],

$$L_{snow} = 4.343 \cdot 10^3 \sigma_{ext} N(r, R_{snow}) r F, \quad (21)$$

where $N(r, R_{snow})$ is the Gunn-Marshall (G-M) snow size distribution [69], R_{snow} is the snowfall rate, and F is the number of snowflakes present in the antenna beam.

The antenna beam's volume, V , is a cylinder with the height equal to the link distance, $H = 70$ m, and the diameter equal to the lens width of the horn antenna, $D = 0.122$ m. The number of snowflakes in the antenna beam at any instance can be calculated based on the snow characteristics. For an estimated constant snow density, $\rho = 0.52$ g/cm³, snowflake terminal velocity, $v = 1.5$ m/s, and snowflake weight, $m = 2.5$ mg [56], the number of snowflakes in the antenna beam becomes a function of the time-varying snowfall rate, $R_{snow}(t)$,

$$F(t) = R_{snow}(t) \left(\frac{\pi}{14.4} \frac{D^2 H \rho}{vm} \right), \quad (22)$$

where $F(t)$ is the time-varying number of snowflakes from (21). The equation calculates the accumulation of snow in the antenna beam's cylindric volume after $\frac{D}{v}$ seconds scaled by the snow density, and it factors in unit conversion. For $R_{snow} = 0.45$ mm/hr (mean snow precipitation in Table 4), there are approximately 15 snowflakes scattering the antenna beam.

Received Power Model The theoretical received power from (18) is calculated with the modeled absorption and scattering losses. The modeled $P_{r,T}$ is plotted as a function of time along with the experimentally measured received power, $P_{r,E}$, in Fig. 25. We find that the experimental and modeled P_r are relatively close. Their root mean square error (RMSE) is in the order of 1 dB for all the considered scenarios.

The clear and rain $P_{r,E}$ are mostly flat throughout their runtimes as the weather didn't significantly change during those days. The empirical link budget value matches the analytical one for clear conditions, but there is a constant offset of around 1 dB for the rain $P_{r,T}$. The latter is possibly due to a minor antenna misalignment.

The presented snow scattering model well matches the average received power measured experimentally. The model is also able to follow the trend at around the two hour mark in the plot. The snowfall greatly increased for one hour causing $P_{r,E}$ to drop by over 2 dB, and the developed model captures this effect.

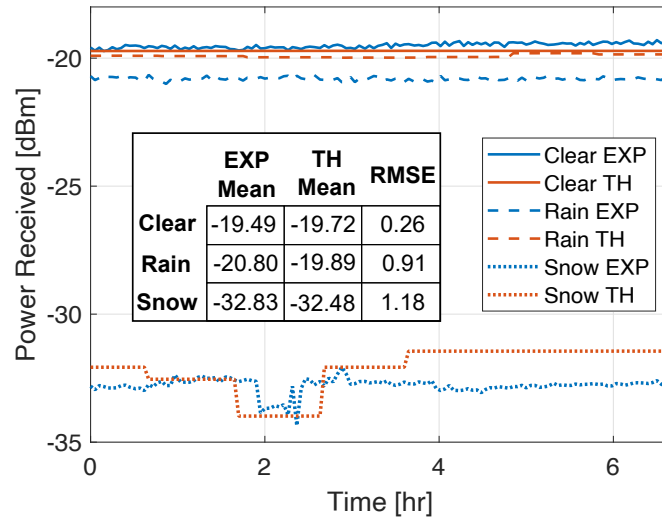


Fig. 25: Experimental and theoretical received power. Table: mean experimental and theoretical received power received and RMSE.

Multipath Profile The sliding correlator generates multipath profiles by detecting delayed copies of the transmitted m-seq. This measured channel impulse response and its multipath components (MPC) are plotted for one instance in Fig. 26. The channel impulses/peaks are estimated through global convergence method based on the time resolution and dynamic range of the system.

Although the system can detect delays up to 409.5 ns, we find that there are not a significant amount of MPCs beyond the 25 ns delay. We notice that for clear and rain scenarios a cluster of MPCs with relatively high power is located around 20 ns. The same MPCs are weaker in the snow's channel impulse response. This is likely due to a reflection off of the roof in front of the receiver, see Fig. 14b, whose power decreases when the roof is covered by a layer of snow. In the next paragraphs, we analyze the MPC profiles of the collected frames to estimate relevant channel metrics.

Channel Metrics The outdoor point-to-point link is characterized in terms of the Rician K-factor and the root mean square (RMS) delay spread, which are among the crucial metrics to design a wireless communication system.

K-factor: The Rician K-factor is the ratio of the power in the LoS path (P_{LoS}) and other NLoS paths ($2 S^2$), and it thus given by

$$= \frac{P_{LoS}}{2 S^2}. \quad (23)$$

In dB scale it is represented by $10 \log_{10}(\cdot)$. The K-factor is a significant component in static point-to-point links, as it characterizes the distribution of the power across LoS and reflected paths.

The cumulative distribution function (CDF) of the K-factor for clear, rain, and snow, obtained from different time instances, is shown in Fig. 27. The K-factor is reasonably steady for all three weather conditions, as we consider a static point-to-point link. Further, the metric variation is significantly low with the rain and snow rate changes. The K-factor in dB scale can be approximated by the normal distribution with mean, μ , of 35.2 dB, 33.5 dB, and 29.8 dB, accompanied by

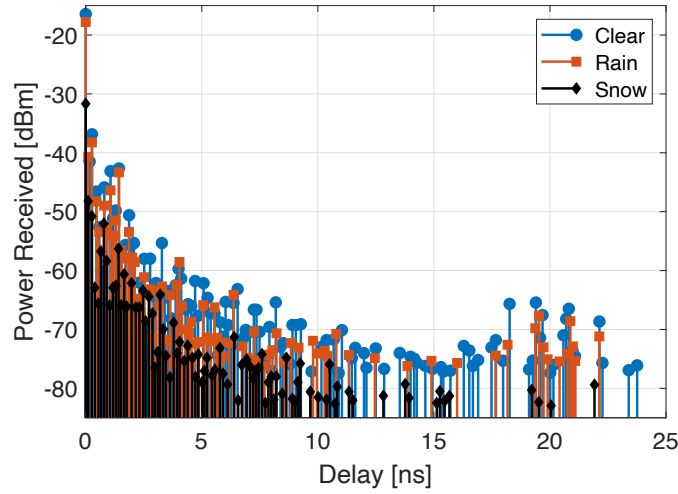


Fig. 26: Multipath profiles for clear, rain, and snow weather.

the variance, σ^2 , of 0.25, 0.19, and 0.22 for clear, rain, and snow, respectively. The statistical measurements are obtained by the maximum likelihood estimation (MLE). Moreover, the K-factor decreases by 1.7 dB and 5.4 dB in case of rain and snow, respectively, compared to clear weather. The latter suggests that wireless THz communications and networks will need adaptive link modulation and coding scheme techniques to react to the different channel parameters with varying weather conditions.

Delay Spread: To describe the multipath richness of the link, the root mean square (RMS) delay spread, τ_{RMS} , is calculated in different weather conditions as follows

$$\tau_{RMS} = \sqrt{\frac{\sum_i (d_i - \hat{d})^2 p_i^2}{\sum_i p_i^2}}, \quad (24)$$

where d_i and p_i are the delay and received power of propagation path i , respectively, and $\hat{d}$ is the mean delay represented by

$$\hat{d} = \frac{\sum_i d_i p_i}{\sum_i p_i}. \quad (25)$$

For our point-to-point link setup, the value for the delay spread is notably small (in the range of sub-nano seconds). The CDF of the RMS delay spread (captured at different time instances) for three weather conditions is shown in Fig. 28. The empirical CDF can be approximated by the CDF of the Rician distribution with the non-centrality values $\nu = 0.29, 0.23, 0.39$, and the scale parameter values $\sigma = 0.024, 0.020, 0.048$ for clear, rain, and snow conditions, respectively.

However, a closer fit is found with the stable distribution [70], which is defined by a family of distributions capturing certain heavy-tail behaviors. In fact, for the point-to-point link with a high K-factor (as shown in Fig. 27), the RMS delay spread exhibits different behaviors on the left and the right sides of the peak of its probability density function. Decaying on the left-hand side is significantly faster than on the right, which is observed from the CDF curves in Fig. 28. Therefore,

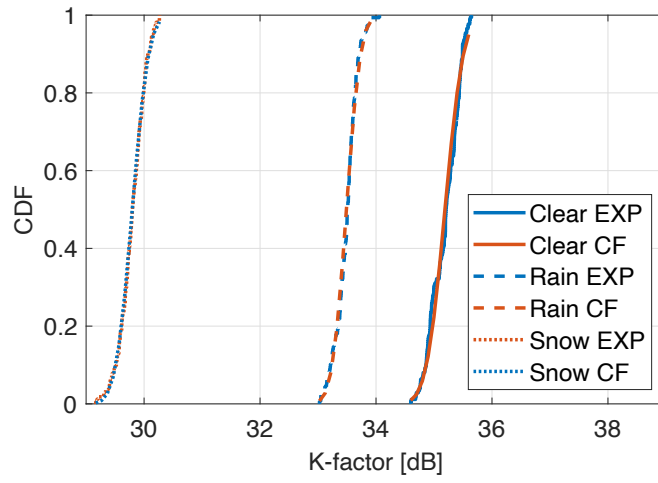


Fig. 27: The CDF of K-factor [dB] in different weather conditions for experimental (EXP) results and curve fitting (CF) values.

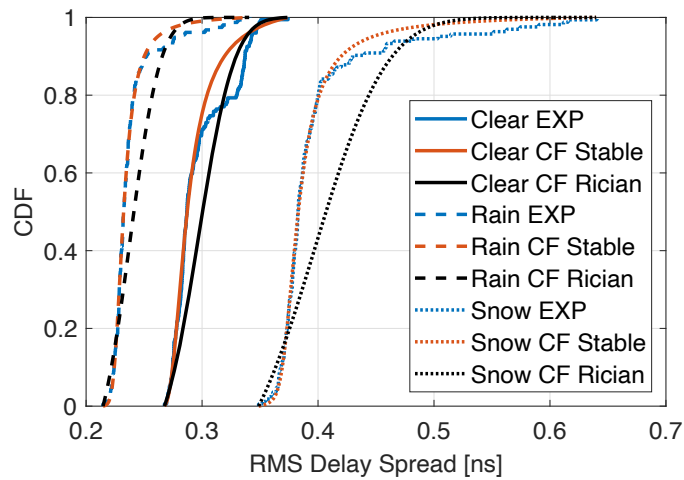


Fig. 28: The CDF of RMS delay spread in different weather conditions for experimental (EXP) result and curve fitting (CF) value with the Stable and the Rician distributions.

Table 5: The α (first shape), β (second shape), γ (scale), and δ (location) parameters of the stable distribution for clear, rain, and snow.

	α	β	γ	δ
Clear	1.07	1.0	0.0086	0.28
Rain	1.28	0.9	0.0052	0.23
Snow	1.14	0.8	0.0092	0.38

the stable distribution presents a closer fit for RMS delay spread, being able to model the complex nature of the tail and non-tail regions. The calculated distribution parameters are shown in Table 5.

Considering the 20 GHz bandwidth (symbol duration of 0.1 ns), the delay spread is large enough to create ISI. This aspect should be further considered in designing physical- and link-layer mechanisms for reliable communications. Further, the RMS delay spread value goes above 0.5 ns for more than 10% of the cases in snow weather, which is due to the scattering of the signal by snowflakes. Therefore, we need to expand and/or adapt the state-of-the-art link adaptation techniques to still maintain reliable data exchange during harsh weather conditions.

Overall, our results show that the impact of weather is limited and does not disrupt the wireless link in the urban context where we deployed our transceivers. The snow introduces the highest path loss increase – around 13 dB when compared to clear weather – while rain only increases loss by 1.3 dB. In addition, the delay spread with snow exceeds 0.5 ns for more than 10% of the measured frames, which is enough to create ISI when communicating with 4 GHz of bandwidth or more. This analysis provides insights and takeaways that will be pivotal for the design of reliable communications and link layer solutions above 100 GHz.

5 ULTRABROADBAND MODULATIONS FOR TERAHERTZ COMMUNICATIONS

5.1 Motivation

As mentioned in the introduction, there are many ongoing efforts aimed at determining *the best waveform* for future terahertz communication systems. As part of this effort, besides the OFDM implementation discussed in Sec. 3, we have looked into low-complexity single-carrier modulations that can overcome the limitations of terahertz devices and even leverage the peculiarities of the terahertz channel. In particular,

- We have proposed, modeled and experimentally studied the performance of M-APSK modulations for low-PAPR spectrally-efficient ultrabroadband terahertz communication. The modulation design is based on arranging the constellation points in a number of concentric rings, similar to PSK. The number of points in each ring proportionally varies with its amplitude, maintaining the same minimum distance within adjacent points. Thus, there are fewer points in an inner circle compared to the outer circle. Concentric ring amplitude and phase modulations were introduced over fifty years ago [71], and several works have been done subsequently, suggesting differential encoding as well as error-correcting coding techniques [72–74] to increase the reliability of the scheme. In most cases, the study is confined to APSK utilized in digital video broadcasting for satellite applications (i.e., DVB-S2) [75], which have a specific constellation design considering the fixed number of rings. Instead, in this work, we have optimized the constellation of M-APSK modulations with the number of rings as a constraint, developed a generalized closed-form SER expression, and studied the PAPR performance.
- we have conceived, modeled and experimentally tested innovative Hierarchical Bandwidth Modulations (HBM). HBM is an extension of traditional hierarchical modulation (HM) that introduces a hierarchy in bandwidth by simultaneously servicing users at different symbol rates. The concept of changing the symbol rate in multi-resolution constellations is a novel approach motivated by the distance-dependent bandwidth in the THz band. First proposed in [76], HBM uses the THz band’s molecular absorption losses to increase the data rate of single-transmitter multiple-receiver (STMR) systems and to multiplex users in space. To our knowledge, it is the only work that leverages the THz channel’s high path loss to increase the system’s overall data rate and offer a solution for high propagation loss scenarios across a range of frequencies. As part of this effort, we have derived the error rate performance for 4/M-QAM HBM receivers; defined the HBM *functional region* necessary for designing successful HBM systems; suggested a *transition region* for receivers experiencing fluctuating channels; verified our error rate derivations with full end-to-end simulations of a THz transmitter, channel, and receiver in MATLAB; and experimentally verified the existence of the HBM functional region in the THz band.

5.2 Methods, Assumptions and Procedures

5.2.1 Design of M-ary Amplitude Phase Shift Keying Modulations

The proposed M-APSK modulation is adopted to enhance the THz band communication spectral efficiency by increasing the modulation order. The PAPR and SER performance of higher-order modulation are two crucial limiting factors to implementing the schemes for THz ultrabroadband communication. By the proposed method, the PAPR is reduced compared to M-QAM by arranging the constellation points astutely over fewer constant amplitude rings. In addition, by arranging the constellation points into separate rings, the SER is improved significantly in contrast of the M-PSK

Constellation Design The purpose of designing the constellation points is to optimize the minimum distance within the different concentric rings, i.e., keeping it same, while the M-ary points are divided into those separate amplitude rings similar to the PSK system. To satisfy the constant minimum distance between constellation points as well as between the rings (Fig. 29a)), we keep the distance $d_{min1} = d_{min2} = R_{min}$ for the two-ring system. Therefore, the system design following the minimum distance equation of the M-PSK we have

$$\sin\left(\frac{\pi}{X}\right)2A = \sin\left(\frac{\pi}{M-X}\right)2A(A - \sin\left(\frac{\pi}{X}\right)2A), \quad (26)$$

where M is the modulation index, X and $M - X$ is the number of constellation points on the outer, i.e., a ring with maximum amplitude, A , and inner ring, respectively. The separation between the rings is maintained by $A - \sin\left(\frac{\pi}{X}\right)2A$.

Similarly, the constellation design equation for the 3 ring system (Fig. 29(b)) is based on $d_{min1} = d_{min2} = d_{min3} = R_{min1} = R_{min2}$ and given by

$$\sin\left(\frac{\pi}{X}\right)2A = \sin\left(\frac{\pi}{Y}\right)2A(A - \sin\left(\frac{\pi}{X}\right)2A) = \sin\left(\frac{\pi}{M-X-Y}\right)2A(A - 2 \times \sin\left(\frac{\pi}{X}\right)2A), \quad (27)$$

where Y is the number of constellation points in the middle ring.

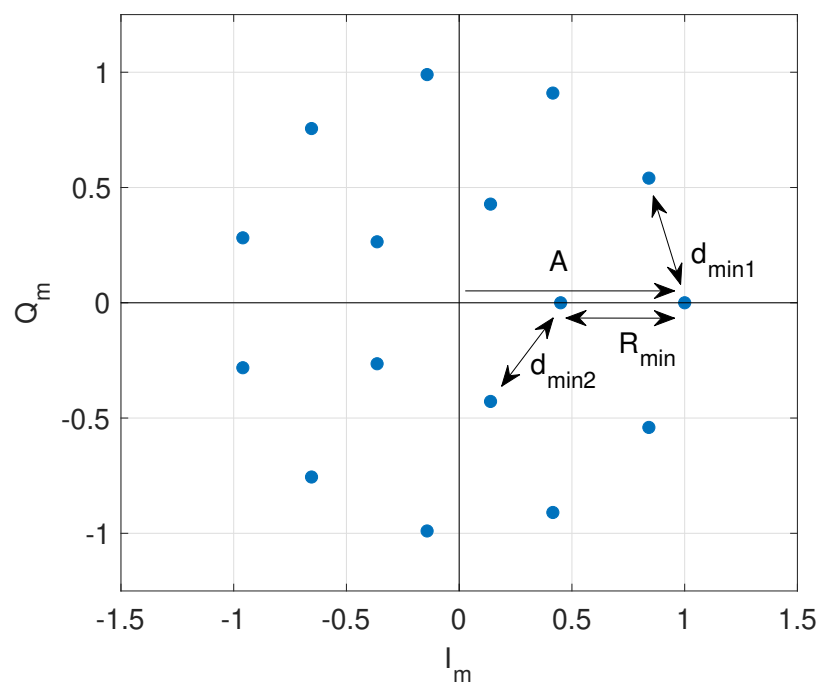
To find the solution for X , we use the approximation $\sin\left(\frac{\pi}{K}\right) \approx \left(\frac{\pi}{K}\right)$, where K is the number of constellation points in any ring. The approximation satisfies closely if $K \geq 4$ (i.e., the approximation error < 0.1). In other words, the innermost ring should contain at least four constellation points. The solution for X with a two-ring system is obtained by solving (26), which can be written as

$$X = \left(\frac{M}{2}\right) + \pi, \quad (28)$$

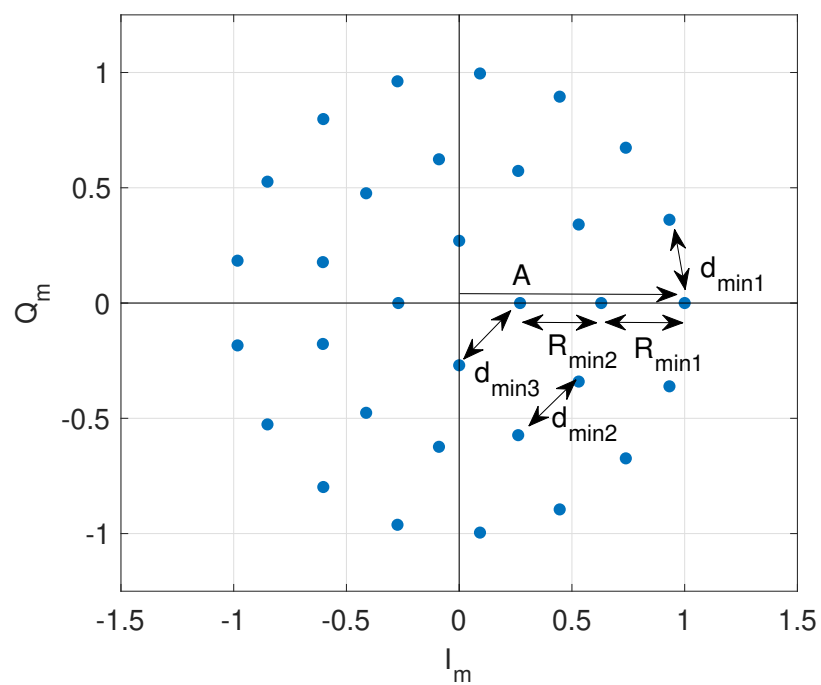
and the solution for constellation points in the innermost ring is $M - X$. In a similar fashion, (27) can be solved for the solution for X and Y for three-ring case which yields to

$$X = \left(\frac{M}{3}\right) + 2\pi, Y = X - 2\pi. \quad (29)$$

Here, the constellation points in the innermost ring are $M - X - Y$. Following these equations, the two-ring 16-APSK and 32-APSK correspond to $X = 11$, $M - X = 5$ and $X = 19$, $M - X = 13$, respectively. Similarly, the three-ring 32-APSK yields to $X=17$, $Y=11$ with $M-X-Y=4$



(a) Two-ring 16-APSK



(b) Three-ring 32-APSK

Fig. 29: Constellation design with design criteria: a) 16-APSK with two rings; (b) 32-APSK with three rings to depict the different scenarios.

Approved for Public Release; Distribution Unlimited.

constellation points in the innermost ring. While the derived optimum constellation design is asymmetric, it can easily be changed to a symmetric one by rounding up the constellation points in each ring to the closest even number with a slight increase in SER.

Further, the equations impose limitations in the case of the minimum M requirement while utilizing a specific number of rings, R . This constraint comes from i) comparing the maximum amplitude with the separation between the rings that need to be maintained, and ii) the number of constellation points in the innermost ring. Precisely, the constraint can be expressed as

$$A - \sin\left(\frac{\pi}{X}\right) 2(R-1)A > 0 \quad (30)$$

and

$$C_{innermost} \geq 4, \quad (31)$$

respectively, where R is the number of rings we want to utilize and $C_{innermost}$ is the number of constellation points in the innermost ring. The system design should jointly obey both equations. Therefore, by substituting the value for X , Y and $C_{innermost}$ (i.e., $M-X$ and $M-X-Y$ for 2-ring and 3-ring, respectively) and solving them, we have the minimum number of M for a specific value for R . For example, the minimum number required for modulation index, M , is 16 and 32 for two and three-ring systems, respectively, considering that M is an integer of power of 2. Moreover, the solution can be extended for higher-order modulation, M , with added rings, R , in the same way. The scheme does not provide a proper fit for $M \leq 8$ schemes, making this complementary to the IEEE 802.15.3d standard schemes for low order modulation [38].

Symbol Error Rate The theoretical SER for M-APSK can be obtained by solving the SER for the X -ary PSK, P_{XPSK} (largest ring), with a minimum distance of $d_{min} = \sin(\frac{\pi}{X})2A$, which is kept equal for any two adjacent points in any ring by design and the error introduced probabilistically due to the interaction of constellation points between the neighboring rings, P_{ring} . Therefore, the SER depends on the number of rings utilized for the system and decreases as the number of ring increases, which is directly related to the decrease in the number of X . The expression for P_{XPS} (i.e., intra-ring error probability) is given by

$$P_{XPSK} = 2Q\left(\frac{\sin(\pi/X)2A}{\sqrt{2N_0}}\right), \quad (32)$$

where N_0 is noise power spectral density. Furthermore, there are contributing terms in the effective/total SER due to the error incorporated by one ring's constellation points to the other rings. It can be statistically shown that the inter-ring error probability decreases with the decrease in constellation points of the inner rings compare to the outermost ring's constellation point, X . So, the expression for the two-ring system is given by

$$P_{2ring} \approx \frac{1}{2 \times 2} \left[2 \frac{M-X}{X} Q\left(\frac{\sin(\pi/X)2A}{\sqrt{2N_0}}\right) + 2 \frac{M-X}{X} Q\left(\frac{2(A - \sin(\pi/X)2A)}{\sqrt{2N_0}}\right) \right]. \quad (33)$$

Similarly, for the three-ring system, the SER due to interaction between the rings is given by

$$P_{3ring} \approx \frac{1}{2 \times 3} \left[2 \frac{Y}{X} Q \left(\frac{\sin(\pi/X) 2A}{\sqrt{2N_0}} \right) + 2 \frac{M - X - Y}{X} Q \left(\frac{\sin(\pi/X) 2A}{\sqrt{2N_0}} \right) + 2 \frac{M - X - Y}{X} Q \left(\frac{2(A - \sin(\pi/X) 4A)}{\sqrt{2N_0}} \right) \right]. \quad (34)$$

The above equation can be extended for higher-order modulation with more number of rings in same way. Therefore, the effective/total SER for M-APSK, P_{eff} , is obtained by

$$P_{eff} = 1 - (1 - P_{XPSK})(1 - P_{ring}), \quad (35)$$

which can be approximated as

$$P_{eff} = P_{XPSK} + P_{ring}. \quad (36)$$

The comparison of SER of M-QAM, M-PSK, and M-APSK is shown in Fig. 30 for $M = 16$ and $M = 32$. It is observable that the simulated SER closely matches the analytical result, which confirms our numerical calculation. The slight mismatch is due to the approximation of constellation points in a particular ring to the nearest integer value. It is depicted that the performance of M-APSK is better than M-PSK in both cases. However, there is a small improvement compared to M-QAM if the maximum number of rings is utilized for M-APSK, i.e., $R = 2$ for $M = 16$ and $R = 3$ for $M = 32$. M-QAM functions better than M-APSK if the ring number is reduced. This is a trade-off with PAPR.

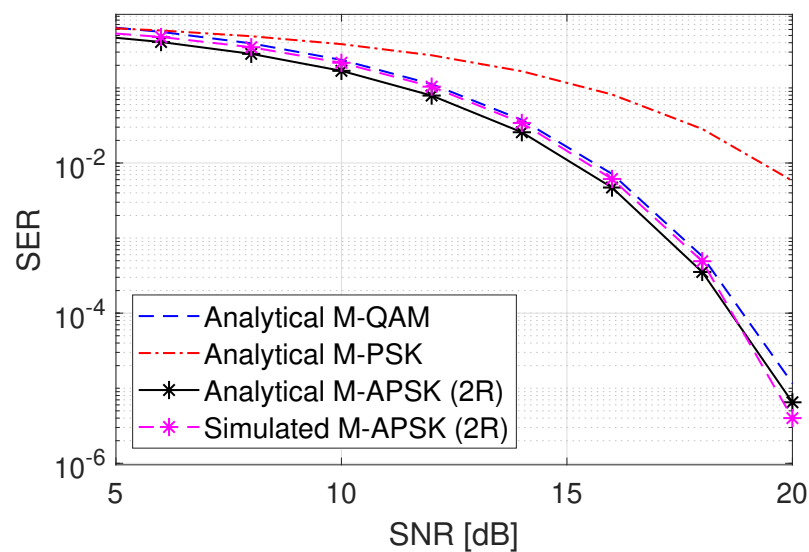
Peak to Average Power Ratio The PAPR is one of the essential metrics that restricts the utilization of spectral efficient modulation techniques, such as conventional high-order QAM, OFDM for THz band communication, which arises mainly due to the front-end electronics' efficiencies at THz frequencies dominated by the PAs. Literature shows that the P_{sat} , gain and efficiency decrease by following $\frac{1}{f_c^2}$ relation [77]. Therefore, in most cases, the single-carrier communication waveform with a low modulation index is proposed for THz band communication. The M-APSK scheme has a comparatively low- PAPR than other amplitude modulation schemes. The constellation points are suitably divided into fewer constant amplitude rings to reduce the PAPR as well as SER. However, the PAPR is increased with the number of rings utilized to split the points. The average power, P_{avg} , of the scheme with two/three-rings is given by,

$$P_{avg-2R} = \frac{1}{M} [XA^2 + (M - X)(A - \frac{\pi}{X} 2A)^2], \quad (37)$$

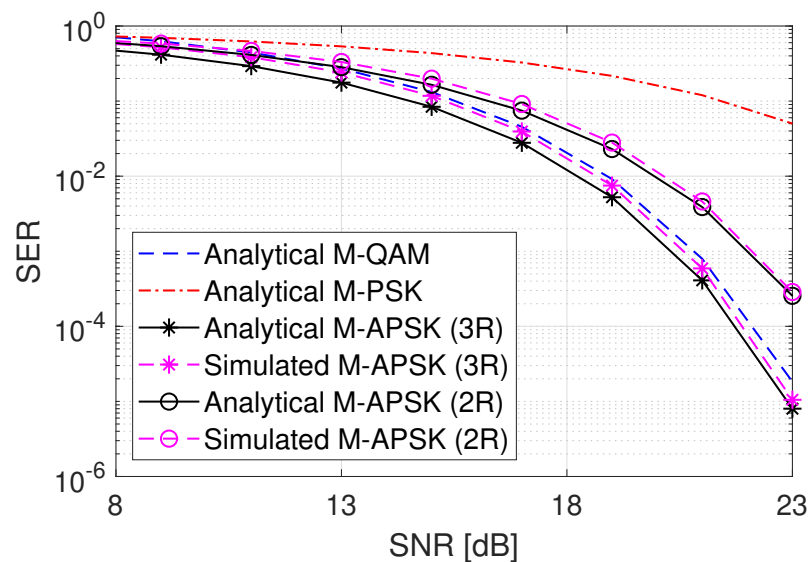
$$P_{avg-3R} = \frac{1}{M} [XA^2 + Y(A - \frac{\pi}{X} 2A)^2 + (M - X - Y)(A - \frac{\pi}{X} 4A)^2], \quad (38)$$

respectively, where, A^2 is the peak power.

In Table 6, the comparison of PAPR between M-APSK, M-PSK and M-QAM is shown. It is observable that the PAPR is reduced considerably compared to the M-QAM counterpart. Also, the PAPR decreases with the number of rings. However, the PAPR for the M-PSK scheme is the lowest among the three and is 0 dB. Therefore, the system design supports utilizing the higher-order



(a) $M=16$



(b) $M=32$

Fig. 30: SER comparison between M-QAM, M-PSK, and M-APSK with increasing SNR: a) $M=16$; (b) $M=32$.

Table 6: Comparison of PAPR.

Modulation	PAPR (dB)	
	M=16	M=32
PSK	0	0
QAM	2.5	3.2
M-APSK 2R/3R	1.22/NA	1.08/1.73

schemes while keeping the PAPR in a considerable range and could be utilized to increase the spectral efficiency without performance degradation due to PA limitations. Besides, as the number of rings provides a trade-off between PAPR and SER, it can be utilized as a design metric depending on the system requirement.

5.2.2 Design of Hierarchical Bandwidth Modulations

Fundamentals of HBM An HBM system uses a hierarchical constellation similar to HM, but HBM also introduces a hierarchy in bandwidth and therefore in the symbol duration as well. Traditional HM [78, 79] is a modulation scheme for STMR systems that employs a hierarchical constellation. It was presented for broadcast backwards compatible digital television systems [80], has been optimized for RF broadcast scenarios [81], and has recently been used for millimeter wave communications [82]. A similar concept can offer the flexibility and higher aggregate data rates we desire for the THz band. In traditional HM, receivers that observe a higher signal-to-noise ratio (SNR) demodulate a higher order constellation than users with a lower SNR [78]. An example of a 4/16-QAM hierarchical constellation is shown in Fig. 31. Notice that the points are not equidistant from each other. If a system were to use the multi-resolution 4/16-QAM shown in Fig. 31 receivers experiencing a high SNR would demodulate each of the 16 individual constellation points. Meanwhile, receivers farther away experiencing a lower SNR would only demodulate the 4-QAM corresponding to the four-point clusters in each quadrant. Thus, these receivers would see a lower resolution of information. In a THz communications context, this flexibility for receivers with respect to SNR allows users to adjust to their perceived channel while still successfully demodulating necessary information. However, traditional HM does not consider the distance-dependent bandwidth of the THz channel into account.

The differences between an HM and an HBM signal are shown in Fig. 32a and Fig. 32c. Although both the HM symbols and HBM symbols contain base level of information (the blue symbols) sent with a larger amplitude than the secondary information symbols (the yellow symbols), the HBM high-resolution symbols occur at twice the rate as those for HM. Thus, HBM can achieve faster aggregate data rates. This hierarchy in bandwidth can also be seen in the signal spectrum itself. Fig. 32b and Fig. 32d show the baseband spectrum of an HBM and traditional HM signal respectively. The HBM signal has a high-resolution bandwidth of 10 GHz and a low-resolution bandwidth of 5 GHz. The low-resolution bandwidth is seen as a main lobe from 2.5-7.5 GHz, and on either side of it there are smaller lobes bringing the entire high-resolution bandwidth from 0-10 GHz. These additional lobes on either side of the main lobe do not contain any low-resolution information, and a low-resolution system would not see them. In fact, if the low-resolution receiver could see them at all, it would likely filter them out using a bandpass filter. Even without those portions of

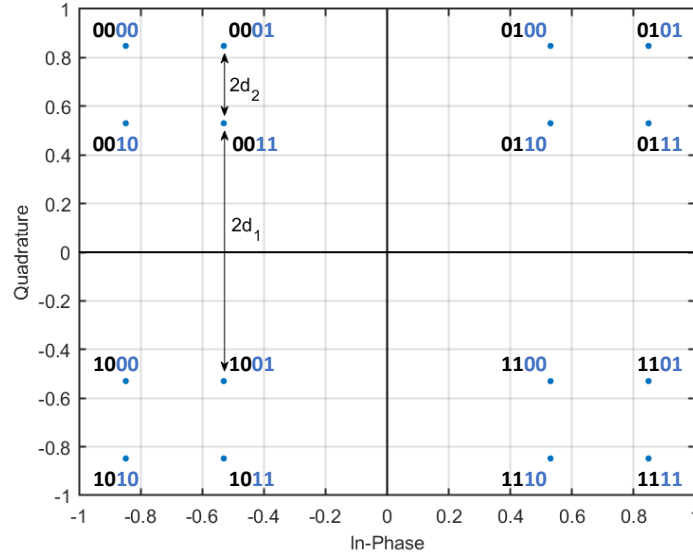


Fig. 31: Hierarchical 4/16-QAM Constellation

the signal, the low-resolution receiver would still receive a full lobe of bandwidth information due to the hierarchical design of the signal. The corresponding HM spectrum only has one main lobe corresponding to the smaller bandwidth; its hierarchy only exists in amplitude whereas for HBM there is also a hierarchy in bandwidth and in time. The frequency spectra shown in Fig. 32b and Fig. 32d also show that HBM can use an overall broader bandwidth than traditional HM can use, which enables HBM to achieve the higher aggregate data rate.

This hierarchy in bandwidth is useful in the THz band, especially if the transmission is close to an absorption line. The channel between the transmitter and users at farther transmission distances will filter out the high-resolution sidelobe(s). When absorption is not present, HBM can still improve performance by allowing users with a lower SNR to use a smaller bandwidth and therefore reduce the observed noise power. Furthermore, although we focus on the LoS scenario, non-LoS and multi-path channels are being explored for THz communications [83, 84]. In the LoS case in the presence of multi-path, the multi-path components are traveling a longer distance than the LoS, which will result in the absorption lines filtering part of the signal. In the case where the primary path is a non-LoS path, HBM analysis is still valid as long as the total path length is considered.

Modulator Design In order for this modulation scheme to work, the adjacent high-resolution HBM symbols within one low-resolution symbol shown in Fig. 32 must exist within the same cluster of the constellation shown in Fig. 31. This interrelation can be achieved by using the block-wise modulator in Fig. 33. Each block represents a corresponding level of resolution in the constellation and in time. The symbol duration for the modulator block 1, 2, ... M is given by $T_{s1} = n_1 T_s, T_{s2} = n_2 T_s, \dots, T_{sM} = n_M T_s$, respectively. Here, T_s is the base symbol duration, and $\frac{n_1}{n_2}, \frac{n_2}{n_3}, \dots, \frac{n_{M-1}}{n_M}$ should all be integers. Thus the number of transmitted bits per the base symbol time are given by $\frac{\log_2(m_1)}{n_1} + \frac{\log_2(m_2)}{n_2} + \dots + \frac{\log_2(m_M)}{n_M}$. The total modulation order achieved after level M is $m_T = m_1 \times m_2 \times \dots m_M$, where m_i is the modulation index for the i^{th} modulator block. For

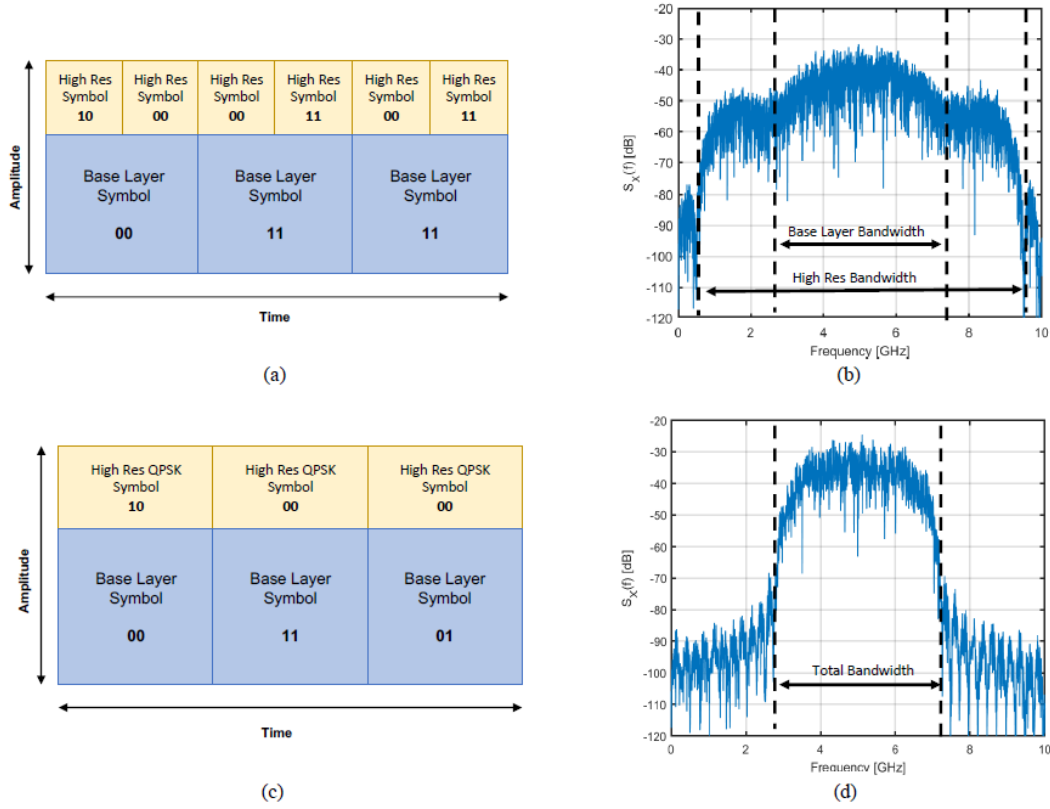


Fig. 32: Amplitude and Time Distribution of (a) an example 4/16-QAM HBM System and (b) its frequency spectrum compared with (c) the amplitude and time distribution of a traditional 4/16-QAM HM System and (d) its frequency spectrum.

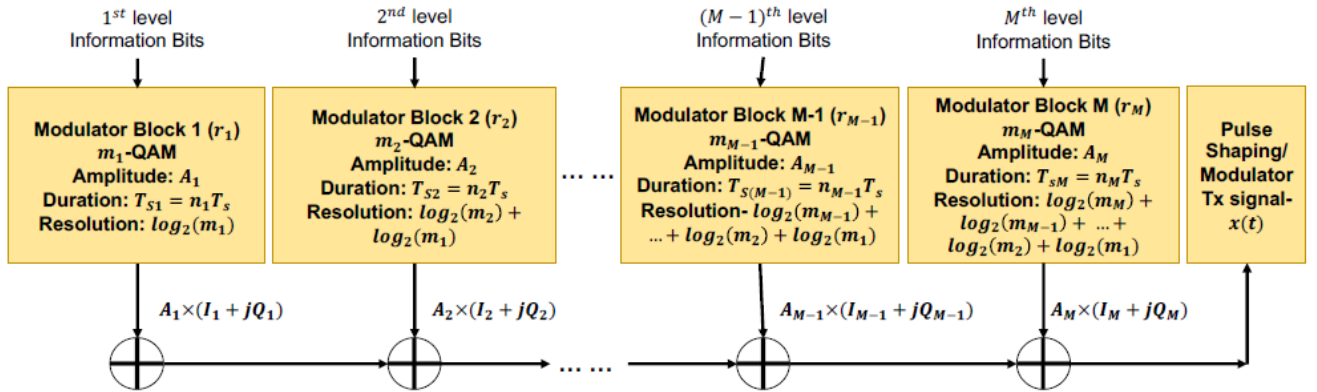


Fig. 33: Block diagram for hierarchical bandwidth modulator

efficient utilization of HBM in terms of bandwidth usage the ratio would often be 2, which implies $\frac{T_{S1}}{T_{S2}} = \frac{n_1}{n_2} = 2$. This indicates that for fixed transmission power, higher level modulations are held for shorter times, resulting in a lower energy per bit E_b . However, over the shorter distances at which these symbols are expected to be received, the channel path loss is lower which results in higher E_b/N_o values at these receivers.

The m_i -QAM constellation at each block is scaled by A_i such that $A_1 > A_2 > \dots A_M$ and such that $\sum_{m=1}^M A_m$ satisfies the system energy requirements. It is also important to choose the amplitudes such that constellation points from subsequent modulators do not overlap. This can be achieved using the conditions given by

$$\frac{A_1}{(\sqrt{m_1} - 1)} \geq A_2 + A_3 + \dots A_M, \quad (39)$$

$$\frac{A_2}{(\sqrt{m_2} - 1)} \geq A_3 + A_4 + \dots A_M, \quad (40)$$

...

$$\frac{A_{M-1}}{(\sqrt{m_{M-1}} - 1)} \geq A_M. \quad (41)$$

These amplitude values and the constellation orders of each block generate a hierarchical constellation. To generate a hierarchy in bandwidth, each block operates with a different symbol duration.

The bits at each block are mapped to the required I and Q symbols within the m_i -QAM constellation. Thus the transmitted I and Q values for the highest resolution symbol time can be represented by

$$I + jQ = A_1(I_1 + jQ_1) + A_2(I_2 + jQ_2) + \dots A_M(I_M + jQ_M), \quad (42)$$

where $I_1 + jQ_1, I_2 + jQ_2, \dots I_M + jQ_M$ are normalized I and Q components generated by the modulator blocks 1 through M . In the next instance of T_{sM} , $I_1 + jQ_1, I_2 + jQ_2, \dots I_{M-1} + jQ_{M-1}$ would be held constant while the M^{th} modulation block would produce a new $I_M + jQ_M$ symbol. In the following T_{sM} time instance, $I_1 + jQ_1, I_2 + jQ_2, \dots I_{M-2} + jQ_{M-2}$ would be held constant while the M^{th} and $M - 1^{th}$ modulation blocks would produce new symbols and so on.

Thus the base I and Q symbols generated at 1^{st} modulator block are intended for all receivers. They have the lowest resolution symbols in terms of the constellation and occupy the longest symbol duration (i.e., smallest bandwidth). The higher-resolution symbols are modulated on top of the base modulation for the closer receivers. Therefore, the end modulator block modulates the symbol based on all the previous modulators and has the shortest symbol duration (i.e., the largest bandwidth). The modulator output signal is based on a bitstream that is the combination of the bits from each modulator block, and at the end of the block-wise modulator, the signal is shaped with the highest resolution I and Q symbols by a pulse shaping block.

On the receiver side, the signal is demodulated according to the SNR and bandwidth observed at the receiver to maintain a low SER. Therefore, in LoS communication, the farthest receiver with the lowest SNR and narrowest bandwidth would demodulate the base or low-resolution bitstream, and the closest receiver with the highest SNR and bandwidth have the highest-resolution demodulating capability. We focus on LoS communications in this paper only because it allows for simple estimation of the receivers' SNRs and bandwidths. As long as receivers are able to evaluate their bandwidth and received SNR, HBM would also work in non-LoS scenarios.

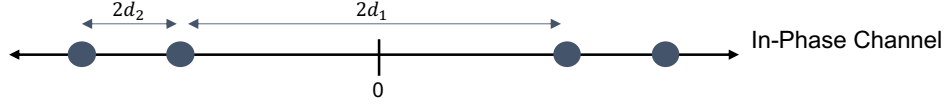


Fig. 34: Constellation diagram for the non-uniform 4PAM modulation.

An example of a two-level HBM encoding scheme was shown in Fig. 31. In that case, both m_1 and m_2 were set to 4, resulting in a non-uniform 16-QAM for the highest level of resolution. Hence the name of the HBM scheme is 4/16-QAM. The first pair of bits, corresponding to the base layer information, were given a certain amplitude and held for the symbol time T_S ($n_1 = 1$). The following two pairs of bits corresponding to high-resolution information were given a smaller amplitude, and each held for time $\frac{1}{2}T_S$ ($n_2 = \frac{1}{2}$) in succession.

Even though we focus on QAM constellations here, multi-resolution constellations can come in a variety shapes and sizes [78], and any hierarchical constellation would work for HBM. Going forward, we focus on the two-level HBM scheme, where the base or first level modulation is the 4-QAM and the second level modulation is M-QAM. The results, such as constellation design, theoretical SER calculation, obtained from the two-level system can be extended for the higher-level schemes and higher-order modulations.

Error Rate Performance Next, we characterize HBM's SER performance by deriving the expected error rates for the 4/16-QAM and 4/M-QAM HBM system.

The constellation diagram for the 4/16-QAM system was shown in Fig. 31 along with the variables d_1 and d_2 . d_1 is the distance from the In-Phase or Quadrature axis to the closest symbol and $2d_2$ is the minimum distance between two high-resolution symbols. λ is defined as the ratio of $\frac{d_2}{d_1}$ which helps characterize the ratio of power used for the high resolution and low resolution information. These will be key parameters in finding the SER as a function of $\frac{E_s}{N_0}$.

For the closer or high-resolution receiver, the constellation looks like a non-uniform 16-QAM and the resulting SER can be derived considering the 16-QAM as the Cartesian product of a one-dimensional 4-PAM constellation with itself [85]. The corresponding non-uniform 4-PAM constellation is shown in Fig. 34. Considering equally probable symbols, the likelihood of symbol error for the non-uniform 4-PAM constellation shown is given by:

$$P_e = Q\left(\frac{d_2}{\sqrt{\frac{N_0}{2}}}\right) + \frac{1}{2}Q\left(\frac{d_1}{\sqrt{\frac{N_0}{2}}}\right), \quad (43)$$

where $Q(\cdot)$ is the standard Q-function. This expression can be used to find the probability of correct detection for the corresponding 16-QAM

$$P_c = (1 - P_e)^2. \quad (44)$$

Hence, the probability of symbol error for 16-QAM is

$$SER_{HighRes} = 1 - P_c \approx 2P_e = 2Q\left(\frac{d_2}{\sqrt{\frac{N_0}{2}}}\right) + Q\left(\frac{d_1}{\sqrt{\frac{N_0}{2}}}\right), \quad (45)$$

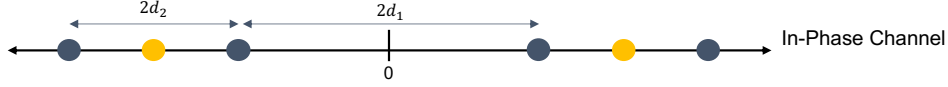


Fig. 35: Constellation diagram for the non-uniform 6-PAM modulation observed by a low-resolution receiver.

where $SE_{HighRes}$ is the probability of symbol error for the closer receiver demodulating the high-resolution constellation.

The low-resolution SER expression calculated for traditional HM in [79] offers a consistent upper bound for the low-resolution HM receiver. It is given by

$$SE_{LowRes} = Q\left(\frac{d_1 + 2d_2}{\sqrt{\frac{N_0}{2}}}\right) + Q\left(\frac{d_1}{\sqrt{\frac{N_0}{2}}}\right). \quad (46)$$

To find a tighter bound for HBM, we must consider what exactly the corresponding low-resolution receiver will observe, which is different from an HM receiver. The same 16-QAM constellation that was used for the high-resolution receiver is transmitted, but the low-resolution receiver operates with a longer symbol duration (i.e. smaller bandwidth). Thus, instead of considering one symbol at a time, the low-resolution receiver demodulates $\frac{n_1}{n_2}$ consecutive high-resolution symbols at once. These symbols exist in the same quadrant of Fig. 31. For efficient utilization of the bandwidth, we take $\frac{n_1}{n_2} = 2$ for low-resolution symbols that are twice as long as the high-resolution symbols. Thus, unlike the high-resolution case, which corresponds to the non-uniform 4-PAM in Fig. 34, the low-resolution receiver observes the transmitted constellation as binary PAM with the three distinct points on each shown in Fig. 35. In the case of consecutive transmission of two identical high-resolution symbols, the receiver detects the constellation points on one of the grey points. Whereas for the case of transmission of two different high-resolution symbols (i.e., symbol at d_1 and $d_1 + 2d_2$ for PAM system on I/Q axis) within one low-resolution symbol duration, the middle point is detected by the low-resolution receiver as an average value, i.e, $d_1 + d_2$ denoted by the yellow points. These points are not truly being transmitted but are the observed constellation point when an outer and inner point are transmitted within one low-resolution symbol time. Therefore, the error probability for the low-resolution receiver is given by

$$P_e = \frac{2}{8} \left(Q\left(\frac{d_1 + 2d_2}{\sqrt{\frac{N_0}{2}}}\right) + Q\left(\frac{d_1}{\sqrt{\frac{N_0}{2}}}\right) + 2Q\left(\frac{d_1 + d_2}{\sqrt{\frac{N_0}{2}}}\right) \right), \quad (47)$$

where $Q\left(\frac{d_1 + 2d_2}{\sqrt{\frac{N_0}{2}}}\right)$ is the probability that an outer edge symbol is received in error, $Q\left(\frac{d_1}{\sqrt{\frac{N_0}{2}}}\right)$ is the probability an inner symbol is received in error, and $2Q\left(\frac{d_1 + d_2}{\sqrt{\frac{N_0}{2}}}\right)$ is the probability a yellow symbol is received in error. We multiply by 2 to account for symbols on either side of the origin, and we divide by 8 because there are 8 possible pairs of high-resolution symbols that are transmitted within a low-resolution symbol.

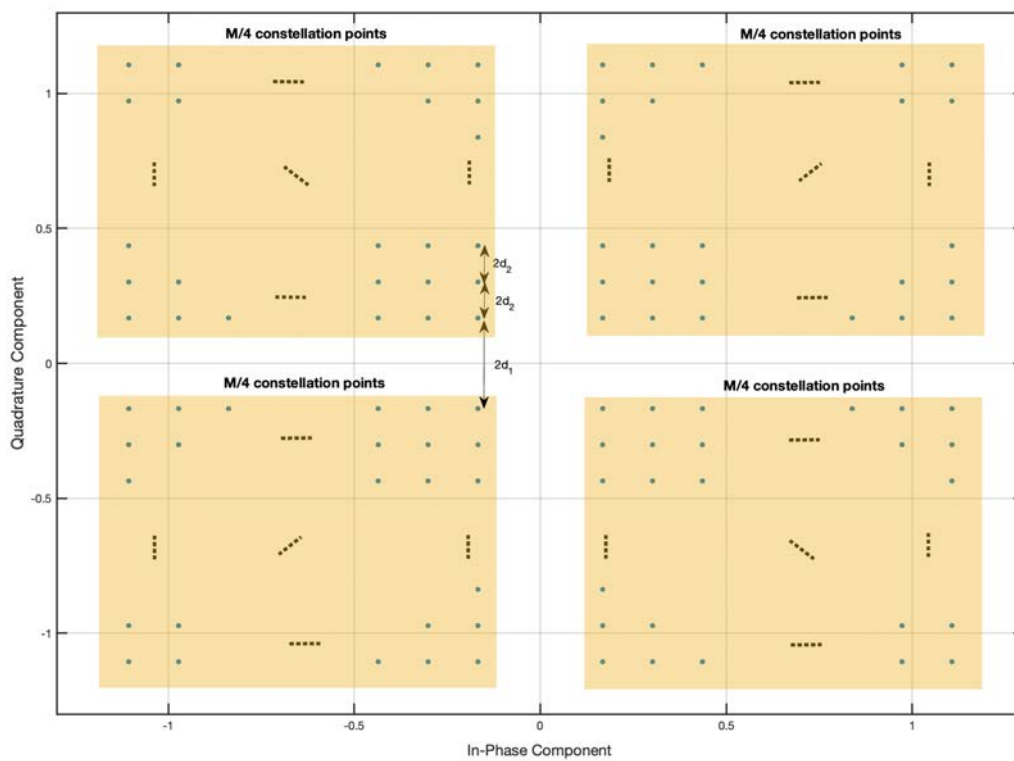


Fig. 36: 4/M-QAM Constellation Diagram with $\lambda = \frac{d_2}{d_1}$.

Using (44) again and recalling that $\lambda = \frac{d_2}{d_1}$ the expression for the low- and high-resolution error rates can be rewritten in terms of d_1 and λ which gives us

$$SER_{LowRes} = \frac{1}{2} \left(Q \left(\frac{d_1(1+2\lambda)}{\sqrt{\frac{N_0}{2}}} \right) + Q \left(\frac{d_1}{\sqrt{\frac{N_0}{2}}} \right) + 2Q \left(\frac{d_1(1+\lambda)}{\sqrt{\frac{N_0}{2}}} \right) \right) \text{ and} \quad (48)$$

$$SER_{HighRes} = 2Q \left(\frac{d_1\lambda}{\sqrt{\frac{N_0}{2}}} \right) + Q \left(\frac{d_1}{\sqrt{\frac{N_0}{2}}} \right) \quad (49)$$

for the low and high resolutions respectively. The average symbol energy can be found from Fig. 31 in terms of d_1 and d_2 to be

$$E_s = \frac{1}{16} \left(4(2d_1^2) + 8(d_1^2 + (d_1 + 2d_2)^2) + 4(2(d_1 + 2d_2)^2) \right). \quad (50)$$

Simplifying, substituting $d_2 = \lambda d_1$, and rearranging, we find

$$d_1 = \sqrt{\frac{E_s}{2(1+2\lambda+2\lambda^2)}}. \quad (51)$$

Using this expression in (49) the symbol error rate as a function of $\frac{E_s}{N_0}$ for the low resolution receiver is determined to be

$$SER_{LowRes} = \frac{1}{2} Q \left((1+2\lambda) \sqrt{\frac{\frac{E'_s}{N_0}}{1+2\lambda+2\lambda^2}} \right) + Q \left(\sqrt{\frac{\frac{E'_s}{N_0}}{1+2\lambda+2\lambda^2}} \right) + 2Q \left((1+\lambda) \sqrt{\frac{\frac{E'_s}{N_0}}{1+\lambda+2\lambda^2}} \right). \quad (52)$$

While for the high-resolution receiver it is given by

$$SER_{HighRes} = 2Q \left(\lambda \sqrt{\frac{\frac{E_s}{N_0}}{1+2\lambda+2\lambda^2}} \right) + Q \left(\sqrt{\frac{\frac{E_s}{N_0}}{1+2\lambda+2\lambda^2}} \right). \quad (53)$$

For the low-resolution receiver, we use E'_s to indicate the symbol energy in order to differentiate it from the high-resolution receiver's symbol energy. $E'_s = E_s \frac{n_1}{n_2}$ because the low resolution symbols are held $\frac{n_1}{n_2}$ times longer than the high resolution symbols. We emphasize again that in our simulations and calculations, we take $\frac{n_1}{n_2} = 2$ for low-resolution symbols that are twice as long as the high-resolution symbols. Decreasing the value of n_2 corresponds to including more high-resolution symbols in each low-resolution symbol. This would result in additional yellow symbols between the inner and outer constellation points in Fig. 35, essentially reducing the probability that the outer- and inner-most constellation points are observed by the low-resolution receiver. Thus we can infer that decreasing n_2 would improve the low-resolution receiver's performance.

For the 4/M-QAM case, the high and low-resolution expected error rates can be derived similarly. The parameters are the same as they were for the 4/16-QAM case, and there are now more

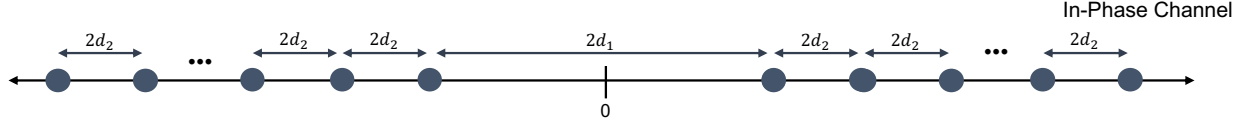


Fig. 37: Constellation diagram for the non-uniform PAM modulation for the 4/M-QAM constellation.

points in the high-resolution portion of the constellation. d_1 and d_2 still characterize the amount of power used for the low and high-resolution information respectively, and λ is still used to represent the ratio $\frac{d_2}{d_1}$. The base non-uniform PAM constellation shown in Fig. 37 can be extrapolated to create the 4/M-QAM constellation shown in Fig. 36.

The resulting error rates are given by

$$\begin{aligned}
 SER_{HighRes} &= \frac{4}{\sqrt{M}} \left((\sqrt{M} - 2) \right. \\
 &Q \left(\lambda \sqrt{\frac{E_s}{N_0} \frac{M}{4 \sum_{i=0}^{\frac{\sqrt{M}}{2}-1} \sum_{j=0}^{\frac{\sqrt{M}}{2}-1} (1 + 2i\lambda + 2j\lambda + 2i^2\lambda^2 + 2j^2\lambda^2)}} \right) \\
 &\left. + Q \left(\sqrt{\frac{E_s}{N_0} \frac{M}{4 \sum_{i=0}^{\frac{\sqrt{M}}{2}-1} \sum_{j=0}^{\frac{\sqrt{M}}{2}-1} (1 + 2i\lambda + 2j\lambda + 2i^2\lambda^2 + 2j^2\lambda^2)}} \right) \right), \tag{54}
 \end{aligned}$$

and

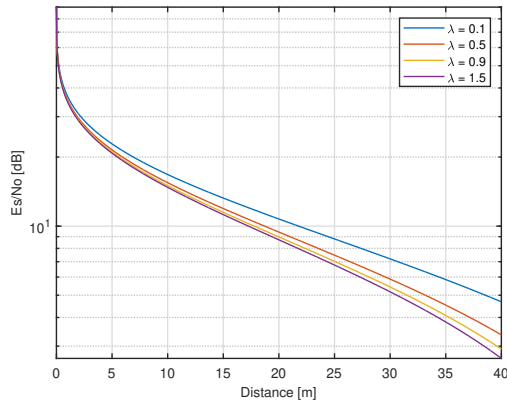
$$\begin{aligned}
SER_{LowRes} = & \frac{8}{M} \left[Q \left(\sqrt{\frac{E'_s}{N_0} \frac{M}{4 \sum_{i=0}^{\frac{\sqrt{M}}{2}-1} \sum_{j=0}^{\frac{\sqrt{M}}{2}-1} (1 + 2i\lambda + 2j\lambda + 2i^2\lambda^2 + 2j^2\lambda^2)}} \right) \right. \\
& + Q \left(2 \left(\frac{\sqrt{M}}{2} - 1 \right) \lambda + 1 \right) \sqrt{\frac{E'_s}{N_0} \frac{M}{4 \sum_{i=0}^{\frac{\sqrt{M}}{2}-1} \sum_{j=0}^{\frac{\sqrt{M}}{2}-1} (1 + 2i\lambda + 2j\lambda + 2i^2\lambda^2 + 2j^2\lambda^2)}} \\
& + \frac{\sqrt{M}}{2} Q \left(\left(\frac{\sqrt{M}}{2} - 1 \right) \lambda + 1 \right) \sqrt{\frac{E'_s}{N_0} \frac{M}{4 \sum_{i=0}^{\frac{\sqrt{M}}{2}-1} \sum_{j=0}^{\frac{\sqrt{M}}{2}-1} (1 + 2i\lambda + 2j\lambda + 2i^2\lambda^2 + 2j^2\lambda^2)}} \\
& + \sum_{n=1}^{\left(\frac{\sqrt{M}}{4}+1\right)^{\frac{1}{2}}} \left\{ ((2n-1)+1) Q \left(((2n-1)\lambda+1) \sqrt{\frac{E'_s}{N_0} \frac{M}{4 \sum_{i=0}^{\frac{\sqrt{M}}{2}-1} \sum_{j=0}^{\frac{\sqrt{M}}{2}-1} (1 + 2i\lambda + 2j\lambda + 2i^2\lambda^2 + 2j^2\lambda^2)}} \right) \right. \\
& + Q \left((((2n-1)+1)\lambda+1) \sqrt{\frac{E'_s}{N_0} \frac{M}{4 \sum_{i=0}^{\frac{\sqrt{M}}{2}-1} \sum_{j=0}^{\frac{\sqrt{M}}{2}-1} (1 + 2i\lambda + 2j\lambda + 2i^2\lambda^2 + 2j^2\lambda^2)}} \right) \\
& + Q \left(\left(\frac{\sqrt{M}}{2} - 1 + (2n-1) \right) \lambda + 1 \right) \sqrt{\frac{E'_s}{N_0} \frac{M}{4 \sum_{i=0}^{\frac{\sqrt{M}}{2}-1} \sum_{j=0}^{\frac{\sqrt{M}}{2}-1} (1 + 2i\lambda + 2j\lambda + 2i^2\lambda^2 + 2j^2\lambda^2)}} \\
& \left. \left. + Q \left(\left(\frac{\sqrt{M}}{2} - 1 + (2n-1) + 1 \right) \lambda + 1 \right) \sqrt{\frac{E'_s}{N_0} \frac{M}{4 \sum_{i=0}^{\frac{\sqrt{M}}{2}-1} \sum_{j=0}^{\frac{\sqrt{M}}{2}-1} (1 + 2i\lambda + 2j\lambda + 2i^2\lambda^2 + 2j^2\lambda^2)}} \right) \right\} \right]. \tag{55}
\end{aligned}$$

Performance Analysis In this section, we explore the error rates' dependence on transmission distance and λ , and we proceed to show that HBM thoroughly outperforms traditional HM in terms of the aggregate data rate.

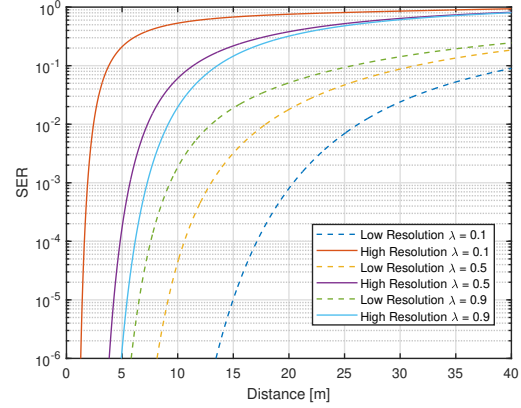
Performance Dependence on Transmission Distance and λ : From the derived error rates, it is interesting to note that the $\frac{E_s}{N_0}$ for each HBM receiver depends on both λ and the transmission distance. All transmitters have finite peak power at which they can transmit, and because increasing λ spreads out the constellation, the average symbol energy will decrease as λ increases. The symbol energy also decreases as the distance increases simply due to the spreading losses. This relation is shown in Fig. 38a. As the distance increases the ratio $\frac{E_s}{N_0}$ seen at the receiver decreases for all values of λ . However, this decrease is slightly less when λ is smaller.

From this result, we can also find the SER as a function of λ and the transmission distance as shown in Fig. 38b. As expected, both error rates increase with distance. The low-resolution error rate also increases with λ while the high-resolution error rate decreases as λ approaches 1.

It is also important to remember that a high-resolution HBM receiver will have a smaller $\frac{E_s}{N_0}$ value than an HM receiver at the same distance due to its shorter symbol duration (i.e. larger bandwidth). Thus, for a given SER threshold, HM can successfully transmit using a higher-order constellation for the high-resolution receivers, but HBM will be transmitting symbols to high-resolution receivers at a faster rate. This phenomenon is shown in Fig. 39. The $\frac{E_s}{N_0}$ of the HM high-resolution receiver is shown on the x-axis while the BER for the various HM and HBM 4/M-QAM schemes are plotted on the y-axis for different values of λ . Looking at the 4/16-QAM

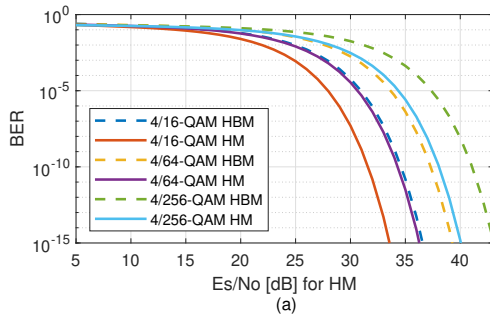


(a) $\left[\frac{E_s}{N_0}\right]$ as a function of λ and distance

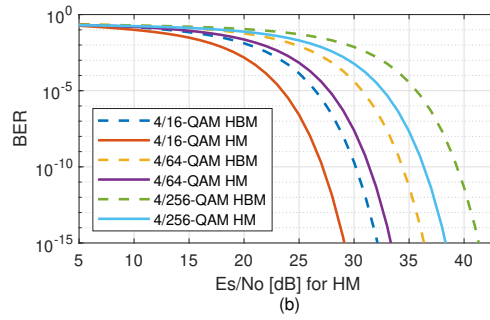


(b) SER as a function of λ and distance

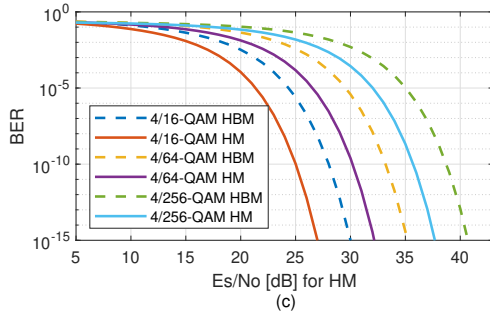
Fig. 38: HBM receiver's $\frac{E_s}{N_0}$ and SER dependence on λ and transmission distance.



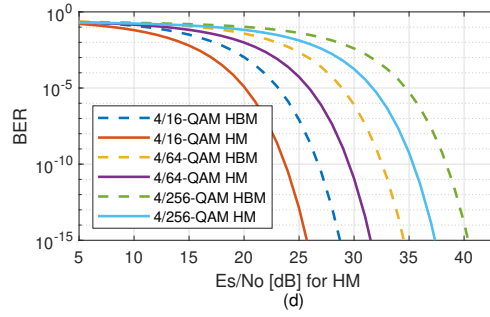
(a)



(b)



(c)



(d)

Fig. 39: A comparison of the high resolution error rates of HM and HBM for (a) $\lambda = 0.2$, (b) $\lambda = 0.4$, (c) $\lambda = 0.6$, (d) $\lambda = 0.8$

HBM constellation as an example, we see that its performance is comparable to that of the 4/64-QAM HM constellation. Similarly, 4/64-QAM HBM constellation has a comparable performance to the 4/256-QAM HM constellation's performance. This is especially the case for smaller values of λ .

From these relations we can see that, for the same transmit power and transmission distance, an HBM system must use a lower-order constellation for its high-resolution information in order to achieve the same performance as a corresponding HM system. Thus we proceed to investigate whether HBM actually improves the aggregate data rate compared to traditional HM.

HBM v. HM Aggregate Data Rates To find the aggregate data rates of the HM and HBM systems we use the following expression

$$R = \frac{\log_2(M_L)}{T_{s1}} + \frac{\log_2(M_H)}{T_{s2}}, \quad (56)$$

where M_H and M_L are the high and low-resolution modulation order respectively, and for traditional HM $T_{s1} = T_{s2} = nT_s$. T_s is the shortest possible symbol time determined by the bandwidth. For HBM, M_H and M_L are still the high and low-resolution modulation order respectively, but $T_{s1} = n_1T_s$ and $T_{s2} = n_2T_s$, where $\frac{n_1}{n_2}$ is an integer, we take to be 2.

To evaluate which system would truly transmit at the fastest data rate, we use a full end-to-end simulation to represent a STMR system operating at 120 GHz with receivers at 3m and 30m. These receivers observe $\frac{E_s}{N_0}$ values of 14dB and 37dB respectively for HM and 14dB and 34dB for HBM. The symbol rate for the low-resolution receivers will be the same for both HM and HBM, but the symbol rates for the high-resolution receivers and the aggregate data rates will differ. With 5 GHz of bandwidth visible for both the high and low-resolution receivers, the HM system can use a 4/256-QAM while keeping its uncoded error rate below 10^{-3} . The bit rate for the high-resolution receiver is given by

$$R_{HighRes} = \log_2(M_{HighRes}) * \frac{B_{HighRes}}{2}, \quad (57)$$

where $M_{HighRes}$ is the highest order of modulation, 256 in this case, and the signal bandwidth B is two times the symbol rate. Thus in the described HM system, the symbol rate is 2.5 Gsps, and the bit rate is 20 Gbps for the high-resolution receiver. The low-resolution receiver's bit rate is very similar, given by

$$R_{LowRes} = \log_2(M_{LowRes}) * \frac{B_{LowRes}}{2}. \quad (58)$$

In the case of traditional HM, $B_{HighRes} = B_{LowRes}$.

An HBM receiver using the same set-up has the same 5 GHz of bandwidth for the low-resolution receiver, but the high-resolution receiver has a bandwidth of 10 GHz, corresponding to a symbol rate of 5 Gsps. Due to its larger bandwidth, this receiver experiences more noise and is therefore only able to operate with a 4/64-QAM while maintaining the same uncoded error rate threshold of 10^{-3} . Using equation (57), the bit rate is found to be 30 Gbps. The low-resolution receiver's bit rate is calculated from (58) and is the same as that of the traditional HM low-resolution receiver: 5 Gbps.

The aggregate data rates for each STMR system are found by adding the high and low-resolution data rates

$$R_{Aggregate} = R_{LowRes} + R_{HighRes}, \quad (59)$$

and Fig. 40a displays the results obtained in this example. The HM system has an aggregate rate of 25 Gbps while HBM has an aggregate data rate of 35 Gbps. Thus, even though the constellation order is smaller, HBM outpaces traditional HM in terms of data rate while achieving the same error rate performance. In other words, even though HM uses a constellation order of 256 for

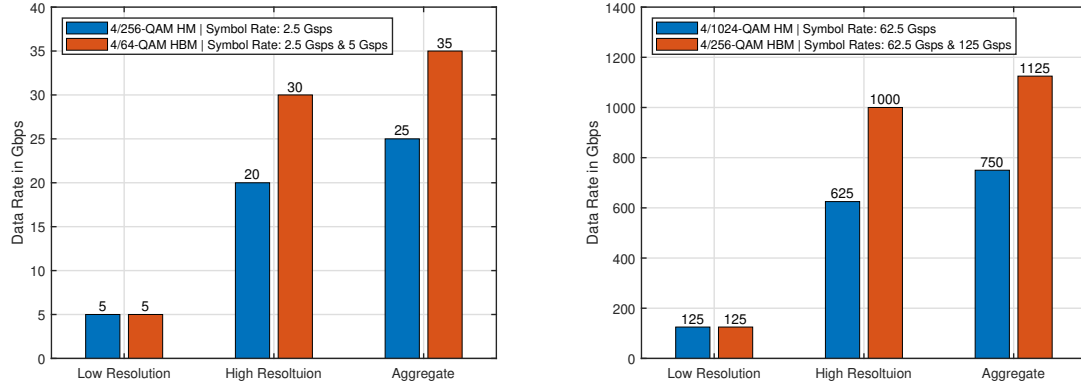


Fig. 40: Achievable data rates for HM and HBM systems.

the high-resolution receiver, HBM achieves a data rate 10 Gbps higher than that of HM while experiencing a comparable error rate. This is possible because HBM uses a faster symbol rate for the high-resolution information. A similar result is shown for higher orders of M and larger bandwidths that would truly take advantage of the THz band's availability in Fig. 40b. In this example, the low-resolution receivers would require an $\frac{E_s}{N_0}$ of at least 14 dB while the HM and HBM high-resolution receivers would require $\frac{E_s}{N_0}$ values of 31 dB and 25 dB respectively. Despite using a lower modulation order and requiring a smaller $\frac{E_s}{N_0}$, HBM reaches an aggregate rate above 1 Tbps before HM by exploiting the distance-dependent bandwidth available in the THz band.

The HBM Functional Region Next, we define the HBM functional region and consider its significance when designing static systems as well as systems with varying channel characteristics.

Now that the SER's dependence on $\frac{E_s}{N_0}$ and λ has been fully characterized, the SER can be represented as a function of λ and transmission distance. Thus, for a pair of distances, the expected $\frac{E_s}{N_0}$ can be found as a function of λ . Then, given a maximum SER threshold, the necessary range of λ s to maintain an SER below the threshold can be found. In other words, to design a functional HBM system, we can determine the range of λ such that

$$SER_{HighRes} \left(\frac{E_s}{N_0}, \lambda \right) < SER_{Th} \cap SER_{LowRes} \left(\frac{E'_s}{N_0}, \lambda \right) < SER_{Th} \quad (60)$$

where SER_{Th} is the threshold SER and $\frac{E_s}{N_0}$ values are determined using a link budget analysis for the transmission distances. We call the range of λ values which satisfy (60) the HBM *functional region*. As an example, the scenario of two stationary receivers at 3m and 12m in a 4/16-QAM HBM system is considered. To find the functional region, the expected $\frac{E_s}{N_0}$ values for the experimental system is calculated for 3m and 12m, and the SER is calculated as a function of lambda for each receiver. The results are shown in Fig. 41. The SER curves for each resolution follow what intuition would expect. For the low-resolution receiver, the SER is minimized when λ is equal to 0. This corresponds to the uniform 4-QAM that the low-resolution receiver is trying to demodulate. From the low-resolution receiver's perspective, increasing λ increases unwanted deviations from the base layer symbols, thus when these deviations do not occur, the low-resolution receiver demodulates at

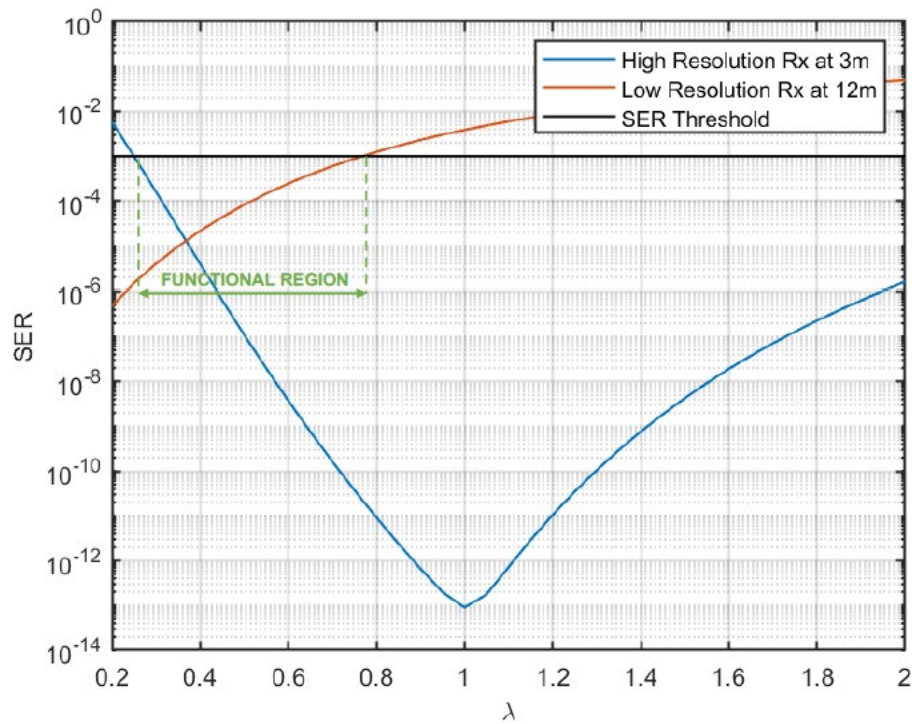


Fig. 41: The Functional Region for the 4/16 QAM HBM System

its best. Conversely, for the high-resolution receiver, when λ is equal to zero, it is impossible to differentiate the high-resolution symbols, which results in the highest error rate for this receiver. The high-resolution receiver instead does best when λ is equal to 1, which corresponds to a uniform 16-QAM constellation. We would expect the high-resolution receiver to prefer this scenario because all the symbols are equidistant from each other, which minimizes the possibility that symbols are mistaken for each other. In order to find a value of λ for which both receivers are functional we define a threshold SER that both receivers must operate below. From this threshold, it is easy to identify the range of λ values that allow both receivers to operate below the threshold. As is labeled in Fig. 41, for an SER threshold of 10^{-3} the functional region for the system described above with receivers at 3m and 12m exists between about 0.25 and 0.75. Although most wireless communication systems require an $\text{SER} < 10^{-6}$, this threshold was picked so that the error rate could be experimentally represented within one frame.

To consider the generalized expressions for SER, the functional region was also explored for a 4/64-QAM HBM system with receivers at 3m and 6m. For $M = 64$, the expected error rates can be found using (55) and (54).

The functional region can be found for a 4/M-QAM system as described above for the 4/16-QAM. For a given SER threshold of 10^{-3} , the functional region of a 4/64-QAM is shown in Fig. 42 to exist between about 0.3 and 0.8.

In LoS scenarios, the HBM functional region depends heavily on the receiver distances, and Fig. 43 shows the SER as a function of λ for various distances. The performance of receivers demodulating the low resolution information and the high resolution information is shown by the

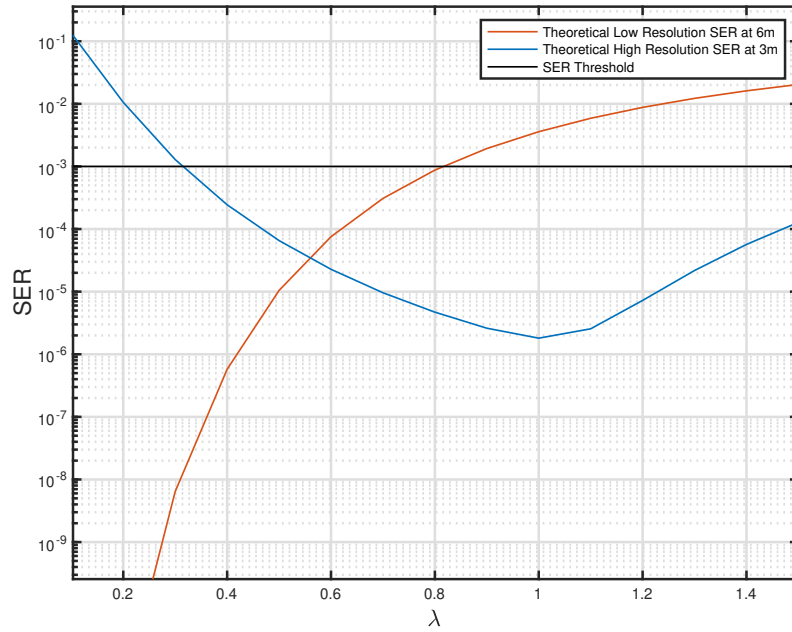


Fig. 42: Functional Region for the 4/64-QAM HBM System

dashed and solid lines respectively. In both instances, an increased transmission distance results in a worse performance at all values of λ . Thus λ s providing an SER below the desired threshold decreases. In other words, as receivers move farther away from the transmitter, the functional region shrinks to a smaller range of λ values.

HBM Systems with Mobile Receivers The functional region offers an important design consideration for implementing an HBM system when the transmission distances or range of transmission distances are known. After determining the proper value of λ for the transmitter, receivers still must determine which resolution they should demodulate. One advantage HBM systems have over non-hierarchical systems is their flexibility in this case. When a receiver's bandwidth or SNR changes, the receiver can adjust accordingly to receive the maximum amount of information with an acceptable error rate. Any number of changes in the surrounding environment can cause such fluctuations, and the receiver will need to determine which resolution is best to achieve acceptable SER values. Assuming a system that performs equalization, the equalization process must be done on a per-frame basis and for the receiver's chosen resolution. For this reason, receivers should avoid switching resolutions within a frame.

Although more research needs to be done for receivers to be able to estimate their available bandwidth, a simple SNR or $\frac{E_s}{N_0}$ estimate would provide a strong indication of which resolution can be successfully demodulated in a given channel. Since these estimates must be performed for each frame, the strength of the header can be used to determine the $\frac{E_s}{N_0}$. From there, the $\frac{E_s}{N_0}$ estimate should be compared to thresholds determined by the target SER. The $\frac{E_s}{N_0}$ threshold can be found

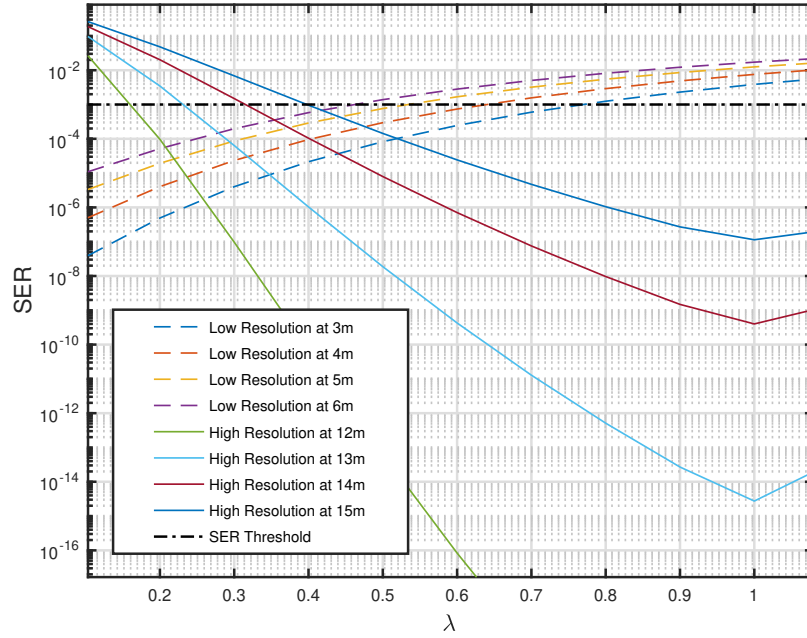


Fig. 43: Functional Region for the 4/16-QAM HBM System for various distances

using equation 54. The goal is to find the range of $\frac{E_s}{N_0}$ such that

$$SER_{HighResTh} > SER_{HighRes}\left(\frac{E_s}{N_0}, \lambda\right). \quad (61)$$

The lower bound of this region can be called $\frac{E_s}{N_0ThHigh}$ and will correspond to the maximum distance the high-resolution information can be successfully transmitted. λ would be kept constant in these expressions since it is a physical layer parameter that would have already be established.

Theoretically, if the estimated $\frac{E_s}{N_0}$ is above $\frac{E_s}{N_0ThHigh}$, the receiver should use the high resolution, while if it is below the threshold, the receiver should use the low resolution. However, there could be a few areas of concern when a receiver nears this threshold. First, given that the $\frac{E_s}{N_0}$ estimate is taken at the beginning of the frame, there is a possibility that the $\frac{E_s}{N_0}$ is just above the threshold when the estimate is taken, but falls below the threshold as the rest of the frame is received. This would likely result in an SER above the determined maximum. To combat this issue, the $\frac{E_s}{N_0}$ threshold should be adjusted by a parameter α that depends on how quickly the channel is expected to change and on the frame duration.

Another possibly problematic situation could occur if a receiver's $\frac{E_s}{N_0}$ is fluctuating rapidly around $\frac{E_s}{N_0ThHigh}$. In this case, the receiver might switch resolutions very frequently, or after switching to the high resolution the $\frac{E_s}{N_0}$ might fall back below $\frac{E_s}{N_0ThHigh}$ in the middle of a frame resulting in a high SER. These problems can be addressed by establishing a *transition region* in which a receiver could use either resolution. Thus while a receiver will switch from the high

Table 7: BER comparison for different modulation schemes with 5 GHz baseband bandwidth.

Distance (m)		3		9		15	
Modulation	Data Rate (Gbps)	SNR (dB)	BER	SNR (dB)	BER	SNR (dB)	BER
16-PSK	20	29	0*	20.9	0*	16.5	8×10^{-5}
16-QAM	20	27.5	0*	18	0*	13.1	0*
16-APSK 2R	20	28.7	0*	19.1	0*	13.5	0*
32-PSK	25	29.6	1.6×10^{-3}	21	5×10^{-3}	15.8	8×10^{-3}
32-QAM	25	28.1	0*	18.2	2.5×10^{-4}	13.2	5×10^{-4}
32-APSK 2R	25	27.7	0*	19.5	1.6×10^{-4}	14	5.8×10^{-4}
32-APSK 3R	25	27.7	0*	19.5	0*	14	8×10^{-5}

*0 BER: no bit error within 10 frames with 1,200 bits each.

resolution to the low resolution at

$$\frac{E_s}{N_0}_{HighLowTh} [dB] = \frac{E_s}{N_0}_{ThHigh} [dB] + \alpha [dB], \quad (62)$$

receivers should switch from the low to the high resolution at

$$\frac{E_s}{N_0}_{LowHighTh} [dB] = \frac{E_s}{N_0}_{ThHigh} [dB] + \alpha [dB] + \beta [dB], \quad (63)$$

where β is a function of how quickly the channel fluctuates, the frame duration, and the number of subsequent frames that should be demodulated with the same resolution.

5.3 Results and Discussion

5.3.1 Experimental Validation of Optimized APSK Modulations

In this section, experimental results are provided by comparing the proposed optimized M-APSK modulation with M-PSK and M-QAM modulations. The offline 32-GHz signal processing back-end with 120-140 GHz front-ends presented in [46] are utilized for the purpose.

To compare the experimental bit error rate (BER) performance of the M-APSK, M-PSK, and M-QAM schemes at THz frequencies, the modulated IF signals are up-converted to 130 GHz. It is noted here that, for the experimental results, BER is considered instead of SER to show the performance at the bit level. Furthermore, in the case of M-PSK and M-QAM, grey coding is used, whereas for M-APSK binary coding is implemented since the grey code conversion is challenging [86]. In this case, IF modulated signals of up to 20 GHz bandwidth with data-rate up to 50 Gbps are considered varying distances until 15 m. In Table 7 and 8, we summarize the results of different schemes. To calculate the BER, ten frames are considered (each 1,200 bits long).

It is observable from the experimental result that the BER of 16-PSK and 32-PSK are comparatively higher than the corresponding M-QAM and M-APSK counterparts. Although the PAPR is low, the M-PSK is not a suitable candidate for high-order modulation schemes when utilized for ultrabroadband communication for THz frequencies. It is noticeable that the BER of two-ring 16-APSK and three-ring 32-APSK are better than 16-QAM and 32-QAM, respectively.

Table 8: BER comparison for different modulation schemes with 10 GHz baseband bandwidth.

Distance (m)		3		9		15	
Modulation	Data Rate (Gbps)	SNR (dB)	BER	SNR (dB)	BER	SNR (dB)	BER
16-PSK	40	26.2	1.9×10^{-2}	17.3	3.7×10^{-2}	13.9	4.4×10^{-2}
16-QAM	40	24	4.1×10^{-3}	15.1	2×10^{-2}	11.2	3.2×10^{-2}
16-APSK 2R	40	25	0*	16.5	4.1×10^{-3}	12	1.5×10^{-2}
32-PSK	50	25.9	7×10^{-2}	17.5	1.5×10^{-1}	13.2	1.8×10^{-1}
32-QAM	50	23.1	1.8×10^{-2}	15.5	7×10^{-2}	11.1	8×10^{-2}
32-APSK 2R	50	25.1	1.5×10^{-2}	17.5	1.1×10^{-1}	12	1.3×10^{-1}
32-APSK 3R	50	24.9	2.5×10^{-3}	17.1	9×10^{-2}	11.5	1×10^{-2}

*0 BER: no bit error within 10 frames with 1,200 bits each.

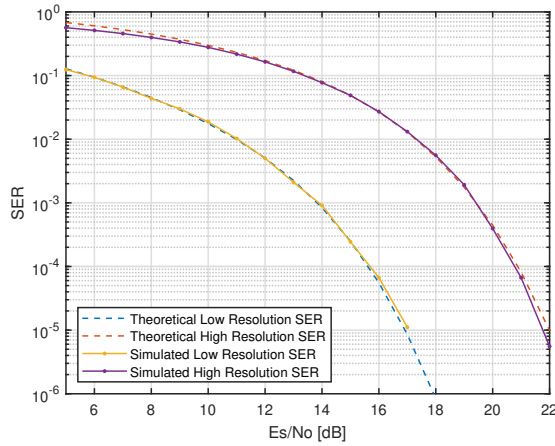
Furthermore, the two-ring 32-APSK performance is quite close to the 32-QAM system. The reason behind the reduced performance of the M-QAM is the higher PAPR compared to the M-APSK scheme. It is depicted in the corresponding captured SNR for the schemes mentioned above, where PSK has a maximum captured SNR followed by M-APSK and QAM. Due to the damage limit constraint of the front-end's input power, we have maintained the peak power for all the schemes same. Moreover, the transceiver chains have a PA to boost the received signal's power, which imposes restrictions on the utilization of signals with high PAPR. Added, the non-linearity and frequency selective response due to the ultrabroadband devices, the performance of all the schemes are relatively impoverished compared to the normal AWGN channel and relatively prominent in the case of 10 GHz baseband signals than the 5 GHz one. Thus, the experimental results motivate us to utilize the M-APSK for ultrabroadband THz band communication when utilizing the constellation scheme with high modulation index, M .

Next Steps The optimized M-APSK constellations have been contributed to the revised IEEE 802.15.3d standard. As part of the next steps, we will implement this single-carrier scheme in the real-time platform presented in Sec. 3. In addition, as we discuss next, we might consider an APSK-HBM system.

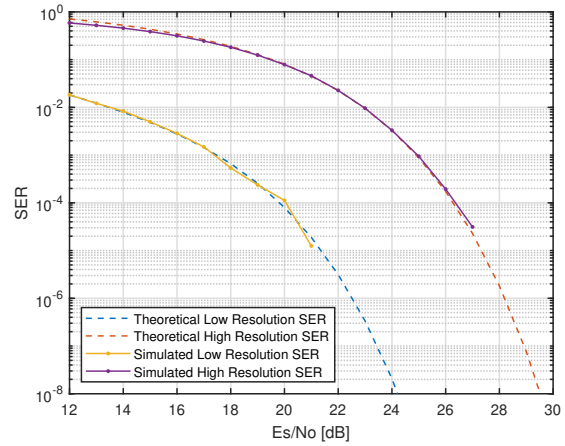
5.3.2 Experimental Validation of HBM

Simulated Results First, we verify the expressions obtained in Sec. 5.2.2 through a full end-to-end MATLAB simulation of an HBM system. The results are compared with our expressions to obtain the plot shown in Fig. 44a and Fig. 44b. The simulated system involves encoding of a bitstream into a grey-coded HBM constellation, modulating the constellation symbols onto a 130 GHz digital signal, adding Additive White Gaussian Noise (AWGN), and implementing separate correlating demodulators for the high-resolution and low-resolution receivers. Because 130 GHz exists inside an absorption-free window, the absorption loss is negligible. For this simulation, a constant value of $\lambda = 0.6$ is used, and the SER is calculated by sweeping the $\frac{E_s}{N_0}$ values. As can be seen in both Fig. 44a and Fig. 44b, the high-resolution simulated and theoretical values align nearly perfectly.

Now that we have verified our error rate expressions through simulations, we proceed to use experiments to verify the functional region's existence.



(a) 4/16-QAM with $\lambda = 0.6$



(b) 4/64-QAM with $\lambda = 0.6$

Fig. 44: Comparison of the SER for simulated HBM systems and the theoretical equations.

Experimental Results Once again, we leverage the TeraNova testbed with the non-real-time DSP engine at 130 GHz. Our setup is shown in Fig. 45, operates at 130 GHz where absorption losses are negligible. Thus, we do not see the distance-dependent bandwidth in our experiments, but we use the test bed to validate the concepts presented.

The frame structure of the data starts with an 18-bit maximal merit factor (MF) header for time synchronization at the receiver. 240 pilot bits are modulated as base-layer signals used for equalization by both receivers, with an additional 480 high-resolution pilot bits modulated on top for the high-resolution receiver's equalization. The data portion of the frame consists of 12000 base-layer bits and 24000 high-resolution bits modulated onto a non-uniform QAM as was shown in Fig. 31. For this experiment, a low-resolution bandwidth of 5 GHz and a high-resolution bandwidth of 10 GHz are used for symbol rates of 2.5 Gps and 5 Gps at the low- and high-resolution receivers respectively.

The goal of these experiments is to verify the calculated functional region in Fig. 41 and to prove its existence. Reviewing the experimental results for the 4/16-QAM, summarized in Table 9, we observe what we expected; for values of λ outside the functional region (i.e. 0.2 and 0.8), one receiver is unable to demodulate with an SER below the established threshold, and for values of λ within the functional region (i.e. 0.4 and 0.6), SER values are all below the desired threshold. The results also show that outside the functional region, the receivers we expected to be functional and non-functional are behaving as expected. For instance, when λ is equal to 0.2, the low-resolution receiver is expected to do well while the high-resolution receiver struggles. Table 9 verifies that this is the case, since for the signal with $\lambda = 0.2$, the low-resolution SER is very low, while the high-resolution SER is above the desired threshold at 0.0028 in Fig. 46.

In the same way, on the other side of the functional region, when λ is equal to 0.8, Fig. 47 shows that the high-resolution receiver is still successfully demodulating its information, but the symbols have spread out enough that the low-resolution receiver is struggling. Meanwhile for λ

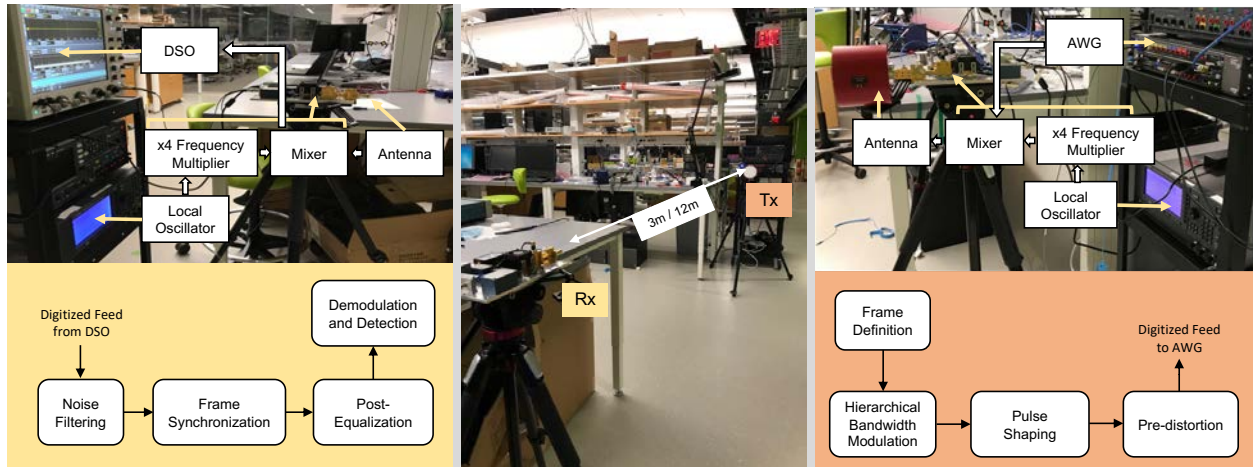


Fig. 45: TeraNova Testbed

Table 9: Experimental Results for Demonstrating the HBM Functional Region for 4/16QAM System

λ	Low Resolution SER	High Resolution SER
0.2	$8.33 * 10^{-4}$	0.0028
0.4	$< 4.17 * 10^{-4}$	$1 * 10^{-4}$
0.6	$< 8 * 10^{-5}$	$< 4 * 10^{-5}$
0.8	0.005	$2 * 10^{-4}$

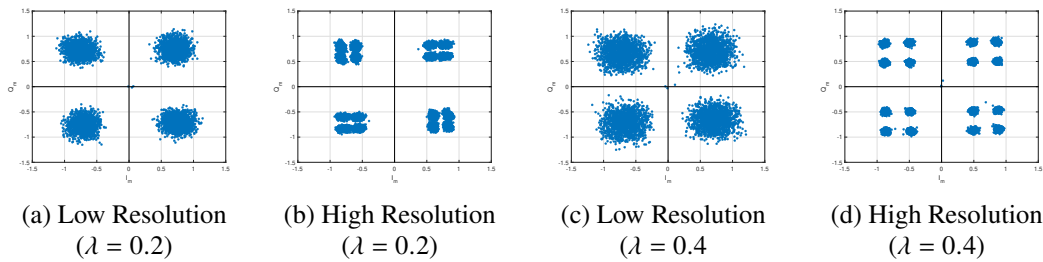


Fig. 46: Received constellations for high- and low-resolution receivers at 3m and 12m respectively for a 4/16-QAM HBM system.

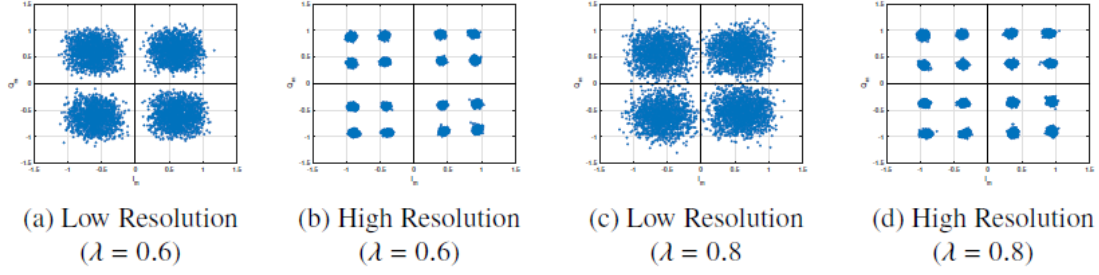


Fig. 47: Received constellations for high and low-resolution receivers at 3m and 12m respectively for a 4/16-QAM HBM system.

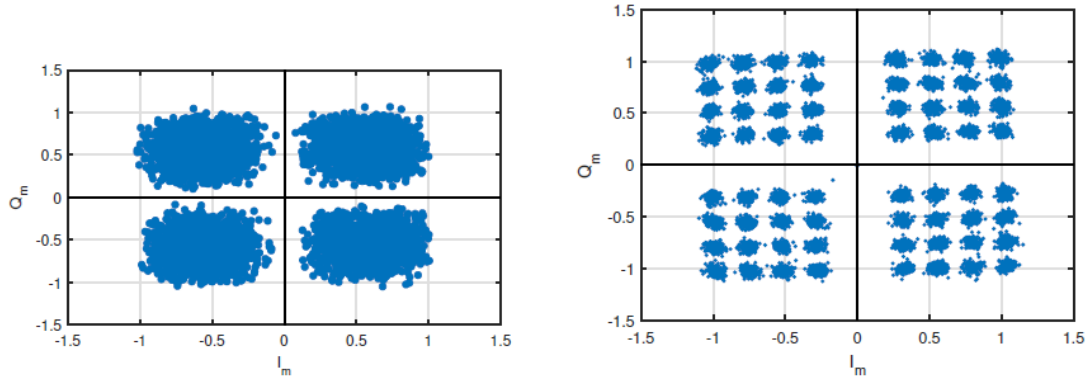


Fig. 48: Received Constellation by the far and near receivers in the 4/64-QAM system with $\lambda = 0.4$.

values within our functional region, both receivers are able to successfully demodulate their symbols.

In Fig. 46 and Fig. 47 the constellations are all successfully demodulated with symbol error rates below the 10^{-3} threshold. These results correspond with our analysis of Fig. 41.

The experimental results using a 4/64-QAM HBM system are summarized in Table 10, and they also verify the expected functional region shown in Fig. 42. As expected, the near receiver cannot demodulate below the desired threshold of 10^{-3} when $\lambda = 0.2$, while the far receiver can demodulate successfully. For λ values of 0.4 and 0.6, both receivers are able to demodulate successfully, but for $\lambda = 0.9$, which is outside the functional region, the low-resolution receiver cannot demodulate at an SER below the desired threshold. An example of the received constellation for this experiment is shown in Fig. 48.

Next Steps Continuing this work would involve generalizing the design expressions for more constellations and more than two resolutions. Additionally, the effect λ has on the $\frac{E_s}{N_0}$ ratio could be characterized further. Peak-to-average-power ratio is a drawback to high-order QAM systems. A properly designed HBM signal would never perform worse than a uniform QAM in terms of PAPR, because recalling our analysis, receivers will always prefer λ values less than 1. Still, a deeper

Table 10: Experimental Results for Demonstrating the HBM Functional Region for 4/64-QAM System

λ	Low Resolution SER	High Resolution SER
0.2	$1.25 * 10^{-4}$	0.0112
0.4	$< 8.3 * 10^{-5}$	$2.167 * 10^{-4}$
0.6	$6.667 * 10^{-4}$	$8.33 * 10^{-5}$
0.9	0.0038	$1 * 10^{-4}$

study of how the PAPR can be considered in designing and analyzing the performance of HBM systems would be important. Besides QAM modulations, other modulations can be used to build the hierarchy. For example, in [87], we have designed hierarchical PSK modulations with minimize the PAPR.

6 MEDIUM ACCESS CONTROL FOR TERAHERTZ NETWORKS

6.1 Motivation

The very large available bandwidth at terahertz frequencies, which leads to very high data-rates and very short packet transmission times, combined with the requirement for high-gain directional antennas (DAs) simultaneously at the transmitter and the receiver at all times, motivate the design of new MAC protocols. MAC protocols designed for lower frequencies cannot be reused, as the use of omnidirectional antennas in at least one of the two ends, even if just during an initial access phase [88], renders them useless for terahertz networks over meaningful communication distances.

As part of this effort,

- We have developed ADAPT, an Adaptive Directional Antenna Protocol for Terahertz Networks. This is a receiver-initiated MAC protocol for ultra-high-speed centralized networks at macro-scale distances, which builds upon our earlier work [89], where both the Access Point (AP) and the client nodes use narrow-beam DAs. The protocol relies on a turning AP that divides the space into sectors, periodically providing coverage to the whole area, and on a 3-way receiver-initiated handshake for link-layer synchronization and collision avoidance. The electronic beam steering capability of the AP can be realized with dense antenna array technology. Adaptive sector time and adaptive modulation mechanisms reduce cycle time and maximize the performance of the protocol. A 1-way handshake variant with reduced control overhead is also presented. We evaluate the performance of ADAPT both analytically and with simulations in *ns-3*. We present an analytical model that studies the stability condition and achievable throughput. In *ns-3*, the protocol is implemented using *TeraSim* [39], an extension that models the frequency-selective channel and physical layers of the THz band. The performance is investigated in terms of throughput and packet discard rate with extensive simulations, using the physical layer parameters of the IEEE 802.15.3d standard. It is compared to a non-adaptive version of the 1-way handshake protocol, and to Carrier Sense Multiple Access with and without collision avoidance (CSMA, CSMA/CA).
- We have designed and experimentally demonstrated the first dual-band MAC protocol able to dynamically switch between two transmission windows within the terahertz band while minimally impacting on the user experience. Our work has been presented in the context of dynamic spectrum sharing between on-the-ground communication infrastructure (particularly, terahertz wireless backhaul applications for rural connectivity) and Earth Exploration Satellite Services (EESS) that conduct atmospheric studies from a low-Earth orbit. Passive users can be negatively impacted by interference from active transmissions [90]. Therefore, they retain exclusive access to relatively narrow portions of the above-100-GHz spectrum, preventing the allocation of contiguous chunks with tens of GHz of bandwidth for communications [91, 92]. For example, in the U.S., the largest allocations for active transmissions between 100 and 275 GHz are 32.5 GHz (116 – 148.5 GHz) and 18.5 GHz (231.5 – 250 GHz), but with only 12.25 GHz (non-contiguous) earmarked for unrestricted fixed or mobile terrestrial usage [91]. Transmissions are strictly forbidden in 33.5 GHz of spectrum, and conditioned to the protection of coexisting passive users in the remaining spectrum. These conservative regulations apply even without passive users exploiting the spectrum for sensing. This prevents multiplexing of unused resources, making this spectrum less appealing for wireless backhaul [93]. With our

solution, we demonstrate that effectively and dynamically sharing the spectrum is possible, without ever leaving the terahertz band.

6.2 Methods, Assumptions and Procedures

6.2.1 Design and Performance Analysis of ADAPT

System Model

Network Architecture: We design a protocol for a centralized architecture in macro-scale networks. It is governed by a central AP that provides connectivity to client nodes, which are randomly distributed in a two-dimensional plane. The AP coordinates all transmissions, maximizing efficiency and guaranteeing fairness. In this paper, only up-link communication is defined and analyzed, but the protocol can be extended to include down-link. This type of architecture is relevant for WPAN, WLAN or small cells in beyond-5G/6G cellular networks.

Turning Directional Access Point: We assume that nodes and AP are equipped with dense antenna arrays that support dynamic beamforming and beam steering. With these capabilities, we define a turning AP that divides space in sectors by discretizing the antenna azimuth, as in Fig 49. We assume that nodes have already discovered the AP and are always directionally listening toward it. When the AP turns into a sector, it will initiate the handshake by sending a Call-to-Action (CTA) packet, which can trigger a data transmission. We consider that communication occurs only in LoS D-to-D links. We do not take into account multi-path propagation, as THz communications rely mainly on LOS due to the very high attenuation that non-LOS signals suffer. Nonetheless, if the LOS path was blocked and a node discovered the AP through a reflective surface, using another sector, it would make no difference from a MAC design perspective.

Terahertz Channel Model: The high path loss is the most important constraint in THz macro-scale communications. The channel model used in this work is a simplified version of the work introduced in Sec. 5. The channel frequency response is given by

$$H_c(f, d) = \left(\frac{c}{4\pi f d} \right) \exp \left(-\frac{k_{\text{abs}}(f) d}{2} \right), \quad (64)$$

where f stands for frequency, d for distance, c refers to the speed of light and k_{abs} to the molecular absorption coefficient of the medium. With this frequency response, power at reception P_r is given by

$$P_r(d) = \int_B S_t(f) |H_c(f, d)|^2 G_t G_r df, \quad (65)$$

where S_t is the single-sided power spectral density of the transmitted signal, B is the bandwidth and G_t, G_r are the gain of the transmitter and receiver antennas, respectively.

Noise also has two components, thermal and molecular absorption noise. However, if signals are transmitted in absorption-defined windows, we can neglect the effect of molecular absorption on noise. We will consider only thermal noise, calculated as

$$N = k_B T B, \quad (66)$$

where k_B is the Boltzmann constant, T is temperature and B is bandwidth. It is important to note that, since thermal noise depends on bandwidth, ultra-broadband communications in the THz band

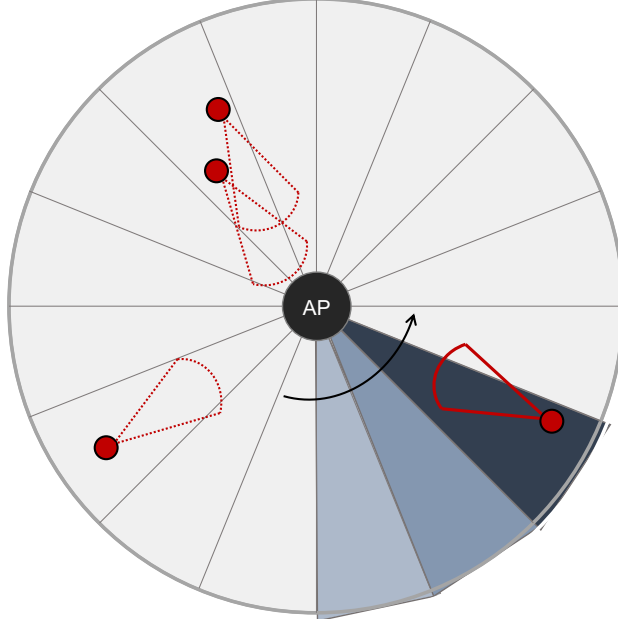


Fig. 49: Network architecture: A central turning AP divides space in sectors, and nodes are randomly distributed in a 2-D plane.

will have very high noise values. For reference, at 75 GHz of bandwidth and 300 K of temperature, thermal noise is of -65 dBm.

Range/Performance trade-off: The AP beamwidth chosen plays a key role in the performance and range of the network. A narrower beamwidth leads to more gain and longer range, but also results in a larger number of sectors, slowing the turning speed and effectively worsening the overall protocol performance.

Beamwidth ($\Delta\theta$) is defined as the beam angle overture for which the gain is at most 3 dB weaker than the maximum gain. Each AP full cycle is divided into a number of even sectors (N_{sec}), given by

$$N_{\text{sec}} = \frac{2\pi}{\Delta\theta}. \quad (67)$$

The antenna gain is inversely proportional to the beam effective area. For highly directional antennas, the gain can be approximated as [94]

$$G \approx \frac{4\pi}{\Omega_A} = \frac{4\pi}{\theta_h \phi_h}, \quad (68)$$

where Ω_A is the array solid beam angle, and θ_h and ϕ_h refer to the half power beam width (HPBW) in the elevation and azimuthal plane respectively. We will assume an identical angle in both planes, so that $\Delta\theta = \theta_h = \phi_h$. We also consider that nodes and AP use the same beamwidth, so $G = G_t = G_r$. The relation between number of sectors, beamwidth and gain is captured in Tab. 11.

In order to properly decode a signal, P_r must be high enough to satisfy the Signal-to-Noise-Ratio threshold (SNR_{th}). We can use Eq. (65) to calculate the necessary gain to achieve a certain distance, and adjust beamwidth accordingly. Nonetheless, a high number of sectors is detrimental to

the protocol performance, as it will increase cycle time. Moreover, with a longer range comes an increase in propagation time, which is a limiting factor in THz networks. Therefore, a compromise has to be made to find an optimal balance between range and network performance.

Protocol Design

In this section, we present and discuss the design of ADAPT. A variant of the protocol using a 1-way handshake is also proposed. We conclude with a note on neighbor discovery.

3-way Handshake and Adaptive Sector Time: The most challenging question in directional networks is how to overcome the deafness problem. In ADAPT, we rely on a turning AP and a receiver-initiated handshake to establish link-layer synchronization, which prevents unnecessary transmission attempts when the AP and transmitter node are not facing each other.

When the AP turns into a new sector, it sends a Call-to-Action (CTA) packet to let client nodes know it is available to start a communication. Nodes that receive the CTA will answer with a Request-to-Send (RTS) packet if they have data to transmit. The AP will wait a certain wait time (T_{wait}) which is long enough to let the furthest node reply, and then decide on the next action. There are 3 possible scenarios:

- **0 RTS received:** If there are no client nodes in the sector, or they have an empty queue, the CTA will not trigger any response. After T_{wait} , the AP will skip this sector and turn to the next one, as in Fig. 50.
- **1 RTS received:** When a node has data to transmit, it will reply to the CTA with an RTS. If only one RTS has been received when T_{wait} expires, the AP will grant access to the channel with a Clear-to-Send (CTS) packet, completing the 3-way handshake. Upon the CTS reception, the node will send data (DATA), which will be acknowledged (ACK) if successfully received, as illustrated in Fig. 51. Only one DATA packet per node can be sent in a sector. If the node has more data in queue, it should be sent in subsequent cycles. This design choice limits the maximum sector time and reduces the delay for nodes waiting to transmit in other sectors, increasing the fairness of the proposed solution.
- **>1 RTS received.** When more than one RTS is received, all of them will be answered with a CTS each. The AP defines a transmission (Tx) slot for every node so that DATA packets do not collide, while making an efficient use of the channel. The timing information is included

Table 11: Sectors, Beamwidth and Gain Relation

N_{sec}	Beamwidth ($\Delta\theta$)	Max. Gain [dB]
15	24°	18.55
20	18°	21.04
30	12°	24.57
45	8°	28.09
60	6°	30.59
90	4°	34.11
120	3°	36.61

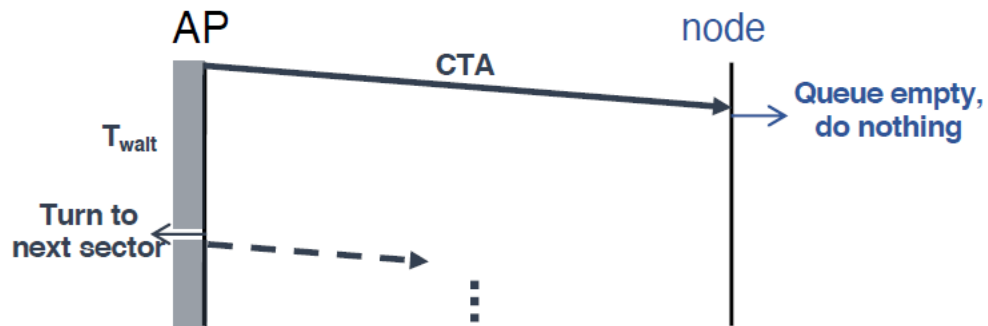


Fig. 50: Packet flow in the case of unanswered CTA.

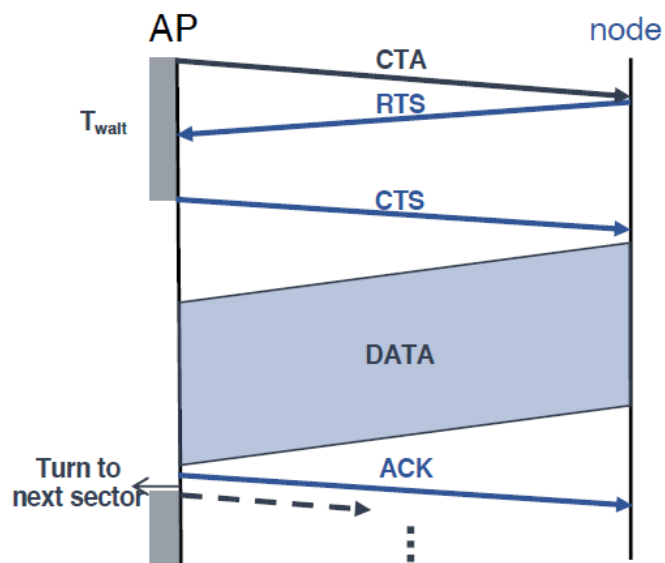


Fig. 51: Packet flow of a 3-way handshake, data transmission and acknowledgement between AP and one client node.

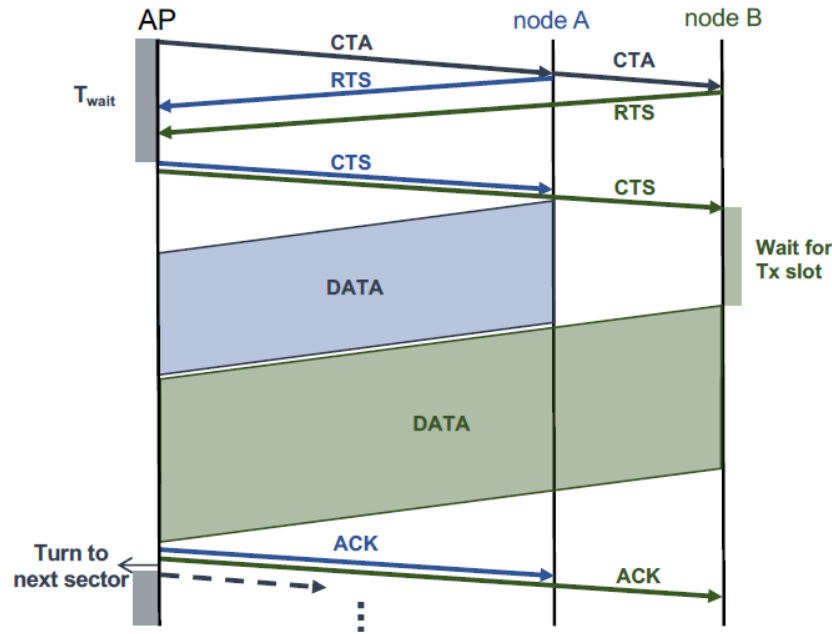


Fig. 52: Packet flow of a 3-way handshake, data transmission and acknowledgement for multiple active nodes in a sector.

in the CTS. To reduce the propagation overhead, all CTS and ACK packets are grouped respectively in bursts. Again, only one DATA packet per node is allowed, the next packet should be sent in the following cycle. Fig. 52 shows the packet flow for the case of two nodes, and it can be extended for a higher number of nodes.

The performance of the protocol is governed by AP's cycle time, which is the time necessary to probe all sectors once and complete data transmissions, if any. Achieving a low cycle time is critical to reduce the latency of the protocol and increase the capacity.

Importantly, the sector time is adapted depending on the number of RTS received. An empty sector will be *skipped*, reducing cycle time, whereas, in busy sectors, the sector time will be extended to let all nodes complete the transmission of one DATA packet. This strategy greatly improves channel utilization compared to a fixed sector time approach.

Another benefit of the 3-way handshake is collision avoidance. DATA packets cannot collide, since exclusive Tx slots are defined; only RTS packets can collide. However, due to the very short transmission time of control packets, of few nanoseconds, RTS collisions are quite unlikely. Propagation delay (3.33 ns/m) multiplexes RTS packets in time, so nodes at different distances to the AP will never collide. Collisions can only occur between two nodes at a similar distance to the AP. To avoid systematic collisions, nodes wait a random backoff time of $[0, 10]$ ns before sending the RTS.

When collisions do occur, the Tx opportunity in the current AP cycle is lost, and nodes should retry in subsequent cycles. A random backoff mechanism will hold off the next retry to avoid colliding again. This is a count-based backoff with an initial life of $[0, 2^r]$, where r is the number of retries, and is decreased with every CTA received.

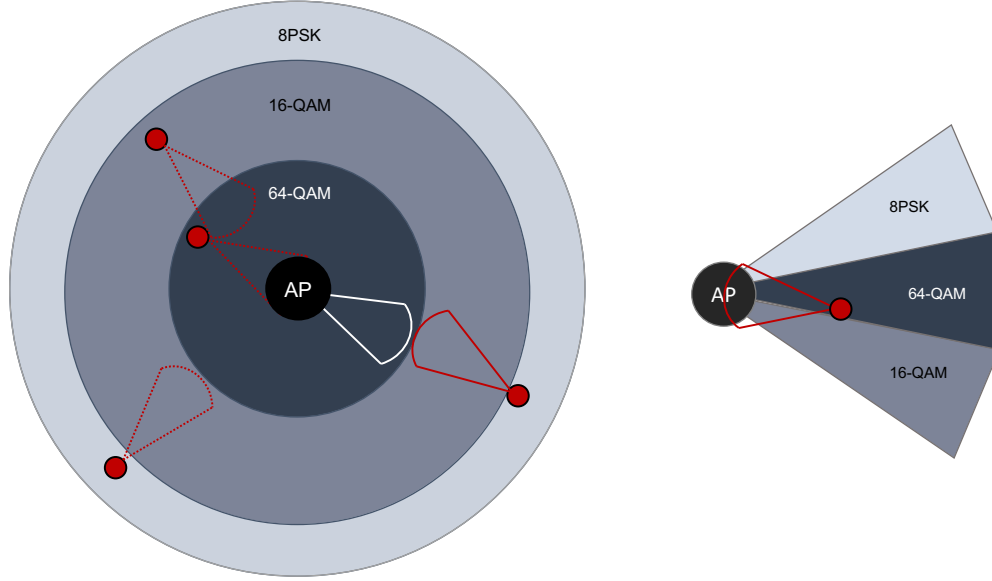


Fig. 53: Example of highest possible MCS depending on distance to the AP when using the primary sector (left), and MCS in the primary sector and adjacent sectors for a node close to the AP (right).

Adaptive MCS and White List: With the use of a 3-way handshake, new opportunities for optimization arise. We have designed an adaptive Modulation and Coding Scheme (MCS) mechanism that dynamically changes the modulation depending on signal power.

A higher-order modulation encodes more information by using a more complex signal, but also requires a higher SNR to be decoded. With the pronounced path loss that affects THz signals, power at reception depends highly on distance. Control packets are always sent using the base MCS, which is simple enough for reliable communication with the furthest nodes. After the CTA-RTS exchange, the AP decides which MCS can a node use depending on the signal power received, and indicates it in the CTS. DATA is always transmitted using the highest possible modulation order, thus reducing DATA Tx time.

Since path loss is lower in short distances, nodes closer to the AP can use their primary sector and also adjacent sectors, even outside of the 3 dB beamwidth. However, the gain is lower in adjacent sectors and a sub-optimal MCS would be used. We have implemented a sector's *white list* to avoid this situation. After a sounding phase, the AP creates a white list for every sector with the criteria of only allowing nodes to send data in their highest power sector. Thus, it ensures that the highest possible MCS is always used. The white list is checked before sending a CTS to a node. Another benefit of this strategy is that of guaranteeing that all nodes have one and only one DATA Tx opportunity per cycle, for maximum fairness.

1-way variant: We also propose a variant of ADAPT with a 1-way handshake (ADAPT-1). The benefit of this approach is the reduction of overhead in the handshake. In this case, after receiving a CTA, nodes can directly send DATA.

The concept of adaptive sector time is still partially applied: the AP waits for T_{wait} and will turn to the next sector if it does not sense any DATA being received. If it does sense DATA, the AP waits until reception is finished and sends an ACK. The adaptive MCS strategy and sector's white

list cannot be implemented as they require a 3-way handshake.

Communication with multiple nodes in a sector is not supported. In the case that two or more nodes send DATA, channel sensing is performed for collision avoidance, even though it is not reliable due to the deafness problem.

The comparison between the 3-way version (ADAPT-3) and ADAPT-1 will allow us to evaluate whether or not the benefits of a 3-way handshake outweigh the cost of introducing extra overhead.

A note on neighbor discovery: In this work, we assume that all nodes know the position of the AP and are directionally listening towards it. However, we are aware that neighbor discovery is an open challenge in wireless directional networks. A possible option relies on leveraging the full antenna pattern information and the unique changes in the time-derivatives of the received electric field at the receiver of a turning antenna, to expedite the discovery process, as we showed in [95].

Analytical model In this section, we derive an analytical model of the 3-way ADAPT protocol for the scenario of a stable network and assuming no collisions, as the 3-way handshake greatly reduces collision probability. As we will see in Sec. 6.3.1, the model is accurate despite this simplification.

We consider nodes to be randomly distributed following a Poisson space distribution of rate λ_A nodes/m² around the central AP. For a range r , the total number of nodes is $N_{\text{nodes}} = \lambda_A \pi r^2$.

We model arrivals at every node as a Poisson process of rate λ packets/s, with an exponential distribution $\lambda e^{-\lambda t}$. It can be expressed in terms of the mean inter-arrival time as $\lambda = 1/T_{\text{ia}}$. Assuming no collisions, a node will have exactly one transmission opportunity at every cycle. Therefore, service rate μ depends only on cycle time (T_{cycle}), given by $\mu = 1/T_{\text{cycle}}$ packets/s. System load ρ is defined as the ratio between arrival and service rates, $\rho = \lambda/\mu$.

Cycle time: T_{cycle} has two components, a fixed time to sound all sectors and a variable term that depends on the number of data transmissions n in the cycle. It is expressed as

$$T_{\text{cycle}}(n) = N_{\text{sec}} T_{\text{wait}} + n T_{\text{tx}}. \quad (69)$$

T_{wait} must be long enough to let the furthest node reply with an RTS, and is defined as

$$T_{\text{wait}} = T_{\text{cta}} + \max(T_{\text{b/o}}) + T_{\text{rts}} + 2 \max(T_{\text{prop}}), \quad (70)$$

where $T_{\text{b/o}}$ is a random backoff of $[0, 10]$ ns and T_{prop} is the one-way propagation time. T_{tx} accounts for the CTS-DATA-ACK exchange.

To formulate T_{tx} , we use an average T_{prop} and T_{data} . We acknowledge that T_{prop} depends on the distance from node to AP, different for each node, and the model should account for its distribution. However, we found the protocol to be accurately modeled using an average time (results in Sec. 6.3.1), while complexity is significantly reduced. Thus, T_{tx} is

$$T_{\text{tx}} = T_{\text{cts}} + T_{\text{data}} + T_{\text{ack}} + 2 T_{\text{prop}}. \quad (71)$$

T_{cta} , T_{rts} , T_{cts} , T_{data} and T_{ack} represent transmission time for each type of packet. Each node can send either send zero or one packet in a cycle. The load ρ represents the probability that a node will have data to transmit. With a total of N_{nodes} , the probability of n transmissions in a cycle follows a binomial distribution given by

$$P_n = \binom{N_{\text{nodes}}}{n} (1 - \rho)^{(N_{\text{nodes}} - n)} \rho^n, \quad (72)$$

with an expected value of $E[n] = N_{\text{nodes}} \rho$. Therefore, the average T_{cycle} is

$$T_{\text{cycle}} = N_{\text{sec}} T_{\text{wait}} + N_{\text{nodes}} \rho T_{\text{tx}}. \quad (73)$$

As for the distribution of T_{cycle} , we know that cycle time can only take a set of values corresponding to (69) for $n = 0, 1, 2, \dots$. Therefore, the probability density function (PDF) of cycle time is not smooth, but shall be expressed as a set of Dirac delta functions placed at each of these time values, and weighted by the probability of n from (72), resulting in

$$f_{T_{\text{cycle}}}(t) = \sum_n P_n \delta(t - T_{\text{cycle}}(n)). \quad (74)$$

Stability condition: Stability is studied in terms of load ρ . The system requires that all nodes meet the condition $\rho < 1$ to be stable, meaning that arrival rate is lower than service rate. Otherwise, the number of packets in queue would be ever-growing, resulting in buffer overflow. Since all nodes have the same arrival and service rates, the system stability condition is given by

$$\rho = \frac{\lambda}{\mu} = \frac{T_{\text{cycle}}}{T_{\text{ia}}} < 1. \quad (75)$$

As T_{ia} is a fixed parameter, the system stability is defined by the T_{cycle} achieved. The faster the AP turns, the more transmission opportunities and more robust network. From (75) and (73), we can isolate ρ and express the stability condition as

$$\rho = \frac{N_{\text{sec}} T_{\text{wait}}}{T_{\text{ia}} - N_{\text{nodes}} T_{\text{tx}}} < 1. \quad (76)$$

Packet Delay and Throughput: To prevent buffer overflow and ensure stability the system needs $\rho < 1$, but for good network performance the load should be much lighter. In the packet delay and throughput analysis we will consider networks with $\rho \ll 1$.

We define packet delay T as the time from enqueue to reception of ACK. It is divided in three terms, expressed as

$$T = T_{\text{face}} + T_{\text{w}} + T_{\text{sector}}. \quad (77)$$

T_{face} is the time from arrival of the packet until the AP and node face. T_{w} is the wait time to send packets in queue. Lastly, T_{sector} is the sector time in which the packet is transmitted, from the moment CTA is sent until ACK is received.

For each packet, throughput is defined as $S = L/T$, where L is the payload length. For an accurate throughput estimate that captures the distribution of packet delay, we formulate the PDF of T , $f_T(t)$, and then compute throughput as

$$S = \int \frac{L}{t} f_T(t) dt. \quad (78)$$

T is a random variable formed by the sum of three independent random variables. Thus, its PDF is the convolution of the PDFs of its terms, expressed as

$$f_T(t) = f_{T_{\text{face}}}(t) * f_{T_{\text{w}}}(t) * f_{T_{\text{sector}}}(t). \quad (79)$$

Starting with T_{face} , the time from enqueue until the AP and node face each other is uniformly distributed between 0 and T_{cycle} , given by

$$T_{\text{face}} = U(0, 1) T_{\text{cycle}}. \quad (80)$$

From (74) and (80), we can formulate the PDF of T_{face} as

$$f_{T_{\text{face}}}(t) = \sum_n P_n U[0, T_{\text{cycle}}(n)]. \quad (81)$$

For T_w , the wait time to send N_w packets, one cycle per packet, is

$$T_w = N_w T_{\text{cycle}}. \quad (82)$$

Considering that the service rate is much higher than arrival rate ($\rho \ll 1$), it is unlikely to find many packets in queue at the time of arrival. In our numerical analysis we will only consider the scenarios of $n_w = 0, 1$, with probabilities $P_{n_w}(0) = (1 - \rho)$, $P_{n_w}(1) = \rho$, where n_w is the number of packets in queue. Hence, using (74), we can express the PDF of T_w as

$$f_{T_w}(t) = \sum_{n_w} P_{n_w} n_w f_{T_{\text{cycle}}}(t) = (1 - \rho)\delta(t) + \rho f_{T_{\text{cycle}}}(t). \quad (83)$$

Lastly, T_{sector} corresponds to the time between CTA is sent until ACK is received, given by

$$T_{\text{sector}} = T_{\text{wait}} + 3 T_{\text{prop}} + n_s(T_{\text{cts}} + T_{\text{data}} + T_{\text{ack}}). \quad (84)$$

It depends on the number of packets sent in the sector (n_s). For scenarios with $\rho \ll 1$ it is unlikely to have multiple transmissions in a sector. We will only consider $n_s = 1$. T_{data} is not fixed, it depends on the MCS chosen. Considering M possibilities, each modulation $m \in M$ has a probability P_m and DATA Tx time T_{data}^m . The resulting PDF is expressed as

$$f_{T_{\text{sector}}}(t) = \sum_m^M P_m \delta(t - (T_{\text{wait}} + 3T_{\text{prop}}^m + T_{\text{cts}} + T_{\text{data}}^m + T_{\text{ack}})). \quad (85)$$

At this point, we can derive $f_T(t)$ in (79) from (81), (83), (85). Then, we are able to estimate the system's throughput by evaluating (78).

6.2.2 Dual-band MAC Protocol for Dynamic Spectrum Sharing above 100 GHz

We have designed and implemented a dynamic spectrum access, dual-band backhaul system in the above-100-GHz spectrum, specifically, at 123.5–140 GHz (referred to as lower band throughout the paper) and 210–225 GHz (upper band). This is the first above-100-GHz wireless system that (i) operates a real-time protocol stack in two frequency bands; (ii) performs automated and dynamic spectrum sharing; and (iii) adapts link parameters (e.g., the frequency band) in real-time. The proposed system tracks satellite orbits and automatically switches frequency bands, thus preventing backhaul transmissions from interfering with passive incumbents while enabling the allocation of large, contiguous bands for communication services on shared spectrum. Prototyping this system has required us to address challenges that span from RF design for the dual-band setup, to dynamic control and integration of multiple independent systems through open interfaces.

RF and Communication Components Figure 54 shows the prototype hardware, with transmitting and receiving real-time signal processing back-ends and dual-band sub-terahertz front-ends. At the time this work was conducted, the real-time DSP engine presented in Sec. 3 was not completed and, thus, we had to rely on a millimeter-wave, narrow-band alternative, namely, the NI mmWave platform [96]. This features a hardware-accelerated protocol stack that generates an OFDM waveform inspired by the 5G NR physical layer, aggregating eight component carriers with 100 MHz each [97]. The dual-band front-ends, supporting 123.5–140 GHz and 210–225 GHz, accept an IF input at the up-converters (transmitter-side) and generate an IF output at the down-converters (receiver-side). They are connected to the baseband-to-IF conversion modules of the NI system through a Single-Pole Double Throw (SPDT) RF switch, which supports DC–18 GHz [98]. Four Keysight E8257D signal generators drive the LO of the dual-band front-ends.

The overall system design was hindered by two main challenges. First, the same IF output, with high PAPR, needs to be interfaced with dual-band front-ends with different and limited input power specifications. Second, phase noise and dual sideband front-ends introduce amplitude impairments in the received signal.

The link budget planning aims at having a 10 dB SNR margin at the IF input of the receiver back-end, independently on the lower-band or upper-band front-ends. This is challenged by two factors. First, the back-end OFDM implementation introduces a high peak-to-average power ratio (more than 10 dB), even for the lowest order modulations. This is a known drawback for OFDM systems using large FFT sizes (2048 for each component carrier, in this case). Second, there is an input power mismatch between the lower-band and upper-band transmitter front-ends. The lower-band front-end needs an input signal with power lower than 0 decibel-milliwatts (dBm), and generates a maximum 13 dBm output power. The upper-band front-end needs an input power of 10 dBm to generate a 3 dBm output at the desired harmonic. To address this, we precisely characterized the OFDM waveform power, and reduced the back-end output power to maintain the peak power under the limit of the upper-band front-end. The input to the lower-band front-end was then further attenuated to be below 0 dB. Finally, to close the link margin for the 43 m link, the lower-band and upper-band systems use 38 dBi and 40 dBi antennas, respectively.

The transmitter and receiver at both bands feature two mixing stages, part of the analog front-ends shown in Fig. 54, thus forming a heterodyne system. However, due to the lack of RF filters at the considered frequency ranges, both lower-band and upper-band front-ends lack a sideband selection filter. This, combined with the high phase noise (due to the frequency multiplication process), may lead to drastic amplitude fluctuations in the received signal. To address this, we offset the IF frequency in the up- and down-converting paths by 1 GHz, different from traditional heterodyne systems, which use the same IF frequencies. Thus, only one of the sidebands is selected for receiver-side processing, preventing undesired amplitude fluctuations. To enable quick dynamic switching, the lower-band and upper-band front-ends are set to the same IF frequency.

The frequency offset is compensated by appropriately setting the IF-to-baseband and the IF-to-RF LO frequencies in the NI back-end and the two lower-band and upper-band transmit-receive paths. Notably, the lower-band front-ends perform a $4\times$ frequency multiplication of the LO input, thus the transmitter LO is set to 32.5 GHz (130 GHz at the up-converter), while the receiver LO is set to 32.75 GHz (131 GHz at the down-converter). Similarly, upper-band converters utilize a $9\times$ ($6\times$) LO multiplication at the transmitter (receiver) side. Therefore, the LO input at the transmitter-side was set to 23.88 GHz (corresponding to 215 GHz) and the receiver-side was set to $216/6 = 36$ GHz, selecting the upper sideband for baseband processing.

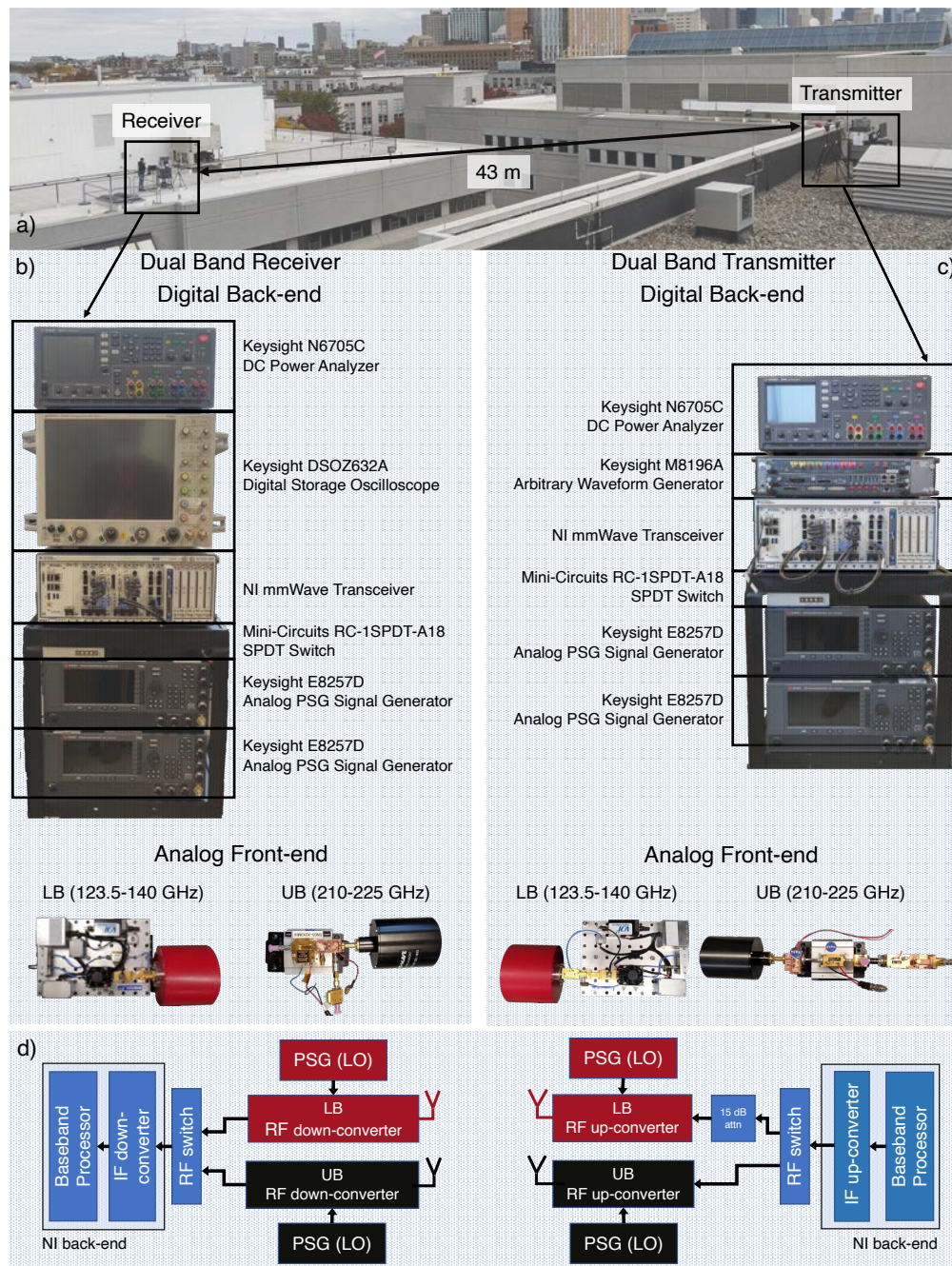


Fig. 54: Prototype schematics and dual-band backhaul link deployment.

Dynamic Frequency Switching Framework The dynamic control of the dual-band front end requires the integration of the hardware components with a programmable, software-based control logic that can automatically switch bands when the active system interferes with a passive incumbent. The framework features (i) a tracking system that identifies when lower-band or upper-band transmissions may be interfering with a passive user (e.g., a satellite orbiting over the link); and (ii) an interface with the SPDT switch.

This flexible architecture can operate in two different modes. The first one is named independent switching and assumes no coordination between the backhaul endpoints, which are only loosely time-synchronized to a common source through the network time protocol and perform frequency switching decisions independently. This configuration represents an ad hoc network deployment, without out-of-band control. The framework runs on the NI back-ends. We call the second deployment mode coordinated switching. It mimics the configuration of backhaul nodes for cellular networks, where the equipment is connected to a control overlay [99]. The tracking framework is hosted by a central controller, which notifies the two endpoints simultaneously.

We considered as passive user the NASA satellite Aura, which studies the Earth climate and air quality [100] and is listed as incumbent in the spectrum adjacent to the upper-band band in the World Meteorological Organization (WMO) Observing Systems Capability Analysis and Review Tool (OSCAR) database [101]. Note that we operate the upper-band backhaul link in the 210–225 GHz range, compliant with FCC regulations and an experimental license. However, its RF front-end could extend up to 240 GHz, thus potentially interfering with the Aura Microwave Limb Scanner (MLS). This instrument scans the ozone (O₃) absorption line around 235.7098 GHz, and the carbon monoxide (CO) line at 230.538 GHz [102]. Therefore, in this specific setup, we consider the lower-band as the band where the upper-band traffic can be offloaded to avoid interference to the Aura MLS. The tracking system retrieves Aura’s orbits from public API [103] and checks whether the satellite is orbiting over the backhaul link deployment. Harmful interference is detected when the satellite and backhaul are in line-of-sight, or the actual link budget exceeds an ITU threshold (as discussed next). Accordingly, the SPDT switch interface triggers an upper-band-to-lower-band band switch through HTTP API [104]. The switching mechanism can be further refined to avoid unnecessary or too frequent band switching that may degrade the system performance.

6.3 Results and Discussion

6.3.1 Simulation-based Study of ADAPT

In this section, we investigate the performance of the protocol proposed, both the 3-way (ADAPT-3) and 1-way variant (ADAPT-1), in terms of throughput and packet discard rate. We compare them to the previous 1-way solution with fixed cycle time (1-way-2019), CSMA (0-way) and CSMA/CA (2-way). Cycle time and optimal packet size are also studied.

The protocols are implemented in the *ns-3* network simulator using TeraSim [39], an extension that models the peculiarities of the THz frequency-selective channel. TeraSim is available and free to download in the *ns-3* store. Extensive simulations are used to investigate the performance and validate the analytical model.

Each data point is the average of 10 simulations of different seed, with a duration of 10 ms each. This time duration is long enough to allow for hundreds of AP cycles, and short enough to consider a fixed scenario, as the impact of mobility is left for future work. Packet arrivals at

Table 12: Modulation schemes supported in 802.15.3d

Modulation	Data rate [Gbps]	SNR _{th} [dB]	Range [m]
BPSK	52.4	10.6	48
QPSK	105.3	12.4	34
8PSK	157.4	17.6	18
16-QAM	210.2	19.2	15.4
64-QAM	315.4	25.4	7.5

every node follow a Poisson process of rate $\lambda = \frac{1}{T_{ia}}$, and a packet is discarded after 5 unsuccessful transmission attempts. More or longer simulations do not have a significant impact on the results, since a permanent state is reached after few microseconds of simulation.

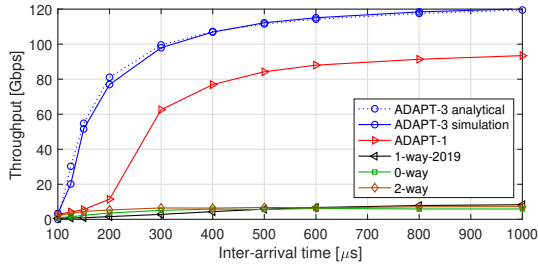
Configuration In our analysis, we use the physical layer proposed in the IEEE 802.15.3d standard for high data rate wireless networks. We select a single carrier signal of $B = 69.12$ GHz, between 252.72 – 321.84 GHz. In this window, the thermal noise for a temperature of 300 K is $N_T = -65.4$ dBm. Noise due to molecular absorption is very low at these frequencies and can be neglected. We assume a noise figure at reception of $N_F = 7$ dB. In the trade-off between gain and sectors, we have chosen a beamwidth of 12° that divides a cycle in 30 sectors and provides 24.57 dB of maximum gain. 802.15.3d supports modulations from BPSK to 64-QAM with a forward error correction (FEC) rate of 14/15. We calculate the SNR_{th} necessary to achieve a BER of 10^{-6} . Transmission power is set to $P_t = 20$ dBm. After the effect of path loss and the gain of both antennas, the resulting power at reception P_r defines the range of communication. Reception will only be considered successful if $P_r/(N_F + N_T) > \text{SNR}_{th}$.

Tab. 12 shows the modulations and data rates supported in 802.15.3d, the SNR_{th} for the proposed BER and the maximum range for each modulation. We simulate a range of 18 m, so nodes will use either 8PSK, 16-QAM or 64-QAM, depending on P_r . Packet length is of 65 000 bytes. DATA Tx time at the highest modulation order of 64-QAM, with a data rate of 315.4 Gbps, is of $1.64 \mu\text{s}$.

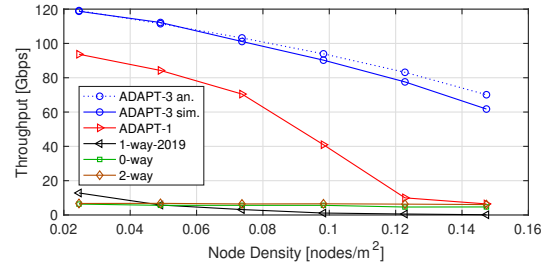
Throughput and Discard Rate Fig. 55a and Fig. 55b show throughput against inter-arrival time and node density, respectively. With the AP covering a radius of 18 m, a node density of 0.05 nodes/m^2 represents 50 nodes.

We observe that ADAPT brings a notorious improvement, reaching throughput peaks of 120 Gbps, compared to the ~ 10 Gbps of other solutions. The main reason for such improvement is the adaptive sector time mechanism, which greatly reduces cycle time and, thus, packet delay. As expected, when the network load is increased with a higher node density or arrival rate, throughput decreases and the network eventually reaches an unstable point.

Also included in Fig. 55 is the analytical throughput for ADAPT-3. We validate the model by comparing analytical to simulation results. The analytical estimate is fairly accurate despite assuming no collisions, while collisions are taken into account in the simulations. Only in the scenarios with a higher network load, as collision probability rises, the analytical model slightly overestimates throughput.

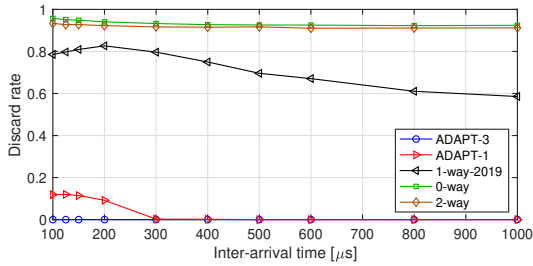


(a) Fixed node density at 0.05 nodes/m²

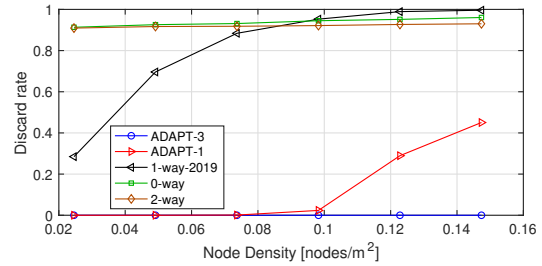


(b) Fixed T_{ia} at 500 μ s

Fig. 55: Throughput simulation and analytical results for a range of 18 m.



(a) Fixed node density at 0.05 nodes/m²



(b) Fixed T_{ia} at 500 μ s

Fig. 56: Packet discard rate simulation results for a range of 18 m.

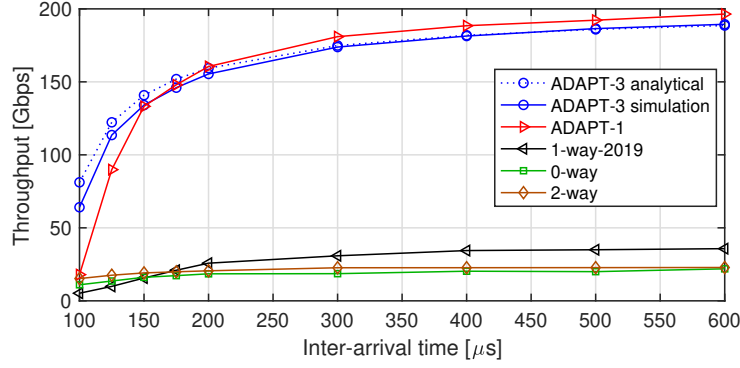


Fig. 57: Throughput simulation and analytical results for a range of 7.5 m. Fixed node density at 0.28 nodes/m².

Improvement is observed as well in packet discard rate, Fig. 56a and Fig. 56b, as ADAPT-3 presents a null discard rate in all of the scenarios simulated. ADAPT-1 also reduces packet discards despite not having reliable collision avoidance, thanks to the faster turning speed. Transmitter-initiated handshakes prove not to be fit for a this network architecture. Since most transmission attempts will find the AP facing elsewhere, 0-way and 2-way protocols have a steadily high discard rate.

Comparing ADAPT-3 and ADAPT-1, we observe that the benefits of the 3-way handshake pay off and it reaches higher throughput rates than the 1-way, despite the larger overhead. This is due to the collision avoidance of the 3-way exchange and the adaptive MCS strategy.

To study protocol performance in smaller areas, a scenario of 7.5 m of range was also investigated. Fig. 57 shows throughput as a function of T_{ia} for 50 nodes. There are two major differences compared to the 18 m range. Firstly, all nodes will be able to use the highest modulation order of 64-QAM, nullifying the adaptive MCS strategy. ADAPT-3 loses its edge over ADAPT-1 and now both perform similarly. And secondly, with the shorter distance comes a lower propagation delay, which reduces cycle time and packet delay, and enables throughput rates up to 200 Gbps.

Cycle time The AP's average cycle time is a key parameter that will define the performance of the protocol. Since nodes have one Tx opportunity per cycle, a faster turning speed means more frequent Tx opportunities, leading to less packet delay and a more stable network. Moreover, the cycle time also determines the latency of the system.

The improvement of ADAPT compared to the 1-way solution comes from the reduction of cycle time. Because of the adaptive sector time and MCS, T_{cycle} is also adaptive, unlike the previously fixed T_{cycle} . The minimum cycle time, in the case of zero data transmissions in a cycle, is the time to sound all sectors, $T_{cycle}^{min} = N_{sec} T_{wait}$. For the 18 m range, T_{cycle}^{min} is of 3.96 μs.

Overall, ADAPT-3 achieves a 20-fold reduction on the average T_{cycle} compared to 1-way-2019. Fig. 58a shows an inverse dependence between T_{cycle} and send rate, asymptotically approaching the minimum as the network load decreases. A similar behaviour is observed in Fig. 58b, as T_{cycle} rises when the network load is increased with a higher number of nodes. The distribution of nodes between the sectors does not affect significantly T_{cycle} . When the average cycle time is very close to the minimum, it effectively means that most sectors hold no DATA Tx and are skipped.

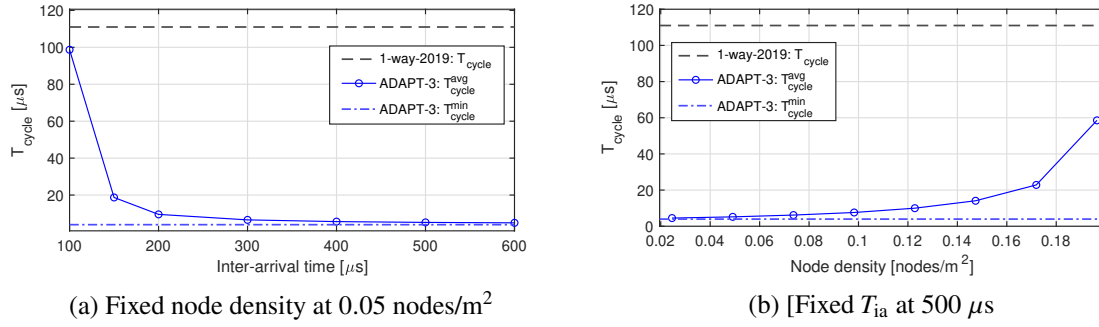


Fig. 58: Cycle time of 1-way-2019 and ADAPT-3 for a range of 18 m.

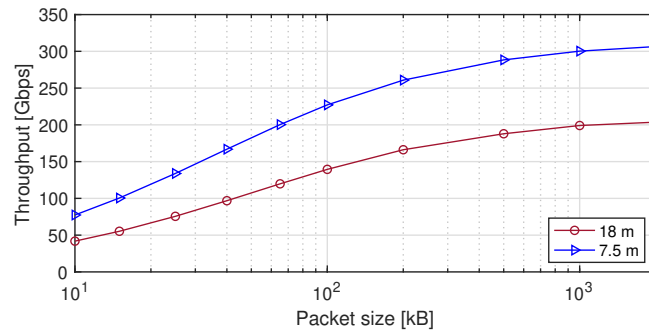


Fig. 59: Analytical throughput for ADAPT-3 as a function of packet size. Parameters: 50 nodes and 10 ms of T_{ia} .

Packet size Very high speed networks in ultra-broadband communications open the door to the use of much larger payloads, as the faster data rates allow to transmit long packets in a very short time. Furthermore, in the network architecture presented, each Tx opportunity has a fixed delay cost, which is the waiting time until AP and node face each other. Sending more data in each Tx opportunity minimizes this cost. In 802.15.3d, payload size limits are set to a minimum of 2 kB and maximum of 2 MB. Our analysis was conducted with DATA payloads of 65 kB.

Fig. 59 shows the achievable throughput for packets between 2 kB - 2 MB, estimated with the analytical model for a scenario of 50 nodes and T_{ia} of 10 ms. We can see how throughput as a function of packet size has a logarithmic growth and it approaches a superior limit set by the physical data rates. However, the use of long payloads also brings some concerns, such as error correction or the negative impact on latency. An optimal packet size, or the use of a dynamic payload, should be studied weighing up these concerns.

Future work In the design of ADAPT we have used a simplified scenario for a more clear presentation of the protocol. Future work can consider extending ADAPT for downlink communication, generalizing from 2-D to a 3-D space, or modeling node mobility. Another interesting question

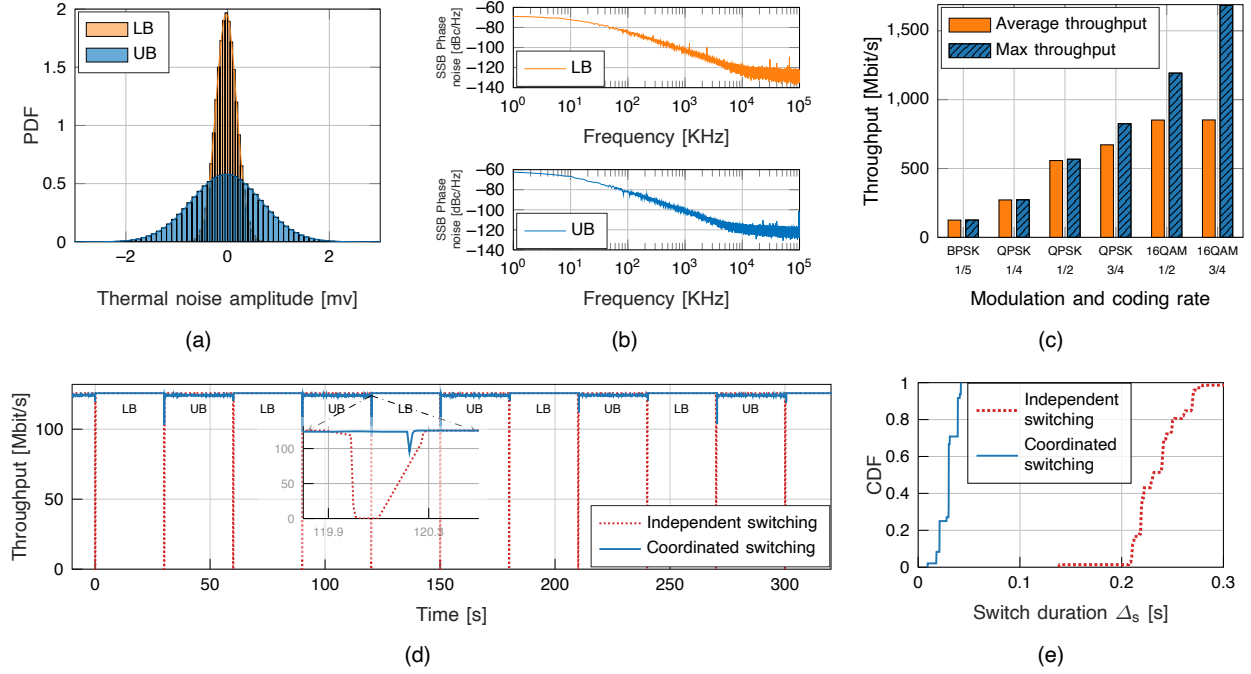


Fig. 60: Experimental result for the dual-band self-adaptive backhaul link.

would be to include machine learning techniques, which could be used to learn the most active sectors or to adapt the beamwidth dynamically. Moreover, it would be of interest to study the use of multiple antennas in an AP, which could operate simultaneously by leveraging directional communication.

6.3.2 Experimental Validation of the Dynamic Dual-band Terahertz MAC Protocol

The dual-band transmitter and receiver were deployed outdoor, on top of two Northeastern University campus buildings, at a distance of 43 m, as shown in Figure 54.

Despite the similar link budget, the higher thermal and single sideband phase noise (Figs. 60a and b) of the upper band front-end (with a ~ 10 dB noise figure increase) affects its throughput when higher MCSs are considered (Figure 2c). The maximum throughput (achieved with the lower band system) and the average throughput (which accounts for both) are comparable up to a QPSK modulation with a 1/2 coding rate. Less robust MCS lead to lower throughput for upper band. As a consequence, the average throughput is below the maximum achievable by the system.

To evaluate the effectiveness of the proposed frequency switching approach, Fig. 60d reports the throughput over time with independent and coordinated switching artificially triggered every 30 s. The system maintains stable throughput with negligible impact of the switching procedure, especially with coordinated switching. This is only an illustrative example to show the effectiveness of the switching mechanism, as the actual time spent in each frequency band varies and depends on the presence of passive users in the upper band. Fig. 60f plots the cumulative distribution function of the switching duration Δ_s over 120 independent switching events, defined as the time interval

between the switch at one endpoint and the switch at the other. Independent switching requires a median value of $\Delta_s = 230$ ms, as the switch may happen at slightly different time instants in the two endpoints, which only rely on the Network Time Protocol for a loose synchronization [105]. The coordinated approach takes less than 42 ms in the worst case (median 30.11 ms), as the switch is triggered by a centralized controller for both endpoints. A short switching time also allows the system to easily scale in case multiple incumbents need to be tracked and avoided. Overall, independent switching—which facilitates deployment in challenging environments—has a trade-off in performance with coordinated switching. Even in case the frequency switch happens infrequently, and thus has limited impact on the average throughput, a downtime of tens of milliseconds may affect the end-to-end performance of traffic flows traversing the link, e.g., it may trigger retransmissions or session resets at the transport or application layers. Nonetheless, both switching mechanisms allow the prototype to promptly avoid harmful interference, while maintaining high average throughput values.

To understand the dynamic spectrum sharing opportunities above 100 GHz (specifically, between 209–241 GHz), Figure 61 illustrates spectrum regulations (bottom) [91, 92], and a link budget for the worst-case interference toward a sensing system on the Aura satellite orbiting over the backhaul link (top). For this analysis, we focus on the 209–241 GHz band, as it is spectrum in which the upper band front-end operates, but similar coexistence issues apply to the ITU-regulated spectrum above 100 GHz [92].

The 209–241 GHz spectrum includes bands where active users (i) cannot transmit, (226–231.5 GHz, according to footnotes FCC US246 [91] and ITU 5.340 [92]); or (ii) need to protect passive radio-astronomy and Earth monitoring users (209–226 GHz and 235–238 GHz, according to the FCC US342, ITU 5.149, and 5.563). The ITU defines harmful interference as received power above the threshold $P_{RX,th} = -194$ dBW in the 226 – 231.5 GHz band [106]. The top left corner of Figure 61 evaluates whether this threshold is exceeded for Aura—which orbits 705 km over the Earth [100]—by reporting the received power P_{RX} for different elevation angles α . We conservatively assume that the backhaul link directional antenna (with a 40 dB gain, and 23 dBm of transmit power) steers toward the satellite in a line-of-sight scenario. This is usually not the case, yet provides a worst-case estimate of the interference. We consider the ITU channel models for the atmospheric attenuation as a function of frequency, height, and elevation angle α [54]. The carrier frequency is set to $f_c = 226$ GHz, i.e., the closest frequency to the upper band carrier actually used in the experiment (224 GHz) with a threshold defined in [106]. While the atmospheric attenuation of the signal prevents interference for small α , for $\alpha \geq 8.4^\circ$ the received power exceeds $P_{RX,th}$. This translates into satellite passes (highlighted in Figure 61) for which the dual-band backhaul link needs to switch from the upper band to the lower band.

This considered, sharing solutions that would enable contiguous, large bandwidth transmissions in the above-100-GHz bands must not be static. Dynamic spectrum sharing would allow regulations to evolve from a firm prohibition of any active transmissions in passive bands, to more flexible schemes where active users could transmit on ultra-large contiguous bandwidths for a portion of time. For example, by considering the 209 – 241 GHz band, only 6.5 GHz (non-contiguous, 231.5 – 235 GHz and 238 – 241 GHz) are earmarked for fixed or mobile transmission without the need to share spectrum and thus protect passive incumbents. Fig. 62 reports the Shannon capacity as a function of the bandwidth B for different spectrum sharing and aggregation configurations. We consider a received power $P_{RX} = -23$ dBm, a noise figure $F = 8.5$ dB for the receiver, a noise temperature $T_0 = 296$ K. The SNR in dB is $\Gamma = P_{RX} - 10 \log_{10}(kT_0B) - F$, and the Shannon capacity is computed

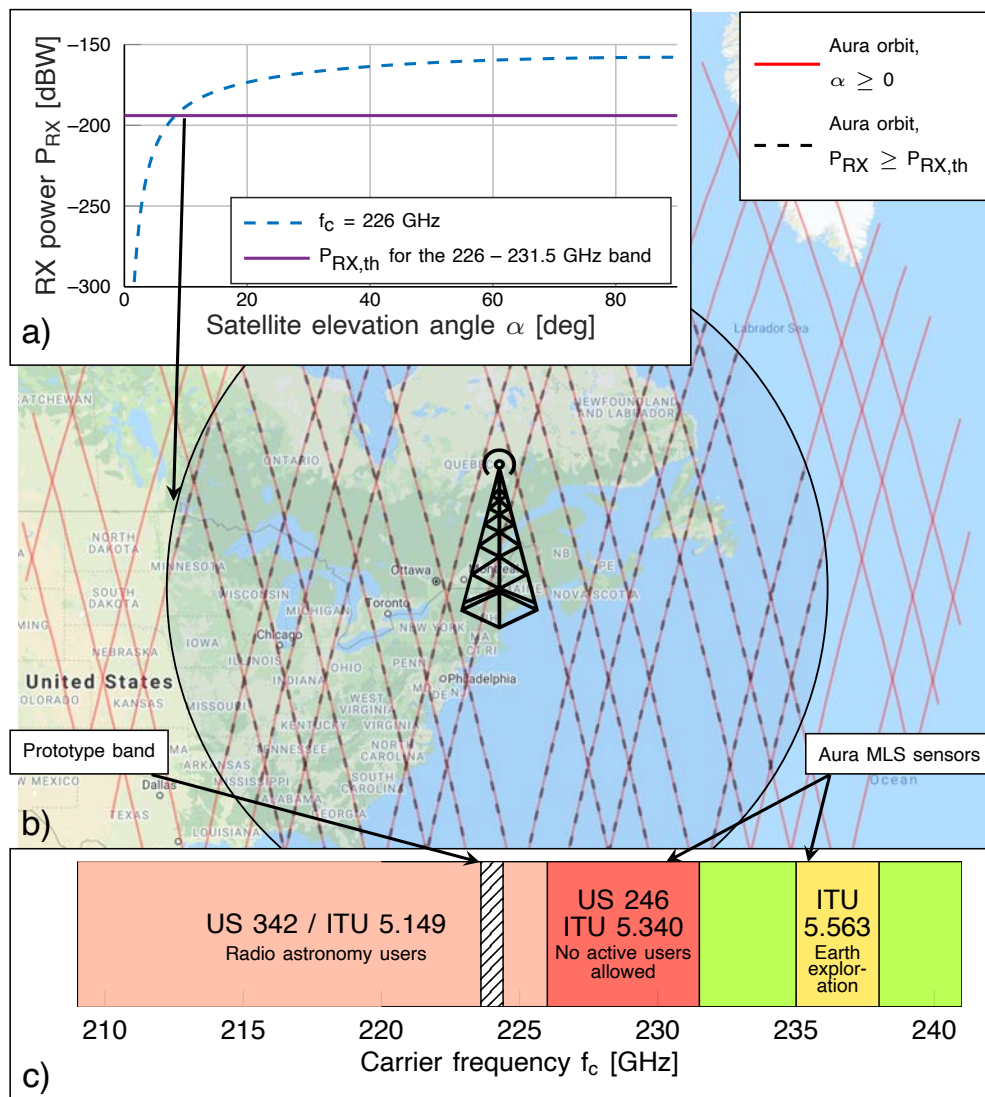


Fig. 61: Link budget analysis, Aura satellite orbits, and spectrum.

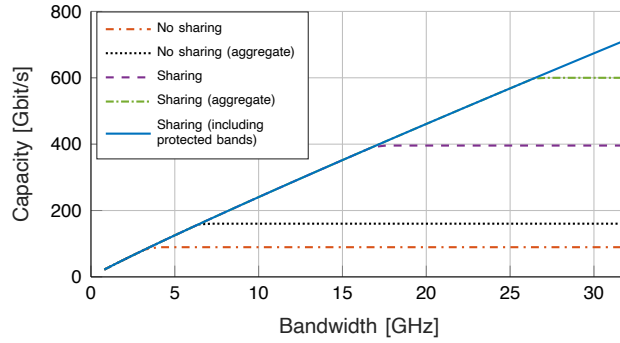


Fig. 62: Capacity in Gbps as a function of the bandwidth B .

as $C = B \log_2(1 + \Gamma)$. According to current regulations, we limit the capacity to the value supported by the maximum bandwidth allowed by the FCC allocations (e.g., 3.5 GHz for the “No sharing” configuration). We also consider the possibility of aggregating multiple frequency bands for a single transmission, in this case, the “No sharing” configuration would exploit $3.5 + 3 = 6.5$ GHz. Note that carrier aggregation may not be feasible for all wireless communication technologies. Figure 62 clearly shows the gap in capacity between a system where dynamic spectrum sharing is allowed and one which follows the current regulations, with a gain of up to 8x (no aggregation) and 4.5x (aggregation). Additionally, results in [107] show that, to reach a target rate of 1 Terabit per second (Tbit/s), future 6G systems above 100 GHz will require more contiguous bandwidth than what is available without spectrum sharing.

Next Steps While this solution has been presented in the context of spectrum sharing and coexistence, it ultimately shows the ability to dynamically change sub-terahertz frequencies in real-time and on demand. Effectively, this is the first THz MAC protocol ever experimentally demonstrated. Moving forward, the ability to dynamically change bands or to, potentially, operate over multiple bands simultaneously will be incorporated in our protocol drawing board. In addition, the much faster real-time DSP engine presented in Sec. 3 will be leveraged for future works.

7 RELAYING AND ROUTING IN TERAHERTZ NETWORKS

7.1 Motivation

Beyond continuously increasing the antenna gains, other strategies to meaningfully increase the communication distance in terahertz networks include the development of multi-hop relaying strategies. While this is a very well-explored topic at lower frequencies [108–110], existing solutions cannot be simply reused, because they do not capture the peculiarities of the aforementioned THz channel and THz devices and their impact on the link layer. More specifically, due to the unique distance-dependent behavior of the available bandwidth, the reduction of the transmission distance results in twofold benefits. On the one hand, as in any other wireless communication system, by reducing the transmission distance between nodes and maintaining the transmission power, the SNR increases and higher order modulations can be utilized to increase the data rates. On the other hand, because of the behavior of molecular absorption, shorter transmission distances also benefit from a substantially larger channel bandwidth, which again contributes to the increase of achievable data rates. However, the need for more hops to cover the same distance and the computational cost of relaying data at multi-Gbps or even Tbps pose major constraints on the system design. In addition, the impact of distance on the required DA gains and, consequently, beam widths, again affect the system performance through the aforementioned MAC and neighbor discovery implications. Moreover, the queuing delay needs to be taken into account, particularly since we are dealing with very fast transmission speeds. The queuing delay plays a key role in the total E2E delay, and is likely to increase when more relays are involved in the communication link.

As part of this effort, we have studied new cross-layer relaying strategies for multi-hop communications in the terahertz band. More specifically, we have derived a mathematical framework to smartly allocate the resources across layers by optimizing the distribution of relays and ultimately achieving the smallest E2E delay. The developed framework takes into account the effects from not only the physics of the hardware and the channel, e.g., transmission power, DA beamwidth, transmission distance and operating bandwidth in the case of line of sight propagation, but also the MAC layer protocol design, e.g., interference and re-transmission, and the queuing delay when transmitting at multi-Gbps rates with buffer/memory limitations. The framework is based on a quasi-birth-death (QBD) Markov process model, which captures the capabilities and the behavior of a node in THz communication networks. This QBD model consists of a series of levels, including boundary level and repeating levels. The repeating levels stand for the number of packets currently queued in one specific node. The boundary level represents the status that the node does not have any packet to transmit. Within one repeating level, the transitions between different states are related to the re-transmission attempts that the packet will undergo. The transitions between inter-levels are further related to the queuing and dequeue probabilities and also the transmission(TX)/reception(RX) mode switching probabilities. We formulate the QBD Markov process model with all aforementioned cross-layer effects and, by utilizing the framework, we derive the distribution of relays to achieve the smallest E2E delay and the highest E2E throughput in different scenarios.

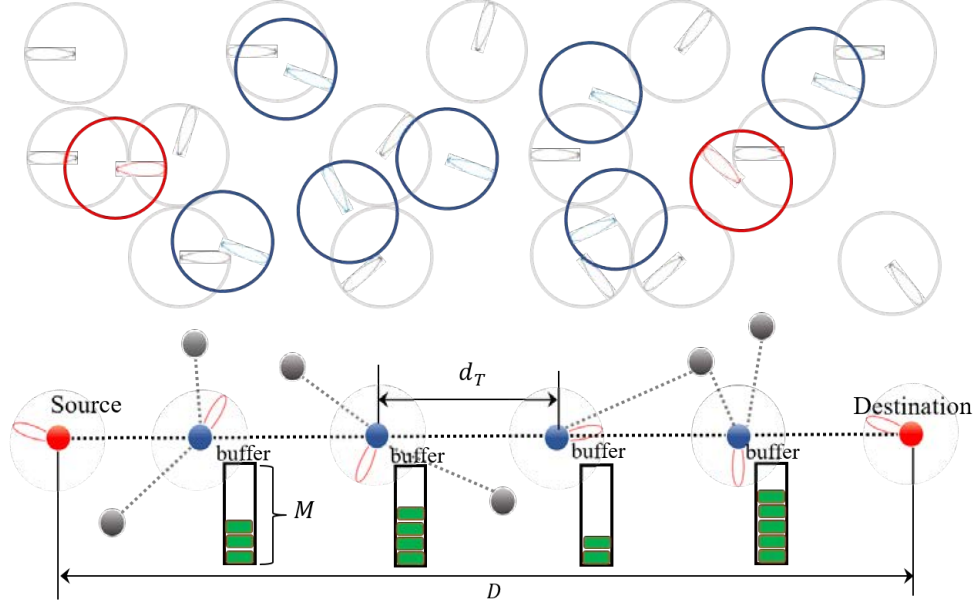


Fig. 63: Network architecture.

7.2 Methods, Assumptions and Procedures

7.2.1 System Preliminaries

In this section, we first describe the scenario under analysis as well as the network topology and operation. Second, we introduce the THz-band channel model and present the energy consumption, THz-band channel frequency response and the molecular absorption noise power as functions of the transmission distance. Then, we introduce the DA model utilized in our analysis. We derive the DA characteristics, which include antenna gain and antenna beamwidth, as functions of the transmission distance. Finally, we describe the MAC protocols considered in our analysis.

Network Topology and Operation We consider that all the nodes are randomly distributed in a flat area following a spatial Poisson distribution. Nodes operate using the receiver-initiated MAC protocol for directional THz networks that we proposed in [89], further extended in [111], and discussed in Sec. 6 which was shown to outperform the more conventional transmitter-initiated MAC protocols in terms of throughput, drop probability and E2E. Accordingly, each node operates in half-duplex mode and periodically switches between TX mode and RX mode. A node in TX mode, i.e., with data to transmit, checks whether a CTS frame from the intended receiver has been recently received. We consider that a CTS frame is valid for the duration that the transmitter and receiver are within LoS. The transmitter who has data to send to a specific receiver can only proceed with the data transmission using a DATA frame after receiving a valid CTS frame from that receiver, otherwise, the transmitter listens to the channel until the reception of a new CTS frame. A node in RX mode periodically broadcasts its status with CTS frames.

We illustrate the network topology in Fig. 63. The network is formed as an ad-hoc architecture, whose coverage range can be extended with the help of the relays. Relays nodes, which help to relay

Table 13: Parameters and Constants

Symbol	Quantity
D	Distance between source-destination pair
d_T	Single hop transmission distance
P_r	Received signal power
S_t	Single-side power spectral density of the transmitted signal
B	3dB bandwidth
f	Operation frequency
H_c	THz-band channel frequency response
c	Speed of light
k_{abs}	Molecular absorption coefficient
N_r	Molecular absorption noise power
S_{NB}	Background atmospheric noise power spectral density
S_{NI}	Self-induced noise power spectral density
SNR_{min}	Minimum SNR threshold
G_t and G_r	Gain of transmitter antenna and receiver antenna, respectively
G	Desired antenna gain by considering $G_t(d_T) = G_r(d_T) = G(d_T)$
Ω_A	Array solid beam angle of DAs
θ_h and ϕ_h	Half Power Beam Width of DAs in the elevation plane and azimuthal plane respectively
$\Delta\theta$	Beam width of the DAs
θ_{tx}^{start} and θ_{rx}^{start}	Initial angle of transmitter DA and receiver DA respectively
ω_{tx} and ω_{rx}	Turning speed of transmitter DA and receiver DA respectively

other nodes' packets in addition to generating their own packets, are marked as blue nodes and are uniformly distributed in the whole network to guarantee network connectivity for any node in any direction. The remaining nodes, which never relay other nodes' packets, are marked as grey nodes. Within these, we have selected two as the source and the destination for the multi-hop E2E link analysis, colored in red. The goal of this paper is to determine the relay density and transmission range needed to achieve the maximum E2E throughput for such a network with multi-hop links. A summary of the parameters in our analysis is provided in Table 13.

Channel Model The unique distance-dependent characteristic of THz communication is caused by the physics of the channel. The propagation of electromagnetic waves at THz-band frequencies is mainly affected by molecular absorption, which is transmission-distance dependent and frequency selective, as we discussed in Sec. 4 and leveraged in Sec. 5. In this work, we consider that nodes operate in the first absorption-defined window above 1 THz. The 3 dB window bandwidth varies with the transmission distance d_T , i.e., longer transmission distance results in a narrower 3 dB bandwidth window, see Fig. 64.

Molecular absorption results in both molecular absorption loss and molecular absorption noise, both of which, are transmission-distance-dependent parameters. In particular, the signal

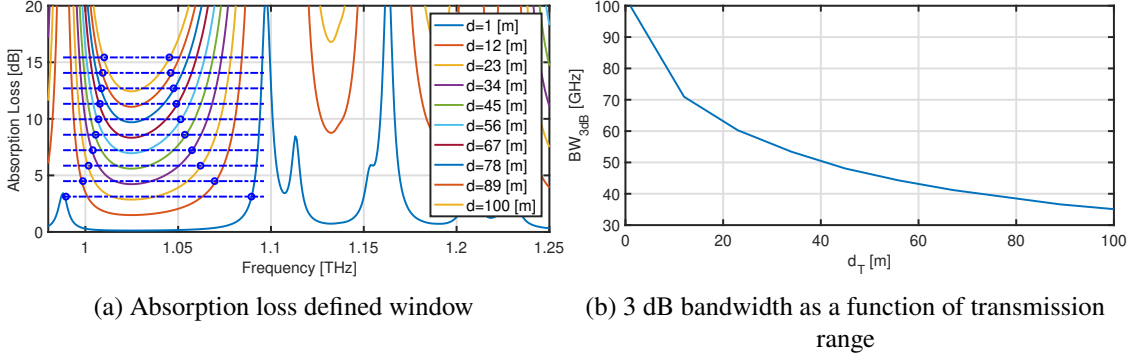


Fig. 64: Absorption loss defined 3 dB bandwidth.

power at a transmission distance d_T from the transmitter is given by:

$$P_r(d_T) = \int_{B(d_T)} S_t(f) |H_c(f, d_T)|^2 G_t(d_T) G_r(d_T) df, \quad (86)$$

where S_t is the single-sided power spectral density (p.s.d) of the transmitted signal, B stands for the 3 dB bandwidth in respect of different transmission range. f refers to the operation frequency and H_c refers to the THz-band channel frequency response, which is given by:

$$H_c(f, d_T) = \left(\frac{c}{4\pi f d_T} \right) e^{-\frac{k_{abs}(f)d_T}{2}}, \quad (87)$$

where c refers to the speed of light and k_{abs} is the molecular absorption coefficient of the medium, which is a frequency-dependent value and is calculated based on the high-resolution transmission molecular absorption (HITRAN) database. The value of k_{abs} depends on the molecular composition of the transmission medium, i.e., the type and concentration of molecules found in the channel. G_t and G_r refer to the gain of the transmitter antenna and receiver antenna, respectively, which in our study we consider can be adapted according to the transmission distance (see next section).

Similarly, the molecular absorption noise power N_r at a distance d_T from the transmitter, can be modeled as additive, Gaussian, colored and correlated to the transmitted signal [112], and is given by:

$$N_r(d_T) = \int_{B(d_T)} (S_{NB}(f) + S_{NI}(f, d_T)) |H_r(f)|^2 df, \quad (88)$$

where it is taken into account that the total molecular absorption noise is contributed by the background atmospheric noise p.s.d., S_{NB} and the self-induced noise p.s.d., S_{NI} , and are computed as described in [112].

Directional Antenna Model To overcome the very high molecular absorption loss over long distances and to increase the SNR, DAs are used at both the transmission and the reception with gains as G_t and G_r respectively. Such antennas can be realized either by means of directional antenna designs, such as horn antennas, or by utilizing beamforming antenna arrays [113].

The gains need to be calculated by taking into account the condition that the received signal strength should surpass the received signal power threshold, i.e.,

$$\int_{B(d_T)} S_t(f) \frac{c^2}{(4\pi d_T f)^2} e^{-k_{abs}(f)d_T} G_t(d_T) G_r(d_T) df \geq N_r(d_T) SNR_{min}, \quad (89)$$

where SNR_{min} stands for the minimum SNR threshold (10 dB in our analysis). Without loss of generality, the antenna gains of the transmitter and the receiver are considered to be identical and constant over the 3 dB frequency window, i.e., $G_t(d_T) = G_r(d_T) = G(d_T)$. In this case, the desired antenna gain can be expressed as:

$$G(d_T) \geq \sqrt{\frac{N_r(d_T) SNR_{min}}{\int_{B(d_T)} S_t(f) \frac{c^2}{(4\pi d_T f)^2} e^{-k_{abs}(f)d_T} df}}. \quad (90)$$

The directivity gain of highly DAs can be approximated as [94]:

$$G(d_T) \approx \frac{4\pi}{\Omega_A(d_T)} \approx \frac{4\pi}{\theta_h(d_T)\phi_h(d_T)}, \quad (91)$$

where Ω_A refers to the array solid beam angle, θ_h and ϕ_h are the Half Power Beam Width (HPBW) in the elevation plane and azimuthal plane, respectively. If we assume the HPBW in the elevation plane and azimuthal plane are identical, i.e., $\theta_h = \phi_h = \Delta\theta$, the beamwidth of the DA can be calculated as:

$$\Delta\theta(d_T) \leq \sqrt{4\pi \sqrt{\frac{\int_{B(d_T)} S_t(f) \frac{c^2}{(4\pi d_T f)^2} e^{-k_{abs}(f)d_T} df}{N_r(d_T) SNR_{min}}}}. \quad (92)$$

7.2.2 Overview of the Quasi Birth Death Markov Process

In this section, we provide an overview of the QBD Markov process for THz-band communication network model introduced in Sec. 7.2.1. In a THz communication network, each node is represented by a queuing model, which accounts for the packet arrivals and departures as well as the service process of each packet. The QBD Markov process, as a generalization of the birth-death process, is utilized to comprehensively model the service behavior and analyze the distribution of the E2E delay of each packet. On the one hand, the QBD Markov process moves up and down between levels one at a time to capture the enqueue or dequeue behavior introduced by the cross-layer nature of the network under analysis. On the other hand, the delay time of transitions between levels can have a more complicated distribution in the blocks of each level of the QBD Markov process. The block models the transmission/reception process of each packet taking into account the effects from the physics of the channel and the devices, namely, transmission power, communication distance, bandwidth and antenna beamwidth, or from the MAC layer protocol design, namely, interference and re-transmission.

The QBD Markov process model shown in Fig. 65 consists of two types of levels, including boundary level and repeating levels. The repeating levels are the levels that satisfy the condition $1 \leq m \leq M$. We denote the maximum buffer capacity of each node as M . The index m represents the number of packets currently queued in node. The boundary level, satisfying $m = 0$, represents

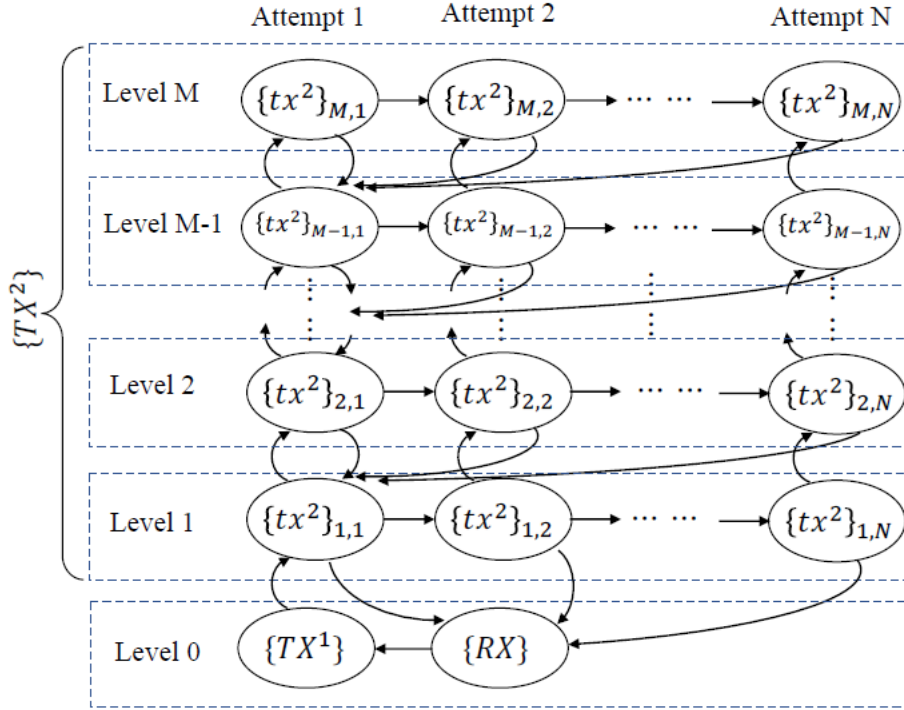


Fig. 65: QBD Markov process model of the system.

the status that the node does not have any packet to transmit. $\{RX\}$ represents all states in RX mode. $\{TX^1\}$ and $\{TX^2\}$ are two components of the TX mode. tx^2 is the basic component of $\{TX^2\}$. More details are introduced in Sec. 7.2.3.

Within one repeating level, the transitions between different states are related to the retransmission attempts that the DATA packet will undergo. The number of the maximum DATA packet re-transmissions is defined in the MAC protocol and denoted as N . The transitions between inter-levels further relate to the queuing and dequeue probabilities and also the TX/RX mode switching probabilities. In all the transitions within a level and cross the levels, the cross-layer peculiarities of the THz-band communication networks are taken into account, including the channel, the antenna, the buffer and the physical, link and network layers. The detailed transitions within a level and cross levels are described in Sec. 7.2.3 and Sec. 7.2.4.

7.2.3 Intra-level states and transition model based on THz MAC Protocol

In this section, we describe the states within a level of the QBD Markov process. They are defined by the receiver-initiated MAC protocol described in Sec. 7.2.1. Then, we incorporate the unique distance-dependent properties, the directional THz communication feature and the THz channel effects into the designed transitions between states inside one level.

Overview of the Intra-level States in the QBD Process In Fig. 66, the intra-level states of the QBD process model, consisting of different element states, are shown. The duration of each state (summarized in Table 14) differs from each other because of the different processing times associated

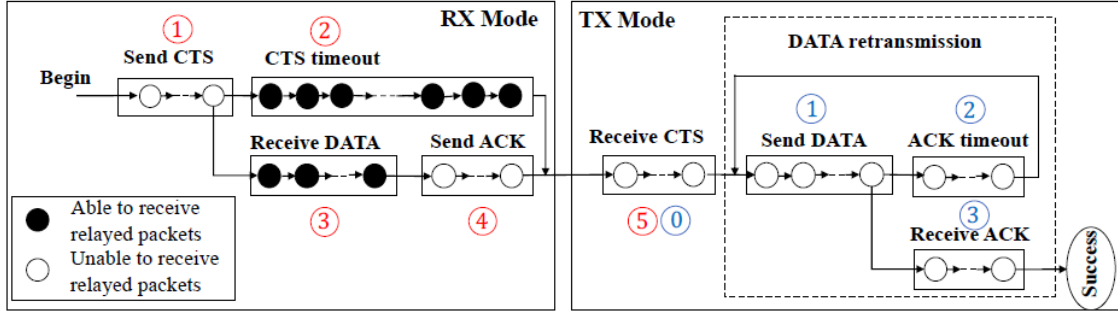


Fig. 66: MAC protocol.

Table 14: States Delay Duration

Notation	Parameter	Function
t_1^{rx}	Delay in "Send CTS"	$t_1^{rx} = T_{CTS} + T_{prop}$
t_2^{rx}	Delay in "CTS timeout"	$t_2^{rx} = N(T_{DATA} + T_{prop} + T_{ACK} + T_{prop})$
t_3^{rx}	Delay in "Receive DATA"	$t_3^{rx} = N(T_{DATA} + T_{prop}) + (N - 1)(T_{ACK} + T_{prop})$
t_4^{rx}	Delay in "Send ACK"	$t_4^{rx} = T_{ACK} + T_{prop}$
t_0^{rx}	Delay in "Receive CTS"	$t_0^{rx} = T_{proc}^{CTS}$
t_1^{tx}	Delay in "Send DATA"	$t_1^{tx} = T_{DATA} + T_{prop}$
t_2^{tx}	Delay in "ACK timeout"	$t_2^{tx} = T_{ACK} + T_{prop}$
t_3^{tx}	Delay in "Send ACK"	$t_3^{tx} = T_{ACK} + T_{prop}$

$T_{CTS}, T_{DATA}, T_{ACK}$: Transmission delay of a CTS frame, DATA frame and ACK frame, respectively.

T_{prop} : Propagation delay.

T_{proc}^{CTS} : Processing delay of CTS frame.

to different packet types. The local DATA packet generation rate λ_l of each node is determined by the demands of the application layer. In this paper, we consider that λ_l follows a Poisson distribution and affects all states of the Markov process. Because of the half-duplex operation, a node cannot transmit and receive packets simultaneously. Thus, when nodes are in TX mode, the majority of TX duration is designed to send DATA packets. Even though there are time slots for receiving CTS and ACK frames during TX mode, these time slots are too short to accomplish a complete DATA packet reception. In this case, we consider that nodes cannot receive other nodes' DATA packets when they operate in TX mode. In contrast, nodes can help receive and relay other nodes' packets in most states of RX mode, except "Send CTS" and "Send ACK" states. Thus, only those blackened block-represented-states are affected by the relayed packets. The effects of packets arrival rate to the Markov process will be analyzed later in this section. A summary of the parameters in this section is shown in Table 15.

The Markov process starts at "Send CTS" state in RX mode. After one sector duration T_{sector} of DA, the node switches from RX mode to TX mode. T_{sector} largely depends on N , which is the maximum number of DATA packet retransmission attempts during TX mode. When the buffer is

Table 15: Parameters and Constants of Intra-level Analysis

Symbol	Quantity
λ_l	Local DATA generation packet rate
R	Local DATA generation bit rate
λ_r	Relayed DATA packet generation rate
λ_g	Total DATA packet generation rate, $\lambda_g = \lambda_l + \lambda_r$
λ_n	Node density
T_{sector}	Sector duration of DA
N_{sector}	Number of sectors in one DA turning circle
N	Maximum number of DATA packet re-transmission attempts
M	Maximum buffer capacity of each node
$P_{i,j}^{rx}$ and $P_{i,j}^{tx}$	Probability that process change from i^{th} state to j^{th} state in RX mode and TX mode, respectively
$A_{coll}^{x,y}$	Interference area formed by the DA's coverage area
$n_I^{x,y}$	Expected nodes number within the interference area
π	State probability vector
P_{RX}	Transition probability matrix for the RX process
α_{RX}	Initial probability vector for RX process
P_{RX}^{suc} and P_{RX}^{fail}	Probability vector for RX process to complete the transmission process successfully and unsuccessfully, respectively
P_{tx^2}	Transition probability matrix of tx^2 process
α_{tx^2}	Initial probability vector for tx^2 process
$P_{tx^2}^{suc}$ and $P_{tx^2}^{fail}$	Probability vector for tx^2 process to complete the transmission process successfully and unsuccessfully, respectively
$P_{\lambda_{tx^2}}$	Packet arrival probability vector for tx^2 process
P_{tx}	Transition probability matrix for a single repeating level
α_{tx}	Initial probability vector of a single level of Markov Queue
P_{tx}^{suc} and P_{tx}^{fail}	Probability vectors to complete a single level of Markov Queue successfully and unsuccessfully respectively
$P_{\lambda_{tx}}$	Packet arrival probability vector for single repeating level states

full, the node starts to drop packets from the end of its buffer. A newly arrived DATA packet will push the buffer towards its capacity limitation, and the successful or the failed transmission of a DATA packet will release one DATA packet space from the buffer. Once there are no DATA packets left in the buffer, the node switches from TX mode to RX mode.

In Fig. 65, all states in RX mode are represented by $\{RX\}$. TX mode consists of $\{TX^1\}$ and $\{TX^2\}$, which stand for the “Receive CTS” state and the DATA retransmission states, respectively. The basic component of $\{TX^2\}$ is denoted as tx^2 , which represents the process of a DATA packet retransmission.

RX Process We consider that the $\{RX\}$ states and the $\{TX^1\}$ states together construct the boundary level of the QBD Markov process model. For the purpose of simplification, we denote the combination boundary level as **RX** process. As shown in Fig. 66, the **RX** process consists of 5 states including “Send CTS,” “CTS timeout,” “receive DATA,” “send ACK,” and “Receive CTS,” which are marked by an ascending order sequence from 1 to 5 for future notation. Based on the **RX** process illustrated in Fig. 66, the transition probability matrix is developed as:

$$P_{RX}(d_T) = \begin{bmatrix} 0 & P_{1,2}^{rx}(d_T) & P_{1,3}^{rx}(d_T) & 0 & 0 \\ 0 & 0 & 0 & 0 & P_{rx}^{tx}(d_T)P_{0,1}^{tx}(d_T) \\ 0 & 0 & 0 & P_{3,4}^{rx} & 0 \\ 0 & 0 & 0 & 0 & P_{rx}^{tx}(d_T)P_{0,1}^{tx}(d_T) \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (93)$$

where the format of $P_{i,j}^{rx}$ or $P_{i,j}^{tx}$ represents the probability that the **RX** process or the tx^2 process changes from i^{th} state to j^{th} state, respectively. These probabilities are derived as follows:

- $P_{1,3}^{rx}$: is the probability that the node sends a CTS frame successfully and receives at least one DATA packet successfully within N maximum retransmission limitation, which is calculated as:

$$P_{1,3}^{rx}(d_T) = \underbrace{\left(1 - \prod_{y=1}^{n_l^{x,y}(d_T)} \left(1 - P_{S_{tx}^{0,0}}^y\right)\right)}_{(94-a)} \underbrace{\frac{1}{(N_{sector}(d_T))^2}}_{(94-b)} \underbrace{P_{CTS}^{suc}(d_T)}_{(94-c)} \underbrace{P\{n_{DATA}^{S_{cts}} \in T_{CTS}^{life}(d_T)\}}_{(94-d)} \underbrace{\left(1 - (1 - P_{DATA}^{suc}(d_T))^N\right)}_{(94-e)}. \quad (94)$$

The component (94-a) is the probability that there is at least one transmitter node y currently able to establish one hop communication link with the node x , which is the receiver that just sent a CTS frame. To establish the link, y is currently in $S_{tx}^{0,0}$ state, which is the state of “Receive CTS”. The probability that y is at state $S_{tx}^{0,0}$ is denoted as $P_{S_{tx}^{0,0}}^y$, which can be found in the state vector π of

the Markov chain. $n_I^{x,y}$ represents the expected number of nodes in the interference area of one hop communication range and is calculated as:

$$n_I^{x,y}(d_T) = \sum_{i=1}^{\infty} i \frac{\left(\lambda_n A_{coll}^{x,y}(d_T)\right)^i}{i!} e^{-\lambda_n A_{coll}^{x,y}(d_T)}, \quad (95)$$

where λ_n is node density, $A_{coll}^{x,y}$ represents the interference area formed by the DA's coverage area of node x and node y , which can be calculated as:

$$A_{coll}^{x,y}(d_T) = 2d_T^2 \tan\left(\frac{\Delta\theta(d_T)}{2}\right). \quad (96)$$

The component (94-b) is the facing probability of node x and node y , where N_{sector} stands for the number of sectors in DA's complete turning circle, thus:

$$N_{sector}(d_T) = \frac{2\pi}{\Delta\theta(d_T)}. \quad (97)$$

The component (94-c) is the probability that node x sends CTS frame successfully, which is calculated as:

$$P_{CTS}^{suc}(d_T) = \left(1 - P_{CTS}^{coll}(d_T)\right) \left(1 - PER_{CTS}^{QPS}(d_T)\right), \quad (98)$$

where PER_{CTS}^{QPS} refers to the packet error rate of the CTS frame by using QPSK. P_{CTS}^{coll} refers to the collision probability of sending a CTS frame, which is calculated as:

$$P_{CTS}^{coll}(d_T) = \left(1 - \overbrace{P\{0 \in A_{coll}^{x,y}(d_T)\}}^{(99-a)}\right) \left(1 - \overbrace{P\{0 \in 2FT_{CTS}(d_T)\}}^{(99-b)}\right). \quad (99)$$

The component (99-a) represents the probability that no node is located in the DAs' interference area of node x and node y , thus:

$$P\{0 \in A_{coll}^{x,y}(d_T)\} = e^{-\lambda_n A_{coll}^{x,y}(d_T)}. \quad (100)$$

The component (99-b) represents the probability that there is no DATA packet generated during the CTS frame vulnerable time, which is calculated as:

$$P\{0 \in 2FT_{CTS}(d_T)\} = e^{-\lambda_g(d_T)2FT_{CTS}(d_T)}, \quad (101)$$

where λ_g is the total DATA packet generation rate consists of the local DATA packet generation rate λ_l and the relayed DATA packet generation rate λ_r . FT_{CTS} is the transmission time of a CTS frame.

The component (94-d) represents the probability that node y generates $n_{DATA}^{S_{cts}}$ DATA packets towards node x during the CTS life time, given that $T_{CTS}^{life} = T_{sector}$, thus:

$$\begin{aligned} P\{n_{DATA}^{S_{cts}} \in T_{CTS}^{life}(d_T)\} &= P\{n_{DATA}^{S_{cts}} \in T_{sector}(d_T)\} \\ &= \sum_{n_{DATA}^{S_{cts}}=0}^{\infty} \frac{\left(\frac{\lambda_l T_{sector}(d_T)}{N_{sector}(d_T)}\right)^{n_{DATA}^{S_{cts}}}}{n_{DATA}^{S_{cts}}!} e^{-\frac{\lambda_l T_{sector}(d_T)}{N_{sector}(d_T)}}, \end{aligned} \quad (102)$$

where the DATA packet generation rate towards the direction of node x is calculated as λ_l/N_{sector} .

The component (94-e) is the probability that at least one DATA packet is successfully received with N maximum retransmission limitation. P_{DATA}^{suc} is the probability that a DATA packet has been successfully transmitted and received, thus:

$$P_{DATA}^{suc}(d_T) = \left(1 - P_{DATA}^{coll}(d_T)\right) \left(1 - PER_{DATA}^{QPSK}(d_T)\right), \quad (103)$$

where P_{DATA}^{coll} and PER_{DATA}^{QPSK} refer to the collision probability and the packet error rate of DATA packets, respectively, which are calculated by using the same methodology as we applied in (98).

- $P_{1,2}^{rx}$: refers to the probability of CTS timeout caused by the failure of either CTS or DATA transmission, thus:

$$P_{1,2}^{rx}(d_T) = 1 - P_{1,3}^{rx}(d_T). \quad (104)$$

- $P_{3,4}^{rx}$: is the probability that after receiving DATA packet, the node replies an ACK frame. Since it is a strict rule of the MAC protocol, we have:

$$P_{3,4}^{rx} = 1. \quad (105)$$

- P_{rx}^{tx} : stands for the probability that node x transfers from RX mode to TX mode, which can be calculated as:

$$P_{rx}^{tx} = P_{1,2}^{rx}(d_T) + P_{1,3}^{rx}(d_T)P_{3,4}^{rx} = 1. \quad (106)$$

- $P_{0,1}^{tx}$: represents the probability of the node, now performs as a transmitter, received a CTS frame from a receiver node successfully and generates at least one DATA packet to the receiver node during the CTS frame life time. Since the node under study switches from RX mode to TX mode, we mark it as y , and the receiver node as x , thus:

$$P_{0,1}^{tx}(d_T) = \overbrace{\left(1 - \prod_{x=1}^{n_I^{x,y}(d_T)} \left(1 - P_{S_{rx}^{1,1}}^x\right)\right)}^{(107-a)} \frac{1}{(N_{sector}(d_T))^2} P_{CTS}^{suc} P\{n_{DATA}^{S_{cts}} \in T_{CTS}^{life}(d_T)\}, \quad (107)$$

where the component (107-a) is the probability that at least one node x , that able to establish one hop communication link with node y , and is currently in the states of “send CTS”, which is the 1st state, $S_{rx}^{1,1}$, in the RX mode. $P_{S_{rx}^{1,1}}^x$ is the probability of the receiver node in states $S_{rx}^{1,1}$, which is indicated in the state probability vector π of the Markov chain.

We consider that the initial probability vector of RX process is $\alpha_{RX} = [1, 0, 0, 0, 0]^T$, the probability vector for RX process to complete the transmission process successfully/unsuccessfully are $P_{RX}^{suc} = [0, 0, 0, 1, 1]$ and $P_{RX}^{fail} = [0, 1, 0, 0, 1]$, respectively.

tx^2 Process As shown in Fig. 65, the states in $\{tx^2\}$ construct the repeating levels of the QBD Markov process model, which we denote as tx^2 process. The tx^2 process represents a complete retransmission process, which consists of “Send DATA, ACK timeout” and “Receive ACK” states and marked by an ascending order sequence from 1 to 3 as shown in Fig. 66. In light of the communication process illustrated in Fig. 66, the transition probability matrix of tx^2 process is:

$$P_{tx^2}(d_T) = \begin{bmatrix} 0 & P_{1,2}^{tx}(d_T) & P_{1,3}^{tx}(d_T) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ where} \quad (108)$$

- $P_{1,3}^{tx}$ is the probability that node y successfully send one DATA packet and receive an ACK frame.

$$P_{1,3}^{tx}(d_T) = P_{DATA}^{suc}(d_T)P_{AC}^{suc}(d_T) \quad (109)$$

- $P_{1,2}^{tx}$ stands for the probability that node y failed to send DATA or to receive ACK frame.

$$P_{1,2}^{tx}(d_T) = 1 - P_{1,3}^{tx}(d_T). \quad (110)$$

We consider the initial probability vector of tx^2 process is $\alpha_{tx^2} = [0, 0, 0]^T$, the probability vector for tx^2 process to complete the transmission process successfully/unsuccessfully are $P_{tx^2}^{suc} = [0, 0, 1]$ and $P_{tx^2}^{fail} = [0, 1, 0]$, respectively. For the reason that the packet arrival probability vector for each state in tx^2 process affects the buffer occupancy rate and, thus, the performance of the Markov process. The packet arrival probability vector can be calculated as:

$$P_{\lambda_{tx^2}}(d_T) = \left[P_{\lambda_g}^{tx_1^2}(d_T), P_{\lambda_g}^{tx_2^2}(d_T), P_{\lambda_g}^{tx_3^2}(d_T) \right], \quad (111)$$

where the format $P_{\lambda_g}^{tx_k^2}$ is the probability that at least one DATA packet arrives the transmitter node y at the k^{th} state in tx^2 process, giving that the DATA packets arrival rate is λ_g , thus:

$$P_{\lambda_g}^{tx_k^2}(d_T) = 1 - P\{0 \in t_k^{tx}(d_T)\} = 1 - e^{-\lambda_g(d_T)t_k^{tx}(d_T)}, \quad (112)$$

where t_k^{tx} is the duration of the k^{th} state of the Markov process, which is described in Table 14.

Construction of Single Repeating Level of the QBD Markov Process In Fig. 65, each repeating level consists of N retransmission processes tx^2 . Each retransmission process is triggered by the failure of previous DATA packet transmission. Thus, the transition probability matrix for a single repeating level of the QBD Markov process is calculated as:

$$P_{tx}(d_T) = \begin{bmatrix} P_{tx^2}(d_T) & \alpha_{tx^2} P_{tx^2}^{fail} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & P_{tx^2}(d_T) & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & P_{tx^2}(d_T) & \alpha_{tx^2} P_{tx^2}^{fail} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & P_{tx^2}(d_T) \end{bmatrix}. \quad (113)$$

Table 16: Parameters and Constants of Inter-level Analysis

Symbol	Quantity
P_{tx}^{up}	Probability of the transition matrix of the retransmission process goes to the upper level
P_{tx}^{same}	Probability of the transition matrix of the retransmission process stays at the same level
P_{tx}^{drop}	Probability of the transition matrix of the retransmission process drops packets at the highest level
P_{tx}^{down}	Probability of the transition matrix of the retransmission process goes to the lower level
P_{rx}^{up}	Probability that the Markov process transmits from RX process to TX^2 process
P_{rx}^{same}	Probability of the transition matrix stays at the RX level
P_{rx}^{down}	Probability that the Markov process transmits from TX^2 to RX process

We further develop the initial probability vector of a single repeating level as $\alpha_{tx} = [\alpha_{tx^2}, \mathbf{0}, \dots, \mathbf{0}]^T$, the probability vectors to complete a single repeating level successfully/unsuccessfully are $P_{tx}^{suc} = [P_{tx^2}^{suc}, P_{tx^2}^{suc}, \dots, P_{tx^2}^{suc}]$ and $P_{tx}^{fail} = [P_{tx^2}^{fail}, P_{tx^2}^{fail}, \dots, P_{tx^2}^{fail}]$, respectively. Considering that once the system is in the n^{th} retransmission state ($1 < n \leq N$), the previously failed DATA packets must be retransmitted, regardless if there is any new arrived DATA packet in the queue. The possibility of having a packet ready to transmit is one. The packet arrival probability vector for single repeating level states is thus developed as $P_{\lambda_{tx}}(d_T) = [P_{\lambda_{tx^2}}(d_T), \mathbf{1}, \dots, \mathbf{1}]$.

7.2.4 Inter-level analysis of the QBD Markov process

In this section, the inter-level transitions of the QBD Markov process model are further analyzed. First, we conduct the states transition analysis according to the enqueue and dequeue process as well as the TX/RX mode switching. Then, we provide the solution of the QBD Markov process model. A summary of the parameters in our analysis is shown in Table 16.

States Transition with Buffer Capacity Limitation To analyze the transition process between levels in the QBD Markov process model with buffer capacity limitation, we adopt the similar methodology introduced in [114]. For any level m , satisfying $1 \leq m \leq M - 1$, the retransmission process transits to a higher level when a new packet arrives. The probability matrix of such transition is:

$$P_{tx}^{up}(d_T) = (\mathbf{1}P_{\lambda_{tx}}(d_T)) \otimes P_{tx}(d_T), \quad (114)$$

where $\mathbf{1}$ is a properly dimensioned matrix containing all 1s, $\otimes$ is the Hadamard product. $P_{\lambda_{tx}}$ and P_{tx} , as defined in Sec. 7.2.3, refer to the packet arrival probability vector and the transition probability matrix for a single repeating level, respectively.

When the node attempts to retransmit a DATA packet, the process stays at the same level m ,

with $1 \leq m \leq M - 1$. The transition matrix of the retransmission process is calculated as:

$$\begin{aligned} \overbrace{P_{tx}^{same}(d_T)}^{(115-a)} &= \overbrace{(1P_{\lambda_{tx}}(d_T)) \otimes \left(\alpha_{tx} (P_{tx}^{suc} + P_{tx}^{fail}) \right)}^{(115-b)} + \\ &\overbrace{(1 - 1P_{\lambda_{tx}}(d_T)) \otimes P_{tx}(d_T)}^{(115-b)}, \end{aligned} \quad (115)$$

where the first component (115-a) represents the probability that a new DATA packet arrives to the buffer at the same time when the recently transmitted DATA packet is removed from the buffer because of either successful or failed transmission. The second component (115-b) represents the probability that there is no DATA packet arrives to or been removed from the buffer.

At the highest level $m = M$, the queue is full. Any newly arrived DATA packet or recently transmitted DATA packet that meets the retransmission limitation is directly dropped, accordingly, we have:

$$P_{tx}^{drop}(d_T) = P_{tx}^{up}(d_T) + \alpha_{tx} (P_{tx}^{suc} + P_{tx}^{fail}). \quad (116)$$

For any level m , with $1 \leq m \leq M$. When no DATA packet arrives and the current DATA packet service is completed, the transmission process transits to a lower level. The transition matrix of this process is:

$$P_{tx}^{down}(d_T) = (1 - 1P_{\lambda_{tx}}(d_T)) \otimes \left(\alpha_{tx} (P_{tx}^{suc} + P_{tx}^{fail}) \right). \quad (117)$$

To complete the entire QBD Markov process, we extend our analysis to interrelate the boundary level with the repeating levels as follows. We consider that when a new DATA packet arrives, if RX mode time has not expired, the arrival of the new DATA packets will not affect the operation of RX mode. When RX mode time has expired, i.e., the system has experienced a complete process of RX mode, the Markov process transmits from RX process to tx^2 process, and thus:

$$P_{rx}^{up}(d_T) = \left(P_{RX}^{suc} + P_{RX}^{fail} \right)^T \alpha_{tx}^T. \quad (118)$$

When the node in tx^2 process completes the last DATA packet transmission, the buffer is empty, the Markov process transmits from tx^2 to RX process, thus, we have:

$$P_{rx}^{down} = \left(P_{tx}^{suc} + P_{tx}^{fail} \right)^T \alpha_{RX}^T. \quad (119)$$

If there is no arriving DATA packet, the transition process stay in RX level, thus:

$$P_{rx}^{same}(d_T) = P_{RX}(d_T). \quad (120)$$

With the above analysis, we develop our entire QBD Markov process model for the cross-layer

$$\begin{aligned}
& \begin{bmatrix} \pi_0(d_T), & \pi_1(d_T) \end{bmatrix} \begin{bmatrix} P_{rx}^{same}(d_T) - I & P_{rx}^{up}(d_T) & \mathbf{1} \\ P_{rx}^{down} & P_{tx}^{same}(d_T) - I + R(d_T)P_{tx}^{down}(d_T) & (I - R(d_T))^{-1} \mathbf{1} \end{bmatrix} \\
& = \begin{bmatrix} \mathbf{0}, & \mathbf{0}, & 1 \end{bmatrix}.
\end{aligned} \tag{125}$$

system as P_{sys} , shown in (121).

$$P_{sys}(d_T) = \begin{bmatrix} P_{rx}^{same}(d_T) & P_{rx}^{up}(d_T) & \mathbf{0} & \mathbf{0} & \dots & \dots & \mathbf{0} & \mathbf{0} \\ P_{rx}^{down} & P_{tx}^{same}(d_T) & P_{tx}^{up}(d_T) & \mathbf{0} & \dots & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & P_{tx}^{down}(d_T) & P_{tx}^{same}(d_T) & P_{tx}^{up}(d_T) & \dots & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & P_{tx}^{down}(d_T) & P_{tx}^{same}(d_T) & \dots & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & P_{tx}^{down}(d_T) & \ddots & & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \dots & P_{tx}^{same}(d_T) & P_{tx}^{up}(d_T) \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \dots & P_{tx}^{down}(d_T) & P_{tx}^{drop}(d_T) \end{bmatrix} \tag{121}$$

Solution of Quasi-Birth-Death Markov Model In this paper, we only consider the case with maximum buffer capacity $M \gg 2$. Based on this consideration, we summarize the conditions to calculate the rate matrix R and present them as follows:

$$\begin{cases} \pi_{m-1}(d_T)P_{tx}^{up}(d_T) + \pi_m(d_T)P_{tx}^{same}(d_T) + \pi_{m+1}P_{tx}^{down}(d_T)\pi_m(d_T) \\ \pi_m(d_T) = \pi_{m-1}(d_T)R(d_T) \end{cases} \tag{122}$$

Thus, the rate matrix R is calculated as:

$$R(d_T)_{(n)} = - \left(P_{tx}^{up}(d_T) + R(d_T)_{(n-1)}^2 P_{tx}^{down}(d_T) \right) (P_{tx}^{same}(d_T) - I)^{-1}. \tag{123}$$

We derive R iteratively by successive substitution making use of $R_{(1)} = \mathbf{0}$. The iteration is repeated until two consecutive iterates satisfy $\|R_{(n+1)} - R_{(n)}\| < \epsilon$, where ϵ is the predefined tolerance and $\|\cdot\|$ is an appropriate matrix norm.

By solving $\pi P_{sys} = \pi$ and $\pi \mathbf{1} = \sum_i \pi_i = 1$, we derive the equations as:

$$\begin{cases} \pi_0(d_T)P_{rx}^{same}(d_T) + \pi_1(d_T)P_{rx}^{down} = \pi_0(d_T) \\ \pi_0(d_T)P_{rx}^{up}(d_T) + \pi_1(d_T)P_{tx}^{same}(d_T) + \pi_2P_{tx}^{down}(d_T)\pi_1(d_T) \\ \pi_0(d_T)\mathbf{1} + \pi_1(d_T)(I - R(d_T))^{-1}\mathbf{1} = 1 \\ \pi_{M-1}(d_T)P_{tx}^{up}(d_T) + \pi_M(d_T)P_{tx}^{drop}(d_T) = \pi_M(d_T) \end{cases}, \tag{124}$$

which can be rearranged as (125).

Therefore, we can calculate π_0 and π_1 . Moreover, we solve π_m according to the second condition in (122), given that $2 \leq m \leq M - 1$. For the highest level M , we calculate π_M based on the last condition of (124). In light of the above illustrations, the entire probability vector of the QBD Markov process model for our cross-layer system can be derived and denoted as:

$$\pi(d_T) = [\pi_0(d_T), \pi_1(d_T), \dots, \pi_M(d_T)]. \tag{126}$$

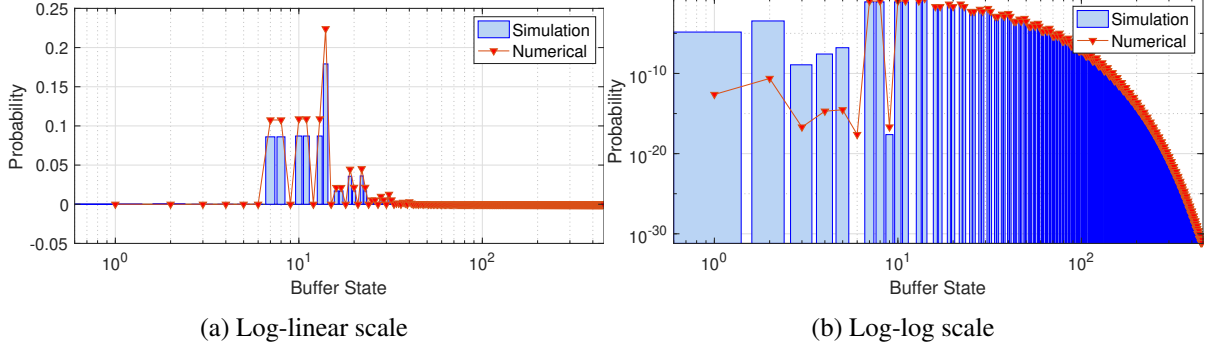


Fig. 67: Entire probability vector of the QBD Markov process model in (126) ($D = 100$ m, $R = 100$ Gbps, $M = 50$ packets, $N = 3$).

The steady-state probability vector is achieved by updating the queuing system with the latest result until the difference between the latest consecutive results is negligible.

7.3 Results and Discussion

In this section, first, we validate the analytical QBD Markov process model described in the previous sections by means of simulation. Then, we use this model to analyze the E2E performance of THz communication networks.

7.3.1 Validation of the QBD Markov process Model

We use MATLAB to simulate a multi-hop THz communication network which is composed of THz nodes. Unless otherwise stated, we define the network with the following system parameters: the THz nodes are randomly distributed in an interference area following a Poisson distribution and with node density λ_n as 1 node per m^2 . The distance D between the source node and the destination node is 100 m, the interference area is calculated as $\pi (\frac{D}{2})^2$. The local DATA packet generation rate λ_l is $\frac{R}{L_{\text{DATA}}}$ Packets per Second (PPS), where the DATA generation rate R is 100 Gbps. The maximum number of DATA packet re-transmissions N is set to 3, the maximum buffer capacity M of each node is set to 50 packets. The central frequency f_c is 1.03 THz, the transmission power P_t is 0 dBm, while the background noise N_r equals -90 dBm. QPSK is used at the physical layer. The control frame length is 17 bytes and the DATA packet length is 2,000 bytes.

To validate the QBD Markov process model described in sections 7.2.2, 7.2.3 and 7.2.4, we compare the normalized histogram of each state's probability evolution over time in the simulation, with the entire probability vector of the QBD Markov process model in (126), which is obtained from the proposed analytical model. We account only for the steady state of the network. Thus, the initial probabilities of all states at the beginning stage of the evolution are discarded. Besides, more than 200 evolution times are considered, which guarantees that the network achieved stability. The comparison results are shown in Fig. 67, from which, the simulation results and the numerical

results from the steady state analysis of the QBD Markov process model match appropriately. From Fig. 67a, it clearly indicates that the most vulnerable buffer levels that suffer ongoing DATA packet re-transmissions are the lower buffer levels. In light of these results, next, we analyze the E2E performance of THz communication networks by means of the proposed QBD Markov process model.

7.3.2 Performance Metrics

The packet delay of the entire QBD Markov process of any single node is denoted as:

$$\mathbf{T}_{delay}(\mathbf{d}_T) = [\mathbf{T}_0(\mathbf{d}_T), \mathbf{T}_1(\mathbf{d}_T), \dots, \mathbf{T}_M(\mathbf{d}_T)]^T, \quad (127)$$

where the format of $\mathbf{T}_m$ represents the delay duration in buffer level m , and is calculated by accumulating the duration that one DATA packet spent in all experienced states as partially indicated in Table 14.

The single hop throughput is calculated as:

$$S_{hop}(\mathbf{d}_T) = L_{DATA} / (\boldsymbol{\pi}(\mathbf{d}_T) \mathbf{T}_{delay}(\mathbf{d}_T)). \quad (128)$$

When multi-hop E2E communication link is considered to satisfy longer distance D between source-destination pair, the throughput of E2E communication link is calculated as:

$$S_{E2E}(\mathbf{d}_T) = \frac{L_{DATA}}{(\boldsymbol{\pi}(\mathbf{d}_T) \mathbf{T}_{delay}(\mathbf{d}_T)) E_{hop}(\mathbf{d}_T)}, \quad (129)$$

where $E_{hop}(\mathbf{d}_T) = \lceil D/d_T \rceil$ is the expected hop number between source and destination.

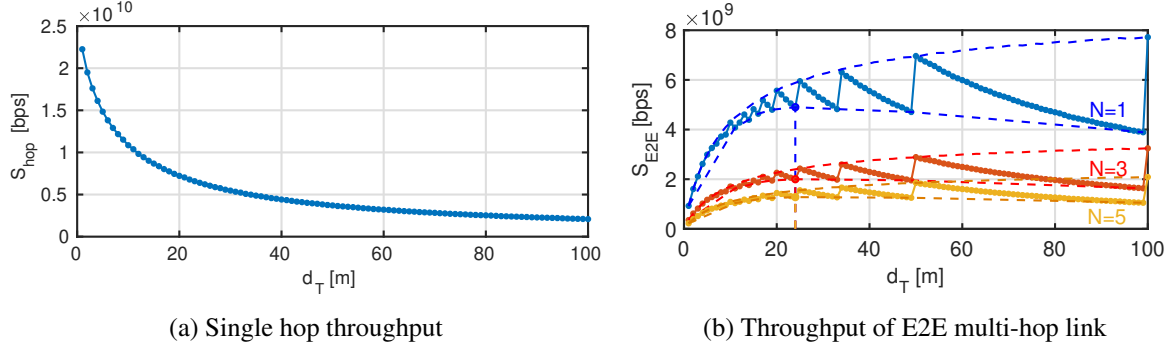


Fig. 68: Single hop throughput v.s. throughput of E2E multi-hop link
($D = 100$ m, $R = 100$ Gbps, $M = 500$ packets).

7.3.3 Numerical Analysis

Throughput Analysis As illustrated in Fig. 68a and 68b, we test the single hop throughput and the E2E multi-hop throughput for a fixed distance D between source and destination. The local packet generation rate R is set as an expected high rate to target the applications of THz communication, the buffer size of each node is 500 packets. As shown in Fig. 68a, the single hop throughput is a function of the maximum number of DATA packet re-transmission N . Small N effectively reduces the single-level delay of the QBD Markov process model. Fig. 68a also suggests that the maximum single hop throughput occurs in the shortest transmission distance case. This is because transmitted packets suffer a huge propagation delay from operating under long transmission distance conditions. However, as illustrated in Fig. 68b, the maximum throughput of the E2E multi-hop link shifts away from the shortest distance case. Since when nodes operate at longer transmission distances, reduced queueing delay for less hops compensates for the longer propagation delay in a single hop. The zigzag pattern of E2E multi-hop throughput is caused by the *ceil* function of E_{hop} . We curve-fit the E2E multi-hop throughput with the lower boundary to estimate the bottom line of the E2E throughput performance. For the parameters chosen in this section, the relaying distance to achieve the maximum E2E throughput is approximately 22 m, which would be well-aligned with backhaul-type applications.

Buffer Status Analysis As illustrated in Fig. 69, we inspect the buffer status of each node. We check both the buffer occupancy rate of different groups of buffer levels, which is classified every 100 packets buffer levels, and the expected delay distribution of a DATA packets in the corresponding buffer groups. With an increasing single-hop transmission distance d_T , longer propagation delay is triggered, which results in the growing of the expected packets delay in each buffer group. In Fig. 69a, we observe that without any DATA packet re-transmission ($N = 1$), the DATA packets only occupy the first buffer group, i.e., 0 packets $< m \leq 100$ packets. While with more DATA packet re-transmissions allowed ($N = 3, 5$), as shown in Fig. 69b, 69c, more time is needed for a node to finish the complete process of DATA packet transmission/re-transmissions. Thus, with the same DATA packet generation rate R , DATA packets are gradually pushed up towards the higher buffer

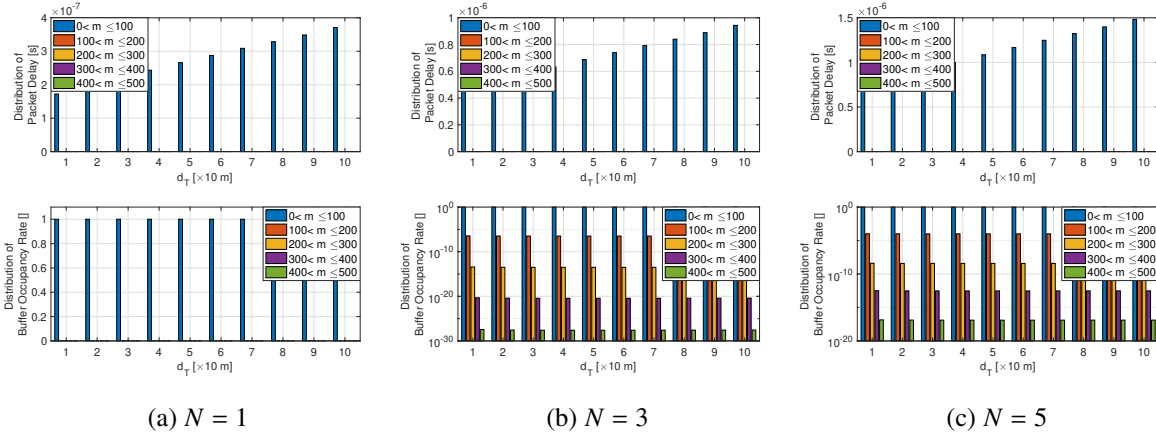


Fig. 69: Inspection of single node buffer status ($D = 100$ m, $R = 100$ Gbps, $M = 500$ packets).

levels. However, the majority of the DATA packets still stored in the first buffer group.

Performance Analysis with Varying Packet Generation Rate With an increasing local packet generation rate R , as shown in Fig. 70a, the maximum buffer occupancy rate grows almost at the same rate in all transmission distance cases. This is because, comparing with the removal rate of DATA packets, the packets arrival rate is not fast enough to drastically change the buffer occupancy rate distribution. Moreover, the probability vector of a DATA packet to achieve higher level states of our QBD Markov process model is too small to make a distinguished difference in the total delay. Thus, as indicated in Fig. 70b, with an increasing local packet generation rate, the throughput curve remains almost the same.

Performance Analysis with Varying Source-Destination Distance As shown in Fig. 71a, the maximum buffer occupancy rate is affected by the distance D between source and destination pair. More specifically, longer D requires more relays, thus, tremendous relayed packets from more neighbor nodes push into the buffer much faster than the packet removal rate, and thus, leads to a higher buffer occupancy rate. As indicated in Fig. 71b, with the increasing D , the E2E throughput drops gradually because of both the longer delay in each relay and the increased queuing delay for more hops. The maximum throughput shifts to the longer transmission distance d_T side, which suggests to apply longer d_T to each relay.

Performance Analysis for long distance transmission As shown in Fig. 72a, the distance D between the source and the destination is 1000 m. In the long communication distance case, since more hops are needed to establish the communication between the source and the destination, the buffer is easily saturated and the queuing delay increases. Thus, the throughput drastically drops compared with the short distance transmission cases. Also, longer single hop transmission distance d_T is always preferred to reduce the number of hops and to achieve higher E2E throughput. Other than that, the increased local packet generation rate does not change the throughput curve apparently, which is for the same reason as we analyzed in Fig. 70a. As a comparison, the throughput curve changes apparently with the variation of the distance D between the source and the destination,

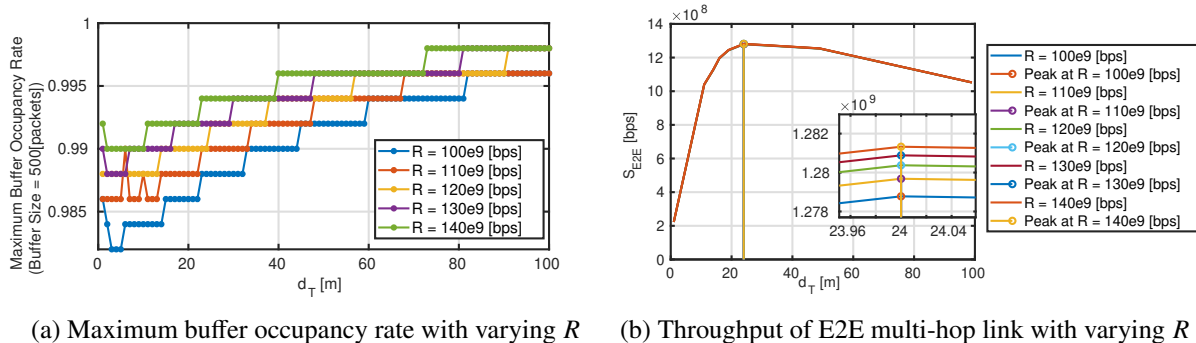


Fig. 70: The maximum buffer occupancy rate and the throughput of E2E multi-hop link with varying R ($D = 100$ m, $N = 5$, $M = 500$ packets).

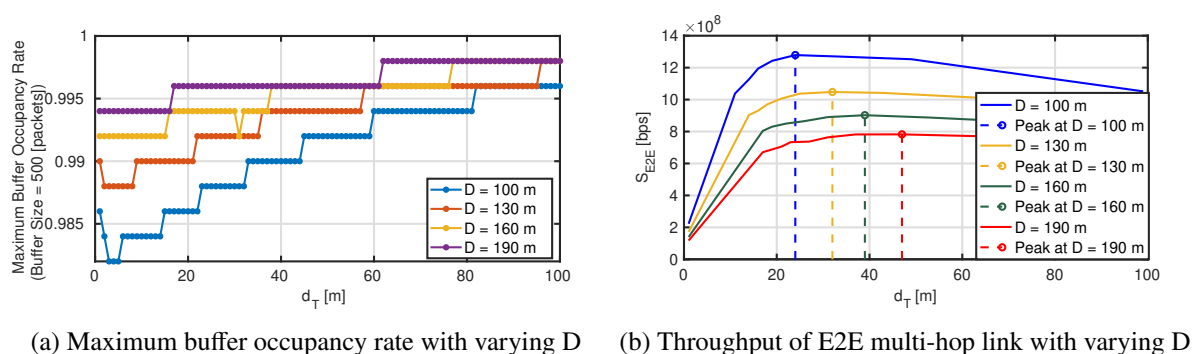
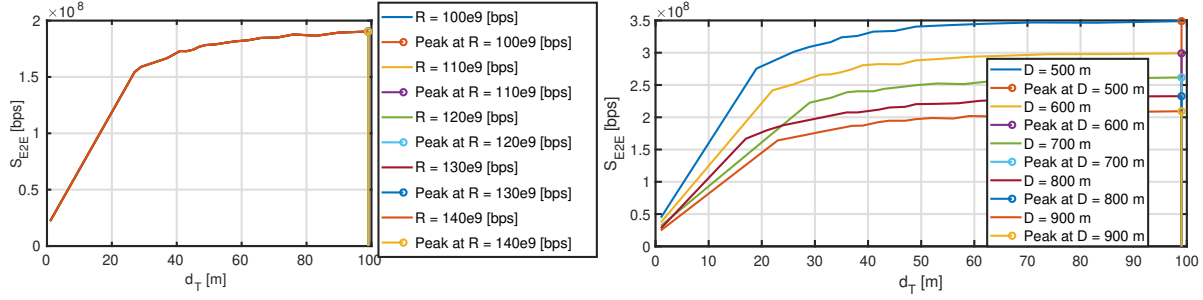


Fig. 71: The maximum buffer occupancy rate and the throughput of E2E multi-hop link with varying D ($R = 100$ Gbps, $N = 5$, $M = 500$ packets).



(a) Throughput of E2E multi-hop link with varying R (b) Throughput of E2E multi-hop link with varying D

Fig. 72: The throughput of E2E multi-hop link with varying R ($D = 1000$ m, $N = 5$, $M = 500$ packets) and with varying D ($R = 100$ Gbps, $N = 5$, $M = 500$ packets).

Table 17: Summary of Testing Parameters

Increasing Parameter	Effect on single node throughput	Effect on E2E throughput
Re-transmission Limitation (N)	Decreasing	Decreasing
Local DATA packet generation bit rate (R)	Slightly increasing (not fully buffered)	Slightly increasing (not fully buffered)
Single hop transmission distance (d_T [m])	Decreasing	Optimal exist (first increase then decrease)
E2E distance (D [m])	Decreasing	Decreasing

which is indicated in Fig. 72b. With the increasing D , the throughput drops gradually because of both the longer delay in each hop and the drastically increased queuing delay for more hops. The higher throughput is achievable with the longer single hop transmission distance d_T , which suggests to apply less number of hops between the source and the destination.

In the end, Table 17 summarized the effects on single hop throughput and the E2E throughput from independently increasing each key parameter of the studied scenario.

Next Steps This work sets a mathematical framework, which can lead to very different conclusions depending on the technology parameters being used. Therefore, these results need to be updated as terahertz technology evolves. In addition, this framework was done assuming that nodes were not moving. An extension to accommodate some or all the nodes to be mobile should be investigated. Moreover, given a relaying infrastructure, routing across it is not a trivial task. As we discussed in [115], the application of highly DAs introduces many challenges for multi-hop routing. Among others, the best routing path dynamically changes since the directional communication links are

periodically on and off, as determined by the DAs' current directions. Another challenge for routing protocol design comes from the limited memory or buffer size of THz devices, which is filled quickly when concurrent Tbps transmissions are handled. The buffer will be easily blocked by a locally stored packet that keeps waiting for the availability of the "best" route. This issue becomes even worse in directional networks, where such route may not be available shortly, and severely affects the network performance. These challenges require the adoption of new ideas, such as moving away from the store-and-forward approach and instead offloading packets as soon as they come, even if not along the optimal path.

8 CONCLUSIONS

In the last five years, the field of terahertz communications has dramatically evolved. Today, terahertz communications are no longer a fundamental research topic, but a key technology for the upcoming sixth-generation (6G) of cellular networks. Still, the majority of the work developed to this date is primarily focused on device technology development, propagation and channel modeling, and modulation design. This work is largely tailored to terrestrial networks on the ground and, with the exception of device development and channel modeling, remains largely theoretical.

In this project, we have paved the way toward airborne terahertz communications networks through pioneering theoretical and experimental contributions ranging from experimental platform development and channel modeling to physical, link and network layers design. The impact of the proposed solutions goes beyond the US Air Force, and includes for example upcoming non-terrestrial networks for commercial and scientific applications, for example. This project has also had non-traditional outcomes for an academic-driven effort, including a direct contribution to an active IEEE standard, and the generation of intellectual property that sets the foundations of a potential spin-out company.

9 REFERENCES

- [1] R. Piesiewicz *et al.*, “Short-range ultra-broadband terahertz communications: Concepts and perspectives,” *IEEE Antennas and Propagation Magazine*, vol. 49, no. 6, pp. 24–39, 2007.
- [2] J. Federici and L. Moeller, “Review of terahertz and subterahertz wireless communications,” *Journal of Applied Physics*, vol. 107, no. 11, p. 6, 2010.
- [3] H.-J. Song and T. Nagatsuma, “Present and future of terahertz communications,” *IEEE transactions on terahertz science and technology*, vol. 1, no. 1, pp. 256–263, 2011.
- [4] T. S. Rappaport *et al.*, “Wireless communications and applications above 100 ghz: Opportunities and challenges for 6g and beyond,” *IEEE access*, vol. 7, pp. 78 729–78 757, 2019.
- [5] I. F. Akyildiz, C. Han, Z. Hu, S. Nie, and J. M. Jornet, “Terahertz band communication: An old problem revisited and research directions for the next decade,” *IEEE Transactions on Communications*, vol. 70, no. 6, pp. 4250–4285, 2022.
- [6] K. Sengupta, T. Nagatsuma, and D. M. Mittleman, “Terahertz integrated electronic and hybrid electronic–photonic systems,” *Nature Electronics*, vol. 1, no. 12, pp. 622–635, 2018.
- [7] J. V. Siles, K. B. Cooper, C. Lee, R. H. Lin, G. Chattopadhyay, and I. Mehdi, “A new generation of room-temperature frequency-multiplied sources with up to 10× higher output power in the 160-ghz–1.6-thz range,” *IEEE Transactions on Terahertz Science and Technology*, vol. 8, no. 6, pp. 596–604, 2018.
- [8] W. R. Deal, K. Leong, A. Zamora, B. Gorospe, K. Nguyen, and X. B. Mei, “A 660 ghz up-converter for thz communications,” in *Compound Semiconductor Integrated Circuit Symposium (CSICS), 2017 IEEE*. IEEE, 2017, pp. 1–4.
- [9] M. Sung *et al.*, “Design considerations of photonic thz communications for 6g networks,” *IEEE Wireless Communications*, vol. 28, no. 5, pp. 185–191, 2021.
- [10] P. Lu *et al.*, “Mobile thz communications using photonic assisted beam steering leaky-wave antennas,” *Optics Express*, vol. 29, no. 14, pp. 21 629–21 638, 2021.
- [11] J. Crabb, X. Cantos-Roman, J. M. Jornet, and G. R. Aizin, “Hydrodynamic theory of the dyakonov-shur instability in graphene transistors,” *Physical Review B*, vol. 104, no. 15, p. 155440, 2021.
- [12] J. Crabb, X. Cantos-Roman, G. R. Aizin, and J. M. Jornet, “Amplitude and frequency modulation with an on-chip graphene-based plasmonic terahertz nanogenerator,” *IEEE Transactions on Nanotechnology*, vol. 21, pp. 539–546, 2022.
- [13] S. Abu-Surra *et al.*, “End-to-end 140 ghz wireless link demonstration with fully-digital beamformed system,” in *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 2021, pp. 1–6.

- [14] A. A. Farid, A. S. Ahmed, A. Dhananjay, P. Skrimponis, S. Rangan, and M. Rodwell, "135ghz cmos/lccc mimo receiver array tile modules," in *2021 IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS)*. IEEE, 2021, pp. 1–4.
- [15] P. Sen, D. A. Pados, S. N. Batalama, E. Einarsson, J. P. Bird, and J. M. Jornet, "The teranova platform: An integrated testbed for ultra-broadband wireless communications at true terahertz frequencies," *Computer Networks*, vol. 179, p. 107370, 2020.
- [16] K. Technologies, "Early insights into the 220 to 330 ghz sub-terahertz band," *White Paper*, 2022.
- [17] C. Han *et al.*, "Terahertz wireless channels: A holistic survey on measurement, modeling, and analysis," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1670–1707, 2022.
- [18] J. M. Jornet and I. F. Akyildiz, "Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the terahertz band," *IEEE Transactions on Wireless Communications*, vol. 10, no. 10, pp. 3211–3221, 2011.
- [19] C. Han, A. O. Bicen, and I. F. Akyildiz, "Multi-ray channel modeling and wideband characterization for wireless communications in the terahertz band," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2402–2412, 2014.
- [20] Y. Chen, Y. Li, C. Han, Z. Yu, and G. Wang, "Channel measurement and ray-tracing-statistical hybrid modeling for low-terahertz indoor communications," *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 8163–8176, 2021.
- [21] S. Ju and T. S. Rappaport, "Sub-terahertz spatial statistical mimo channel model for urban microcells at 142 ghz," in *2021 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2021, pp. 1–6.
- [22] N. A. Abbasi *et al.*, "Thz band channel measurements and statistical modeling for urban d2d environments," *IEEE Transactions on Wireless Communications*, vol. 22, no. 3, pp. 1466–1479, 2022.
- [23] S. Chakraborty, C. Parisi, D. Saha, and N. Thawdar, "A case for ofdm in ultra-broadband terahertz communication: An experimental approach," in *Proceedings of the 5th ACM Workshop on Millimeter-Wave and Terahertz Networks and Sensing Systems*, 2021, pp. 1–6.
- [24] Y. Wu, F. Lemic, C. Han, and Z. Chen, "Sensing integrated dft-spread ofdm waveform and deep learning-powered receiver design for terahertz integrated sensing and communication systems," *IEEE Transactions on Communications*, 2022.
- [25] Y. Bello, S. Barnola, D. Demmer, and J.-B. Doré, "Otf waveform with phase noise in sub-thz," in *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*. IEEE, 2022, pp. 1–5.
- [26] Y. Wu, C. Han, and Z. Chen, "Dft-spread orthogonal time frequency space system with superimposed pilots for terahertz integrated sensing and communication," *IEEE Transactions on Wireless Communications*, 2023.

- [27] C. Bosso, P. Sen, X. Cantos-Roman, C. Parisi, N. Thawdar, and J. M. Jornet, "Ultrabroadband spread spectrum techniques for secure dynamic spectrum sharing above 100 ghz between active and passive users," in *2021 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. IEEE, 2021, pp. 45–52.
- [28] P. Sen, H. Pandey, and J. M. Jornet, "Ultra-broadband chirp spread spectrum communication in the terahertz band," in *Next-Generation Spectroscopic Technologies XIII*, vol. 11390. SPIE, 2020, pp. 7–18.
- [29] I. F. Akyildiz and J. M. Jornet, "Realizing ultra-massive mimo (1024× 1024) communication in the (0.06–10) terahertz band," *Nano Communication Networks*, vol. 8, pp. 46–54, 2016.
- [30] B. Ning *et al.*, "Beamforming technologies for ultra-massive mimo in terahertz communications," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 614–658, 2023.
- [31] J. Hall, D. Bodet, and J. M. Jornet, "Experimental demonstration of multiple input multiple output communications above 100 ghz," in *IEEE INFOCOM 2022-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2022, pp. 1–2.
- [32] D. M. Bodet and J. M. Jornet, "Impact of antenna element directivity and reflection-interference on line-of-sight multiple input multiple output terahertz systems," in *2022 3rd URSI Atlantic and Asia Pacific Radio Science Meeting (AT-AP-RASC)*. IEEE, 2022, pp. 1–4.
- [33] P. Sen, J. V. Siles, N. Thawdar, and J. M. Jornet, "Multi-kilometre and multi-gigabit-per-second sub-terahertz communications for wireless backhaul applications," *Nature Electronics*, vol. 6, no. 2, pp. 164–175, 2023.
- [34] S. K. Moorthy and Z. Guan, "Beam learning in mmwave/thz-band drone networks under in-flight mobility uncertainties," *IEEE Transactions on Mobile Computing*, vol. 21, no. 6, pp. 1945–1957, 2020.
- [35] D. Mishra, A. M. Vegni, V. Loscr , and E. Natalizio, "Drone networking in the 6g era: A technology overview," *IEEE Communications Standards Magazine*, vol. 5, no. 4, pp. 88–95, 2021.
- [36] I. Mehdi, J. Siles, C. P. Chen, and J. M. Jornet, "Thz technology for space communications," in *2018 Asia-Pacific Microwave Conference (APMC)*. IEEE, 2018, pp. 76–78.
- [37] A. J. Alqaraghuli, J. V. Siles, and J. M. Jornet, "The road to high data rates in space: Terahertz vs. optical wireless communication," *IEEE Aerospace and Electronic Systems Magazine*, 2023.
- [38] IEEE, "Ieee standard for high data rate wireless multi-media networks–amendment 2: 100 gb/s wireless switched point-to-point physical layer," *IEEE Std 802.15.3d-2017 (Amendment to IEEE Std 802.15.3-2016 as amended by IEEE Std 802.15.3e-2017)*, pp. 1–55, 2017.
- [39] Z. Hossain, Q. Xia, and J. M. Jornet, "Terasim: An ns-3 extension to simulate terahertz-band communication networks," *Nano Communication Networks*, vol. 17, pp. 36–44, 2018.

- [40] X. Chen *et al.*, “All-electronic 100-ghz bandwidth digital-to-analog converter generating pam signals up to 190 gbaud,” *J. Lightwave Technol.*, vol. 35, no. 3, pp. 411–417, Feb 2017.
- [41] Keysight Technologies. M8199A arbitrary waveform generator. Accessed: March 30, 2023. [Online]. Available: <https://www.keysight.com/us/en/assets/3120-1465/data-sheets/M8199A-128-256-GSa-s-Arbitrary-Waveform-Generator.pdf>
- [42] Keysight Technologies. Infiniium UXR-series oscilloscopes. Accessed: March 30, 2023. [Online]. Available: <https://www.keysight.com/us/en/assets/7018-06242/data-sheets/5992-3132.pdf>
- [43] Xilinx. Zynq UltraScale+ RFSoc. [Online]. Available: www.xilinx.com/products/silicon-devices/soc/rfsoc.html
- [44] RFSoc product selection guide. [Online]. Available: www.xilinx.com/support/documentation/selection-guides/zynq-usp-rfsoc-product-selection-guide.pdf
- [45] V. Ariyaratna, A. Madanayake, and J. M. Jornet, “Real-time digital baseband system for ultra-broadband THz communication,” in *2020 45th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)*. IEEE, 2020, pp. 1–2.
- [46] P. Sen, V. Ariyaratna, A. Madanayake, and J. M. Jornet, “A versatile experimental testbed for ultrabroadband communication networks above 100 ghz,” *Computer Networks*, vol. 193, p. 108092, 2021.
- [47] Hitech Global. Htg-zrf8: Xilinx zynq ultrascale+ rfsoc development platform. [Online]. Available: [www.hitechglobal.com/Boards/Zynq\\$_RFSoc.htm](http://www.hitechglobal.com/Boards/Zynq$_RFSoc.htm)
- [48] J. Y. Suen, M. T. Fang, S. P. Denny, and P. M. Lubin, “Modeling of terabit geostationary terahertz satellite links from globally dry locations,” *IEEE Transactions on Terahertz Science and Technology*, vol. 5, no. 2, pp. 299–313, March 2015.
- [49] J. F. Federici, J. Ma, and L. Moeller, “Review of weather impact on outdoor terahertz wireless communication links,” *Nano Communication Networks*, vol. 10, pp. 13 – 26, 2016, terahertz Communications.
- [50] Y. Balal and Y. Pinhasi, “Atmospheric effects on millimeter and sub-millimeter (thz) satellite communication paths,” *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 40, pp. 219–230, March 2019.
- [51] S. Paine, “The *am* atmospheric model,” Smithsonian Astrophysical Observatory, Tech. Rep. 152, version 11.0, 2019.
- [52] A. Berk, P. Conforti, R. Kennett, T. Perkins, F. Hawes, and J. van den Bosch, “MODTRAN6: a major upgrade of the MODTRAN radiative transfer code,” in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XX*, M. Velez-Reyes and F. A. Kruse, Eds., vol. 9088, International Society for Optics and Photonics. SPIE, 2014, pp. 113 – 119.

- [53] J. R. Pardo, J. Cernicharo, and E. Serabyn, "Atmospheric transmission at microwaves (ATM): An improved model for millimeter/submillimeter applications," *IEEE Trans. Antennas Propag.*, vol. 49, no. 12, pp. 1683–1694, Dec. 2001.
- [54] ITU, "Attenuation by atmospheric gases and related effects," Recommendation ITU-R P.676-12, Tech. Rep., 08 2019.
- [55] ITU, "Radio noise," Draft revision of Recommendation ITU-R P.372-13, Tech. Rep., 04 2019.
- [56] J. Ma, J. Adelberg, R. Shrestha, L. Moeller, and D. Mittleman, "The effect of snow on a terahertz wireless data link," *Journal of infrared, millimeter and terahertz waves*, vol. 39, 03 2018.
- [57] E.-B. Moon, T.-I. Jeon, and D. R. Grischkowsky, "Long-Path THz-TDS Atmospheric Measurements Between Buildings," *IEEE Transactions on Terahertz Science and Technology*, vol. 5, no. 5, pp. 742–750, Sep. 2015.
- [58] *ITU-R (2005) Recommendation P.838-3, Specific attenuation model for rain for use in prediction methods*, International Telecommunication Union Radiocommunication Sector Std.
- [59] *ITU-R (2013) Recommendation P.840-6, Attenuation due to clouds and fog*, International Telecommunication Union Radiocommunication Sector Std.
- [60] J. Kokkonen, J. M. Jornet, V. Petrov, Y. Koucheryavy, and M. Juntti, "Channel modeling and performance analysis of airplane-satellite terahertz band communications," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 3, pp. 2047–2061, 2021.
- [61] J. M. Jornet, P. Sen, and V. Ariyaratna, "Ultra-broadband networking systems testbed at northeastern university," in *THz Communications*. Springer, 2022, pp. 473–476.
- [62] NASA, "U.s. standard atmosphere 1976," Tech. Rep., 1976.
- [63] "NASA EarthData GES DISC." [Online]. Available: disc.gsfc.nasa.gov/
- [64] S. D. Slobin, "Microwave noise temperature and attenuation of clouds: Statistics of these effects at various sites in the united states, alaska, and hawaii," *Radio Science*, vol. 17, no. 6, pp. 1443–1454, Nov. 1982.
- [65] D. Renaud and J. Federici, "Terahertz attenuation in snow and sleet," *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 40, no. 8, pp. 868–877, Aug. 2019.
- [66] F. Ulaby *et al.*, *Microwave Radar and Radiometric Remote Sensing*, 01 2014.
- [67] T. Oguchi, "Electromagnetic wave propagation and scattering in rain and other hydrometeors," *Proceedings of the IEEE*, vol. 71, no. 9, pp. 1029–1078, 1983.
- [68] J. H. Jiang and D. L. Wu, "Ice and water permittivities for millimeter and sub-millimeter remote sensing applications," *Atmospheric Science Letters*, vol. 5, no. 7, pp. 146–151, Oct. 2004. [Online]. Available: <https://onlinelibrary.wiley.com/doi/10.1002/asl.77>

- [69] K. L. S. Gunn and J. S. Marshall, "The Distribution with Size of Aggregate Snowflakes," *Journal of Atmospheric Sciences*, vol. 15, no. 5, pp. 452–461, Oct. 1958.
- [70] J. P. Nolan, *Univariate stable distributions*. Springer, 2020.
- [71] J. Hancock and R. Lucky, "Performance of combined amplitude and phase-modulated communication systems," *IRE Transactions on Communications Systems*, vol. 8, no. 4, pp. 232–237, 1960.
- [72] W. Weber, "Differential encoding for multiple amplitude and phase shift keying systems," *IEEE Transactions on Communications*, vol. 26, no. 3, pp. 385–391, 1978.
- [73] H. Rohling and V. Engels, "Differential amplitude phase shift keying (dapsk)-a new modulation method for dtvb," 1995.
- [74] X. Xiang and M. C. Valenti, "Closing the gap to the capacity of apsk: Constellation shaping and degree distributions," in *2013 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, 2013, pp. 691–695.
- [75] D. ETSI, "Document etsi en 302 307 v1. 2.1 (2009-08)," *Digital Video Broadcasting (DVB)*.
- [76] Z. Hossain and J. M. Jornet, "Hierarchical bandwidth modulation for ultra-broadband terahertz communications," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, 2019, pp. 1–7.
- [77] H. Wang *et al.*, "Power amplifiers performance survey 2000-present," *Georgia Tech Electronics and Micro-System Lab (GEMS)*, 2018.
- [78] K. Ramchandran, A. Ortega, K. M. Uz, and M. Vetterli, "Multiresolution broadcast for digital hdtv using joint source/channel coding," *IEEE Journal on Selected Areas in Communications*, vol. 11, no. 1, pp. 6–23, 1993.
- [79] K. Fazel and M. Ruf, "Combined multilevel coding and multiresolution modulation," in *Proceedings of ICC '93 - IEEE International Conference on Communications*, vol. 2, 1993, pp. 1081–1085 vol.2.
- [80] H. Jiang and P. A. Wilford, "A hierarchical modulation for upgrading digital broadcast systems," *IEEE transactions on broadcasting*, vol. 51, no. 2, pp. 223–229, 2005.
- [81] S. Wang, S. Kwon, and K. Y. Byung, "On enhancing hierarchical modulation," in *2008 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*. IEEE, 2008, pp. 1–6.
- [82] A. Riaz, S. Ghafoor, and R. Ahmad, "Integration of millimeter-wave and optical link for duplex transmission of hierarchically modulated signal over a single carrier and fiber for future 5g communication systems," *Telecommunication Systems*, vol. 72, no. 2, pp. 221–229, 2019.

- [83] S. Ju, Y. Xing, O. Kanhere, and T. S. Rappaport, "Millimeter wave and sub-terahertz spatial statistical channel model for an indoor office building," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 6, pp. 1561–1575, 2021.
- [84] S. Priebe and T. Kurner, "Stochastic modeling of thz indoor radio channels," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4445–4455, 2013.
- [85] S. S. Haykin, *Digital communication systems*. Wiley, 2014, vol. 1.
- [86] Z. Liu, Q. Xie, K. Peng, and Z. Yang, "Apsk constellation with gray mapping," *IEEE Communications Letters*, vol. 15, no. 12, pp. 1271–1273, 2011.
- [87] D. M. Bodet and J. M. Jornet, "Hierarchical bandwidth phase modulation for reduced peak-to-average power ratio in ultrabroadband terahertz communications," in *European Wireless 2022; 27th European Wireless Conference*. VDE, 2022, pp. 1–6.
- [88] Y. Ghasempour, C. R. da Silva, C. Cordeiro, and E. W. Knightly, "Ieee 802.11 ay: Next-generation 60 ghz communication for 100 gb/s wi-fi," *IEEE Communications Magazine*, vol. 55, no. 12, pp. 186–192, 2017.
- [89] Q. Xia, Z. Hossain, M. J. Medley, and J. M. Jornet, "A link-layer synchronization and medium access control protocol for terahertz-band communication networks," *IEEE Transactions on Mobile Computing*, 2019.
- [90] M. J. Marcus, "Harmful interference and its role in spectrum policy," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 265–269, March 2014.
- [91] FCC - Office of Engineering and Technology Policy and Rules Division, "FCC Online Table Of Frequency Allocations," 2020. [Online]. Available: transition.fcc.gov/oet/spectrum/table/fcctable.pdf
- [92] ITU, "Radio regulations," 2020. [Online]. Available: <http://handle.itu.int/11.1002/pub/814b0c44-en>
- [93] Ofcom, "Unlocking the potential of Terahertz radio spectrum - The role of spectrum management."
- [94] C. A. Balanis, *Antenna theory: analysis and design*. John wiley & sons, 2016.
- [95] Q. Xia and J. M. Jornet, "Expedited neighbor discovery in directional terahertz communication networks enhanced by antenna side-lobe information," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7804–7814, 2019.
- [96] NI, "Introduction to the NI mmWave Transceiver System Hardware," White Paper, <http://www.ni.com/product-documentation/53095/en/>, 2019.
- [97] 3GPP, "NR and NG-RAN Overall Description - Rel. 15," TS 38.300, 2018.
- [98] Mini-Circuits, "RC-1SPDT-A18 specifications," 2020. [Online]. Available: www.minicircuits.com/pdfs/RC-1SPDT-A18.pdf

- [99] O-RAN Working Group 1, “O-RAN Architecture Description - v2.00,” O-RAN.WG1.O-RAN-Architecture-Description-v02.00 Technical Specification, July 2020.
- [100] M. R. Schoeberl *et al.*, “Overview of the EOS Aura mission,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 5, pp. 1066–1074, May 2006.
- [101] World Meteorological Organization (WMO). (2011) Observing Systems Capability Analysis and Review Tool (OSCAR). [Online]. Available: space.oscar.wmo.int
- [102] J. W. Waters *et al.*, “The Earth observing system microwave limb sounder (EOS MLS) on the Aura Satellite,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 5, pp. 1075–1092, May 2006.
- [103] N2YO, “N2YO.com REST API v1,” 2020. [Online]. Available: www.n2yo.com/api/
- [104] Mini-Circuits, “RC-1SPDT-A18 APIs,” 2020. [Online]. Available: www.minicircuits.com/softwaredownload/Prog_Manual-2-Switch.pdf
- [105] D. C. Mazur, R. A. Entzminger, J. A. Kay, and P. A. Morell, “Time synchronization mechanisms for the industrial marketplace,” *IEEE Transactions on Industry Applications*, vol. 53, no. 1, pp. 39–46, Jan 2017.
- [106] ITU-R, “Performance and interference criteria for satellite passive remote sensing,” Recommendation ITU-R RS.2017-0, Aug 2012.
- [107] M. Polese *et al.*, “Coexistence and Spectrum Sharing Above 100 GHz,” *arXiv:2110.15187 [cs.NI]*, 2021. [Online]. Available: arxiv.org/abs/2110.15187
- [108] H. Sun, M. Naraghi-Pour, W. Sheng, and R. Zhang, “A hop-by-hop relay selection strategy in multi-hop cognitive relay networks,” *IEEE Access*, vol. 8, pp. 21 117–21 126, 2020.
- [109] A. Ikhlef, D. S. Michalopoulos, and R. Schober, “Max-max relay selection for relays with buffers,” *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 1124–1135, 2012.
- [110] C. Chang, C. Chang, T. Wang, and M. Li, “Throughput-enhanced relay placement mechanism in wimax 802.16j multihop relay networks,” *IEEE Systems Journal*, vol. 9, no. 3, pp. 728–742, 2015.
- [111] D. Morales and J. M. Jornet, “Adapt: An adaptive directional antenna protocol for medium access control in terahertz communication networks,” *Ad Hoc Networks*, vol. 119, p. 102540, 2021.
- [112] J. M. Jornet and I. F. Akyildiz, “Femtosecond-long pulse-based modulation for terahertz band communication in nanonetworks,” *IEEE Transactions on Communications*, vol. 62, no. 5, pp. 1742 – 1754, May 2014.
- [113] A. Singh, M. Andrello, N. Thawdar, and J. M. Jornet, “Design and operation of a graphene-based plasmonic nano-antenna array for communication in the terahertz band,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 9, pp. 2104–2117, 2020.

- [114] Y. Wang, M. C. Vuran, and S. Goddard, “Cross-layer analysis of the end-to-end delay distribution in wireless sensor networks,” *IEEE/ACM Transactions on networking*, vol. 20, no. 1, pp. 305–318, 2012.
- [115] Q. Xia and J. M. Jornet, “Routing protocol design for directional and buffer-limited terahertz communication networks,” in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*. IEEE, 2020, pp. 1–7.

10 LIST OF ACRONYMS

5G	Fifth Generation
6G	Sixth Generation
A2A	Airplane to Airplane
A2E	Airplane to Earth
A2G	Airplane to Ground
A2S	Airplane to Satellite
ACK	Acknowledgment
ADAPT	Adaptive Directional Antenna Protocol for Terahertz networks
ADC	Analog to Digital Converter
AFOSR	Air Force Office of Scientific Research
AFRL	Air Force Research Laboratory
AP	Access Point
API	Application Programming Interface
APSK	Amplitude and Phase Shift Keying
ASIC	Application-Specific Integrated Circuits
ATM	Atmospheric Transmission at Microwaves
AWG	Arbitrary Waveform Generator
AWGN	Additive White Gaussian Noise
AXI	Advanced eXtensible Interface
BB	Baseband
BER	Bit Error Rate
BNC	Berkeley Nucleonics
BPSK	Binary Phase Shift Keying
CA	Collision Avoidance
CDF	Cumulative Distribution Function
CF	Curve Fitting
CMOS	Complementary Metal-Oxide Semiconductor
CP	Cyclic Prefix
CSMA	Carrier Sense Multiple Access
CTA	Call to Action
CTS	Clear to Send
DA	Directional Antenna
DAC	Digital to Analog Converter
DC	Direct Current
DFT	Discrete Fourier Transform
DFT-S-OFDM	Discrete Fourier Transform-Spread Orthogonal Frequency-Division Multiplexing
DSO	Digital Storage Oscilloscope
DSP	Digital Signal Processing
DVB-S2	Digital Video Broadcasting Satellite Second Generation
E2A	Earth to Airplane
E2E	End-to-End
E2S	Earth to Satellite
EIRP	Equivalent Isotropic Radiated Power

ES	Symbol Energy
EVM	Error Vector Magnitude
EXP	Experimental
FCC	Federal Communications Commission
FEC	Forward Error Correction
FFT	Fast Fourier Transform
FPGA	Field-Programmable Gate Array
FSPL	Free-space Path Loss
GBPS	Gigabit-per-second
GEO	Geostationary Orbit
GMAO	Global Modeling and Assimilation Office
GOES	Geostationary Operational Environmental Satellite
GSPS	Giga-sample-per-second
HBM	Hierarchical Bandwidth Modulation
HM	Hierarchical Modulation
HPBM	Half Power Beam Width
HTTP	Hypertext Transfer Protocol
IEEE	Institute of Electrical and Electronics Engineers
IF	Intermediate Frequency
IP	Intellectual Property
IQ	In-phase and Quadrature
ISEC	Interdisciplinary Science and Engineering Complex
ISI	InterSymbol Interference
ITU	International Telecommunications Union
JPL	Jet Propulsion Laboratory
LB	Lower Band
LEO	Low Earth Orbit
LNA	Low-Noise Amplifier
LO	Local Oscillator
LOS	Line Of Sight
LPF	Low Pass Filter
LTF	Long Training Field
MAC	Medium Access Control
MCS	Modulation and Coding Scheme
MERRA	Modern-Era Retrospective Analysis for Research and Applications
MIMO	Multiple-Input Multiple-Output
MLE	Maximum Likelihood Estimation
MLS	Microwave Limb Scanner
MODTRAN	MODerate resolution atmospheric TRANsmission
MPC	Multi Path Components
MSPS	Megasample-per-second
NASA	National Aeronautics and Space Administration
NI	National Instruments
NLOS	Non-Line of Sight
NO	Noise Power Density
NOAA	National Oceanic and Atmospheric Administration

NR	New Radio
NSF	National Science Foundation
NTP	Network Time Protocol
NU	Northeastern University
OFDM	Orthogonal Frequency Division Modulation
OTA	Over-the-Air
OTFS	Orthogonal Time Frequency Space
PA	Power Amplifier
PAM	Pulse Amplitude Modulation
PAPR	Peak to Average Power Ratio
PC	Personal Computer
PCIE	Peripheral Component Interconnect Express
PD	Probability of Detection
PDF	Probability Distribution Function
PER	Packet Error Rate
PI	Probability of Intercept
PL	Programmable Logic
PLL	Phase-locked loop
PS	Processing System
PSG	Performance Signal Generator
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QBD	Quasi-Birth-Death
QPSK	Quadrature Phase Shift Keying
QSFP	Quad Small Form-Factor Pluggable
RAM	Random Access Memory
RF	Radio Frequency
RFSOC	Radio Frequency System on
Chip RMS	Root Mean Square
RMSE	Root Mean Square
Error RTS	Request to Send
RX	Receiver
S2A	Satellite to Plane
S2E	Satellite to Earth
S2S	Satellite to Satellite
SDR	Software-defined Radio
SER	Symbol Error Rate
SISO	Single Input Single Output
SMA	Sub-Miniature Version A
SMC	Sub-Miniature Version C
SNR	Signal-to-Noise Ratio
SoC	System-on-Chip
SPDT	Single-Pole Double Throw
STF	Short Training Field
STMR	Single Transmitter Multiple Receiver
SW	Software

TBPS	Terabit-per-second
TH	Theoretical
TX	Transmitter
UB	Upper Band
UE	User Equipment
US	United States
USAF	United States Air Force
VDI	Virginia Diodes Inc.
WMO	World Meteorological Organization
WLAN	Wireless Local Area Network
WPAN	Wireless Personal Area Network