



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

DEPARTMENT OF COMPUTER ENGINEERING AND
INFORMATICS

UNIVERSITY OF PATRAS

DOCTORAL DISSERTATION

ADVANCED SIGNAL PROCESSING TECHNIQUES FOR NEXT GENERATION MOBILE
NETWORKS IN THE MMWAVE SPECTRUM

MARIA TRIGKA

SUPERVISOR: KOSTAS BERBERIDIS, PROFESSOR

PATRAS - JULY 2022



Operational Programme
Human Resources Development,
Education and Lifelong Learning
Co-financed by Greece and the European Union



DEPARTMENT OF COMPUTER ENGINEERING AND
INFORMATICS

UNIVERSITY OF PATRAS

DOCTORAL DISSERTATION

ADVANCED SIGNAL PROCESSING TECHNIQUES FOR NEXT GENERATION MOBILE
NETWORKS IN THE MMWAVE SPECTRUM

MARIA TRIGKA

DISSERTATION COMMITTEE:

KOSTAS BERBERIDIS, PROFESSOR (SUPERVISOR)

CHRISTOS BOURAS, PROFESSOR (COMMITTEE MEMBER)

VASILEIOS PALIOURAS, PROFESSOR (COMMITTEE MEMBER)

DIMITRIOS AMPELIOTIS, ASSISTANT PROFESSOR (COMMITTEE MEMBER)

ARISTEIDIS LALOS, PRINCIPAL RESEARCHER (COMMITTEE MEMBER)

EVANGELOS VLACHOS, RESEARCH ASSISTANT (COMMITTEE MEMBER)

EMMANOUIL PSARAKIS, ASSOCIATE PROFESSOR (COMMITTEE MEMBER)

MARIA TRIGKA
JULY 2022

ABSTRACT

The ever-increasing need for wireless data transmission at a very high rate has resulted in the study of systems that can occur in the 5G and 6G mobile networks with frequencies up to over 60 GHz. A key feature of the millimeter frequency bands is that, due to the severe path loss, the channel is very sparse, not only in time but also in the frequency and angular space, and the received signals at the antennas are correlated. High attenuation and high sensitivity to blocking require architectures and protocols to address these challenges. Hence, massive hybrid antenna arrays have been proposed in the relevant literature for transmitting these signals. Their hybrid structure seems to be necessary since a purely digital structure has high implementation costs and power consumption in massive MIMO antenna arrays. This structure consists of an analogue and a digital part, which, in the transmitter, is called precoder and, in the receiver, combiner. The basic architectures for implementing the analogue part of the hybrid antenna systems are divided into fully or partially connected and sub-connected (the antennas are organized into non-overlapping groups), which can be implemented through either phase shifters or switches. Channel knowledge is required for the design of these structures. For this reason, this thesis focuses on channel estimation techniques to identify open issues and propose new approaches. The application of conventional channel estimation techniques is prohibitive due to the large number of transmitter-receiver antennas and required training symbols due to the high bandwidth and low SNR before transmitter-receiver alignment.

The channel estimation in the mmWave spectrum can be achieved either by scanning the angle space (e.g., through an exhaustive, hierarchical search) to find the best pair of beams between transmitter-receiver or by formulating a sparse recovery problem. The propagation environment is typically modeled by employing a geometric channel model that includes the path loss and angles of departure (AoD) and arrival (AoA) on the transmitter and receiver sides, respectively. The basic idea of parametric modeling is to estimate the involved angles and the corresponding attenuations rather than estimating the entire impulse response of the channel.

In array signal processing, various AoA estimation algorithms have been designed for classical ULAs and are of particular interest to researchers for several decades in,

e.g., radar and smart antennas. In the first part of the dissertation, assuming a hybrid architecture antenna array at the receiver, a novel preprocessing scheme is suggested which enables the full snapshot recovery as if we had a ULA, using measurements from the radio-frequency (RF) chains of the hybrid structure. Then, any classical high-resolution power spectrum technique can be employed to estimate the involved AoAs and thus the channel.

Also, a plethora of efficient sparse coding approaches have been suggested to directly estimate the geometric parameters by taking advantage of the sparsity properties of the mmWave channel to reduce the computational cost of the spatial scanning methods. Moreover, given that 5G networks will consist of small and dense cells, cooperative communication among proximal devices that share similar channel characteristics will be a promising technology. In wireless communications, sparse representations have been utilized to efficiently model and estimate a device channel with the help of sparse coding approaches, enhancing the spectrum and energy efficiency of fifth-generation (5G) and Internet of Things (IoT) networks. Hence, the second part of this thesis focuses on joint sparsity models for the representation of the channels of proximal devices and distributed sparse coding-based estimation methods to formulate a fully distributed voting-based (adaptive weights) cooperative approach suitable for networks of multiple tasks with partially overlapping parameters.

Keywords: *5G, 6G, MIMO, mmWave, Channel Estimation, MUSIC, H-MUSIC, Distributed Estimation, Sparse Representations/Coding, OMP.*

ACKNOWLEDGEMENT

This dissertation signifies the end of an academic journey as a PhD student. At this point, I would like to sincerely thank all those who have supported me all these years. First and foremost, with immense pleasure and deep sense of gratitude, I wish to express my sincere thanks to my supervisor Dr. Kostas Berberidis, Professor, University of Patras, without his motivation and the opportunity to undertake this thesis, the current research would not have been successfully completed.

I would also like to express my sincere thanks to Dr. Christos Mavrokefalidis for his kind support, cooperation and critical suggestions in several ways throughout my research work.

I wish to extend my profound sense of gratitude to my husband Elias Dritsas for his constant encouragement and moral support along with patience and understanding.

Last but not the least, I would like to thank my parents for their support during my studies.

This research was co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme 'Human Resources Development, Education and Lifelong Learning' in the context of the project 'Strengthening Human Resources Research Potential via Doctorate Research' (MIS-5000432), implemented by the State Scholarships Foundation (IKY). This work was also supported by the University of Patras.

Place: Patra

Date: 1/07/2022

Maria Trigka



TABLE OF CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENT	iii
LIST OF FIGURES	vii
LIST OF TABLES	ix
LIST OF ACRONYMS	x
1 Introduction	1
2 Mobile Networks-Evolution	8
2.1 1G Networks	8
2.2 2G Networks	8
2.3 3G Networks	9
2.4 4G Networks	10
2.5 5G Networks	11
2.6 6G Networks	13
3 Key enablers for 5G: mmWave and massive MIMO	15
3.1 The mmWave Spectrum	15
3.1.1 Types of Attenuation	15
3.1.1.1 Pathloss	15
3.1.1.2 Rain, Atmospheric and Foliage Attenuation	16
3.1.1.3 Penetration Loss	16
3.1.2 mmWave Modelling	17
3.2 Beamforming	18
3.3 Full Duplex Technology	18
3.4 MIMO Antenna Systems	19
3.4.1 General Characteristics	20

3.4.2	Massive MIMO Potentials in mmWave	21
3.5	MIMO Architectures	23
3.5.1	Analog and Digital	24
3.5.2	Hybrid	25
3.5.2.1	Fully-connected network	27
3.5.2.2	Sub-connected network	27
3.6	Hybrid MIMO Systems and Channel Estimation	28
4	mmWave Channel Estimation with Direction-of-Arrival Estimation Techniques for Hybrid Architectures	31
4.1	Introduction	31
4.2	Our Contribution	33
4.3	The System Model	33
4.3.1	The classical ULA case	35
4.3.2	MUSIC - A Conventional ULA DoA Estimation Method	35
4.3.3	The hybrid case	36
4.4	Recovering the Full Snapshot	38
4.4.1	Description of the procedure for recovering the full snapshot	38
4.4.2	Problem formulation for the RF combiner design	43
4.4.2.1	Constant Source Signals	44
4.4.2.2	Non-constant source signals	47
4.4.3	Hybrid MUSIC (H-MUSIC)	50
4.5	Results and Discussion	52
4.5.1	Noise Power in Reconstructed Snapshots	52
4.5.2	Application of Full Snapshot Reconstruction	54
4.6	Conclusions	57
5	mmWave Channel Estimation with Distributed Sparse Representation Techniques	58
5.1	Introduction	58
5.2	Motivation	62
5.3	The main contribution	63
5.4	The System Model	66

5.4.1	Single Antenna Devices	66
5.4.2	Multi-Antenna Devices	67
5.5	Joint Sparse Channel Modeling	69
5.5.1	Single Antenna Devices	70
5.5.2	Multi-Antenna Devices	71
5.6	Multi-task mmWave Channel Estimation	74
5.7	Voting-based Multi-task Distributed Solution	75
5.7.1	Proposed Solution	76
5.7.2	Adaptive Weights Voting Mechanism	78
5.7.2.1	Weight method based on support set as single entity	79
5.7.2.2	Weight method based on, per-index, support set similarities	80
5.8	Performance Evaluation	80
5.8.1	Required iterations before termination and Computational Complexity	81
5.8.2	Transmission Overhead	83
5.8.3	Impact of Support Set on Reconstruction Quality	85
5.8.4	Impact of Dictionary Resolution	88
5.9	Simulations	92
5.9.1	Multi-Antenna devices: simulation on real data	92
5.9.2	Single antenna devices: simulation on synthetic support sets	96
6	Conclusions and Future Works	101
	REFERENCES	103
	LIST OF PUBLICATIONS	115

LIST OF FIGURES

3.1	Beamforming [18]	19
3.2	Fully Analog and Fully Digital [40]	24
3.3	Hybrid Architectures in transmitter and receiver side [40]	25
3.4	Analog part with Phase shifters ((a) fully- and (b) sub-connected) and Lens antenna array [40]	27
3.5	Analog part with switches (a) fully- and (b) sub-connected antenna ar- ray in receiver side [40]	28
3.6	Analog part with phase shifters and switches in receiver side [62]	28
4.1	The receiver in the conventional case.	34
4.2	The receiver in the partially connected hybrid case.	36
4.3	The receiver in the fully connected hybrid case.	37
4.4	Total Noise Power Evaluation for a Combiner Matrix based on i) DFT matrix and ii) Random Phase Shifters for $\sigma_{nc}^2 = \{0, 0.01, 0.1\}$	53
4.5	RMSE of AoA estimation for the partially connected architecture, $M_c =$ $500, S_1 = 20, S_2 = 80, \lambda = 1cm$	55
4.6	RMSE of AoA estimation for the fully connected architecture, $M_c =$ $500, S_1 = 20, S_2 = 80, \lambda = 1cm$	55
4.7	RMSE of AoA estimation for the sub- and fully-connected architecture, $M_c = 500, SNR = 20dB, S_1 = 1 : 20, S_2 = 4 : 80$ snapshots ($S_2 = TS_1$ where $T = 4$), $\lambda = 1cm$ (sub- and fully-connected).	56
4.8	RMSE of AoA estimation of PREMUSIC for a combiner matrix based on i) DFT and ii) Random Phase Shifters, $M_c = 500, S_1 = 20, S_2 = 80,$ $\lambda = 1cm$	57
5.1	Downlink transmission scenario: $\mathcal{C}_1 = \{1, 2, 3, 4, 5, 6\}, \mathcal{C}_2 = \{5, 6, 7, 8, 9, 10\}$ [92].	70

5.2	Sparse Channel matrices of the sided devices $\bar{\mathbf{H}}_1^v, \bar{\mathbf{H}}_2^v, \bar{\mathbf{H}}_7^v, \bar{\mathbf{H}}_8^v$ and devices in the overlapping area $\bar{\mathbf{H}}_5^v, \bar{\mathbf{H}}_6^v$ assuming $N_r = 2$, $G_r = N_r$ and $G_t = 256$ where red, green and blue bins stand for the global and partially common atoms with total sparsity level $L_{AoD} = 6$	73
5.3	Local Information Sharing and Processing per device k	78
5.4	Transmission overhead in Kbits of SSOMP, (W)DiSOMP for $K = 10$, $T = [12 : 60]$, $L_{AoD} = 6$, $N_r = \{1, 2, 4\}$, $G_t = 512$, $q = 32$ bits	85
5.5	Average NMSE over all nodes for $T = [12 : 2 : 40]$, $SNR = 20$ dB . . .	94
5.6	NMSE of devices 1-5, devices 7-10 and device 6 for $T = [12 : 2 : 40]$, $SNR = 20$ dB	95
5.7	Average NMSE of all devices for $SNR = [0 : 5 : 30]$ dB, $T = 30$	96
5.8	Average NMSE of devices 1-5, 7-10 and device 6 for $SNR = [0 : 5 : 30]$ dB, $T = 30$	97
5.9	Average NMSE of all devices for $SNR = [0 : 5 : 30]$, $T = \{30, 60\}$. . .	98
5.10	(left) $L_g = 5$, $L_c = 0$, $SNR = \{10, 20\}$ dB, (right) Different SNR conditions per device. $L_g = 5$, $L_c = 0$	98
5.11	Two groups of devices facing global and common support sets. $L_g = 5$, $L_c = 3$, $SNR = 20$ dB.	100

LIST OF TABLES

2.1	Performance comparison of 4G, 5G and 6G networks [18]	14
5.1	Comparison with Previous Works	62
5.2	Transmission overhead in Kbits of centralized SSOMP for $K = 10$, $G_t = 512, T = 30, \mathcal{L}_k = K - 1, L_{AoD} = 6$	84
5.3	Transmission overhead in Kbits of (W)DiSOMP and centralized SSOMP for $K = 10, N_r = 2, G_t = \{256, 512, 1024\}, T = 30, \mathcal{L}_k^{out} = K - 1$, $L_{AoD} = 6$	84
5.4	Transmission overhead in Kbits of SSOMP, (W)DiSOMP for $K = 10$, $T = 30, L_{AoD} = [5 : 10], N_r = \{2, 4\}, G_t = 512, q = 32$ bits	85
5.5	Right term in (5.49) of the instantaneous channel error	92
5.6	Multitask device grouping according to AoDs sparsity patterns for $L_g =$ $2, L_c = 4$	93

LIST OF ACRONYMS

1G	First Generation.
2G	Second Generation.
3G	Third Generation.
4G	Fourth Generation.
5G	Fifth Generation.
6G	Sixth Generation.
AMPS	Advanced Mobile Phone System.
FM	Frequency Modulation
FDD	Frequency Division Duplexing
FDMA	Frequency Division Multiple Access.
OFDM	Orthogonal Frequency Division Multiplexing.
GSM	Global System for Mobile Communication.
EDGE	Enhanced Data GSM Evolution.
GPRS	General Packet Radio Service
HSDPA	High Speed Downlink Packet Access.
HSPA+	Evolved High Speed Packet Access.
HSUPA	High Speed Uplink Packet Access.
HSCSD	High Speed Circuit Switched Data.
ITU	International Telecommunication Union.
LTE-A Pro	Long Term Evolution Advanced Pro.
LTE-A	Long Term Evolution Advanced.
LTE	Long Term Evolution.
UMTS	Universal Mobile Telecommunications System.
CDMA	Code Division Multiple Access.

TDMA Time Division Multiple Access.

WCDMA Wideband Code Division Multiple Access.

AR Augmented Reality.

VR Virtual Reality.

SMS Short Message Service.

D2D Device to Device.

M2M Machine to Machine.

IoT Internet of Things.

Kbps kilobit per second.

SISO Single Input Single Output.

MISO Multiple Input Single Output.

MIMO Multiple Input Multiple Output.

SU-MIMO Single-User MIMO.

MU-MIMO Multi-User MIMO.

OFDM Orthogonal Frequency Division Multiplexing.

QoS Quality of Service.

LoS Line of Sight.

NLoS Non-Line-of-Sight

DoA Direction of Arrival.

AoA Angle of Arrival.

AoD Angle of Departure.

RF Radio Frequency.

mmWave millimeter Wave.

NMSE Normalized Mean Squared Error.

STNR Signal-to-Total-Noise-Ratio.

RMSE Root Mean Squared Error.

MUSIC Multiple Signal Classification.

ESPRIT Estimation of Signal Parameters via Rotational Invariant Techniques.

H-MUSIC Hybrid Multiple Signal Classification.

H-ESPRIT Hybrid Estimation of Signal Parameters via Rotational Invariant Techniques.

CS Compressive Sensing.

OMP Orthogonal Matching Pursuit.

SOMP Simultaneous Orthogonal Matching Pursuit.

DiOMP Distributed Orthogonal Matching Pursuit.

WDiOMP Weighted Distributed Orthogonal Matching Pursuit.

WDiSOMP Weighted Distributed Simultaneous Orthogonal Matching Pursuit.

CHAPTER 1

Introduction

As the need for mobile broadband capacity increases dramatically every year, wireless networks need to evolve to support better service for the ever-increasing volume of data traffic [78], [76], [70]. This requires researchers to look for new systems with greater capacity as well as methods for exploiting new areas of the wireless spectrum, in addition to those used in 4th generation (4G) systems and the developments in view of 5th generation (5G). Therefore, this requirement paves the way for the development and implementation of networks beyond the fifth generation (B5G). The use of millimeter frequency bands (30-300GHz or 10-1mm wavelength) [78], [76], gives the prospect of utilizing greater bandwidth and transmitting data at higher rates. However, existing approaches need to be reviewed and redesigned to meet the new requirements. It raises many open issues for study in order to make the networks operational beyond the fifth generation. These include the development of new techniques for assessing the status of the involved channels, the efficient use of dense antennas with a large number of multiple inputs and multiple outputs [96], [113] (Massive Multiple Input - Multiple Output, Massive - MIMO), as well as the possibility of simultaneous multi-user service, in combination with the oriented transmission methods of various formats, such as analogue, digital, hybrid.

The development of wireless technologies in the last decade has led to a rapid expansion in the use of mobile devices. Users also expect that all their devices will be able to be online simultaneously to access data and new services. Therefore, wireless service providers seek new technologies to meet the need for even greater data volume, higher quality services and lower latency. To achieve these goals, several techniques have been developed that make the best use of the limited available frequency range. However, it is now undeniable that the scope for spectacular improvement is very limited.

Thus, a wide-ranging debate has recently begun on the need to extend to frequencies far beyond those currently in use. Internationally, there is a belief that the millimeter frequency spectrum will play an important role in the 5G and B5G cellular systems [76]. In combination with the lack of spectrum, [70], it is of great scientific and technological interest to study it. Some of its characteristic advantages are the following. First, the use of frequencies in the range of 30-300GHz allows for larger bandwidth channels compared to current 4G systems and the planned 5G [78], [76]. Thus, the available frequency spectrum can be up to 200 times larger than the total spectrum used in 4G systems. There is a great interest in the researchers, and for this reason, relevant experiments have already been carried out to test, in principle, the feasibility of using these frequency bands. These experiments were conducted in densely populated urban and suburban areas to confirm the propagation characteristics of these frequencies. The signals experience significant path loss [30], as well as attenuation either due to adverse weather conditions or because they are blocked by buildings, or by the human body moving between transmitter and receiver [78], [76], [102].

In a wireless telecommunication network, data is transmitted via electromagnetic waves with a bandwidth depending on the wireless channel and the desired transmission rate. Its basic parts are separated into layers to ensure a more efficient design and implementation of a telecommunication system. In this research study, we are interested in the so-called physical layer, in which the way the transmitter and receiver are connected plays an important role. The transmitted signal, as it travels from the transmitter to the receiver through the air, is subject to a rate of attenuation influenced by the channel characteristics. Many factors affect the transmission of a signal. A key factor is the path loss [88] which is associated with the gradual attenuation of electromagnetic waves as they propagate through space. The phenomenon of attenuation is further affected by the terrain: hills, vegetation, buildings, etc., which is called shading. In addition, weather conditions (rain, snow) [78], [76], [102] as well as building materials and surfaces play an important role. The frequencies in the mmWave spectrum will meet the users' demand for a larger volume of data. However, experiments have proved its high sensitivity to the above factors limiting the transmission range.

Hence, it has been raised the need for advanced signal processing techniques [40], new antenna architectures and transmission methods to address the attenuation that sig-

nals at mmWave confront, as well as intermittent operation through these channels. Systems with dense, high-gain multiple antennas and high-directional transmissions are required to meet these challenges. Specifically, massive MIMO antenna systems of various arrays (e.g., linear, spherical, distributed) [113] is a very promising solution that appeared in the literature, either in the transmitter or in the receiver. It can serve many users simultaneously and improve the bandwidth efficiency through the high gain of multiplexing it offers, as well as the reliability of the transmission due to the many degrees of freedom, reducing the interference between the users due to the small transmission beam. The short wavelengths of these frequencies will allow the use of multiple antennas in a limited space and, in combination with beamforming techniques [81], will direct the beam energy in a specific direction to mitigate the path loss [14], [88]. The beamforming techniques are analogue, digital, and hybrid and implemented through precoding methods.

This doctoral dissertation focuses on the physical layer and suggests new estimation techniques for the mmWave channel based on power spectrum methods, sparse representation/coding theory and distributed optimization. It is expected that the proposed study will have a significant impact and potential contribution to the achievement of high-speed wireless communication between devices, whose interaction is critical, given that we are in the age of the IoT.

The core subject of this dissertation is the channel estimation for MIMO antenna systems operating in the mmWave frequency spectrum, the so-called 5th generation and beyond mobile networks (5G and B5G). The research is divided into two parts. The main research topics and the basic techniques studied and developed in each part of the doctoral dissertation include i) Multiple Antenna Systems of Hybrid Architecture and Power Spectrum Estimation Techniques for finding the Angles/Direction of Arrival (DoA or AoA) and ii) Sparse Coding techniques with predefined dictionaries combined with distributed adaptive methods for the estimation of the sparse channel representation.

The key research results of the dissertation are summarized as follows:

- The problem of DoAs (or AoAs) estimation on hybrid antenna arrays is studied using power spectrum techniques. More specifically, an efficient pre-processing scheme, called PRE+MUSIC, is proposed to retrieve the received signal as if a

conventional (non-hybrid) antenna array is used.

- The efficiency of the proposed scheme is compared with the conventional MUSIC algorithm and its hybrid version (H-MUSIC).
- The DoA estimation problem is studied for two types of architectures (sub-connected and fully-connected) of hybrid antennas on the receiver side, where the hybrid architecture exploits in a new way all the antenna elements of a uniform linear arrangement.
- The pre-processing scheme has been appropriately extended to be applicable in cases of non-constant signal sources (which occur when, for example, there is amplitude variation within a symbol period).
- Also, the estimation problem is tackled in the angular domain taking advantage of the fact that the channel matrices of neighboring users show a common sparse structure due to the common scatterers on the side of the Base Station in a multi-user scenario.
- The involved users collaboratively estimate their channel support sets before individually estimating channel complex values, assuming that there are global and common support subsets. We initially considered single-antenna users, and the performance of the proposed WDiOMP algorithm was compared with the already established method DiOMP, the local OMP, and a centralized solution based on SOMP in terms of support set recovery error in various experimental scenarios. The distributed schemes, WDiOMP and DiOMP, are the first to be applied to a channel estimation problem. The effectiveness of WDiOMP is verified in different scenarios, considering that the support set structure is unknown.
- The WDiOMP algorithm has been extended to be functional for MIMO systems. Its main feature is that each user locally exploits the measurements from all its antennas while performing SOMP from which the distributed WDiSOMP version emerged. Respectively, DiOMP was extended to DiSOMP, and centralized SOMP (SSOMP) was extended to simultaneously utilize the multi-antenna measurements of multiple users knowing the channel sparsity structure.

- It has been proposed an adaptive weights scheme where the desired weights can be estimated in two ways i) at the support set level as a single entity and ii) per index. These weights are involved in the majority voting part of the WDi(S)OMP, through which the joint support set is estimated from each device locally. The second scheme was proposed to increase the efficiency of channel evaluation in cases where there are users in the network whose support set is partially common with different user groups, known as multi-tasking.
- The transmission overhead (or communication cost) of the distributed and centralized methods for the channel estimation was captured as a function of the problem parameters (i.e., the number of devices, training symbols, receiving antennas, sparsity level, dictionary atoms) and was evaluated under different parameter settings. Finally, the efficiency of the methods was validated by exploiting the oracle LS estimator, which helped us determine the lower performance bound.

The structure of the dissertation is as follows. In Chapter 2, we present the evolution of mobile networks, from 1G to 5G and 6G. Chapter 3 presents the key enablers of 5G, focusing on the mmWave spectrum, the beamforming and full-duplex technologies, the MIMO antenna systems, and the respective architectures (analogue, digital, hybrid) and their structural elements. Also, the main parts of a hybrid MIMO system and the mmWave channel model are described. Chapter 4 treats the mmWave channel estimation as a DoA estimation problem assuming a hybrid architecture receiver. In the following, Chapter 5 tackles the channel estimation problem via a cooperative distributed approach whose aspects will be analytically presented in this chapter. Finally, Chapter 6 summarizes the conclusions and future directions of the dissertation.

The following notations are used in this dissertation. Uppercase bold letters are matrices, lowercase bold letters are vectors, letters with a hat are estimations, $(\cdot)^T$ denotes the transposition, $(\cdot)^H$ denotes the complex conjugate transposition, $E[\cdot]$ means statistical expectation, $\text{Tr}\{\cdot\}$ denotes the trace of a matrix and $\mathbf{I}$ is the identity matrix, $\lambda_{\min}(\cdot)$, $\lambda_{\max}(\cdot)$ are the minimum and maximum eigenvalue of a matrix; $\text{supp}(\mathbf{h})$, $\text{supp}(\mathbf{H})$ are the index set of the non-zero entries of vector $\mathbf{h}$ and the non-zero entries columns of matrix $\mathbf{H}$, respectively; $\mathbf{A}_{\mathcal{S}}$ denotes the sub-matrix formed by collecting the columns of $\mathbf{A}$ whose indexes are in set $\mathcal{S}$; $\|\mathbf{A}\|_F$, $\|\mathbf{A}\|$ and $\|\mathbf{a}\|$ denote the Frobenius norm, spectrum norm of $\mathbf{A}$ and Euclidean norm of vector $\mathbf{a}$ respectively.

CHAPTER 2

Mobile Networks-Evolution

Up to date, four generations of cellular communication systems have been adopted, while we are going through the era of fifth-generation systems [78].

2.1 1G Networks

First generation mobile networks were deployed in the 1980s. Their operation was based on analogue technology for voice-only services with data rates up to 2.4Kbps. Specifically, it was used FM (Frequency Modulation), FDMA (Frequency Division Multiple Access) and FDD (Frequency Division Duplex enabling two-way communication). 1G systems included Advanced Mobile Phone Systems (AMPS), Total Access Communication System (TACS), and Nordic Communication System (NMTS) [18]. A limitation of the 1G systems was that they did not provide secure transmission, due to the application of analogue technologies.

2.2 2G Networks

In 1990, 2G cellular networks were first launched to the Global System for Mobile Communications (GSM) standard and supported, along with voice services, Short Message Service (SMS) and basic email service. 2G systems have been established on multi-access digital technology and make use of TDMA (time division multiple access) and CDMA (code division multiple access). TDMA allows many users to share the same frequency channel by dividing the signal into different time intervals. With CDMA, several transmitters can send information simultaneously through a communication channel. Hence, multiple users can share a frequency band either by code division (CDMA) or by division of time (TDMA). Standards which were developed for the 2G networks were the IS-136 (TDMA), IS-95 (CDMA) [17], e.t.c.

The 2G networks provide wireless data transmission and voice speed thanks to the highest quality spectrum and advanced roaming services. It is characterized by three times higher capacity and better voice quality than the 1G. In addition, it offers higher security for communication thanks to encrypted information.

The advances in digital technology paved the way to the 2.5G systems with data rates up to 384 Kbps. 2.5G systems included General Packet Radio Service (GPRS), Enhanced Data GSM Evolution (EDGE). 2.5G systems use the switching packet technique and is the primary difference between GPRS and GSM, which is a circuit-switching system, thus achieving higher rates for data transfer. The 2.5G CDMA standard is the IS-95B, while the TDMA standards for improving cellular systems are the HSCSD (High Speed Circuit Switched Data), GPRS and EDGE (Enhanced Data rates for GSM Evolution also termed as Enhanced GPRS). The GPRS which uses switching protocols resulting in the freeing of network resources only when needed to send data. In the HSCSD standard, the user can operate consecutive slots for data transfer connection. Its basic disadvantage is the waste of network resources due to the use of switching circuits [18].

2.3 3G Networks

The 3G mobile networks are wireless communications systems with voice and data capabilities. Unlike systems in previous generations, it is the first international standardized system to come out on the market by International Telecommunication Union (ITU). 3G takes advantage of Wideband Code Division Multiple Access (WCDMA) multi-access technology using 5MHz bandwidth and High Speed Packet Access (HSPA). The WCDMA is a radio spectrum modulation technique that exploits channels with a bandwidth much larger than what is needed for transmission data. WCDMA channels share a very large band that meets the estimated maximum data rate instead of having an exclusive frequency band. It also can work with Time Division Duplex (TDD) and Frequency Division Duplex (FDD) techniques.

UMTS (Universal Mobile Telecommunications System) is a standard of 3G that has prevailed in European countries and was developed by the 3rd Generation Partnership Project (3GPP). It is about the evolution of the 2G mobile networks in capacity, speed of data transmission and existence of new services. UMTS uses multi-access radio

technology (W-CDMA) to provide higher spectrum performance and bandwidth to mobile operators. Due to the high speeds offered by UMTS, it provides access to internet services. The most serious disadvantage is that UMTS has higher power consumption compared to GSM networks.

CDMA-2000¹ is a family of 3G mobile technology standards for sending voice, data, and signalling data between mobile phones and cell locations. It was developed by 3GPP2 as a compatible with the previous 2G CDMAONE system (IS-95). 3.5G technology included protocols High-Speed Downlink Packet Access (HSDPA), High-Speed Uplink Packet Access (HSUPA), and 3.75G the High-Speed Packet Access (HSPA+) which were introduced in 3G networks to increase data rates up to 2 Mbps [9].

2.4 4G Networks

The increased use of mobile devices and consequently the large volume of transmission data led to the development of technologies and the next generation mobile communications with very high and reliable prerequisite speeds.

In March 2008, the ITU-R defined a set of requirements for 4G standards called International Advanced Mobile Telecommunications Advanced (IMT Advanced). These standards determine the cutting edge requirements for 4G services at 100 Mbit/s for high communication mobility (such as cars) and 1 Gbit/s for low mobility communication (such as pedestrians and regular users). The primary purpose for the development of 4G networks is high efficiency and reliable communication. The general characteristics that are consistent with this purpose are the comparatively higher rates of data transmission with a high level of Quality Of Service (QoS) and lower costs. At the same time, the spectrum distribution is such that high transmission rates are allowed on medium-sized cells. Networks serving such purposes consist of switching circuits, and all their components are digital. The 4G mobile telephony networks apply the LTE and WiMax technologies. LTE (developed by the 3GPP) is orthogonal frequency-division multiplexing (OFDM)-based radio access technology that supports multi-antenna transmission and bandwidth up to 20 MHz. Multiple-Input Multiple-Output (MIMO) is a key technology for high-speed wireless LAN, 3G and 4G mobile systems to enable multi-stream transmission with high spectrum efficiency [78].

¹<https://en.wikipedia.org/wiki/CDMA2000>

2.5 5G Networks

Now, we have entered the first era of 5G. By 5G we refer to the 5th generation of wireless mobile networks, which are a significant development of today's 4G LTE networks. They are designed to meet the vast data growth and connectivity of modern society on the internet as over time and due to complex human needs, there will be billions of devices connected to it. In the first phase, 5G will operate in conjunction with existing 4G networks, before evolving into a fully standalone network.

The standards family that characterizes the 5G networks is IMT-2020, issued by the ITU. Apart from ITU, there are other bodies responsible for the development of fifth-generation networks. 3GPP and IEEE are the parties that will play an important role. In addition to the existing bodies, the most important and largest is the 5G PPP (5G Public-Private partnership), which has been created and operates in Europe in collaboration among companies operating in the telecommunications sector, either as service providers or as manufacturers.

A characteristic advantage of 5G, except for providing faster connections and greater capacity, is the quick response time and minimization of latency. Latency is the time it takes for devices to respond to each other over the wireless network. 3G networks had a standard response time of 100 milliseconds, 4G reduced the time to about 30 milliseconds, and 5G is said to have a latency of as low as one millisecond. Essentially, 5G will be a new ecosystem that will make possible the fully connected world.

The 5G network will be highly heterogeneous with multiple access technologies, such as WiFi and LTE networks, in both licensed and unlicensed spectrum, providing an enhanced user experience. 5G networks will be more efficient than previous generations, as they are expected to be faster, offer higher data rates, and significantly support more users and connected devices. Also, they will be able to provide customized real-time services. In addition, the higher network coverage and the extended lifespan of the connected devices are among the advantages of 5G networks which will play a key role in transforming the current cities into smart ones and supporting the realization of fully digital age with completely new services. The activation of Internet of Things (IoT) is highly tied with the rise in the number of connected devices in combination with the launch of the 5G cellular mobile communications [3] that are able meet the communi-

cation needs of billions devices with improved performance and reliability by achieving explosive speeds, eliminating latency and synchronization [31].

The constantly increasing use of smart devices and multimedia services, as well as the demand for continuous and uninterrupted connectivity, high data rates, high energy efficiency, low latency and high quality of service (QoS), have made fifth-generation (5G) networks and beyond more imperative than ever. 5G networks address the emerging challenges by considering i) massive MIMO systems in the mmWave range, ii) full-duplex (FD) transmissions, iii) Device-to-Device (D2D) communication protocols and iv) network densification with heterogeneous networking architectures, where various types of cells (e.g., small, large) and access technologies are expected to coexist. In the near future, direct D2D communication is regarded as essential and promising technology, e.g., for extending capacity, enabling scalability and traffic offloading in various short-range applications, including Internet-of-Things (IoT), Machine-to-Machine (M2M) and emergency networks [8]. In the vision of 5G networks, users that are organized into small groups, will exchange information, e.g., about the network and channel status, cooperate in various operations of common interest and employ new technologies in favour of coordinated communication [97].

Moreover, the incorporation of mmWave MIMO antenna systems in the physical layer of 5G networks [39] will increase their transmission efficiency (spectral and power), throughput and network coverage ensuring the successful support of various evolving technologies. The 5G systems are assumed to provide consistent coverage and high-quality connectivity between various devices and behave well under multiple network topologies (multihop networks, mobile networks), such as Vehicular-to-Vehicular (V2V) or Vehicular-to-Everything (V2X), Device-to-Everything (D2E) communication systems, Augmented Reality (AR), Virtual Reality (VR), Smart City, etc. [84]. mmWave will be a key technology in the deployment of Heterogeneous Networks (HetNets) in 5G IoT (such as smart homes, smart cities, e-health) [19].

All these technologies have set new prerequisites for 5G channel modeling [100] to support broad channel bandwidths, mobility and beam tracking, spatial consistency (namely, proximal located transmitters or receivers should have similar channel characteristics) and be applicable in a wide range of scenarios such as indoor, urban, suburban, rural areas, high-speed train scenarios, etc.. To exploit the full potential of the 5G capa-

bilities and applications, channel knowledge is required among all the involved antenna pairs of the participating devices. Also, motivated by the above aspects of 5G, the second part of this dissertation proposes novel and efficient, fully distributed schemes that enable D2D communication, where the devices cooperate for estimating the involved channels.

2.6 6G Networks

The 6G [18] mobile networks are currently under development, and they will provide incredible transmission speed in the terabit range. This technology would require smart antennas, large memory in mobile devices, and optical networking. The 6G networks will be cell-free, and it would enable the incorporation of artificial intelligence (AI) in wireless networks. 6G will be associated with a much higher frequency in THz bands (300 GHz to 3 THz) to increase the required data rate. The commercial use of the THz spectrum for 6G will become until 2030. Some of the applications for 6G networks are connected robotics and autonomous systems, block-chain technology, multi-sensory extended reality, tactile internet, and industrial internet. Some of the advantages of 6G networks are:

1. Data rate is expected to provide data rate up to 10 Tbps, which is almost 100x better than 5G networks.
2. Latency will provide as low as 0.1 ms compared to 1 ms latency provided by 5G networks.
3. Efficient signalling for massive IoT connectivity and M2M communication.
4. User experience in extended reality, augmented reality, virtual reality, and artificial intelligence will be enhanced.
5. Spectral and network efficiency will be 10 x more compared to 5G networks.
6. Energy efficiency will be 100 times more compared to 5G network usage.
7. Ubiquitous Connection will allow huge broadcasting data, namely more than 1 million connections, which is almost 100 x more than 5G networks.

Performance metric	4G	5G	6G
Peak data rate	100Mbps	10Gpbs	upto 10Tbps
Latency	10ms	1ms	upto 0.1ms
Energy Efficiency	$1 \times$	$100 \times 4G$	$100 \times 5G$
Spectral Efficiency	$1 \times$	$100 \times 4G$	$100 \times 5G$
Spectrum	Up to 6 GHz	Up to 300 GHz	Up to 3 THz
Mobility	$200m/h$	$300m/h$	$600m/h$
Connection Density	0.1 million devices/km ²	1 million devices/km ²	10 million devices/km ²
Artificial Intelligence	No	Partial	Fully

Table 2.1 Performance comparison of 4G, 5G and 6G networks [18]

Finally, in Table 2.1, we present a comparison of 4G, 5G and 6G networks in terms of various performance metrics in order to highlight the benefits that the last two provide against predecessors.

CHAPTER 3

Key enablers for 5G: mmWave and massive MIMO

3.1 The mmWave Spectrum

The millimeter wave (mmWave) spectrum plays a key role in the implementation of 5G (and upcoming generations) communication systems. The available bands are between 30 GHz and 300 GHz (i.e., wavelengths between 1 mm to 10 mm), and the large unoccupied frequencies in these bands can be exploited to meet the increasing demand for transmission capacity [6, 109, 40, 4]. Although, from theoretical viewpoint, the mmWave provides potentials for extremely high data rates, the transmitting signals in the mmWave bands confront high attenuation, restricting the communication range [4]. Hence, there are several key technical challenges for the application of 5G in mobile networks. Massive Multi-Input, Multi-Output (MIMO) and directional beamforming (facilitated by the short wavelength of mmWave signals) have been leveraged in the relevant literature to moderate this loss and ensure reliability in communication.

On one hand, the large available bandwidth at mmWave will provide substantial throughput gains and low latency. On the other hand, mmWave channels exhibit reflection and losses in absorption and poor diffraction, severe pathloss, high penetration loss, high power consumption, blockage due to shadowing, hardware impairment, etc. As a result, the mmWave channel experiences much higher channel correlation, sensitivity to blockage and signal attenuation. In the following, we will briefly present the last topics.

3.1.1 Types of Attenuation

3.1.1.1 Pathloss

In free-space transmission, the power of the received signal can be determined by Friis transmission formula

$$P_r(d) = P_t G_t G_r \left(\frac{\lambda}{4\pi} \right)^2 d^{-n}, \quad (3.1)$$

where P_t is the transmit power and G_t and G_r are the antenna gains of the transmitter and receiver, respectively. Furthermore, λ , d are the wavelength, transmission distance, and the path loss exponent $n \in [2, 6]$. In free space $n = 2$ but yet it can be smaller than 2 in indoor environments [105].

Without directional antennas, mmWave propagation would confront a higher path loss relative to conventional microwave frequencies (e.g., below 6GHz), assuming that all other conditions including the antenna gains are the same. However, due to smaller wavelength more antenna elements can be fit in the same physical area, and thus, the maximum directional gains scaling as $G_r, G_t \propto \frac{1}{\lambda^2}$, for a given physical antenna aperture, compensates for the increased free-space path-loss at mmWave frequencies. The compensation of path loss with directional transmissions and high-dimensional antenna arrays is a fundamental characteristic of mmWave communication that ensures reliable communication a few hundreds of meters or even a few kilometers. At this point, it should be noted that directional antennas demonstrated 10km communication ranges under clean air conditions [40].

3.1.1.2 Rain, Atmospheric and Foliage Attenuation

The rain attenuation and atmospheric/molecular absorption increase the pathloss and limit the communication range [78]. For wireless communications signals operating above 10 GHz, rain attenuation is usually the dominant propagation impairment since the size of the raindrops is approximately identical to wavelength at mmWave frequencies inducing scattering effect on the propagating signal. Transmission loss is also induced when mmWave signals traversing through the atmosphere are absorbed in the air. Moreover, attenuation induced by foliage is one of the critical impairments in mmWave propagation [98].

3.1.1.3 Penetration Loss

The pathloss presupposes line-of-sight (LoS) communications, but the high penetration loss is exacerbated in non-line-of-sight (NLoS) scenarios. In indoor areas, the penetration losses may stemmed from buildings, human body and trees. For 28 GHz signals,

the losses for clear glass and dry walls are relatively low compared to microwave bands, while the penetration losses for brick and tinted glass are high (approximately 28 dB and 40 dB), which is much higher than in microwave bands. Materials such as brick can attenuate mmWave signals by as much as 40 to 80 dB and the human body itself can result in a 20 to 35 dB loss. Penetration losses are usually higher at higher frequencies. The human body (depending on the material of the clothing) and most building materials are reflective. This allows them to be important scatterers to enable coverage via NLOS paths for cellular systems. For example, measurements in New York City [78] confirm that even in extremely dense urban environments, coverage is possible up to 200m from a potential cell site. This is feasible since diffraction is not significant at mmWave frequencies, even if it is a primary means of coverage in sub 6 GHz systems. Moreover, the degree of penetration loss imposed by a tree depends both on its type and the vegetation density which changes with seasons (e.g., summer, winter).

3.1.2 mmWave Modelling

Since the wavelength at mmWave is shorter than the one in microwaves, material penetration will incur greater attenuation, thus increasing the significance of Line-Of-Sight (LOS) propagation and reflection. Due to severe path loss, a few reflecting paths could arrive at the receiver [109]. In this case, the channel is strongly sparse in the angle domain. The high propagation attenuation, the increased sensitivity to blockage, and the possible mobility of users, require mature signal processing techniques and novel insights in architectures and protocols to combat these challenges. Hence, severe propagation losses occurring at the mmWave band could be balanced by hybrid precoding (combining) techniques using massive MIMO technology, in order to provide high antenna array gain, through narrow directional beamforming, and sufficient spatial coverage [105].

The mmWave propagation environment is typically modeled via a geometric channel model which involves the Angles of Departure (AoD) and the Angles of Arrival (AoA) at the transmitter and receiver sides, respectively. It is noted, here, that the problem of estimating of AoD/AoA in the frame of array signal processing is closely related to the problem of channel estimation and beamforming in the frame of mmWave wireless communications. The core idea of parametric massive MIMO channel modeling is

to estimate the involved angles and the respective complex factors, instead of estimating the channel impulse response (communication viewpoint). The AoA estimation has been of particular interest to researchers for several decades and still remains an active area in wireless and mobile communication (e.g., radar, smart antenna) [54].

3.2 Beamforming

Beamforming [18] or spatial filtering is a signal processing method which adjusts the signal amplitude and phase (beam steering), to form a powerful beam toward a specific direction. There are three beamforming techniques, depending on the implementation method adopted. These are the analog, digital and hybrid techniques. It can be applied to either the transmitter or receiver. As for the base station (BS), it allows to find the appropriate path to transmit data to a user, the so-called beam alignment [36], which is quite challenging due to the very low signal-to-noise ratio available before beamforming. This technique will be quite beneficial for 5G networks and beyond. In massive MIMO systems, it helps to increase spectral efficiency. In this case, the base station can transmit data to the user from different paths and beamforming coordinates this transmission so that BS can transmit data simultaneously to more users. Also, beamforming increases the received signal strength, which allows for an increase in the transmission rates through the use of better (for each case) modulation techniques. At the same time, it helps the base station to find an appropriate path to transmit the data reducing interference with nearby users along the path. Radio waves in the mmWave spectrum cannot penetrate through natural obstacles, while, at the same time, they are characterized by a large propagation loss due to their short wavelength. Beamforming communication allows the user to receive a stronger signal without interference, reducing the relevant attenuation that could be subjected to due to the transmission environment.

3.3 Full Duplex Technology

Full duplex communication is the simultaneous transmission and reception of information using the same frequency band. Usually, downlink channels and uplink channels are separated into either the frequency domain or the time domain to produce orthogonal signals that do not interfere with each other. This is because, until recently, fully



Fig. 3.1 Beamforming [18]

duplex communication wasn't considered viable due to the strong interference caused at the emission stage due to the high power of the signal emitted compared to what is ultimately received by the receiver [58].

Recent developments in the field of self-interference cancellation, which is a signal processing technique that allows simultaneous transmission and reception on a channel, have made full-duplex communication at an acceptable cost attractive. The benefits offered are significant. In particular, there may be an increase in data rate and a reduction in latency, up to 100%, compared to conventional non-void communication, which allows communication in both directions but not simultaneously. For this reason, full-duplex communication technology will play an important role in the 5G networks and beyond [18].

3.4 MIMO Antenna Systems

Massive Multiple-Input-Multiple-Output (MIMO) radio systems have become a necessary communication solution for the implementation of 5G and B5G networks. A higher number of antenna elements (six and above) are required to achieve high multiplexing and spatial diversity, and benefit from the high bandwidth and data rate that 5G technology provides. Here, our aim is to introduce the basic concepts around the fundamentals of MIMO systems and their potentials in 5G and B5G networks.

3.4.1 General Characteristics

The capacity of a system can be further increased if we use an antenna arrangement not only in the base station, but also in the mobile device. In this case, a Multiple Input Multiple Output (MIMO) telecommunication system consists of N_t antennas in the transmitter and N_r antennas in the receiver. MIMO wireless channels have significantly higher capacity than SISO channels. The basic idea is to reuse the same physical channel, creating multiple virtual and independent subchannels, so that we can transmit different data streams in parallel. Therefore, through a MIMO structure, we can significantly increase the capacity of a wireless channel.

Multiple antennas can be used to increase the rate of transmission through multiplexing or to improve system performance through diversity. The term diversity refers to techniques that utilize variants of the same transmitted signal in order to improve the quality of the received one. Diversity techniques are used to reduce distortions that occur even in flat attenuation channels. Multiplexing utilizes the channel coefficients matrix to acquire independent transmission paths of the signal that can be used to transmit independent data. Also, MIMO systems enhance spectrum efficiency. In order to achieve the aforementioned gains, it is often necessary to know the channel matrix coefficients in the receiver and sometimes in the transmitter as well. Diversity techniques take advantage of the random nature of wireless propagation environment, finding at least highly uncorrelated signal paths for communication. In almost all applications, diversity is exploited by the receiver.

Among the diversity techniques (in space, frequency and time), spatial diversity techniques are the most popular ones in wireless communication systems. In classic cellular wireless systems, BS is located at elevated positions and mobile devices are located close to the ground. There is no guarantee that there will be a direct path between the transmitter and receiver antennas, and the signal attenuation due to scatterers surrounding a mobile device is usually Rayleigh modeled. The base station antennas must be relatively far away for receiving uncorrelated signals.

Conventional systems can ensure that different data streams will be transmitted in different intended users, with minimum interference from the other data streams, by combining multiplexing in time and frequency (namely, TDM and FDM).

MIMO antenna systems are classified as [61]:

1. **Point to Point MIMO:** A base station (transmitter) equipped with an antenna array of N_t elements serves a user equipped with an antenna array of size N_r . Different users are served over different time/frequency blocks.
2. **Multi-User (MU) MIMO:** It is equivalent to a Point-to-Point MIMO, breaking the single the N_r antennas user into N_r independent single-antenna users.
3. **Massive MIMO:** A base station equipped with an antenna array serves multiple users with N_r elements antenna array through precoding methods.

From now on, we will focus on Massive MIMO as it will be a key enabler for 5G networks.

3.4.2 Massive MIMO Potentials in mmWave

Massive MIMO is an important asset for the standardization and successful operation of 5G networks under mmWave, as conventional MIMO technology has proven unable to deliver the spectral efficiency that 5G applications demand.

Actually, the wide use of multimedia services has led to the demand for high data rates to ensure low latency and reliable communication in a wireless communication system. To satisfy these demands, the so-called Massive-MIMO or Large-scale MIMO systems [80], where several transmitting and receiving antennas are employed, have been suggested to exploit the spatial diversity for data transmission with improved throughput and performance.

The application of Massive MIMO may benefit systems operating at millimeter-waves, especially when these ones should support multi-user communication in potentially NLoS scenarios. **The main benefits are the following** [58]:

- **Array Gain:** MIMO technology can be used to direct the signal power spatially, leading to an array gain and providing the flexibility of spatial multiplexing.
- **Spectral efficiency (SE):** Massive MIMO systems achieve higher spectral efficiency with the ability to focus narrow beams on a user. In a Massive-MIMO communication system, the base station (BS) consisting of hundreds of antennas can simultaneously serve multiple users (MU-MIMO) on the same time-frequency resources and multiple data streams can be sent between the BS and each user improving SE.

- **Energy efficiency (EE):** When the beams focus on a small specific area, less radiant energy is required, which reduces the energy requirements. Moreover, the transmit power is essentially decreased when the number of transmit antennas is increased.
- **High data rates:** The antenna gain and the multiplexing gain of massive MIMO systems increase the data rates and the capacity of the wireless systems.
- **User tracking:** As massive MIMO systems direct narrow signal beams to the users making their tracking more reliable and accurate.
- **Less attenuation:** The large number of antennas in the receiver makes this technology more tolerant of attenuation effects [68].
- **Low latency:** Massive MIMO reduces latency at the air interface [42].
- **Reliability:** A large number of antennas offer significant diversity gain that increases the link reliability and elasticity against fading.
- **Robustness:** Massive MIMO systems are more robust against internal interference. Also, they have the ability to avoid antenna failures as a result of large number of antennas.
- **Security:** The massive MIMO offers enhanced security at the physical layer against hackers and passive eavesdropping attacks due to the orthogonal channels and narrow beams [18].
- **Signal processing:** Unlike conventional MIMO, Massive-MIMO systems allow for linear processing approaches such as Matched Filter (MF), Zero Forcing (ZF) and Minimum Mean Squared Error (MMSE) precoding/combining [113].

Massive MIMO uses spatial-division multiplexing such that the different data streams occupy the same frequencies and times. A key element to performing spatial multiplexing is an array of independently-controlled antennas. Under LoS propagation conditions, the data streams are transmitted via directed beams by increasing the number of antennas. However, in a scattering propagation environment, the data streams can arrive from many directions simultaneously. To carry out the multiplexing, the array needs to

know the time or frequency response of the propagation channels between each of its elements and each of the users. The channel state information (CSI) can be acquired operating in either Frequency Division Duplex (FDD) or Time Division Duplex (TDD) mode. The latter is considered a good alternative to FDD since uplink/downlink reciprocity can be exploited to estimate the CSI and control the beamforming operation [37].

MIMO beamforming is feasible in mmWave communications, particularly since the short wavelength at mmWave frequencies makes it possible to fit more antennas into a small area. For sub 6 GHz systems, MIMO arrays are fully digital, where each antenna requires a dedicated radio-frequency (RF) chain including power amplifier (PA), low noise amplifier (LNA), data converter (ADC/DACs), mixer, etc. However, the implementation of a fully digital MIMO at mmWave frequencies is a non-trivial task, using current circuit design technology [40]. The LoS channels in mmWave networks can be estimated using the direction of arrival (DoA) of incident waves. Massive MIMO indirectly offers highly directional and adaptive transmissions required for enhancing signal power and eliminating the interference in the blockage-sensitive environments at mmWave bands.

3.5 MIMO Architectures

Current standards in WLAN (IEEE 802.11n/ac) and cellular (IEEE 802.16e/m, 3GPP cellular LTE, and LTE Advanced) systems at sub-6GHz frequencies [49], [55] have already standardized MIMO technology. These standards support a small number of antennas, with two be the common case and eight be the maximum. However, at mmWave, the antenna arrays need more elements than the ones the lower frequency systems consider (the common number of elements is 32 to 256), but still occupy a small physical size due to the small wavelength.

MIMO communication at sub-6GHz is characterized by important architectural differences with the corresponding at mmWave bands. The operation of MIMO at conventional (lower) frequencies is subjected to fully digital signal processing at baseband. For mmWave communication with many antennas and high bandwidth, the conventional fully digital MIMO architecture requires many energy-consuming RF chains, leading to a prohibitive energy consumption and hardware cost. Therefore, alternative archi-

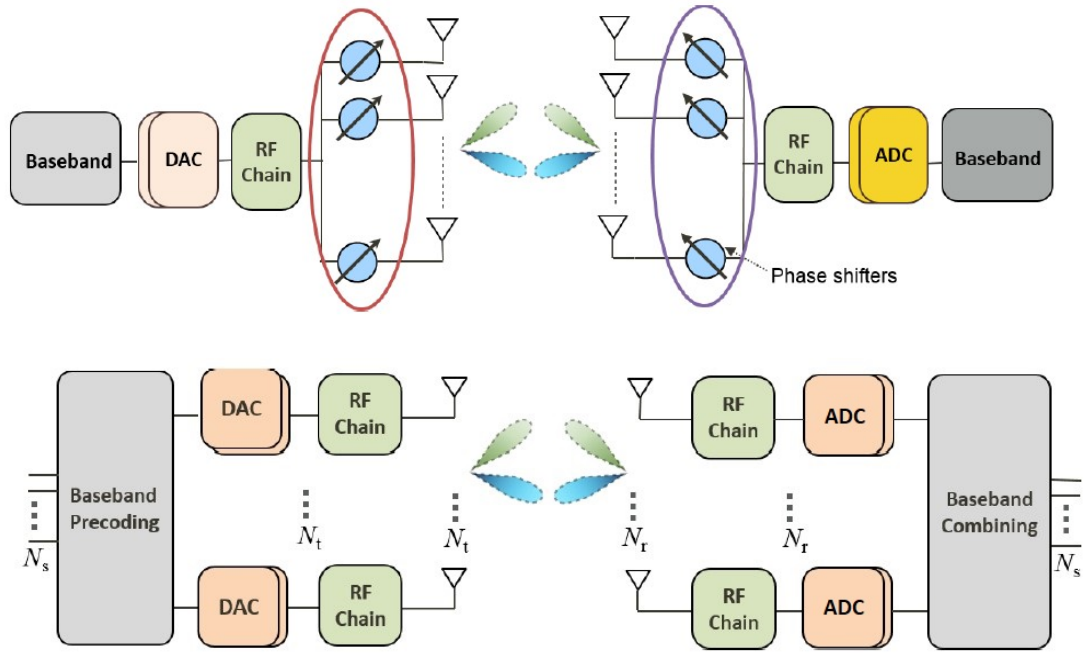


Fig. 3.2 Fully Analog and Fully Digital [40]

techniques are required for the successful establishment of the emerging mmWave mobile networks [105].

Despite the advantages of Massive MIMO antennas, the increased cost, the space and power consumption requirements of the circuits of additional antennas (especially on small devices), and the additional complexity of multi-dimensional signal processing should be taken into account. At higher carrier frequencies and signal bandwidths, there are several hardware constraints that may prohibit or complicate their practical implementation due to the existence of individual RF chain and data converter for each antenna. The hardware constraints have led to several mmWave MIMO architectures where the signal processing is conducted in analog and digital domains or both of them. In the next section, several MIMO architectures for mmWave systems will be presented [40].

3.5.1 Analog and Digital

Analog beamforming is the simplest technique for applying MIMO in mmWave systems. It is a solution supported in the IEEE 802.11ad wireless networking standard, and can be applied at both the transmitter and receiver. Analog beamforming is often

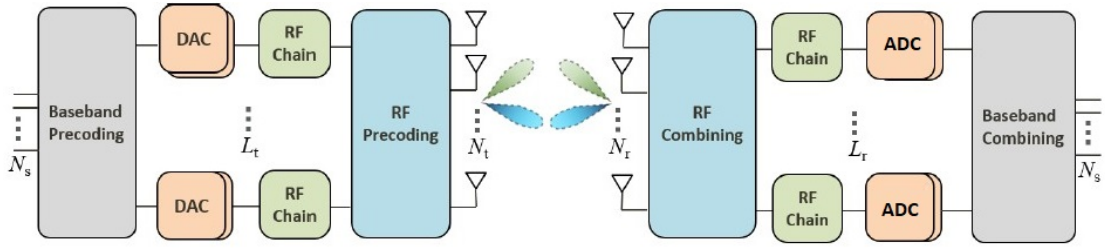


Fig. 3.3 Hybrid Architectures in transmitter and receiver side [40]

implemented using a network of digitally controlled phase shifters. In this configuration, several antenna elements are connected via phase shifters to a single RF chain, as illustrated in Figure 3.2. The phase shifter weights are adaptively adjusted using digital signal processing using a specific strategy to steer the beam and meet a given objective, for example to maximize received signal power.

The performance of analog beamforming based on phased arrays is limited by the use of quantized phase shifts and the lack of amplitude adjustment. This makes it more challenging to finely tune the beams and steer nulls. The power consumed by the phase shifters also depends on the resolution of the quantized phases.

As for the fully analog architecture, it is characterized by low hardware cost and power consumption. Nonetheless, there are several implications of using analog beamforming for mmWave MIMO. Analog beamforming utilizes one RF chain and a single beamformer can be designed, thus it can only support single-user and single-stream transmission. This means that it is impossible to realize multi-stream or multi-user MIMO systems and achieve the multiplexing gain to improve spectral efficiency. Beam steering is not trivial, especially when a communication link has not yet been established. This leads to the need for beam training algorithms and techniques for channel estimation. In general, to achieve the highest performance, the wireless protocol should be designed to support beam steering [105].

3.5.2 Hybrid

The implementation of a conventional MIMO with a large-scale antenna array and a full digital precoding (combining) approach is still prohibitive. This is because each antenna element in the transmitter (receiver) is connected to a digital (analog) to analog (digital) converter and an Radio Frequency (RF) chain of analog elements. Such an option brings high implementation cost and large power consumption at mmWave band

when the antennas are too many. A key solution to jointly balance the benefits of fully digital or fully analog architectures, new MIMO architectures have been designed for their deployment in 5G and B5G communication systems.

Also, hybrid analog/digital architectures are an alternative approach to reduce the number of RF chains, where the processing is split between an analog and a digital part. The analog part of a hybrid precoder (combiner) can be implemented by employing elements such as phase shifters, switches or lens antenna [29], [87], to name some of the most common proposed in the literature. Depending on the involved elements different hardware constraints arise [40]. In the literature, two main MIMO architectures have been proposed [59], [108] for the RF precoding (combining) matrix. In both cases, the aim is to find the optimal number of RF chains thus optimizing the cost, energy consumption and complexity. The first one is called fully-connected architecture, where each of RF chains is connected to all antenna elements. While, in the second one, the so-called partially-connected, the involved RF chains are associated with unique non-overlapping sub-groups of antenna elements [62].

Assuming a network of phase-shifters, the main characteristics and differences between these architectures are the following [105]: (i) A fully-connected network provides full precoding (combining) gain, achieves highly directive transmissions by adjusting the phases of the transmitted signals in all antenna elements with constant modulus phase-shifters, but it has high complexity. (ii) In a partially-connected network, for each RF chain, only the transmitted signals on the corresponding subset of antennas can be adjusted. For practical reasons, the partially-connected structure is often preferable, although it achieves reduced array gain and directivity, proportional to the number of subarrays. Also, it imposes the analog precoder (combiner) to be a block diagonal matrix of unit magnitude nonzero elements. Additionally, it has lower hardware complexity at the cost of beamforming gain, as compared to the fully-connected structure. The authors in [2] claim that there is no distinct solution for the hybrid structure that can offer the best trade-off between complexity and performance. Hence, a dynamic structure is needed depending on the application and channel conditions.

The hybrid architecture extends the fully analog architecture to the multi-stream scenario. The key idea is to divide the processing into two parts: an analog one which is realized by analog circuits and a reduced dimension digital which constitutes the

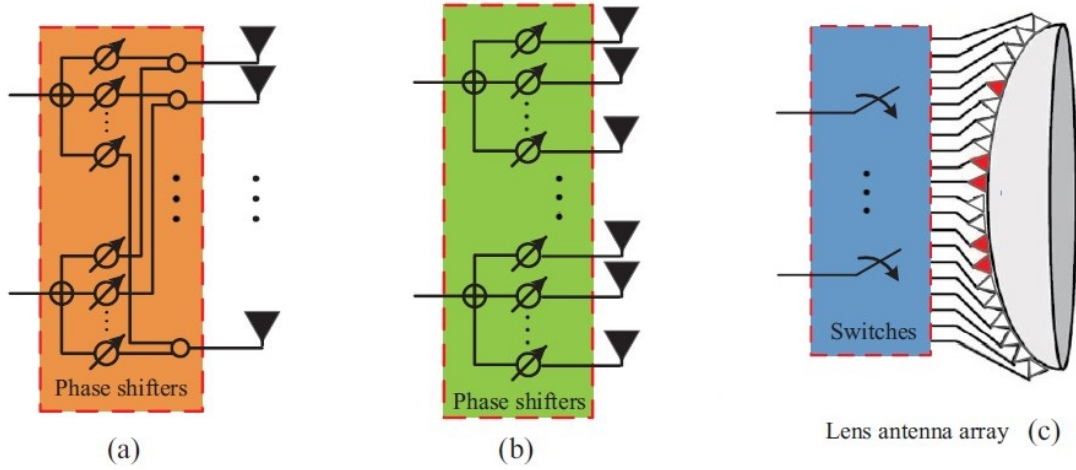


Fig. 3.4 Analog part with Phase shifters ((a) fully- and (b) sub-connected) and Lens antenna array [40]

precoder, in transmitter side, and combiner, in receiver part. The number of data streams determines the minimum required number of RF chains L_t which in hybrid architectures can be significantly reduced, leading to reduced cost and energy consumption [64]. Also, due to the low number of effective scatterers at BS side, only a few significant paths exist. Hence, the MIMO channel at mmWave is sparse with the corresponding matrix be low rank. The implementation of analog structure of the hybrid architecture can be made by different circuit networks (phase-shifters, switches), leading to different hardware constraints. The switching network is a low-cost low-complexity alternative [62]. In the Figures 3.4, 3.5, 3.6 are illustrated some implementation networks .

3.5.2.1 Fully-connected network

Each RF chain is connected to all antennas via phase-shifters and switches, as shown in Figure 3.4(a), 3.5(a). A highly directive signal can be achieved by adjusting the phases of transmitted signals on all antennas. By employing a phase-shifters network, all elements of the analog precoder/combiner have the same fixed amplitude.

3.5.2.2 Sub-connected network

Each RF chain is only connected to a subarray via phase-shifters and switches, as shown in Figures 3.4(b), 3.5(b). As for the phase-shifters, they adjust only the transmitted signals on a subset of antennas. Compared with the previous architecture, it achieves

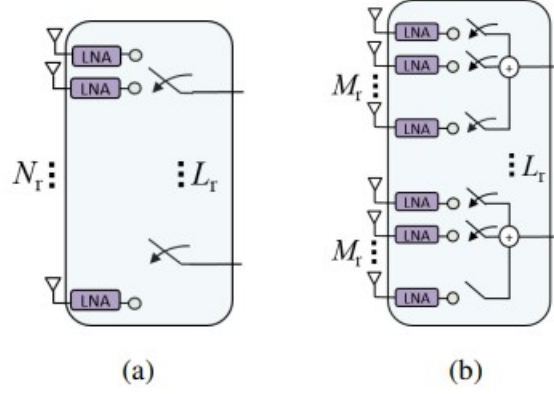


Fig. 3.5 Analog part with switches (a) fully- and (b) sub-connected antenna array in receiver side [40]

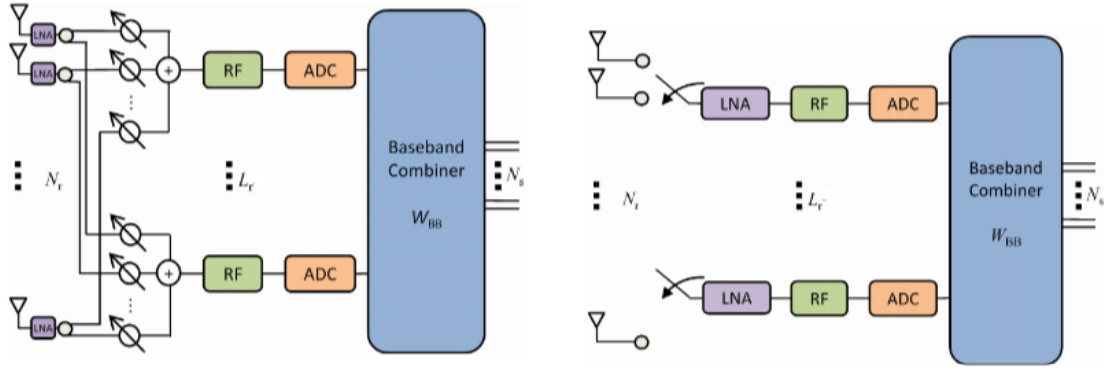


Fig. 3.6 Analog part with phase shifters and switches in receiver side [62]

reduced array gain and directivity proportional to the number of subarrays. Moreover, the network of phase-shifters imposes two hardware constraints i) the analog precoder/combiner should be a block diagonal matrix and ii) all the nonzero elements of the analog precoder/combiner have the same fixed amplitude.

3.6 Hybrid MIMO Systems and Channel Estimation

We assume a mmWave cellular system with a BS of N_t antennas and L_t RF chains that communicates with a user with N_r antennas and L_r RF chains [5], as illustrated in Figure 3.3. The BS transmits N_s data streams to a user for which hold that $N_s \leq L_t \leq N_t$ and $N_s \leq L_r \leq N_r$. The transmitter applies a hybrid precoder $\mathbf{F}$ to the symbol vectors $\mathbf{s} \in \mathbb{C}^{N_s \times 1}$ with $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \frac{1}{N_s}\mathbf{I}$. The hybrid precoder $\mathbf{F} = \mathbf{F}_{BB}\mathbf{F}_{RF}$ is analyzed to an analog $\mathbf{F}_{RF} \in \mathbb{C}^{N_t \times L_t}$ and a baseband precoder $\mathbf{F}_{BB} \in \mathbb{C}^{L_t \times N_s}$. The discrete-time transmitted signal is given by $\mathbf{x} = \mathbf{F}\mathbf{s}$.

We adopt a narrowband frequency-flat channel model and assuming perfect synchronization, the received signal can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{F}\mathbf{s} + \mathbf{n}, \quad (3.2)$$

where $\mathbf{n} \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$.

The matrix $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ captures the mmWave channel. Correspondingly, at the receiver's side, the combiner is composed of the RF and baseband combiners (where $\mathbf{W} = \mathbf{W}_{BB}\mathbf{W}_{RF}$) which processes the received signal $\mathbf{y}$ acquiring the following

$$\mathbf{r} = \mathbf{W}^H \mathbf{H} \mathbf{F} \mathbf{s} + \mathbf{W}^H \mathbf{n}. \quad (3.3)$$

Since mmWave channels are expected to be sparse due to limited scattering at the BS side, the geometric channel model is adopted. Each scatterer may contribute a single or multiple propagation paths between the BS and a user.

The matrix channel $\mathbf{H}$ is assumed to be a sum of the contributions of N_{cl} scattering clusters, each of which contribute N_{ray} propagation paths. Assuming multi-antenna users the channel can be expressed as [95], [62]

$$\mathbf{H} = \sqrt{\frac{N_t N_r}{N_{cl} N_{ray}}} \sum_{l=1}^{N_{cl}} \sum_{m=1}^{N_{ray}} \alpha_{l,m} \mathbf{a}_r(\theta_{l,m}) \mathbf{a}_t^H(\phi_{l,m}), \quad (3.4)$$

To construct the precoding and combining matrices, channel estimation is required, which, in mmWave systems, is difficult due to the use of massive MIMO antenna arrays.

The leverage of traditional techniques (such as LS, LMMSE) that estimate the channel response matrix is quite prohibitive, as the training overhead is high to estimate the channel. Also, the high bandwidth at the mmWave communication channel increases noise power and makes the received SNR very low before beamforming design. To tackle this problem, the base station and mobile device employ training beamforming and combining vectors in the channel estimation phase.

The mmWave channel characteristics can be leveraged to develop efficient training and estimation algorithms using hybrid precoding/ combining architectures [6]. In the angular domain, the mmWave channel is sparse in nature, due to limited scattering on the BS side, and it can be alternatively defined in terms of the AoA/AoD and com-

plex gain of each path. Hence, mmWave channel estimation is equivalent to estimating these parameters. In [26], a challenging DoA-based channel estimation method is suggested for a hybrid MIMO planar array at the BS, exploiting 2D-DFT. Moreover, using sparse representation theory, a channel estimation approach can be formulated using basis pursuit (BP) or matching pursuit approaches. From the matching pursuit methods, OMP for single antenna users or SOMP for multiple receiving antennas, and their variants can be utilized [40]. Also, an adaptive compressive sensing-based channel estimation is suggested in [5] with a multi-resolution hierarchical codebook for the training beamforming vectors. In these cases, the channel estimation is tackled in single user scenarios.

As we see, researchers have posed significant effort to develop effective (namely, low training overhead and/or complexity) and robust mmWave channel estimation algorithms that take into account the sparse nature of the mmWave channel. In the following chapters, we will focus on channel estimation $\mathbf{H}$ with i) Power Spectrum Estimation Techniques for finding the Angles/Direction of Arrival (DoA or AoA) assuming multiple antenna systems of hybrid architecture in the receiver side and ii) Sparse Representation techniques with predefined dictionaries combined with distributed adaptive methods for the estimation of the sparse channel representation. The second part of this thesis aims to develop an efficient multi-user channel estimation algorithm, exploiting distributed joint sparsity models and sparse representation methods to design algorithms that let all users cooperatively estimate the channel with reduced training cost. This method constitutes a promising solution due to the D2D capability that 5G networks provide.

CHAPTER 4

mmWave Channel Estimation with Direction-of-Arrival Estimation Techniques for Hybrid Architectures

4.1 Introduction

The efficient operation of mmWave massive MIMO systems is interwoven with the employment of hybrid architectures where the analog part is implemented by sub-connected and fully-connected antenna networks (e.g., phase-shifters, switches). To fully benefit from the massive MIMO techniques in mmWave communication systems, the accurate channel state information (CSI) estimation is necessary. mmWave channel estimation is equivalent to the accurate estimation of the involved angles information. Here, we will focus on the angles of incoming signal in the receiver side via the employment of DoA estimation techniques.

In array signal processing, many high-resolution subspace-based angle estimation algorithms, such as multiple signal classification (MUSIC), estimation of signal parameters via rotational invariance technique (ESPRIT), and their variants have been widely used due to their high-resolution angle estimation. Their application in massive MIMO systems and full-dimension MIMO systems for azimuth and elevated angles estimation have been extensively studied. However, they have been designed for the classical ULA antennas assuming that the source signals impinging on the receiver array are uncorrelated. If the latter property doesn't hold, spatial smoothing techniques should be employed before applying DoA estimation methods. The work of Chuang et al. [20] was the first that studied and designed the hybrid version of them, the so-called H-MUSIC and H-ESPRIT, to operate for 1-D and 2-D hybrid structure MIMO antenna systems. Moreover, several other methods were developed for DoA estimation in hybrid antenna systems [101], [87], [69], [110], [73] and [86]. In [71], the authors apply a preprocessing step with algebraic manipulation, which transforms the DoA estimation

problem with hybrid beamforming into that of digital beamforming. Hence, a low-complexity subspace-based algorithm is presented for unambiguous DoA estimation that is applicable to multiple signal classification (MUSIC) and its variant. Moreover, beam sweeping algorithms (BSA) have been considered for effectively reconstructing the spatial covariance matrix. In [107], an enhanced BSA method is suggested. Both these methods can then apply any subspace method to estimate the desired DoAs.

In this chapter, we will present the DoA estimation methods for the classical ULA and hybrid structure antenna arrays. Also, we will show the proposed preprocessing scheme that helps recover the full snapshot based on which any traditional DoA method can be applied to estimate the angle of arrival in receiver side. In the context of this dissertation, we study the reconstruction of the initial ULA snapshot, the so-called full snapshot, in hybrid antenna array structures that are utilized in mmWave communication systems. The proposed methodology enables the recovery of the snapshots that would have been obtained if a conventional (non-hybrid) Uniform Linear Antenna (ULA) array was employed. The problem is considered at the receiver side where the hybrid architecture exploits in a novel way the antenna elements of a uniform linear array. To this end, the recommended scheme is properly designed so as to be applicable to overlapping and non-overlapping architectures. Moreover, the full snapshot recoverability is addressed for two cases, namely, for time-varying and constant signal sources. Simulation results are also presented to illustrate the consistency between the theoretically predicted behaviors and the simulated results, and the performance of the proposed scheme in terms Angle-of-Arrival (AoA) estimation, when compared to the conventional MUSIC algorithm and a recently proposed hybrid version of MUSIC (H-MUSIC).

The chapter is organized as follows: In Section 4.3, the system model is defined for both the conventional ULA and the hybrid structure receiver cases. In Section 4.4, the problem formulation for the full snapshot recovery is developed, along with the design issues of the analog combiner in the hybrid structure, investigating the issue for two different phase-shifter architectures. In Section 4.5, we show experimental results for the evaluation of the performance of the proposed scheme. Finally, Section 4.6 concludes the chapter.

4.2 Our Contribution

The main contribution of this chapter of the dissertation is as follows

1. An efficient preprocessing scheme is developed that leads to the acquisition of the baseband snapshot irrespective of the hybrid antenna array architecture, as if a non-hybrid antenna array was employed.
2. The new scheme is studied for two typical hybrid receiver architectures implemented through a network of phase shifters that are employed in order to create suitable combinations of the received signals from which the baseband snapshot is retrieved. By maximizing the Signal-to-Total-Noise-Ratio (STNR) of the restored snapshot, i.e., a signal dependent noisy term, the proposed scheme determines the optimal values of the phase shifters.
3. The full snapshot recoverability problem is treated for both cases of time-varying and constant source signals during sub-snapshots collection time. A different combiner matrix is applied for each sub-snapshot acquisition.
4. The theoretical findings regarding the proposed optimal reconstruction schemes are verified by extensive simulation experiments.
5. As already mentioned, once the full snapshots are reconstructed via the proposed schemes, then any suitable AoA estimation algorithm can be applied. The coherent case, commonly encountered in multi-path channel environments, can be treated by employing appropriate techniques (e.g., spatial smoothing-based algorithm).

4.3 The System Model

Let us consider a ULA in which L far-field, narrowband, bandpass signals impinge on. The L signals arrive at the array from L different directions. Moreover, the antenna array size N is larger than the number of source signals, i.e. $N > L$ [83]. The spatial sampling depends on the wavelength of carrier frequency. In that case, to avoid spatial aliasing, the inter-element spacing should satisfy the condition $d = \frac{\lambda_c}{2} = \frac{c}{2f_c}$, where c is the propagation speed and λ_c, f_c are the carrier wavelength and frequency. Furthermore,

each source signal i is associated with an AoA denoted as θ_i , where $i = 1, 2, \dots, L$ and $\theta_i \in [-\pi, \pi]$. We assume that AoAs do not change during snapshots collection time. Thus, the signal array on the time instant t can be written as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{w}(t), \quad (4.1)$$

where $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_N(t)]$ is a $N \times 1$ vector with the signals received by the N array elements, $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_L]^T$ is an $L \times 1$ vector with the AoA's, $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_L(t)]^T$ is an $L \times 1$ vector containing the source signals and $\mathbf{w}(t) = [w_1(t), w_2(t), \dots, w_N(t)]$ is the $N \times 1$ noise vector. Finally, $\mathbf{A}(\boldsymbol{\theta})$ is the $N \times L$ array response matrix

$$\mathbf{A}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{a}(\theta_1) & \mathbf{a}(\theta_2) & \cdots & \mathbf{a}(\theta_L) \end{bmatrix}, \quad (4.2)$$

where

$$\mathbf{a}(\theta_i) = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & e^{j\frac{2d\pi}{\lambda_c} \sin \theta_i} & \cdots & e^{j\frac{2(N-1)d\pi}{\lambda_c} \sin \theta_i} \end{bmatrix}^T, i = 1, 2, \dots, L. \quad (4.3)$$

The sampled version of $\mathbf{x}(t)$ is usually utilized either by a conventional receiver, where each antenna element is connected to an RF chain (see Figure 4.1), or by a hybrid structure receiver, where the signals of groups of antenna elements are, first, combined and the resulting signal is passed via an RF chain (see Figures 4.2, 4.3). These two cases are analyzed in detail in the following subsections.

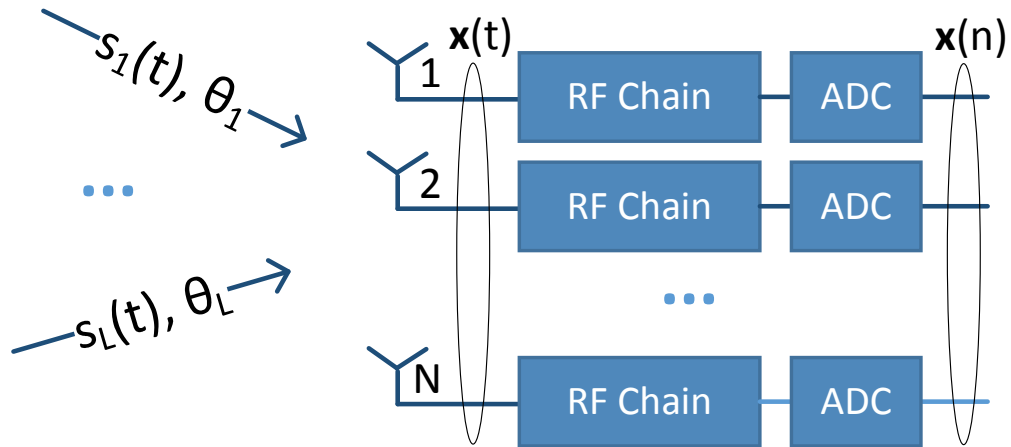


Fig. 4.1 The receiver in the conventional case.

4.3.1 The classical ULA case

Each received signal is processed by an individual RF chain, where, among other, is downconverted, and, then, is sampled by the corresponding Analog-to-Digital-Converter (ADC). At discrete-time n , the baseband signal, called *full snapshot*, is

$$\mathbf{x}(n) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(n) + \mathbf{w}(n), \quad (4.4)$$

where $\mathbf{x}(n), \mathbf{w}(n) \in \mathbb{C}^N$, $\mathbf{s}(n) \in \mathbb{C}^L$ and $\mathbf{w}(n) \sim \text{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$, while $E[\mathbf{w}(n)\mathbf{w}^H(k)] = \mathbf{0}$, for $n \neq k$ and $\mathbf{R}_s = E[\mathbf{s}(n)\mathbf{s}^H(n)]$ the source signals autocorrelation matrix.

4.3.2 MUSIC - A Conventional ULA DoA Estimation Method

MUSIC is a high-resolution direction-finding algorithm based on the eigenvalue decomposition of the covariance matrix observed at a ULA array. The angles at which the pseudospectrum of MUSIC has finite peaks are the desired directions of arrival. It belongs to the family of subspace-based direction-finding algorithms.

Let $\mathbf{R}_x$ be the correlation matrix of the received vector $\mathbf{x}(n)$. Then,

$$\begin{aligned} \mathbf{R}_x &= E\{\mathbf{x}(n)\mathbf{x}^H(n)\} \\ &= E\{\mathbf{A}(\boldsymbol{\theta})\mathbf{s}(n)\mathbf{s}^H(n)\mathbf{A}^H(\boldsymbol{\theta})\} + E\{\mathbf{w}(n)\mathbf{w}^H(n)\} \\ &= \mathbf{A}(\boldsymbol{\theta})\mathbf{R}_s\mathbf{A}^H(\boldsymbol{\theta}) + \sigma_w^2 \mathbf{I}, \end{aligned} \quad (4.5)$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{R}_s$ the source signals correlation matrix. Note that the rank of $\mathbf{A}(\boldsymbol{\theta})$ is L , so is that of $\mathbf{A}(\boldsymbol{\theta})\mathbf{R}_s\mathbf{A}^H(\boldsymbol{\theta})$.

Conducting the singular value decomposition (SVD) of $\mathbf{R}_x$, we can have

$$\mathbf{R}_x = \begin{bmatrix} \mathbf{U}_s & \mathbf{U}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_s & 0 \\ 0 & \boldsymbol{\Sigma}_n \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix} \quad (4.6)$$

where $\boldsymbol{\Sigma}_s \in \mathbb{R}^{L \times L}$ and $\boldsymbol{\Sigma}_n \in \mathbb{R}^{(N-L) \times (N-L)}$ are diagonal matrices with elements that correspond to the L largest and remaining $N - L$ (σ^2) eigenvalues of the $\mathbf{R}_x$, respectively. The columns of the submatrix $\mathbf{U}_s \in \mathbb{C}^{N \times L}$ consist of the eigenvectors of the L largest eigenvalues spanning the signal space, while those of $\mathbf{U}_n \in \mathbb{C}^{N \times (N-L)}$ span the

noise subspace, which is orthogonal to the signal subspace.

The spatial power spectrum of MUSIC algorithm is captured by

$$P(\theta) = \frac{1}{\mathbf{a}^H(f_c, \theta) \mathbf{U}_n \mathbf{U}_n^H \mathbf{a}(f_c, \theta)}. \quad (4.7)$$

The L AoAs in $\boldsymbol{\theta}$ correspond to the L largest peaks of the power spectrum estimated via the employment of $P(\theta)$. The maximum detectable number of the AoAs is $N - 1$ since $\mathbf{U}_n$ must contain at least one column. To use the MUSIC algorithm, we need to know the number of AoAs. In this dissertation, we assume that L is known a priori.

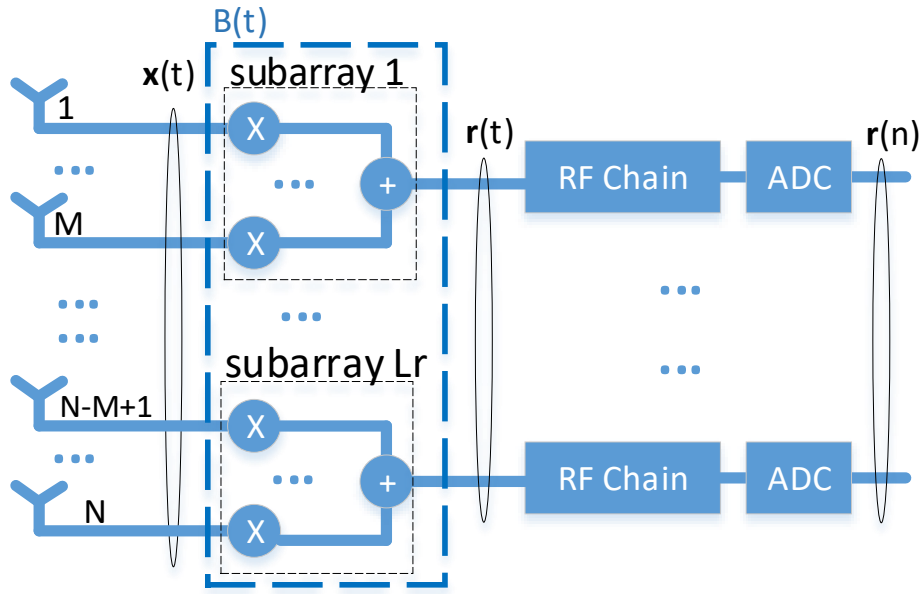


Fig. 4.2 The receiver in the partially connected hybrid case.

4.3.3 The hybrid case

We assume an antenna array with N elements which can be organized into L_r groups considering two widely-used analog beamforming architectures, as shown in Figures 4.2, 4.3. In the partially-connected one, each group, called subarray, consists of M antenna elements and connects to one RF chain, where $N = ML_r$. Furthermore, in the l -th group ($l = 1, 2, \dots, L_r$), the i -th received signal ($i = 1, 2, \dots, M$) is, first, processed by a phase-shifter by $e^{j\phi_{li}(t)}$, $\phi_{li}(t) \in [-\pi, \pi]$, and, then, all of them are added up. In the fully-connected one, all antenna elements form L_r different antenna groups each of which connects to an RF chain. Similarly, each received signal at each antenna element is processed by a phase-shifter, and then, all of them are added up. In

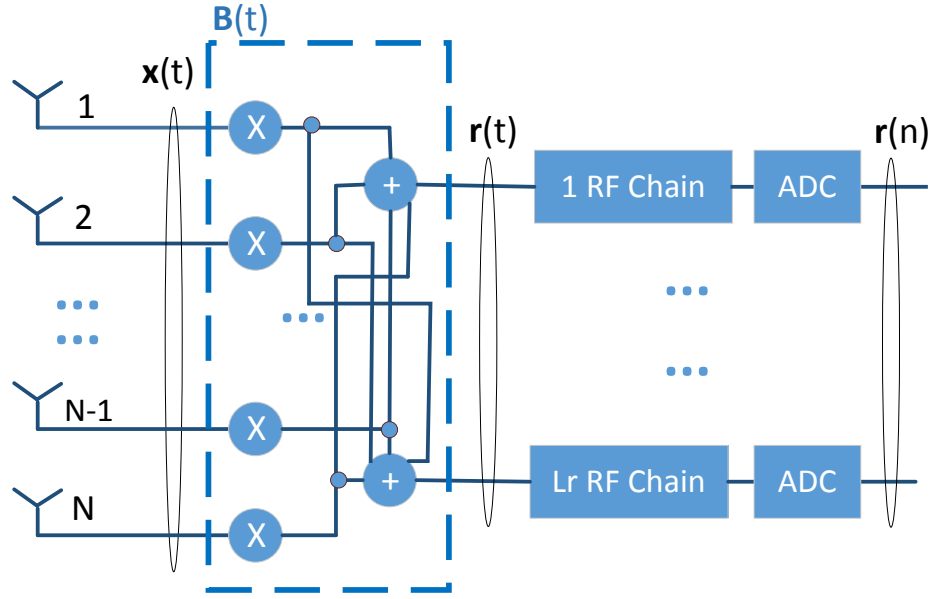


Fig. 4.3 The receiver in the fully connected hybrid case.

both architectures, the matrix $\mathbf{B}(t)$ of size $L_r \times N$, known as analog combiner in the relevant literature [6],[25], constitutes the analog part of the hybrid structure, as well as the involved operations (depicted in Figures 4.2, 4.3).

The receiver first processes the received signals using an $L_r \times N$ combiner, $\mathbf{B}(t)$, implemented using phase shifters such that $|\mathbf{B}_{i,j}| = 1$. Therefore, the matrix $\mathbf{B}(t)$ maps the received signal $\mathbf{x}(t)$ to a reduced size signal vector $\mathbf{r}(t)$, namely, $L_r \times 1$, where

$$\mathbf{r}(t) = \mathbf{B}(t)\mathbf{x}(t). \quad (4.8)$$

Finally, as in the classical case, $\mathbf{r}(t)$ passes through the available RF chains and the corresponding ADCs and the resulting baseband signal can be written as

$$\mathbf{r}(n) = \mathbf{B}(n)\mathbf{x}(n), \quad (4.9)$$

or according to (4.4) in the following compact form,

$$\mathbf{r}(n) = \mathbf{B}(n)\mathbf{A}(\theta)\mathbf{s}(n) + \mathbf{B}(n)\mathbf{w}(n). \quad (4.10)$$

Inspecting (4.9), the reconstruction of the full snapshot in (4.4) depends on the invertibility of matrix $\mathbf{B}(n)$. However, it is non-square with $L_r < N$, i.e., it has more columns than rows, and thus, it is singular. In this case, the initial ULA snapshot $\mathbf{x}(n)$

cannot be restored. Hence, in the following section the adopted methodology for the full snapshot recovery is presented in detail.

4.4 Recovering the Full Snapshot

In this section, the problem of full snapshot reconstruction of the baseband signal in (4.4) is described. The adopted methodology consists of two parts: i) the problem formulation for the full snapshot estimation and ii) the appropriate RF combiner design assuming both cases of time-varying and constant source signals for the hybrid architectures described previously.

4.4.1 Description of the procedure for recovering the full snapshot

Because of the hybrid architecture, as explained in the previous section, at discrete time n , we obtain a sub-snapshot $\mathbf{r}(n)$ of size L_r instead of the desired one in (4.4). To reconstruct the full snapshot of (4.4), a number of sub-snapshots $\mathbf{r}(l)$ are collected and utilized. In the following, the description of this collection, which will permit the reconstruction of the full snapshot, will be presented. As it will be shown, the reconstruction accuracy depends, among other, on the variability of the source signals¹ that impinge on the antenna array, which signals will be denoted as $\mathbf{s}_v(l)$ from now on. Both cases of time varying and constant source signals will be considered in the performance analysis, where the optimal design of the respective combiners will be discussed.

In more detail, let us assume that T signals

$$\mathbf{r}(l) = \mathbf{B}(l)\mathbf{A}(\theta)\mathbf{s}_v(l) + \mathbf{B}(l)\mathbf{w}(l) \quad (4.11)$$

are collected at time instants $l = n, n + 1, \dots, n + T - 1$, which will be referred to as sub-snapshots in the following. By concatenating the T sub-snapshots into the $TL_r \times 1$ vector

¹In many radio communication standards the involved signals exhibit fluctuations [24]

$$\bar{\mathbf{r}}(n) = \begin{bmatrix} \mathbf{r}(n) \\ \mathbf{r}(n+1) \\ \vdots \\ \mathbf{r}(n+T-1) \end{bmatrix}, \quad (4.12)$$

the input-output relation can be formulated as

$$\bar{\mathbf{r}}(n) = \begin{bmatrix} \mathbf{B}(n)\mathbf{A}(\theta)\mathbf{s}_v(n) \\ \mathbf{B}(n+1)\mathbf{A}(\theta)\mathbf{s}_v(n+1) \\ \vdots \\ \mathbf{B}(n+T-1)\mathbf{A}(\theta)\mathbf{s}_v(n+T-1) \end{bmatrix} + \begin{bmatrix} \mathbf{B}(n)\mathbf{w}(n) \\ \mathbf{B}(n+1)\mathbf{w}(n+1) \\ \vdots \\ \mathbf{B}(n+T-1)\mathbf{w}(n+T-1) \end{bmatrix} \quad (4.13)$$

or, in a more compact form,

$$\bar{\mathbf{r}}(n) = \bar{\mathbf{B}}(n)\bar{\mathbf{A}}(\theta)\bar{\mathbf{s}}_v(n) + \bar{\mathbf{w}}(n), \quad (4.14)$$

where

$$\bar{\mathbf{B}}(n) = \begin{bmatrix} \mathbf{B}(n) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}(n+1) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{B}(n+T-1) \end{bmatrix}, \quad (4.15)$$

$$\bar{\mathbf{A}}(\theta) = \begin{bmatrix} \mathbf{A}(\theta) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}(\theta) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{A}(\theta) \end{bmatrix} \quad (4.16)$$

and

$$\bar{\mathbf{s}}_v(n) = \begin{bmatrix} \mathbf{s}_v^T(n) & \mathbf{s}_v^T(n+1) & \dots & \mathbf{s}_v^T(n+T-1) \end{bmatrix}^T. \quad (4.17)$$

In the following, the vectors $\mathbf{s}_v(l)$ involved in (4.17) are modeled as the sum of a constant signal $\mathbf{s}(n)$ and a non-constant signal $\mathbf{s}_{nc}(l)$, i.e., $\mathbf{s}_v(l) = \mathbf{s}(n) + \mathbf{s}_{nc}(l)$, for $l = n, n+1, \dots, n+T-1$. Actually, it is practically infeasible to acquire constant source signals when operating at very high data rates (i.e., multiple Gbps depending on the mmWave carrier frequency), without oversampling. Therefore, here, the general case of non-constant signals is considered, as, the assumption of constant ones in a full snapshot period (usually encountered in the relevant literature [20]) can be treated as a special case, where $\mathbf{s}_v(l) = \mathbf{s}(n)$ for $l = n, n+1, \dots, n+T-1$.

Specifically, the source signals vector in (4.17) can be rewritten as

$$\bar{\mathbf{s}}_v(n) \equiv \begin{bmatrix} \mathbf{s}(n) + \mathbf{s}_{nc}(n) \\ \mathbf{s}(n) + \mathbf{s}_{nc}(n+1) \\ \vdots \\ \mathbf{s}(n) + \mathbf{s}_{nc}(n+T-1) \end{bmatrix} \quad (4.18)$$

Hence, by substituting (4.18) in (4.14) a collection of sub-snapshots of the hybrid array can be obtained as follows

$$\bar{\mathbf{r}}(n) = \begin{bmatrix} \mathbf{B}(n)\mathbf{A}(\theta)\mathbf{s}(n) \\ \mathbf{B}(n+1)\mathbf{A}(\theta)\mathbf{s}(n) \\ \vdots \\ \mathbf{B}(n+T-1)\mathbf{A}(\theta)\mathbf{s}(n) \end{bmatrix} + \begin{bmatrix} \mathbf{B}(n)\mathbf{A}(\theta)\mathbf{s}_{nc}(n) \\ \mathbf{B}(n+1)\mathbf{A}(\theta)\mathbf{s}_{nc}(n+1) \\ \vdots \\ \mathbf{B}(n+T-1)\mathbf{A}(\theta)\mathbf{s}_{nc}(n+T-1) \end{bmatrix} + \begin{bmatrix} \mathbf{B}(n)\mathbf{w}(n) \\ \mathbf{B}(n+1)\mathbf{w}(n+1) \\ \vdots \\ \mathbf{B}(n+T-1)\mathbf{w}(n+T-1) \end{bmatrix} \quad (4.19)$$

or, in a more compact form,

$$\bar{\mathbf{r}}(n) = \mathcal{B}(n)\mathbf{A}(\theta)\mathbf{s}(n) + \bar{\mathbf{B}}(n)\bar{\mathbf{A}}(\theta)\bar{\mathbf{s}}_{nc}(n) + \bar{\mathbf{w}}(n), \quad (4.20)$$

where

$$\mathcal{B}(n) = \begin{bmatrix} \mathbf{B}^T(n) & \mathbf{B}^T(n+1) & \dots & \mathbf{B}^T(n+T-1) \end{bmatrix}^T \quad (4.21)$$

and

$$\bar{\mathbf{s}}_{nc}(n) = \begin{bmatrix} \mathbf{s}_{nc}^T(n) & \mathbf{s}_{nc}^T(n+1) & \dots & \mathbf{s}_{nc}^T(n+T-1) \end{bmatrix}^T. \quad (4.22)$$

The value of parameter T (i.e., the number of sub-snapshots) is selected such that the $TL_r \times N$ matrix $\mathcal{B}(n)$ is invertible, thus, $TL_r \geq N$ should hold. To ensure low sub-snapshots collection overhead the minimum value $T = \frac{N}{L_r}$ is chosen, where the number of antenna elements N and the number of RF chains L_r are considered to be powers of two, i.e., $N = 2^i$ and $L_r = 2^j$, where i, j are integers. By applying different matrices $\mathbf{B}(l)$ for each l , linear independent $\mathbf{r}(l)$'s are collected. Due to the fact that $\mathcal{B}(n)$ is designed to be rectangular and full rank, its pseudo-inverse matrix $\mathcal{B}^\dagger(n)$ is reduced to $\mathcal{B}^{-1}(n)$. In the following, from (4.20) an estimate $\hat{\mathbf{x}}(n)$ of the full snapshot $\mathbf{x}(n)$ can be acquired as

$$\hat{\mathbf{x}}(n) = \mathcal{B}^{-1}(n)\bar{\mathbf{r}}(n). \quad (4.23)$$

Hence, from (4.23) and as outlined in Algorithm 1, we can fully recover the full snapshot as

$$\begin{aligned} \hat{\mathbf{x}}(n) &= \mathbf{A}(\theta)\mathbf{s}(n) + \mathcal{B}^{-1}(n)\bar{\mathbf{w}}(n) \\ &\quad + \mathcal{B}^{-1}(n)\bar{\mathbf{B}}(n)\bar{\mathbf{A}}(\theta)\bar{\mathbf{s}}_{nc}(n). \end{aligned} \quad (4.24)$$

According to (4.24), the $\hat{\mathbf{x}}(n)$ is hindered by a term associated with the additive noise and an additional term which depends on the time variability of the received source signals. As it will be observed in the following, the variability in the source signals is treated as a noise term. Hence, in this case, the total noise term is

$$\mathbf{w}''(n) = \mathcal{B}^{-1}(n)\bar{\mathbf{w}}(n) + \mathcal{B}^{-1}(n)\bar{\mathbf{B}}(n)\bar{\mathbf{A}}(\theta)\bar{\mathbf{s}}_{nc}(n). \quad (4.25)$$

Considering now the special case of constant signals, namely, $\mathbf{s}_v(l) = \mathbf{s}(n)$, for

$l = n, n + 1, \dots, n + T - 1$, $\bar{\mathbf{s}}_v(n)$ is simplified to

$$\bar{\mathbf{s}}_v(n) = \begin{bmatrix} \mathbf{s}(n) \\ \mathbf{s}(n) \\ \vdots \\ \mathbf{s}(n) \end{bmatrix}. \quad (4.26)$$

Hence, $\bar{\mathbf{r}}(n)$ can be written as

$$\bar{\mathbf{r}}(n) = \mathcal{B}(n)\mathbf{A}(\theta)\mathbf{s}(n) + \bar{\mathbf{w}}(n) \quad (4.27)$$

and the restored full snapshot in this case is given as

$$\hat{\mathbf{x}}(n) = \mathbf{A}(\theta)\mathbf{s}(n) + \mathcal{B}^{-1}(n)\bar{\mathbf{w}}(n), \quad (4.28)$$

where the total noise term $\mathbf{w}''(n)$ in (4.25) is reduced to

$$\mathbf{w}''(n) = \mathcal{B}^{-1}(n)\bar{\mathbf{w}}(n). \quad (4.29)$$

Algorithm 1: Proposed Scheme for the full snapshot reconstruction

- 1: **Input:** Number of antennas N , Number of RF chains L_r
 - 2: **Input:** SNR values, Number of sub-snapshots T
 - 3: **Output:** Recovered snapshot $\hat{\mathbf{x}}(n)$
 - 4: Collect T sub-snapshots from L_r RF chains by applying T different phase-shifter matrices $\mathbf{B}(l)$, for $l = n, n + 1, \dots, n + T - 1$, to formulate $\bar{\mathbf{r}}$ in (4.12)
 - 5: Apply matrix $\mathcal{B}^{-1}$ to (4.23)
 - 6: Recover the full snapshot $\hat{\mathbf{x}}$ as shown in (4.24) and (4.28), respectively
-

From (4.24) and (4.28), it is evident that the matrix $\mathcal{B}(n)$ has an impact on the power of the noise terms thus affecting the subsequent utilization of the recovered signals, i.e., in AoA estimation. Hence, the matrix $\mathcal{B}(n)$ should be determined appropriately so that it is invertible and well conditioned and, at the same time, it doesn't increase the power of the noise term. Such a discussion and the relevant analysis are presented in the following section.

4.4.2 Problem formulation for the RF combiner design

In this section, the analysis focuses on the appropriate design of the RF combiner $\mathcal{B}(n)$ matrix, i.e., the analog part of the hybrid structure for the minimum value of $T = \frac{N}{L_r}$.

The problem of interest in this section is to design the $\mathcal{B}(n)$, i.e., the analog part of the hybrid structure, such that the Signal-to-Total Noise-Ratio, after the application of the preprocessing scheme, to be maximized. The STNR is defined as

$$\text{STNR} = \frac{\text{Tr}\{\mathbf{C}\}}{\text{Tr}\{\mathbf{C}_{\mathbf{w}''}\}}, \quad (4.30)$$

where the power of the signal component of ULA snapshot $\mathbf{x}(n)$ in (4.4), defined as

$$\begin{aligned} \text{Tr}\{\mathbf{C}\} &= \text{Tr}\{\mathbf{A}(\theta)E[\mathbf{s}(n)\mathbf{s}^H(n)]\mathbf{A}(\theta)^H\} \\ &= \text{Tr}\{\mathbf{A}(\theta)\mathbf{R}_s\mathbf{A}(\theta)^H\}, \end{aligned} \quad (4.31)$$

is kept constant and identical to the one in the reconstructed full snapshot. Hence, the STNR maximization is equivalent to the minimization of the total noise power $\text{Tr}\{\mathbf{C}_{\mathbf{w}''}\}$ which entails that the total noise term

$$\mathbf{w}''(n) = \mathcal{B}^{-1}(n)\bar{\mathbf{w}}(n) + \mathcal{B}^{-1}(n)\bar{\mathbf{B}}(n)\bar{\mathbf{A}}(\theta)\bar{\mathbf{s}}_{nc}(n)$$

in (4.25) and the noise term in (4.4) have equivalent statistical properties (e.g., as to covariance matrix).

It is noted, here, that in the special case of constant signals, $\mathbf{w}''(n)$ is equal to $\mathcal{B}^{-1}(n)\bar{\mathbf{w}}(n)$. The desired minimization problem, which lies in the minimization of the power of the total noise term $\mathbf{w}''(n)$, is written as,

$$\min_{\mathcal{B}(n)} \text{Tr}\{\mathbf{C}_{\mathbf{w}''}\}, \text{ subject to } \mathcal{B}(n) \in \mathcal{S}, \quad (4.32)$$

where $\mathbf{C}_{\mathbf{w}''} = E\{\mathbf{w}''(n)(\mathbf{w}''(n))^H\}$ is the covariance matrix of the noise term in (4.24) and $\mathcal{S}$ denotes the feasibility set of $\mathcal{B}(n)$, namely, all matrices with elements being equal to exponentials that denote the phase-shifting operation of the hybrid architecture (either fully- or partially-connected). In particular, the set $\mathcal{S}$ consists of all $N \times N$ matrices, of the form of (4.21).

The form of sub-matrices $\mathbf{B}(n)$ in (4.21) captures both the structure and the operations, described in Section 4.4.1, and will be elaborated in the following. If the hybrid architecture is organized into non-overlapping sub-arrays as depicted in Figure 4.2, in (4.21) the sub-matrices are of the form

$$\mathbf{B}(l) = \begin{bmatrix} \mathbf{b}_{l1}^T & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{b}_{l2}^T & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{b}_{lL_r}^T \end{bmatrix} \quad (4.33)$$

where $\mathbf{b}_{lj}$ is an $M \times 1$ vector with the phase-shifters' values to be used by the j th subarray at time $l = n, n+1, \dots, n+T-1$. Assuming a fully-connected architecture, as in Figure 4.3, the sub-matrices in (4.21) are of the form

$$\mathbf{B}(l) = \begin{bmatrix} \mathbf{b}_{l1}^T \\ \mathbf{b}_{l2}^T \\ \vdots \\ \mathbf{b}_{lL_r}^T \end{bmatrix} \quad (4.34)$$

and $\mathbf{b}_{lj}$ is an $N \times 1$ vector with the exponentials to be used by the j th subarray at time $l = n, n+1, \dots, n+T-1$. Here, it should be noted that, the same notation $\mathbf{b}_{lj}$ for two different sized vectors is utilized so as to keep the symbols simple and the dimension is implied by the context.

In the following, the minimization problem of (4.32) will be solved for the two architectures (Figures 4.2 and 4.3) under study and the two signal cases (i.e., for constant and non-constant signals described in (4.26) and (4.18), respectively). As it will be shown later, when constant signals are considered, the problem can be solved optimally, whereas for the non-constant signals the selection of the appropriate design becomes a more intriguing task.

4.4.2.1 Constant Source Signals

First, the special case of constant source signals will be considered. Under this case, the noise term $\mathbf{w}''(n)$ is equal to $\mathcal{B}^{-1}(n)\bar{\mathbf{w}}(n)$, as shown in (4.28), and the covariance

matrix can be written as $\mathbf{C}_{\mathbf{w}''} = E\{\mathcal{B}^{-1}(n)\bar{\mathbf{w}}(n)(\mathcal{B}^{-1}(n)\bar{\mathbf{w}}(n))^H\}$. In order to solve the problem in (4.32), first, $\mathbf{C}_{\mathbf{w}''}$ is written in the following equivalent form,

$$\mathbf{C}_{\mathbf{w}''} = \mathcal{B}^{-1}(n)E\{\bar{\mathbf{w}}(n)(\bar{\mathbf{w}}(n))^H\}(\mathcal{B}^{-1}(n))^H. \quad (4.35)$$

In the following, using (4.35), (4.32) will be solved for the two hybrid architectures. First, the partially-connected case will be treated.

In the partially-connected architecture, it can be proved, by employing (4.21), (4.33), that $E\{\bar{\mathbf{w}}(n)(\bar{\mathbf{w}}(n))^H\} = \sigma^2 M \mathbf{I}_{N \times N}$, and, thus, $\mathbf{C}_{\mathbf{w}''}$ is written as

$$\mathbf{C}_{\mathbf{w}''} = \sigma^2 M \left(\mathcal{B}^H(n) \mathcal{B}(n) \right)^{-1}. \quad (4.36)$$

Moreover, the following inequality will be used for determining a lower bound for the cost function in (4.32),

$$\text{Tr}\{\mathbf{A}^{-1}\} \geq \sum_i \frac{1}{[\mathbf{A}]_{ii}}, \quad (4.37)$$

where $[\mathbf{A}]_{ii}$ is the i th diagonal element of $\mathbf{A}$ and the equality holds when $\mathbf{A}$ is diagonal [1]. By substituting (4.36) in (4.32) and applying (4.37), the following lower bound can be derived,

$$\text{Tr}\{\mathbf{C}_{\mathbf{w}''}\} \geq \sigma^2 M \sum_{i=1}^N \frac{1}{[\mathcal{B}^H(n) \mathcal{B}(n)]_{ii}} = \sigma^2 N. \quad (4.38)$$

Inspecting (4.38), the lower bound is independent of the unknown $\mathbf{B}(n)$ we are looking for. Therefore, any $\mathcal{B}(n) \in \mathcal{S}$ can be considered as the optimal solution in (4.32). In our case, an optimal choice of $\mathcal{B}(n)$ is the one for which the following condition is fulfilled

$$\mathcal{B}^H(n) \mathcal{B}(n) = M \mathbf{I}, \quad (4.39)$$

which subsequently makes the covariance matrix $\mathbf{C}_{\mathbf{w}''}$ diagonal and equal to the one of the noise term in (4.4), i.e., $\mathbf{C}_{\mathbf{w}''} = \sigma^2 \mathbf{I}$ (as observed by 4.36). In the following, such an optimal, in this setting, matrix will be described.

In order to determine a suitable $\mathcal{B}(n) \in \mathcal{S}$, first, $\mathcal{B}^H(n) \mathcal{B}(n)$ is written in the following more convenient form

$$\mathcal{B}^H(n) \mathcal{B}(n) = \mathcal{B}^H(n) \mathbf{P}^T \mathbf{P} \mathcal{B}(n), \quad (4.40)$$

where $\mathbf{P}$ is an $N \times N$ permutation matrix that changes the position of the rows of $\mathcal{B}(n)$ in the following manner. The first row of each $\mathbf{B}(l)$ goes at the top (see (4.21), (4.33)). Then, the second row of each $\mathbf{B}(l)$ follows and so on. This matrix can be created from the identity matrix if its rows are moved similarly. The matrix $\mathbf{P}\mathcal{B}(n)$ can be written as

$$\mathbf{P}\mathcal{B}(n) = \begin{bmatrix} \mathbf{F}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{F}_{L_r} \end{bmatrix}, \quad (4.41)$$

where $\mathbf{F}_i$, $i = 1, 2, \dots, L_r$, is an $M \times M$ matrix with its l th row equal to $\mathbf{b}_{li}^T$, $l = n, n+1, \dots, n+M-1$.

After substituting (4.41) in (4.40), $\mathcal{B}^H(n)\mathcal{B}(n)$ is a block diagonal matrix, as shown in (4.42), with the i th diagonal element being $\mathbf{F}_i^H \mathbf{F}_i$, $i = 1, 2, \dots, L_r$.

$$\mathcal{B}^H(n)\mathcal{B}(n) = \begin{bmatrix} \mathbf{F}_1^H \mathbf{F}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_2^H \mathbf{F}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{F}_{L_r}^H \mathbf{F}_{L_r} \end{bmatrix} \quad (4.42)$$

Thus, the desired condition depicted by (4.39) can be fulfilled if equivalently $\mathbf{F}_i^H \mathbf{F}_i = M\mathbf{I}$.

An appropriate choice for the $\mathbf{F}_i$'s, $i = 1, 2, \dots, L_r$, could be based on the Fourier matrix $\mathbf{F}_{N \times N}$, whose (p, q) th element is given as $[\mathbf{F}]_{p,q} = e^{-j2\pi pq/N}$. Then, only the first M columns are kept and the resulting matrix is denoted as $\mathbf{F}_{N \times M}$. Finally, rows of $\mathbf{F}_i$, $i = 1, 2, \dots, L_r$, are set equal to the rows $i, L_r + i, 2L_r + i, \dots$ of $\mathbf{F}_{N \times M}$. These matrices fulfill the desired condition, namely, $\mathbf{F}_i^H \mathbf{F}_i = M\mathbf{I}$. Hence, by utilizing these matrices, the phase shifters of the receiver in the partially-connected, hybrid case can be determined. It is noted, here, that the $\mathbf{F}_i$'s are actually the beamforming matrices used by H-MUSIC in [20] in order to obtain the overall received vector, assuming $T = M$, e.g. rectangular beamformers. It is worth mentioning that H-MUSIC can work with non-rectangular beamformers $\mathbf{F}_i$'s, i.e. with $T < M$, thus, with smaller size sub-snapshots and snapshots whose size will be less than the so-called full snapshot.

Here, the case of the fully-connected architecture is investigated. In order to solve

the problem in (4.32) by employing (4.34), it can be shown that $E\{\bar{\mathbf{w}}(n)(\bar{\mathbf{w}}(n))^H\} = \sigma^2 \mathcal{D}(n)$, where matrix $\mathcal{D}(n)$ is block diagonal with diagonal elements of the form $\mathbf{B}(l)\mathbf{B}^H(l)$ for $l = n, n+1, \dots, n+T-1$. Using above, the trace of $\mathbf{C}_{\mathbf{w}''}$ in (4.35) is written in the following equivalent form,

$$\begin{aligned} \text{Tr}\{\mathbf{C}_{\mathbf{w}''}\} &= \text{Tr}\{\sigma^2 \mathcal{B}^{-1}(n) \mathcal{D}(n) (\mathcal{B}^{-1}(n))^H\} \\ &= \text{Tr}\{\sigma^2 \mathcal{D}(n) (\mathcal{B}^{-1}(n))^H \mathcal{B}^{-1}(n)\} \\ &= \text{Tr}\{\sigma^2 \mathcal{D}(n) (\mathcal{B}(n) \mathcal{B}^H(n))^{-1}\} \\ &= \text{Tr}\{\sigma^2 (\mathcal{B}(n) \mathcal{B}^H(n) \mathcal{D}^{-1}(n))^{-1}\}. \end{aligned} \quad (4.43)$$

By substituting (4.43) in (4.32) and applying (4.37), the following lower bound can be derived,

$$\text{Tr}\{\mathbf{C}_{\mathbf{w}''}\} \geq \sigma^2 \sum_{i=1}^N \frac{1}{[\mathcal{B}(n) \mathcal{B}^H(n) \mathcal{D}^{-1}(n)]_{ii}} = \sigma^2 N, \quad (4.44)$$

where, for deriving the last equality, $[\mathcal{B}(n) \mathcal{B}^H(n) \mathcal{D}^{-1}(n)]_{ii} = 1, \forall i$, is used, which can be proved by employing (4.21), (4.34) and observing the diagonal elements of the result.

Observing (4.44), any matrix could be utilized for deriving the lower bound. However, in this case, the matrix $\mathcal{B}(n)$ is designed based on the Fourier matrix $\mathbf{F}_{N \times N}$, as this one has the properties imposed by S. Then, submatrices of $\mathbf{F}$ of size $L_r \times N$ can be used for the design of the $\mathbf{B}(l)$'s for $l = n, n+1, \dots, n+T-1$ comprise the matrix $\mathcal{B}(n)$ in (4.21). Hence, by utilizing these matrices the phase shifters of the receiver in the fully-connected, hybrid case can be determined, as, in this case, the minimum value in (4.38) is satisfied.

4.4.2.2 Non-constant source signals

In this sub-section, the case of non-constant source signals will be elaborated. In order to solve the problem defined in (4.32), the covariance matrix, $\mathbf{C}_{\mathbf{w}''}$, for the noise term

$\mathbf{w}''(n) = \mathcal{B}^{-1}(n)\bar{\mathbf{w}}(n) + \mathcal{B}^{-1}(n)\bar{\mathbf{B}}(n)\bar{\mathbf{A}}(\theta)\bar{\mathbf{s}}_{nc}(n)$ can be written as

$$\begin{aligned} \mathbf{C}_{\mathbf{w}''} &= E\{\mathbf{w}''(n)(\mathbf{w}''(n))^H\} \\ &= \sigma^2 \mathcal{B}^{-1}(n)\bar{\mathbf{B}}(n)\bar{\mathbf{B}}^H(n) (\mathcal{B}^{-1}(n))^H + \\ &\quad \mathcal{B}^{-1}(n)\bar{\mathbf{B}}(n)\bar{\mathbf{A}}(\theta)\mathbf{R}_{\bar{\mathbf{s}}_{nc}}\bar{\mathbf{A}}^H(\theta)\bar{\mathbf{B}}^H(n) (\mathcal{B}^{-1}(n))^H \\ &= \mathcal{B}^{-1}(n)(\sigma^2\bar{\mathbf{B}}(n)\bar{\mathbf{B}}^H(n) + \\ &\quad \bar{\mathbf{B}}(n)\bar{\mathbf{A}}(\theta)\mathbf{R}_{\bar{\mathbf{s}}_{nc}}\bar{\mathbf{A}}^H(\theta)\bar{\mathbf{B}}^H(n)) (\mathcal{B}^{-1}(n))^H. \end{aligned} \quad (4.45)$$

By substituting (4.45) into the cost function of (4.32), the following equation can be written as

$$\text{Tr}\{\mathbf{C}_{\mathbf{w}''}\} = \text{Tr}\{\mathcal{B}^{-1}(n)\mathbf{U}(\mathcal{B}^{-1}(n))^H\}, \quad (4.46)$$

where $\mathbf{U} = \sigma^2\bar{\mathbf{B}}(n)\bar{\mathbf{B}}^H(n) + \bar{\mathbf{B}}(n)\bar{\mathbf{A}}(\theta)\mathbf{R}_{\bar{\mathbf{s}}_{nc}}\bar{\mathbf{A}}^H(\theta)\bar{\mathbf{B}}^H(n)$ and $\mathbf{R}_{\bar{\mathbf{s}}_{nc}}$ the non-constant signals correlation matrix of size $TL \times TL$.

For the full snapshot reconstruction the preprocessing matrix $\mathcal{B}(n)$ in (??) should not change the noise power, thus, the set of solutions $\mathbf{S}$ is further restricted into a new set that consists of only unitary matrices, i.e., $\mathcal{B}^{-1}(n) = \mathcal{B}^H(n)$. Hence, the trace of (4.46) can be further simplified as

$$\begin{aligned} \text{Tr}\{\mathbf{C}_{\mathbf{w}''}\} &= \text{Tr}\{\mathcal{B}^H(n)\mathbf{U}\mathcal{B}(n)\} \\ &= \text{Tr}\{\mathbf{U}\mathcal{B}(n)\mathcal{B}^H(n)\} \\ &= \text{Tr}\{\mathbf{U}\} \\ &= \text{Tr}\{\sigma^2\bar{\mathbf{B}}(n)\bar{\mathbf{B}}^H(n) \\ &\quad + \bar{\mathbf{B}}(n)\bar{\mathbf{A}}(\theta)\mathbf{R}_{\bar{\mathbf{s}}_{nc}}\bar{\mathbf{A}}^H(\theta)\bar{\mathbf{B}}^H(n)\} \\ &= \text{Tr}\{\bar{\mathbf{B}}(n)(\sigma^2\mathbf{I}_{TN} + \bar{\mathbf{A}}(\theta)\mathbf{R}_{\bar{\mathbf{s}}_{nc}}\bar{\mathbf{A}}^H(\theta))\bar{\mathbf{B}}^H(n)\} \\ &= \text{Tr}\{\bar{\mathbf{B}}(n)\mathbf{V}\bar{\mathbf{B}}^H(n)\}, \end{aligned} \quad (4.47)$$

where $\mathbf{V} = \sigma^2\mathbf{I}_{TN} + \bar{\mathbf{A}}(\theta)\mathbf{R}_{\bar{\mathbf{s}}_{nc}}\bar{\mathbf{A}}^H(\theta)$. Next, according to [1], the lower bound of (4.32), when the cost-function has the form of (4.47), can be written as

$$\text{Tr}\{\mathbf{C}_{\mathbf{w}''}\} \geq \sum_{i=1}^N \lambda_i(\mathbf{V}), \quad (4.48)$$

where the λ_i 's, for $i = 1, 2, \dots, TN$, are the eigenvalues of the matrix $\mathbf{V}$.

Irrespective of the hybrid architecture, the minimum value in (4.48) is attained when the matrix $\bar{\mathbf{B}}(n)$ equals the matrix of eigenvectors of $\mathbf{V}$ that correspond to the N smallest eigenvalues, sorted in an ascending order. It is noted, here, that $\bar{\mathbf{A}}(\theta)\mathbf{R}_{\bar{\mathbf{s}}_{nc}}\bar{\mathbf{A}}^H(\theta)$ of size $TN \times TN$ is a matrix of rank equal to TL . If the number of antenna elements N and the number of sources L satisfy the condition $TN - TL > N$ this entails that the N smallest eigenvalues are equal to σ^2 , and so the minimum value in (4.48) can be written as $\min \text{Tr}\{\mathbf{C}_{\mathbf{w}''}\} = \sigma^2 N$.

The solution in (4.48) that minimizes the problem in (4.32), does not have the desired phase shifters' matrix structure because of the term $\bar{\mathbf{A}}(\theta)\mathbf{R}_{\bar{\mathbf{s}}_{nc}}\bar{\mathbf{A}}^H(\theta)$ in matrix $\mathbf{V}$, which depends on the unknown AoAs.

Under the assumption of a low SNR regime, in the involved cost function in (4.47), the dominant term would be the one related to the additive noise, especially when the time variations of the source signals are not strong enough. Hence, in this case, (4.47) can be approximated by $\text{Tr}\{\sigma^2 \bar{\mathbf{B}}(n)\bar{\mathbf{B}}^H(n)\}$. To design the $\mathcal{B}(n)$, for either partially-connected or fully-connected architecture, a similar process as the one followed in case of constant signals can be employed. As it has been already proved in the previous section, the Fourier matrix is an appropriate choice for the design of matrix $\mathcal{B}(n)$ for both hybrid architectures, as this one has the properties of $\mathbf{S}$ (i.e., all matrices with elements being equal to exponentials that denote the phase-shifting operation of the hybrid architecture) and satisfies the minimum value $\sigma^2 N$ in (4.48).

Finally, in the medium to high SNR regime, the dominant term in matrix $\mathbf{V}$ is $\bar{\mathbf{A}}(\theta)\mathbf{R}_{\bar{\mathbf{s}}_{nc}}\bar{\mathbf{A}}^H(\theta)$. Hence, the solution to the minimization problem in (4.32) is a matrix which actually depends on the unknown AoAs, which constitutes this case intractable, and the statistical properties of the time varying signals $\mathbf{s}_{nc}(l)$. Since, in this case, the solution to the minimization problem under study is related to the unknown AoAs, any predefined matrix is expected to lead to a floor at the cost function of (4.47). This floor depends on how powerful the time-varying nature of the source signals are (namely, on the values of the covariance matrix $\mathbf{R}_{\bar{\mathbf{s}}_{nc}}$, if one considers the model that is adopted here). A simple model for the time-varying signals $\mathbf{s}_{nc}(l)$ is adopted. Specifically, they can be modeled as discrete random variables of Gaussian distribution assuming that their values are statistically independent of equal power, with known first and second order statistics (i.e., $\mathbf{s}_{nc}(l) \sim \mathcal{CN}(\mathbf{0}, \sigma_{nc}^2 \mathbf{I})$). Under this model, the noise power behav-

ior in (4.47) is verified in the simulations.

4.4.3 Hybrid MUSIC (H-MUSIC)

Now, we proceed to an established method that extends the conventional MUSIC to efficiently operate in the hybrid architecture antenna arrays. This method will be considered for comparison in the experiments.

Assuming that $\mathbf{s}(n) = \mathbf{s}(n+1) = \dots = \mathbf{s}(n+T-1)$, the measurement vector $\mathbf{r}_l(n)$ for the l -th subarray is defined as

$$\begin{aligned} \mathbf{r}_l(n) &= \begin{bmatrix} r_l(n) \\ r_l(n+1) \\ \vdots \\ r_l(n+T-1) \end{bmatrix} \\ &= \mathbf{B}_l^H \mathbf{A}_{(f_c, \boldsymbol{\theta})} \mathbf{F}_{(l, \boldsymbol{\theta})} \mathbf{s}(n) + \mathbf{w}_{l,b}(n) \end{aligned} \quad (4.49)$$

where $\mathbf{B}_l = [\mathbf{b}_{l,1}, \mathbf{b}_{l,2}, \dots, \mathbf{b}_{l,T}]$ ($L_r \times T$), $\mathbf{F}_{(l, \boldsymbol{\theta})}$ is a phase-shift matrix that models the phase difference between the first and the l -th subarray, and $\mathbf{w}_{l,b}(n) = [\mathbf{b}_{l,1}^H \mathbf{w}_l(n), \mathbf{b}_{l,2}^H \mathbf{w}_l(n+1) \dots \mathbf{b}_{l,T}^H \mathbf{w}_l(n+T-1)]^T$ is the equivalent noise vector for the l -th subarray. The matrix $\mathbf{F}_{(l, \boldsymbol{\theta})}$ ($L \times L$) is defined as

$$\mathbf{F}_{(l, \boldsymbol{\theta})} = \begin{bmatrix} e^{-j2\pi f_c((l-1)Md) \sin(\theta_1)/c} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & e^{-j2\pi f_c((l-1)Md) \sin(\theta_2)/c} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & e^{-j2\pi f_c((l-1)Md) \sin(\theta_K)/c} \end{bmatrix}. \quad (4.50)$$

Also, $\mathbf{A}_{(f_c, \boldsymbol{\theta})}$ ($M \times L$) is the response matrix at the l -th sub-array. Collecting all the $T \times 1$ received signal vectors from each RF chain into one vector, we obtain an overall received vector, denoted by $\bar{\mathbf{r}}(n)$, as

$$\bar{\mathbf{r}}(n) = \begin{bmatrix} \mathbf{r}_1(n) \\ \mathbf{r}_2(n) \\ \vdots \\ \mathbf{r}_{L_r}(n) \end{bmatrix} = \begin{bmatrix} \mathbf{B}_1^H \mathbf{A}_{(f_c, \theta)} \mathbf{F}_{(1, \theta)} \\ \mathbf{B}_2^H \mathbf{A}_{(f_c, \theta)} \mathbf{F}_{(2, \theta)} \\ \vdots \\ \mathbf{B}_{L_r}^H \mathbf{A}_{(f_c, \theta)} \mathbf{F}_{(L_r, \theta)} \end{bmatrix} \mathbf{s}(n) + \begin{bmatrix} \mathbf{w}_{1,b}(n) \\ \mathbf{w}_{2,b}(n) \\ \vdots \\ \mathbf{w}_{L_r,b}(n) \end{bmatrix}. \quad (4.51)$$

The correlation matrix of $\bar{\mathbf{r}}(n)$ can then be calculated as

$$\begin{aligned} \mathbf{R}_{\bar{\mathbf{r}}} &= E\{\bar{\mathbf{r}}(n)\bar{\mathbf{r}}^H(n)\} \\ &= \bar{\mathbf{A}}_{(f_c, \theta)} \mathbf{R}_s \bar{\mathbf{A}}_{(f_c, \theta)}^H + \sigma_{v,b}^2 \mathbf{I} \end{aligned} \quad (4.52)$$

where $\sigma_{w,b}^2 = E\{\mathbf{w}_l^H(n) \mathbf{b}_{l,i} \mathbf{b}_{l,i}^H \mathbf{w}_l(n)\}$, $1 \leq l \leq L_r$, and $1 \leq i \leq T$. Conducting the SVD of $\mathbf{R}_{\bar{\mathbf{r}}}$, we have $\Sigma_s \in \mathbb{R}^{L \times L}$ and $\Sigma_n \in \mathbb{R}^{(L_r T - L) \times (L_r T - L)}$ are diagonal matrices, and $[\mathbf{U}_s \ \mathbf{U}_n]$ is the matrix with singular vectors. The columns of the submatrix $\mathbf{U}_s \in \mathbb{C}^{L_r T \times L}$ span the signal space while those of $\mathbf{U}_n \in \mathbb{C}^{L_r T \times (L_r T - L)}$ span its orthogonal subspace. Using the MUSIC algorithm we find the AoAs as

$$P(\theta) = \frac{1}{\bar{\mathbf{a}}_{(f_c, \theta)}^H \mathbf{U}_n \mathbf{U}_n^H \bar{\mathbf{a}}_{(f_c, \theta)}} \quad (4.53)$$

where

$$\bar{\mathbf{a}}_{(f_c, \theta)} = \begin{bmatrix} \mathbf{B}_1^H \mathbf{a}_{(f_c, \theta)} \\ \mathbf{B}_2^H \mathbf{a}_{(f_c, \theta)} e^{-j2\pi f_c (Md) \sin \theta / c} \\ \vdots \\ \mathbf{B}_{L_r}^H \mathbf{a}_{(f_c, \theta)} e^{-j2\pi f_c ((L_r - 1)Md) \sin \theta / c} \end{bmatrix}. \quad (4.54)$$

We refer to this AoA estimation algorithm as H-MUSIC. The maximum detectable number of the AoAs is $L_r T - 1$. Note that the overall degrees of freedom is $L_r T$. If $T = M$, the full dimension of $\mathbf{A}_{(f_c, \theta)}$ can be exploited. Note that T is the number of the angles that the phase-shifters beamform. The larger the T , the longer the time we need for the AoA estimation.

4.5 Results and Discussion

In this section, simulations are carried out to demonstrate the performance of the proposed scheme and verify its behavior under the various cases that have been studied previously. To be more specific, the recoverability of the full snapshot is considered for the problem of AoA estimation in 1D mmWave massive MIMO systems by employing MUSIC. The combination of the proposed full snapshot reconstruction scheme with the classical MUSIC algorithm is called hereafter PRE+MUSIC.

In the following, we assume that there are $L = 4$ signal sources. The constant part $s(n)$ of the source signals is assumed to be an i.i.d. sequence of Quadrature Phase Shift Keying (QPSK) symbols. These signals impinge on the array with AoAs equal to 12.3° , 28.1° , 54.6° , 62.8° , respectively. The carrier frequency f_c is selected in the mmWave band and it is set to 30 GHz, which corresponds to a wavelength λ equals to 1 cm and antenna spacing 0.5 cm. An $N = 32$ elements antenna array is deployed in receiver side which, in the hybrid case, is grouped into $L_r = 8$ RF chains. In the following, three experiments will be presented.

In the first experiment, we demonstrate how well the theoretical analysis, presented in Section 5.8, predict the performance of the proposed scheme for the full snapshot reconstruction and validate the agreement between the computed theoretical minimum of the total noise power $\text{Tr}\{\mathbf{C}_{\mathbf{w}''}\}$, under the optimal DFT-based phase-shift combiner matrix in the case of constant signals. Moreover, the performance improvement of the proposed solution over the random phase-shift combiner matrix is also demonstrated. In the following experiments we investigate the performance of the proposed snapshot reconstruction scheme in the context of AoA estimation.

4.5.1 Noise Power in Reconstructed Snapshots

Let recall that, since the signal component power remains constant, the STNR values in (4.30) depend only on the corresponding total noise power $\text{Tr}\{\mathbf{C}_{\mathbf{w}''}\}$, which, in the case of non-constant sources is dependent on the additive noise power σ^2 plus a term which depends on the non-constant signals power σ_{nc}^2 , treated as noise term, as well. In the special case of constant sources, the total noise term consists only of the additive noise. Hence, the power of the noise component ($\mathbf{w}''(n)$) in the reconstructed snapshot

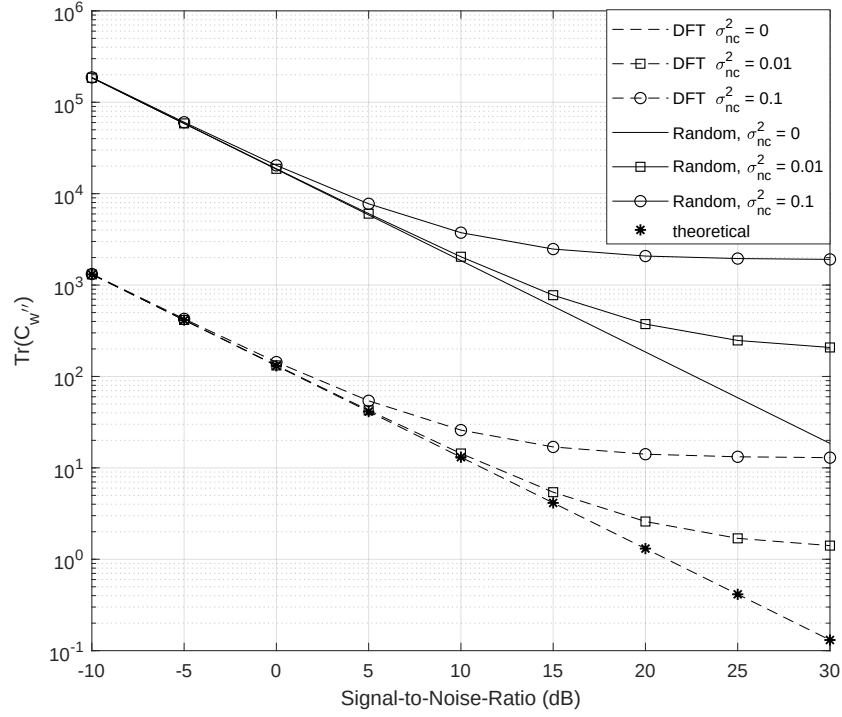


Fig. 4.4 Total Noise Power Evaluation for a Combiner Matrix based on i) DFT matrix and ii) Random Phase Shifters for $\sigma_{nc}^2 = \{0, 0.01, 0.1\}$.

(see (4.25) for non-constant sources or (4.29) for constant ones) is depicted in Figure 4.4, versus the additive noise power σ^2 , for $\sigma_{nc}^2 = \{0, 0.01, 0.1\}$ (denoting the power of $\mathbf{s}_{nc}$), when the Fourier matrix for the phase shifters is utilized, for the fully-connected architecture. It is worth to mention that the behavior of the cost function remains the same if, instead of the fully-connected architecture, the partially-connected architecture is adopted in each case. Furthermore, the theoretical minimum value in (4.38), (4.44) is also depicted in the same figure when the Fourier matrix is considered. Indeed, when $\sigma_{nc}^2 = 0$ the Fourier matrix achieves the minimum value and constitutes an optimal choice. As discussed in Section 5.8, the proposed phase shifters for both architectures, in the case when $\sigma_{nc}^2 \neq 0$, follow the minimum value at the low SNR regime. Additionally, as σ_{nc}^2 increases, the floor of the cost function starts appearing in smaller values of SNR. Also, inspecting Figure 4.4, when considering random phase-shifters combiner matrices $\mathcal{B}(n)$ at the cost function in (4.46), it is observed that the theoretical minimum value of $\sigma^2 N$ can't be achieved because the inequality here is strict. Moreover, the use of random phase-shifters combiner matrix, shifted the total noise power values, however, the form of figures remained the same. Note that, the resulted figures in case of random phases combiner, illustrate the average values of $\text{Tr}\{\mathbf{C}_{w''}\}$ for 500 different

random combiners.

4.5.2 Application of Full Snapshot Reconstruction

In this section, the application of the proposed scheme, under the AoA estimation performance, is evaluated versus the additive noise SNR (see Figures 4.5, 4.6) and the number of required snapshots (see Figures in 4.7) for the covariance matrix estimation. In the first two experiments, the proposed scheme, called PRE+MUSIC, will be compared with H-MUSIC [20] in the same partially- and fully-connected hybrid antenna arrays, when the Fourier matrix is utilized, and MUSIC when the conventional antenna array is used. In the last experiment demonstrated in Figure 4.8, PRE+MUSIC performance is evaluated for both DFT-based and exponentials with random phases combiner matrix. It is interesting to note that, in the case of the PRE+MUSIC (as well as for the H-MUSIC) a number of T sub-snapshots are needed in order to reconstruct a full conventional snapshot as in (4.4). Thus, to make a fair comparison with the conventional MUSIC, we have assumed in the experiments that the number S_2 of snapshots collected by MUSIC is T times the number of snapshots S_1 used by the proposed scheme and H-MUSIC.

The performance measure that is used, i.e. the RMSE (Root-Mean-Square-Error) [79] of the estimated AoAs, is defined as

$$\text{RMSE} = \sqrt{\frac{\sum_{m_c=1}^{M_c} \sum_{l=1}^L (\hat{\theta}_{m_c}(l) - \theta(l))^2}{LM_c}}, \quad (4.55)$$

where $\hat{\theta}_{m_c}(l)$ is the estimate of true AoA $\theta(l)$ in the m_c^{th} Monte Carlo trial and M_c is the total number of trials which is set to 500. Finally, all schemes, in order to estimate the desired AoAs, are employed in the angular range from -90° to 90° which is discretized with a 0.5° step.

In the context of the first two experiments in Figures 4.5, 4.6 and 4.7, first, it is assumed that the full antenna array is organized into non-overlapping subarrays of the same size $M = 4$ as in Figure 4.2. Specifically, we will demonstrate an experiment assuming that $T = M = 4$, i.e., the number of sub-snapshots is equal to the number of antennas in each subarray, and a second experiment considering the fully-connected architecture, where all antenna elements contribute to each RF chain, as shown in Fig-

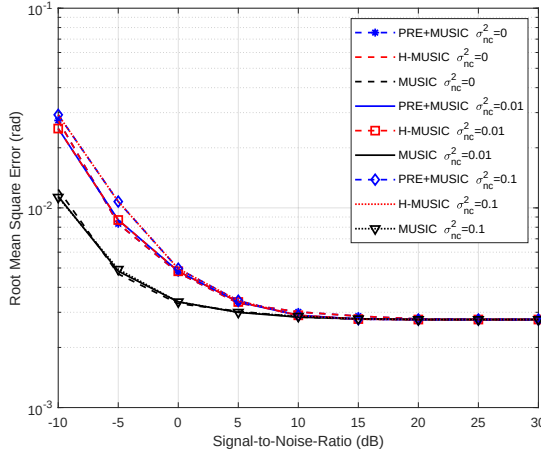


Fig. 4.5 RMSE of AoA estimation for the partially connected architecture, $M_c = 500$, $S_1 = 20$, $S_2 = 80$, $\lambda = 1cm$.

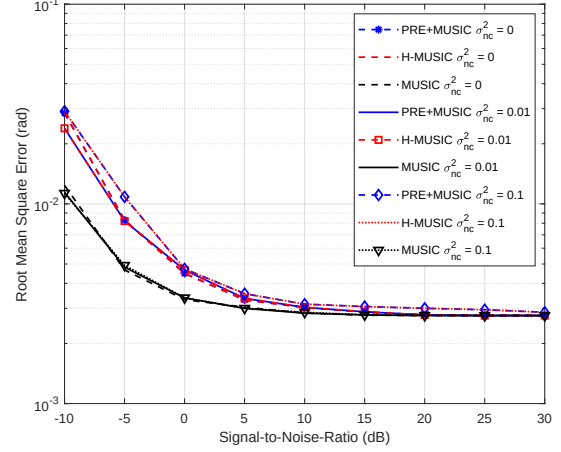


Fig. 4.6 RMSE of AoA estimation for the fully connected architecture, $M_c = 500$, $S_1 = 20$, $S_2 = 80$, $\lambda = 1cm$.

ure 4.3. In the same experiments of this section, the fully connected architecture, as in Figure 4.3, is considered as well, given that $L_r = 8$ and $T = 4$ the desired number of sub-snapshots in order to construct a full snapshot $N = 32$. In particular, in Figures 4.5, 4.6 we demonstrate the RMSE of the three schemes (i.e., PRE+MUSIC, H-MUSIC and MUSIC) versus the SNR for $\sigma_{nc}^2 = \{0, 0.01, 0.1\}$ and $S_1 = 20$, $S_2 = 80$, respectively. It can be observed that both schemes designed for hybrid antennas (i.e., PRE+MUSIC and H-MUSIC) have almost identical performance, which attains that of MUSIC for high SNR values. Also, MUSIC has the best performance at very low SNRs but, it is noted that MUSIC is applicable only to conventional ULAs and, for the same observation time, it utilizes T times more full snapshots, therefore, it is able to have a better estimation of the covariance matrix.

In the context of the second experiment in Figures 4.7, the RMSE performance versus the number of snapshots is also evaluated for $\sigma_{nc}^2 = \{0, 0.01, 0.1\}$ and $SNR = 20dB$. Obviously, as illustrated in Figures 4.7, in the case of both architectures and irrespective of the non-constant signals power σ_{nc}^2 , the classical MUSIC algorithm can achieve better AoA estimation performance than the rest schemes, as, the number of snapshots actually used is larger. Moreover, it is also observed that the RSME performance is similar for all schemes for $S_1 \geq 8$ or $S_2 \geq 32$ despite MUSIC captures greater number of snapshots than the hybrid schemes. Also, concerning the source signals time-variations (in both architectures), it is observed that all schemes are not affected

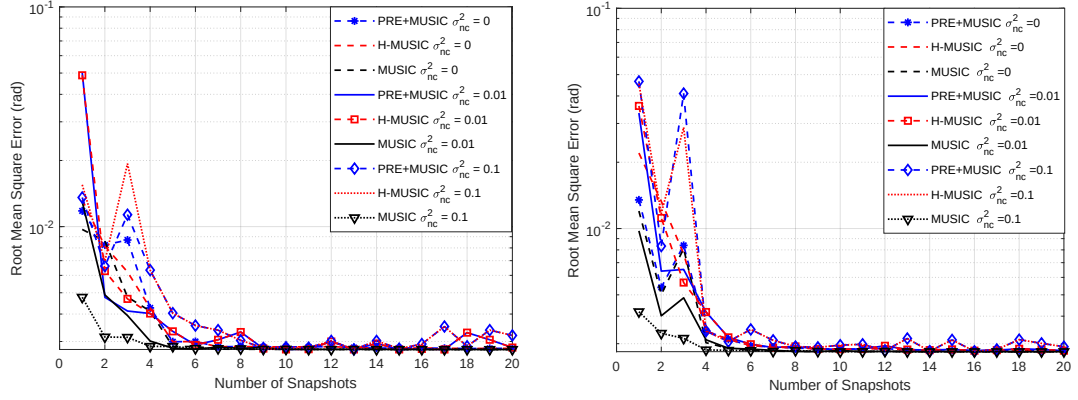


Fig. 4.7 RMSE of AoA estimation for the sub- and fully-connected architecture, $M_c = 500$, $SNR = 20dB$, $S_1 = 1 : 20$, $S_2 = 4 : 80$ snapshots ($S_2 = TS_1$ where $T = 4$), $\lambda = 1cm$ (sub- and fully-connected).

in terms of the AoA estimation and their behavior is approximately the same as in the case of constant sources.

Additionally, in the last experiment in Figure 4.8 the focus is solely on PRE+MUSIC, at which, it is depicted the impact of a random phase-shift combiner matrix in the AoA estimation. Since, the proposed scheme performance is slightly affected by the possible time variations of the source signals and the hybrid architecture type, this experiment is conducted assuming only the fully-connected architecture for constant value sources. As it was expected, the design of the combiner matrix based on DFT optimized the performance of PRE+MUSIC. Indeed, the cost function values and the lower bound impacted by the use of a random phase-shift matrix (as demonstrated in Figure 4.4), which impacted its performance under the AoA estimation problem, as well, as shown in Figure 4.8. This relates with the fact that, in the reconstructed full snapshot the noise term is not white if the $\mathcal{B}(n)$ is not unitary.

We should refer that the use of orthogonal beamforming matrices, e.g. the use of the discrete Fourier matrix achieves the optimal behavior of the performance metric (RMSE) in terms AoA estimation. Non-orthogonal beamforming matrices, e.g., random beamforming matrices can be utilized as well. Notwithstanding, the performance in terms of the AoA estimation accuracy may be considerably degraded in comparison with its optimal behavior, as shown in Figure 4.8.

Ultimately, it should be noted that, irrespective of the adopted hybrid architecture and the power of the non-constant signal component, the performance of PRE+MUSIC in terms AoA estimation, is not affected by the floors that appear in the involved cost

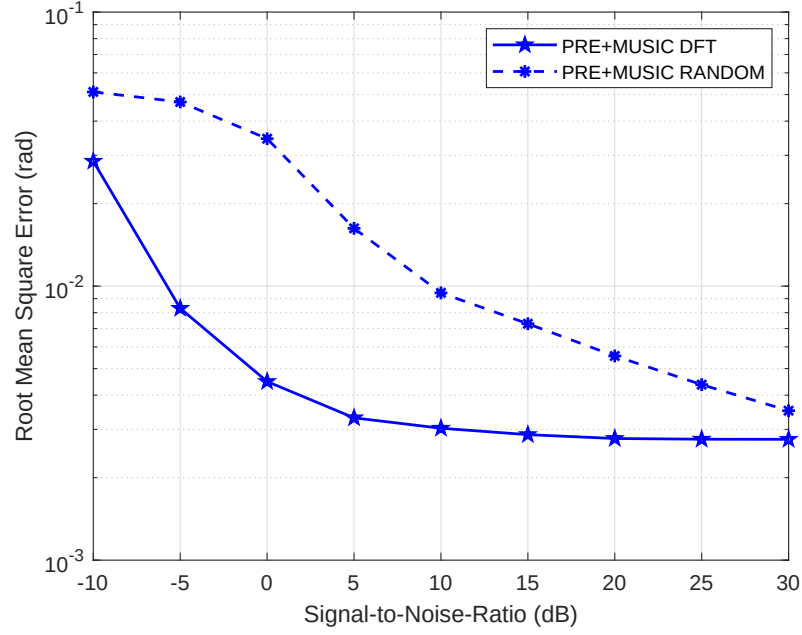


Fig. 4.8 RMSE of AoA estimation of PREMUSIC for a combiner matrix based on i) DFT and ii) Random Phase Shifters, $M_c = 500$, $S_1 = 20$, $S_2 = 80$, $\lambda = 1\text{cm}$.

function in Figure 4.4, revealing that the discrete Fourier matrix is an appropriate choice for the design of the RF combiner.

4.6 Conclusions

In summary, we investigated the recoverability of a ULA snapshot applicable to sub- and fully-connected hybrid antenna arrays. The proposed scheme is able to recover the full snapshot as if we had a conventional antenna array either for time-varying source signals or constant signals. The phase shifters of the hybrid antenna array were designed with the aim to maximize the STNR of the restored snapshot. Finally, typical simulation results have been presented for the AoA estimation problem, confirming the efficacy of the proposed approach under the optimal solution.

CHAPTER 5

mmWave Channel Estimation with Distributed Sparse Representation Techniques

5.1 Introduction

The concept of sparsity in wireless channels isn't new, but many research works have been previously published. In [82], an efficient decision feedback equalizer appropriate for channels with long and sparse impulse response was proposed. Subsequently, in [99], a general framework was suggested for the sparse adaptive equalization problem, under the compressive sensing perspective, and a heuristic scheme was derived assuming unknown the sparsity order.

In this chapter, the problem of sparse channel estimation is investigated exploiting the spatially joint sparsity structure of the involved channels to formulate the channel estimation problem in the angular domain and solve it with the employment of a fully distributed approach. The devices collaboratively estimate their channel sparsity support sets before the local estimation of the channel values, assuming the existence of global and common support subsets. The combination of the proposed distributed scheme with the classical Simultaneous Orthogonal Matching Pursuit (SOMP) algorithm is called Weighted Distributed Simultaneous Orthogonal Matching Pursuit (WDiSOMP). The performance of WDiSOMP is assessed under a multi-task scheme considering a new weighted voting method. The new mechanism is applied at each support set index and enhances its estimation accuracy and, thus, the eventual estimation of the overall channel. The mean squared error (MSE) is utilized to derive the performance bounds and assess the efficacy of the WDiSOMP estimator. Finally, the performance and the theoretical findings are evaluated via the comparison of WDiSOMP with Distributed Simultaneous Orthogonal Matching Pursuit (DiSOMP), local SOMP, and a centralized approach based on Structured SOMP (SSOMP) in terms of the MSE.

In practical scenarios, the mmWave (massive) MIMO channel is sparse in the angular domain due to the narrow-angle spread and limited scattering in the propagation environment at the base station (BS) side [77]. Thus, a few significant paths are actually active (and, essentially, fewer than the number of transmitting/receiving antennas) and contribute to the transmission. In recent years, sparse coding techniques have been widely used for channel estimation tasks [74], [37]. Sparse representation theory can be leveraged to estimate the channel with reduced training overhead by identifying the parameters involved in the commonly adopted geometric channel model (channel gains, Angles of Departure (AoD) and Angles of Arrival (AoA)), related only to those paths, rather than estimating the channel matrix. In the following, related works in the domain of channel estimation with sparse coding-based approaches and the contribution of this research will be described.

Numerous state-of-the-art solutions for mmWave channel estimation are outlined by Hassan et al. in [37]. In the relevant literature, the problem is mainly addressed either in point-to-point [77], [5, 89, 43] or in point-to-multiple points frameworks [21]. In addition, the sparse channel recovery approaches [60] can be employed either individually [5], [51] or centrally (e.g., at the BS) [77], [28], [75]. In the centralized solutions, the CS algorithms exploit the common sparsity patterns among the channels of nearby devices that are manifested in their measurements due to the transmission environment and, thus, enhance the estimation accuracy or reduce the training overhead. The downlink pilot training overhead would be prohibitively in MIMO systems with large-scale antenna arrays. Exploiting already established CS techniques and the hidden sparsity in sparse channel matrices, efficient channel estimation can be accomplished with fewer training symbols and improved performance due to the limited local scattering at the BS.

More specifically, to further enhance the estimation performance and reduce pilot demand, they leverage the so-called joint sparsity either among devices or among different sub-carriers of single devices due to the shared transmission environment. [77] is one of the first works to establish this property with devices of multiple antennas. In [28, 27, 43, 111, 48, 50], channel estimation algorithms exploit common sparsity patterns via a distributed-aided approach employed over different orthogonal frequency division multiplexing (OFDM) sub-carriers of single devices. It should be noted that

in previous algorithms for channel estimation, such as in the Joint Orthogonal Matching Pursuit (J-OMP) in [77] and the Distributed Sparsity Adaptive Matching Pursuit (DSAMP) in [28], the term “distributed” refers to geographically distributed devices that send their data to a central node for processing (e.g., the BS), or engage distributed algorithms executed, however, locally over different sub-carriers of a single device. Here, this term refers to a fully distributed channel estimation approach in which many devices collaborate and exchange information with their neighborhood. Moreover, in [22], Dai et al. impose a non-uniform size burst-sparsity structure in the sparse channel without considering any correlations among the channels of different users. They propose a generic Sparse Bayesian Learning (SBL) framework that jointly detects unknown outliers and bursts and estimates the channel. Furthermore, the authors in [21] introduce a common sparsity model, where the individual non-zero values are treated as outliers, and propose an SBL method that jointly estimates the channel and finds the users’ groups. In [57], the authors exploit common sparsity to build a sparse Bayes tensor and Direction of Arrival (DoA) tracking-based channel estimation algorithm for V2X mmWave massive hybrid MIMO systems.

Recent studies see channel estimation as multi-task, which, however, is interpreted differently from this study. Specifically, in [53], the MIMO channel prediction is based on a multi-task least square support vector machine that exploits the spatial correlations between the channels of different antenna pairs instead of predicting the channel per antenna pair independently. Also, in [106], a joint channel estimation scheme is presented for a reconfigurable intelligent surface-assisted mmWave communication system. In particular, it estimates the direct and cascaded channels at the same coherence time instead of learning them separately. Different from the above works, in [56], a multi-task deep learning-based network is proposed to simultaneously design the hybrid precoding and combining matrices under the geometric mmWave channel model. In Table 5.1, we demonstrate a brief comparative description of our work with recent previous studies. The methods listed in this table are compared, focusing on some key traits, i.e., the assumed properties of the channel, whether multi-task is considered from a sparse representation perspective, the architecture and the type of processing (centralized, distributed, local). Our method is distinct from the sparsity-induced multi-task and processing perspectives.

At this point, we would like to mention that the term multi-task was originally appeared in the relevant literature for distributed estimation as Node-Specific Parameter Estimation (NSPE), where the devices are assumed to have partially overlapping but different estimation interests. In particular, the devices are interested in estimating parameters of local, common and/or global interest. In [15],[72], adaptive LSM-based filtering techniques were exploited to derive novel distributed estimation algorithms, where the devices cooperate under either an incremental or a diffusion strategy to estimate the global and/or common interest parameters. This concept was also extended and studied under a problem in which the devices in a network, each with different data, are interested in computing a common dictionary that is suitable for the efficient sparse coding of all their data [7].

A similar multi-task framework is presented in the current chapter for the mmWave sparse channel estimation. In this thesis, the multi-task term is employed to capture the joint non-uniform sparsity manifested among the channel matrices of proximal devices, which are impacted by the same scattering clusters and thus share the same AoDs or support set indices. The AoDs and the respective support set of each device are divided into globally and/or partially common. Therefore, devices cooperate by exchanging the locally estimated support sets to jointly learn them and the respective part of the MIMO channel.

In the distributed joint sparsity recovery literature [63], [37], several algorithms have been developed for the joint sparsity models JSM-1 (common support and individual values) and JSM-2 (common support), and their generalizations, which, first, were introduced by Baron et al. in [10]. These algorithms either follow a greedy and iterative process (e.g., OMP) or solve an optimization problem (e.g., Least absolute shrinkage and selection operator - Lasso) to estimate the desired sparse vectors or matrices. Here, we will focus on a greedy method, especially OMP, due to its simplicity, fast implementation and lower computational complexity than optimization techniques. However, other sparse coding algorithms can be used, such as matching pursuit (MP) [112]. The distributed versions of subspace pursuit (SP) and orthogonal matching pursuit (OMP), the so-called DiSP and DiOMP, respectively, were first introduced in [90], and they are applicable to both JSM-1 and JSM-2 models (assuming a priori knowledge of the sparsity structure). Besides, two versions of Decentralized and Collaborative

Table 5.1 Comparison with Previous Works

Work	Channel model	Multi-task sparsity induced	Architecture	Processing
[22]	geometric, clustered, multi-path	no	MISO, Multi-user	local
[21]	geometric, clustered, multi-path	no	MISO, Multi-user	centralized
[57]	geometric, clustered, single-path	no	MIMO, Multi-user	local
[106]	geometric, clustered, multi-path	no	MISO, Multi-user	local
[56]	geometric, clustered, single-path	no	MIMO, Single-user	local
[53]	non-parametric, MIMO channel description	no	MIMO, Single-user	local
proposed	geometric, clustered, multi-path	yes	MIMO, Multi-user	distributed

Orthogonal Matching Pursuit (DC-OMP1, DC-OMP2) for JSM-1 have been suggested in [103], taking into account some communication constraints. Compared to DiOMP, DC-OMP 1 recovers the support set with higher accuracy by utilizing a voting mechanism incorporated in the iterations of the OMP. DC-OMP 2, which is based on SOMP, is much more accurate than DC-OMP 1 but at the cost of a higher communication burden. It is noted that in the framework of distributed collaborative algorithms, our recent works [92], [93] (which will present in the following sections) constitute a state of the art solutions for a multi-task joint sparse representation model. Compared to their predecessors, they follow an adaptive process via the employment of adaptive weights majority voting and only need to know the total sparsity level.

5.2 Motivation

The constantly increasing use of smart devices and multimedia services, as well as the demand for high data rates, high energy efficiency, low latency and high quality

of service (QoS), have made fifth-generation (5G) networks and beyond more imperative than ever. 5G networks address the emerging challenges by considering i) massive MIMO systems in the mmWave range, ii) full-duplex (FD) transmissions, iii) D2D communication protocols and iv) network densification with heterogeneous networking architectures, where various types of cells (e.g., small, large) and access technologies are expected to coexist. In the near future, direct D2D communication is regarded as an essential and promising technology, e.g., for extending capacity, enabling scalability and traffic offloading in various short-range applications, including IoT, M2M, and emergency networks [8]. In the vision of 5G networks, users organized into small groups will exchange information, e.g., about the network and channel status, cooperate in various operations of common interest and employ new technologies in favour of coordinated communication [97].

A thorough search of relevant literature shows that the trend toward distributed and cooperative schemes also emerges in the physical layer, with some indicative, yet notable, examples being beamforming [47, 12] and spectrum sensing [38, 91]. For the first time, the present chapter focuses on the problem of fully-distributed channel estimation for mmWave MIMO systems and proposes novel and efficient schemes that enable devices to cooperate for estimating the involved channels.

5.3 The main contribution

In this chapter, a *fully* distributed channel estimation algorithm is presented that efficiently undertakes the multi-task commonalities among the sparsity patterns of adjacent devices. In particular, the main highlights of this research are the following:

- A multi-task framework is elaborated under which the mmWave MIMO channel estimation problem is modelled and investigated. This framework exploits the inherent similarities among the different tasks assuming collaboration among the devices in the network. In particular, the joint channel sparsity among proximal devices surrounded by global and some common scattering clusters, is taken into account to develop methods with improved channel estimation accuracy and reduced training overhead.
- WDiSOMP has been enhanced in order to be more robust when non-uniform

overlapping sparsity groups occur in the network, incorporating a new voting mechanism. Each device applies the new adaptive weights mechanism in order to create suitable combinations of the received support sets of its neighbors, from which the final support set is retrieved. The employed mechanism estimates an individual weight per support set index, not a global one [93]. Our prior scheme determines the weight with which each neighbor participates in the combination part of the mechanism, based on the total number of different elements among the devices' support set. Here, the updated mechanism finds the devices' similarity per index and, thus, can easily identify any overlaps that may occur among the sparsity patterns of proximal devices in a multi-tasking scheme. Also, the deployment of the new voting mechanism in the proposed algorithm further improves the channel estimation accuracy and reduces its uses for training by the BS. Due to the adaptive weights voting, each device only needs to know the number of departing paths or angles.

- The proposed approach's required iterations and overall complexity are discussed, analyzing the computational cost that arises at each stage of the algorithm per communication round. This cost is compared with the algorithms' considered in the simulations.
- The novel approach trades off the efficient channel estimation due to collaboration using an adaptive weights voting mechanism with a communication cost, in contrast to adopting centralized processing at BS. The transmission overhead of the distributed schemes (W)DiSOMP, in relation to centralized SSOMP, is quantified and demonstrated by simulation. This cost in the distributed and collaborative methods depends on the AoDs dictionary resolution and the total sparsity level, which, in the mmWave, is maintained low. On the other hand, the transmission overhead of the centralized approach is subjected to the receiver's antenna array size, the training overhead, the number of devices and the feedback quantization bits.
- We would like to note that from a transmission overhead perspective, the distributed approaches, such as the one presented here, might be preferable in many cases encountered in practice (e.g., cases with low sparsity appearing in the mmWave

range and / or multi-antenna end-users).

- A theoretical performance analysis is conducted where the MSE is analytically computed, assuming an oracle estimator as a benchmark to evaluate how well the suggested estimator behaves. Our purpose is to find the minimum error values of that and identify the parameters that most affect our solution to attain it. Also, given that an overcomplete basis is considered, we aim to determine the properties that impact its quality and, thus, the correct sparse channel recovery via the proposed estimator. The latter analysis will help us to determine an upper error limit, which, as it will be seen, depends on the mutual coherence of the involved dictionary and measurement matrix and the total sparsity level.
- Also, the proposed algorithm is general-purpose since it can be applied to any other estimation problem where the underlying joint sparsity can be exploited (i.e., not only to the mmWave channel estimation).
- Finally, the evaluation of the algorithm is demonstrated via simulations, and it turns out that it performs significantly better than the existing ones. Several experiments have been conducted based on the elaborated multi-task sparsity model to evaluate the channel matrix recovery performance in terms of the number of measurements (i.e., training overhead) and noise conditions (i.e., Signal-to-Noise-Ratio (SNR)).

The rest of this chapter is organized as follows. In Section 5.4, the data model is described. In Section 5.5, the joint sparse channel model, under which our estimator will be designed, is elaborated in detail. Then, Section 5.6 presents the problem formulation for the mmWave channel estimation. In the following, Section 5.7 shows the proposed distributed MIMO downlink channel estimation algorithm with the aid of the CS theory. Moreover, Section 5.8 analyzes the computational complexity (i.e., the required iterations) of the proposed method, the respective transmission overhead, and the impact of support set and resolution on reconstruction quality. In Section 5.9, several experiments are presented to verify the efficiency of the proposed approach, accompanied by the appropriate discussion.

5.4 The System Model

In this paper, we will focus on downlink transmission. BS of N_t antennas communicates with K devices of N_r antennas to transmit (broadcasts) T training symbol vectors which are concatenated in $\mathbf{X} \in \mathbb{C}^{N_t \times T}$ in T time slots via a narrowband, block-fading mmWave MIMO channel.

5.4.1 Single Antenna Devices

In this subsection, it is provided the system model assuming K single antenna devices. The downlink received signal $\mathbf{y}_k \in \mathbb{C}^{T \times 1}$ at the k -th device is given by

$$\mathbf{y}_k = \mathbf{X}^T \mathbf{h}_k + \mathbf{n}_k, \quad (5.1)$$

where $\mathbf{h}_k \in \mathbb{C}^{N_t \times 1}$ and $\mathbf{n}_k \in \mathbb{C}^{T \times 1}$ stands for additive noise which is modeled as random vector with elements being i.i.d (independent and identically distributed) as complex Gaussian random variables with zero mean and variance equal to σ^2 .

Assuming the Geometry-Based Stochastic Channel Model (GSCM) [23], after dropping the index k for clarity, the channel can be written as

$$\mathbf{h} = \sum_{c=1}^{N_c} \sum_{s=1}^{N_{c,s}} \gamma_{c,s} \mathbf{a}(\phi_{c,s}), \quad (5.2)$$

where N_c and $N_{c,s}$ stand for the number of scattering clusters and the number of sub-paths per scattering cluster, respectively, $\gamma_{c,s}$ is the complex gain of the s -th sub-path in the c -th scattering cluster, $\theta_{c,s}$ is the corresponding AoD. From now on, we denote the true AoDs as $\theta_l, l = 1, 2, \dots, L_{AoD}$, with $L_{AoD} = N_c N_{c,s}$.

The channel model is simplified to

$$\mathbf{h} = \sum_{l=1}^{L_{AoD}} \gamma_l \mathbf{a}(\phi_l), \quad (5.3)$$

where L_{AoD} stands for the total number of departure paths, γ_l is the complex gain that captures the attenuation due to scattering clusters of the l -th path, ϕ_l is the corresponding AoD and

$$\mathbf{a}(\phi_l) = \frac{1}{\sqrt{N_t}} \left[1, e^{j \frac{2\pi d}{\lambda} \sin \phi_l}, \dots, e^{j \frac{2(N_t-1)\pi d}{\lambda} \sin \phi_l} \right]^T \quad (5.4)$$

is a steering vector, assuming a Uniform Linear Array (ULA).

The channel can be approximated by $\mathbf{h} \approx \mathbf{D}_t \mathbf{h}^v$, where $\mathbf{h}^v$ is a sparse vector and $\mathbf{D}_t \in \mathbb{C}^{N_t \times G_t}$ is a sparsifying dictionary [23]. A typical $\mathbf{D}_t$ is the normalized square DFT matrix which maps the spatial channel response to the angular domain. Alternatively, an overcomplete dictionary $\mathbf{D}_t$ can be employed via the overcomplete DFT matrix, where $G_t \gg N_t$ is the number of atoms in $\mathbf{D}_t$ corresponding to the AoDs in a finer grid. In this case, $\mathbf{h}^v \in \mathbb{C}^{G_t \times 1}$ is able to approximate more closely the true angles at ϕ_l . Thus, $\mathbf{y}_k$ can be written as

$$\mathbf{y}_k \approx \Phi \mathbf{h}_k^v + \mathbf{n}_k, \quad (5.5)$$

where $\Phi = \mathbf{X}^T \mathbf{D}_t \in \mathbb{C}^{T \times G_t}$ is the measurement matrix.

5.4.2 Multi-Antenna Devices

Each device k , with $k = 1, 2, \dots, K$, receives the signal

$$\mathbf{Y}_k = \mathbf{H}_k \mathbf{X} + \mathbf{N}_k, \quad (5.6)$$

where matrix $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ captures the mmWave downlink MIMO channel from the BS to the k -th device and $\mathbf{Y}_k \in \mathbb{C}^{N_r \times T}$, $\mathbf{N}_k \in \mathbb{C}^{N_r \times T}$ denotes the concatenated received signal and additive noise vectors, respectively. The elements of the latter are modeled as random vectors with elements being i.i.d. and complex Gaussian distribution with zero mean and variance equal to σ^2 .

Since mmWave channels are expected to have limited scattering on the BS side, a geometric channel model is adopted where each cluster of scatterers is assumed to contribute multiple propagation paths between the BS and the k -th device. Under this model, the physical channel matrix $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ between the BS and each device can be given as [95]:

$$\mathbf{H} = \sqrt{\frac{N_t N_r}{L_{AoD} L_{AoA}}} \sum_{p=1}^{L_{AoD}} \sum_{q=1}^{L_{AoA}} \alpha_{p,q} \mathbf{a}_r(\theta_{p,q}) \mathbf{a}_t^H(\phi_p), \quad (5.7)$$

where L_{AoD} represents the number of AoDs each of which contributes L_{AoA} arrival paths on the device side, $\alpha_{p,q}$ is the complex gain of the q -th path on the device side that

departs from the p path on the BS side, $\phi_p, \theta_{p,q}$ are the p -th AoD and its corresponding q -th AoA, and $\mathbf{a}_t(\phi_p), \mathbf{a}_r(\theta_{p,q})$ are the antenna array response vectors at the BS and devices, respectively. The matrix form of the channel model in (5.7) can be written as

$$\mathbf{H} = \mathbf{A}_r \mathbf{H}_a \mathbf{A}_t^H, \quad (5.8)$$

where $\mathbf{A}_t = [\mathbf{a}_t(\phi_1), \mathbf{a}_t(\phi_2), \dots, \mathbf{a}_t(\phi_{L_{AoD}})]$, $\mathbf{A}_r = [\mathbf{a}_r(\theta_1), \mathbf{a}_r(\theta_2), \dots, \mathbf{a}_r(\theta_{L_{AoA}})]$, $\mathbf{H}_a \in \mathbb{C}^{L_{AoA} \times L_{AoD}}$ is a sparse matrix whose non-zero complex values are associated with the channel path gains.

Assuming Uniform Linear Arrays (ULAs), the transmit response vector $\mathbf{a}_t(\phi_p)$ is defined as

$$\frac{1}{\sqrt{N_t}} \begin{bmatrix} 1 & e^{j\frac{2d\pi}{\lambda} \cos \phi_p} & \dots & e^{j\frac{2(N_t-1)d\pi}{\lambda} \cos \phi_p} \end{bmatrix}^T, \quad (5.9)$$

while the receive response vector $\mathbf{a}_r(\theta_{p,q})$ is given as

$$\frac{1}{\sqrt{N_r}} \begin{bmatrix} 1 & e^{j\frac{2d\pi}{\lambda} \cos \theta_{p,q}} & \dots & e^{j\frac{2(N_r-1)d\pi}{\lambda} \cos \theta_{p,q}} \end{bmatrix}^T, \quad (5.10)$$

where λ is the signal wavelength at mmWave frequency and d the inter-element spacing.

Since the quality of CS-based channel estimation is limited by the quantization error in the angular domain representation, the AoDs and AoAs (or the corresponding atoms) are taken from high-density grids of $G_r \gg L_{AoA}$ and $G_t \gg L_{AoD}$ points, obtaining overcomplete dictionaries $\mathbf{D}_t \in \mathbb{C}^{N_t \times G_t}$, $\mathbf{D}_r \in \mathbb{C}^{N_r \times G_r}$ [95], correspondingly. Then, the channel matrix $\mathbf{H}$ can be represented in terms of a non-diagonal sparse matrix $\mathbf{H}^v \in \mathbb{C}^{G_r \times G_t}$, containing the path gains of the quantized angles, as follows

$$\mathbf{H} = \mathbf{D}_r \mathbf{H}^v \mathbf{D}_t^H. \quad (5.11)$$

The representation in (5.11) constitutes a discretized approximation of the channel matrix that can be exploited to cast the estimation of $\mathbf{H}$ to that of detecting some non-zero coefficients in the sparse channel matrix $\mathbf{H}^v$. Focusing on the rows of the channel matrix $\mathbf{H}^v$, they will contain only a few non-zero values corresponding to the AoDs, due to the sparsity of mmWave channels on the BS side. Notice that the number of non-zero values in each of those columns is exactly equal to L_{AoA} . In the special case of ULAs, with $d = \frac{\lambda}{2}$ and quantized AoDs and AoAs selected from uniform grids of resolution

equal to transmitting and receiving antennas ($G_t = N_t$ and $G_r = N_r$), respectively, the dictionaries $\mathbf{D}_r \in \mathbb{C}^{N_r \times N_r}$ and $\mathbf{D}_t \in \mathbb{C}^{N_t \times N_t}$ result in unitary Discrete Fourier Transform (DFT) matrices. This case is known as the *Virtual Channel Model*[62].

In the following section, a more general joint sparsity model than the one that has been established in the literature will be elaborated on, before presenting the proposed collaborative distributed approach. The elaborated model captures the different tasks that occur in the network, which in [67] is called multitasking with overlapping parameters.

5.5 Joint Sparse Channel Modeling

A set of K geospatially distributed devices are interconnected by a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, 2, \dots, K\}$ is the set of devices (or nodes) and $\mathcal{E} = \{(i, j) : \text{Node } i \in \mathcal{V} \text{ is connected with node } j \in \mathcal{V}\}$.

As illustrated in Figure 5.1, the scattering nature of the transmission environment may lead to the devices grouping according to the scatterers getting involved in their channels. As stated in [77], the devices that belong to the same group may face and thus be affected by the same scatterer. It means that the sparse channel matrices $\mathbf{H}_k^v$ share the same sparsity support set on the BS side, but the corresponding path gains may differ due to the surrounding area of each device. To improve the downlink channel estimation performance, we exploit the previously described relationship on the sparsity support sets among the involved channels of the devices [77], [28].

More specifically, the scatterers on the BS side can be clustered into global and common, from now on referred to as g and c_j , which translates the corresponding AoDs into global to all devices in the network and common to particular subgroups. From this observation, we conclude that both the columns (otherwise atoms) of the employed dictionary $\mathbf{D}_t$ are also global and common, respectively.

In the following, $\mathbf{I}_k$ is a set of indices $j = \{1, 2, \dots, J\}$ indicating the maximum number of common interest support sets per device k . Also, $\Omega = \{1, 2, \dots, G_t\}$ denotes the set of atom indices (namely, the number of AoDs) in the overcomplete dictionary $\mathbf{D}_t$.

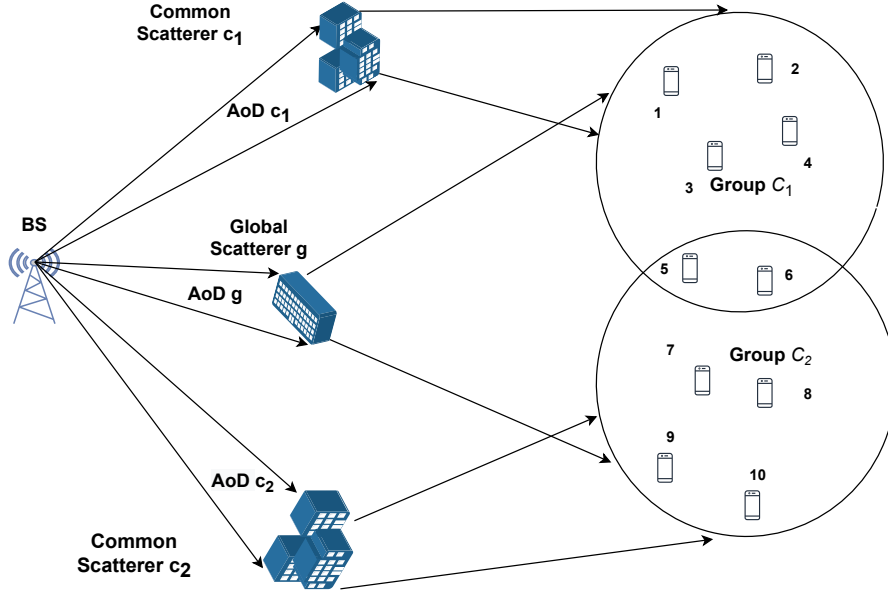


Fig. 5.1 Downlink transmission scenario: $\mathcal{C}_1 = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{C}_2 = \{5, 6, 7, 8, 9, 10\}$ [92].

5.5.1 Single Antenna Devices

Adopting the single antenna sparsity model, the sparse representation vector $\mathbf{h}_k^v$ of device k can be formulated as

$$\mathbf{h}_k^v = \mathbf{h}_k^g + \sum_{j \in I_k} \mathbf{h}_k^{c_j}, \quad (5.12)$$

where $\mathbf{h}_k^g$ stands for the sparse representation vector whose support set is globally shared among all devices in the network and $\mathbf{h}_k^{c_j}$ stands for the sparse representation vector whose support set is commonly shared among devices in a particular group in the network.

Moreover, it is assumed that $\mathbf{h}_k^g$ of each device k consists of L_g non-zero parameters and $\mathbf{h}_k^{c_j}$'s consist of L_{c_j} non-zero parameters. Hence, the sparsity profile of a device k consists of global and multiple common interest support subsets.

Definition 5.5.1. (*Global Interest Support Set*): Let the globally shared sparse representation vector $\mathbf{h}_k^g$ with only L_g nonzero entries. The global interest support set is defined as

$$\mathcal{S}^g = \text{supp}(\mathbf{h}_m^g) = \text{supp}(\mathbf{h}_n^g), \forall m, n \in \mathcal{V}, \quad (5.13)$$

where $\mathcal{S}^g \subset \Omega$ and $|\mathcal{S}^g| = L_g$ (l_0 sparsity).

Definition 5.5.2. (*Common Interest Support Set*): Let the sparse representation vector

$\mathbf{h}_k^{c_j}$ with only L_{c_j} nonzero entries. The common interest support set is defined as

$$\mathcal{S}^{c_j} = \text{supp}(\mathbf{h}_m^{c_j}) = \text{supp}(\mathbf{h}_n^{c_j}), \forall m, n \in \mathcal{C}_j, \quad (5.14)$$

where $\mathcal{S}^{c_j} \subset \Omega$, $|\mathcal{S}^{c_j}| = L_{c_j}$ (l_0 sparsity) and $\mathcal{C}_j$ a set of devices indices that are impacted by the common scattering cluster c_j , for $j = 1, 2, \dots, J$, and are concerned about $\mathcal{S}^{c_j}$.

The channel $\mathbf{h}_k$ of device k is written as

$$\mathbf{h}_k = \mathbf{D}_t(\mathbf{h}_k^g + \sum_{j \in I_k} \mathbf{h}_k^{c_j}). \quad (5.15)$$

5.5.2 Multi-Antenna Devices

Following [63], the physical channel matrix structural sparsity is discriminated into

1. **Sparsity Support within Channel Matrix:** The rows within $\mathbf{H}_k^v$ usually have the same column sparsity support, namely there exists support set $\mathcal{S}_k$ in $\mathbf{H}_k^v$, such that $\mathcal{S}_k = \text{supp}(\mathbf{H}_k^v)$, $|\mathcal{S}_k| = L_{AoD}$.
2. **Global and Multiple Common Supports between Different Channel Matrices:** The channel matrices of different devices are correlated, especially when the devices are physically close to each other. Based on this condition, different devices and their respective (sparse) channel matrices tend to share some global and, partially, common local scatterers on the BS part.

Based on the previously established sparsity cases, the sparse representation matrix $\mathbf{H}_k^v$ of device k can be written as

$$\mathbf{H}_k^v = \mathbf{H}_k^g + \sum_{j \in I_k} \mathbf{H}_k^{c_j}, \quad (5.16)$$

where $\mathbf{H}_k^g$ represents the sparse matrix whose row support set is globally shared among all devices in the network and $\mathbf{H}_k^{c_j}$ captures the sparse matrix whose support set is commonly shared among different subgroups of devices in the network.

Moreover, it is assumed that the rows of $\mathbf{H}_k^g$ and $\mathbf{H}_k^{c_j}$ s of each device k share L_g non-zero columns and L_{c_j} non-zero columns, respectively. So, the sparsity pattern of a

device k is divided into global and multiple common interest support sets. The column sparsity is induced by the assumption that the resolution of the departing angular grid points is much greater than the number of the active transmission paths, i.e., $L_{AoD} \ll G_t$.

In the following, some definitions will be provided.

Definition 5.5.3. (*Global Interest Support Set*): Let the sparse representation matrix $\mathbf{H}_k^g$, whose columns are globally shared with only L_g nonzero entries. The global interest support set is defined as

$$\mathcal{S}^g = \{i \in \Omega, s.t. \text{supp}(\mathbf{H}_m^g) = \text{supp}(\mathbf{H}_n^g)\}, \forall m, n \in \mathcal{V} \quad (5.17)$$

where $\mathcal{S}^g \subset \Omega$ and $|\mathcal{S}^g| = L_g$ (l_0 sparsity).

Definition 5.5.4. (*Common Interest Support Set*): Let the sparse representation matrix $\mathbf{H}_k^{c_j}$ with only L_{c_j} nonzero entries. The common interest support set is defined as

$$\mathcal{S}^{c_j} = \{i \in \Omega, s.t. \text{supp}(\mathbf{H}_m^{c_j}) = \text{supp}(\mathbf{H}_n^{c_j})\}, \forall m, n \in \mathcal{C}_j \subset \mathcal{V}, \quad (5.18)$$

where $\mathcal{S}^{c_j} \subset \Omega$, $|\mathcal{S}^{c_j}| = L_{c_j}$ (l_0 sparsity) and $\mathcal{C}_j$ a set of devices indices that are impacted by the common scattering cluster c_j , for $j = 1, 2, \dots, J$, and are concerned about $\mathcal{S}^{c_j}$.

Based on the previous definitions, the support set $\mathcal{S}_k$ of device k is formulated as $\mathcal{S}_k = \mathcal{S}^g \cup \{\mathcal{S}_k^{c_j}\}$, for all $j \in \mathbf{I}_k$, meaning that the joint sparsity $L_{AoD,k} \leq L_g + \sum_{j \in \mathbf{I}_k} L_{c_j}$. Here, despite $\mathcal{S}^g$ being globally shared among all devices in the network or $\mathcal{S}^{c_j}$ being the same among particular devices that are in the same subgroup $\mathcal{C}_j$, this doesn't hold for the respective non-zero values of $\mathbf{H}_k^g$ and $\mathbf{H}_k^{c_j}$ which remain individual and possibly independent among the devices. Thus, taking into account the aforemen-

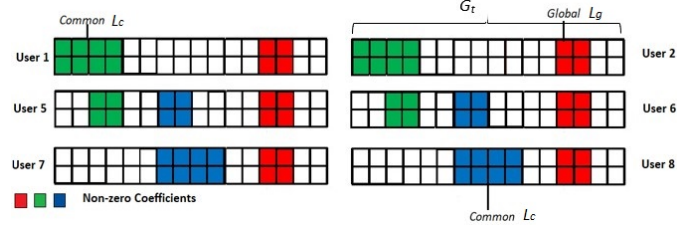


Fig. 5.2 Sparse Channel matrices of the sided devices $\bar{\mathbf{H}}_1^v, \bar{\mathbf{H}}_2^v, \bar{\mathbf{H}}_7^v, \bar{\mathbf{H}}_8^v$ and devices in the overlapping area $\bar{\mathbf{H}}_5^v, \bar{\mathbf{H}}_6^v$ assuming $N_r = 2$, $G_r = N_r$ and $G_t = 256$ where red, green and blue bins stand for the global and partially common atoms with total sparsity level $L_{AoD} = 6$.

tioned sharing of support sets, the channel $\mathbf{H}_k$ of device k can be written as

$$\begin{aligned}
 \mathbf{H}_k &= \mathbf{D}_r (\mathbf{H}_k^g + \sum_{j \in I_k} \mathbf{H}_k^{c_j}) \mathbf{D}_t^H \\
 &= (\mathbf{D}_r \mathbf{H}_k^g + \sum_{j \in I_k} \mathbf{D}_r \mathbf{H}_k^{c_j}) \mathbf{D}_t^H \\
 &= ((\bar{\mathbf{H}}_k^g)^H + \sum_{j \in I_k} (\bar{\mathbf{H}}_k^{c_j})^H) \mathbf{D}_t^H \\
 &= (\bar{\mathbf{H}}_k^v)^H \mathbf{D}_t^H
 \end{aligned} \tag{5.19}$$

From (5.19), we see that the column sparsity in $\mathbf{H}_k^v$ translates into row sparsity in $(\bar{\mathbf{H}}_k^v)^H = \mathbf{D}_r \mathbf{H}_k^v$. According to the adopted formulation, $\mathbf{D}_r$ is hidden in the sparse matrix $\bar{\mathbf{H}}_k^v$ which will be the matrix of interest in the proposed algorithm. Here, the columns of $\mathbf{D}_r$ are fixed to L_{AoA} , namely, as if we had the receiver response matrix $\mathbf{A}_r \in \mathbb{C}^{N_r \times L_{AoA}}$ and the joint sparsity is only observed in the AoDs.

Considering a scenario of multiple devices, the AoDs of the devices' channels impacted by the same scattering clusters are similar [21], [77]. This means that they share the same non-zero columns (i.e., global and common, respectively) in their sparse channels, as shown in Figure 5.2, which motivated us to investigate the channel estimation via a distributed algorithm that estimates the shared support sets in a collaborative manner.

In summary, considering (5.19), our aim is to exploit the joint sparsity sets of neighboring devices to estimate the per device sparse channel in a collaborative and distributed manner. Such an approach will be elaborated on in the following.

5.6 Multi-task mmWave Channel Estimation

In this section, the problem formulation for the multi-task mmWave channel estimation will be presented. Given the geometric channel model representation in (5.19), the problem of interest will be addressed as a sparse reconstruction assuming an overcomplete $\mathbf{D}_t$ for the AoDs sparsifying matrix.

Employing per device k the representation (5.19), we can rewrite the measurements model in (5.6), in compact form, as

$$\bar{\mathbf{Y}}_k = \Phi \bar{\mathbf{H}}_k^v + \bar{\mathbf{N}}_k, \quad (5.20)$$

where $\bar{\mathbf{Y}}_k = [\bar{\mathbf{y}}_{k,1} | \bar{\mathbf{y}}_{k,2} | \dots | \bar{\mathbf{y}}_{k,N_r}] \in \mathbb{C}^{T \times N_r}$, $\Phi = \mathbf{X}^H \mathbf{D}_t \in \mathbb{C}^{T \times G_t}$ and $\bar{\mathbf{N}}_k = (\mathbf{N}_k)^H$. Also, each column of $\bar{\mathbf{Y}}_k$, which constitutes the Single Measurement Vectors (SMV) analogue, is captured by

$$\bar{\mathbf{y}}_{k,r} = \Phi \bar{\mathbf{h}}_{k,r}^v + \bar{\mathbf{n}}_{k,r}, \quad r = 1, 2, \dots, N_r. \quad (5.21)$$

According to the Multiple Measurement Vectors (MMV) model [66], in (5.20), each device k exploits simultaneously the low dimensional measurement vectors, $\{\bar{\mathbf{y}}_{k,r}\}_{r=1}^{N_r}$, from all antenna elements, that share the same measurement matrix Φ and sparsity support sets $\mathcal{S}_k$, to recover the high dimensional $G_t \times 1$ sparse channel vectors between all receiving antennas and BS.

Let $\{\bar{\mathbf{H}}_1^v, \bar{\mathbf{H}}_2^v, \dots, \bar{\mathbf{H}}_K^v\}$ be the collection of parameter matrices across the network. Through (5.20), the objective of each device k is to determine the sparse AoDs, which is equivalent to estimating the rows support set of $\bar{\mathbf{H}}_k^v \in \mathbb{C}^{G_t \times N_r}$. Ignoring the joint sparsity channel modelling, each device k could minimize $\mathcal{J}_k(\bar{\mathbf{H}}_k^v) = \|\bar{\mathbf{Y}}_k - \Phi \bar{\mathbf{H}}_k^v\|_F^2$ on its own, locally (without cooperation) or the aggregated optimization function $\sum_{k=1}^K \mathcal{J}_k(\bar{\mathbf{H}}_k^v)$ could be used in a centralized manner as in [111]. However, there is a third direction that leads to a fully distributed approach which we follow in this chapter. More specifically, by taking into account the joint row sparsity between the $\bar{\mathbf{H}}_k^v$ s of different devices, individual but related cost functions $\mathcal{J}_k(\bar{\mathbf{H}}_k^v)$ can be adopted, thus, allowing us to formulate the problem as a multi-task one [67]. The relatedness lies in the fact that the support set of a device k is $\mathcal{S}_k = \mathcal{S}^g \cup \{\mathcal{S}_k^{c_j}\}$, e.g. with global and partially common

elements. The common elements with the $\mathcal{S}_l$ s of its neighborhood $\mathcal{L}_k^{in}$ will be taken into consideration in the collaboration step of the proposed approach where the (joint) support set is estimated through adaptive weights voting. Hence, the optimization problem can be formulated locally as

$$\underset{\mathcal{S}_k, \bar{\mathbf{H}}_k^v}{\text{minimize}} \quad \mathcal{J}_k(\bar{\mathbf{H}}_k^v) \quad (5.22)$$

$$\text{subject to} \quad \mathcal{S}_k \cap \mathcal{S}_l = \mathcal{S}^g \cup \mathcal{S}_k^{c_j}, \forall l \in \mathcal{L}_k^{in} \quad (5.23)$$

$$|\mathcal{S}_k| = \|\text{diag}((\bar{\mathbf{H}}_k^v)^H \bar{\mathbf{H}}_k^v)\|_0 = L_{AoD} \quad (5.24)$$

Accurate retrieval of the sparse channel matrix $\bar{\mathbf{H}}_k^v$ is only guaranteed when the measurement matrix Φ satisfies specific properties from CS like the Restricted Isometry Property (RIP) and mutual coherence. As this matrix is dependent on both the training symbols and the transmit (overcomplete) sparsifying matrix $\mathbf{D}_t$, the symbols can be properly designed to acquire a more optimized measurement matrix to ensure that the channel is reconstructed successfully with a high probability. Such issues have been addressed, for example, in [51], [23], however, here, the symbols are selected randomly, which still makes the measurement matrix “well-behaved”.

5.7 Voting-based Multi-task Distributed Solution

In this section, the main stages of the proposed algorithm, WDiSOMP, are analysed. Since the dictionary $\mathbf{D}_t$ in (5.19) is predefined, the channel estimation is equivalent to the estimation of $\bar{\mathbf{H}}_k^v$ in (5.20), assuming, in the transmission environment, the existence of sparsity division captured in (5.16) or equivalently in (5.19). More specifically, WDiSOMP was designed to estimate the sparse channel in a collaborative distributed manner. The collaboration stage concerns only the support sets $\mathcal{S}_k$ s, while the non-zero values of the path gains are individual and, thus, estimated via Least-Squares (LS) locally. Our solution employs a weighted majority voting mechanism with adaptive weights that can be estimated either at the support set or index level. With the incorporation of this step, the proposed algorithm only needs to know the total number of departing paths or angles L_{AoD} .

5.7.1 Proposed Solution

The algorithm operates in stages (see Algorithm 2). The first stage initializes the process, and the second one is executed iteratively, consisting of a collaborative (through voting) and a local step (where the SOMP runs to make an individual estimate of the sparsity support set) and the final one that concerns channel estimation is executed locally.

In the *Initialization* stage, the BS, for T time slots, transmits to all devices in the network a number of training symbols, represented by matrix $\mathbf{X}$. Then, each device k collects all measurements in $\mathbf{Y}_k$ and executes SOMP (assuming the overcomplete sparsifying matrix Φ and an empty support set) to get an initial local estimate of its complete support set $\hat{\mathcal{S}}_k$.

In the *Iterative* stage, each device k , $\forall k$, transmits its $\hat{\mathcal{S}}_k$ to its neighbors denoted by the set $\mathcal{L}_k^{out}$, and receives the estimates $\hat{\mathcal{S}}_l$'s from its neighbors, $\forall l \in \mathcal{L}_k^{in}$ (*communication step*). At iteration i of the algorithm, each device exploits the received $\hat{\mathcal{S}}_l$'s and engages a *weighted majority voting* mechanism to determine the i most voted indices. Then, SOMP is executed locally to acquire the remaining non-zero indices, up to L_{AoD} , and again $\hat{\mathcal{S}}_k$ s are communicated over the network. These steps are repeated for L_{AoD} iterations. It is assumed, here, that all devices communicate via a connected network and can exchange their support sets indices by engaging in a suitable D2D communication protocol [46]).

Each device k exploits the estimated set $\hat{\mathcal{S}}_k$ on (5.20) to retain only these specific columns of Φ , apply LS to find the associated non-zero values of $\hat{\mathbf{H}}_k^v$ and through (5.19) its channel matrix.

The above stages and the corresponding steps are also illustrated in Figure 5.3, where the information flow is depicted per device k . The BS broadcasts the same training data to all devices in the network. Then, each device exploits the acquired measurements to initialize the process and iteratively (via sharing - collaboration) estimate the desired support set. Finally, each device k utilizes the estimated support set to recover the sparse channel representation and, thus, the channel matrix.

Algorithm 2: WDiSOMP**Result:** $\{\hat{\mathbf{H}}_k\}$ **Input:** $\Phi = \mathbf{X}^H \mathbf{D}_t$, $\bar{\mathbf{Y}}_k = (\mathbf{Y}_k)^H$; L_{AoD} , $b_{lk,-1} = 0, \forall k \in \{1, \dots, K\}$ and
 $l \in \mathcal{L}_k^{in}$;**Stage I: Initialization Part***Local step***for** $k = 1:K$ **do**

$$\left| \begin{array}{l} [\hat{\mathcal{S}}_{k,0}] = \text{SOMP}(\Phi, \bar{\mathbf{Y}}_k, L_{AoD}, \emptyset); \end{array} \right.$$
end*Transmit:* $\hat{\mathcal{S}}_{k,0}$ to all nodes $\mathcal{L}_k^{out}$ *Receive:* $\hat{\mathcal{S}}_{l,0}$ s from all nodes $\mathcal{L}_k^{in}$ **for** $k=1:K$ **do**

$$\left| \begin{array}{l} \text{Exploit } \hat{\mathcal{S}}_{l,0} \text{ to estimate the weights } a_{lk,0} \end{array} \right.$$

$$\left| \begin{array}{l} \text{Estimate } a_{lk,0} \text{ with either (5.26), (5.27) or (5.28), (5.29), } \forall l \in \mathcal{L}_k^{in} \end{array} \right.$$
end**Stage II: Iterative Part****for** $i = 1 : L_{AoD}$ **do****for** $k=1:K$ **do***Weighted Majority Voting*

$$\hat{\mathcal{S}}_{k,v} = \text{WMVoting}(\{\hat{\mathcal{S}}_{l,i-1}\}_{l \in \mathcal{L}_k^{in}}, i, G_t, a_{lk,i-1});$$
Local Step

$$\left| \begin{array}{l} [\hat{\mathcal{S}}_{k,i}] = \text{SOMP}(\Phi, \bar{\mathbf{Y}}_k, L_{AoD} - i, \hat{\mathcal{S}}_{k,v}); \end{array} \right.$$
end*Transmit:* $\hat{\mathcal{S}}_{k,i}$ to all nodes $\mathcal{L}_k^{out}$ *Receive:* $\hat{\mathcal{S}}_{l,i}$ s from all nodes $\mathcal{L}_k^{in}$ *Weights $a_{lk,i}$ Estimation***for** $k=1:K$ **do**

$$\left| \begin{array}{l} \text{Estimate } a_{lk,i} \text{ with either (5.26), (5.27) or (5.28), (5.29), } \forall l \in \mathcal{L}_k^{in} \end{array} \right.$$
end**end****Stage III: Final****LS:** $\hat{\mathbf{H}}_k^v = \Phi^\dagger(\cdot, \hat{\mathcal{S}}_k) \bar{\mathbf{Y}}_k$; $\hat{\mathbf{H}}_k = (\hat{\mathbf{H}}_k^v)^H \mathbf{D}_t^H, \forall k = 1, \dots, K$

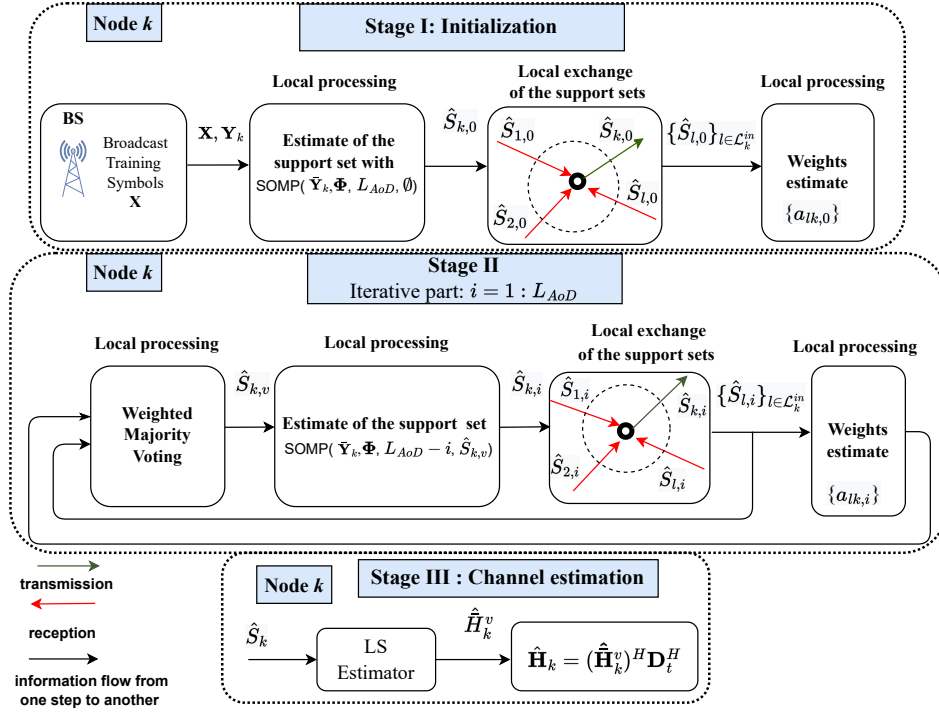


Fig. 5.3 Local Information Sharing and Processing per device k

5.7.2 Adaptive Weights Voting Mechanism

The aim of adopting a weighted voting mechanism was to make the algorithm WDi-SOMP robust to different conditions. In particular, this step (step 1 of stage II) was motivated by the fact that a device may not have the same impact on the support set retrieval, which can happen for several reasons. First, even if the devices have the same sparsity profile, they may face different communication conditions (i.e., noise level). Also, different common support sets among device groups may exist on the unknown sparsity patterns. Finally, the (hidden) sparsity order of the support sets and the grouping of devices are not usually known. For these reasons, two weighting methods are selected in the majority voting mechanism. The first method can cope with these cases with satisfactory performance [92] except for a case [93] where part of the support set of a device in the network is partially shared with different subgroups of devices, such that all devices' support sets have the same cardinality. For this purpose, the second method is suggested.

More specifically, at each iteration i of the algorithm, device k employs a zero elements vector $\mathbf{z}_{k,i}$ of size G_t and activates the positions indexed by each received $\hat{\mathbf{S}}_{l,i}$, adding the weight a_{lk} at the relevant positions, for all $l \in \mathcal{L}_k^{\text{in}}$. Afterwards, the

Algorithm 3: WMVoting

Result: $\hat{\mathcal{S}}_{k,v}$
 $i, \{\hat{\mathcal{S}}_{l,i}\}, l \in \mathcal{L}_k^{in}, \mathbf{z}_k = \mathbf{0}_{G_t \times 1}, \{a_{lk,i}\};$
for $l \in \mathcal{L}_k^{in}$ **do**
 for $j \in \hat{\mathcal{S}}_{l,i}$ **do**
 $\mathbf{z}_k(j) = \mathbf{z}_k(j) + a_{lk,i}(j);$
 end
end
 $\hat{\mathcal{S}}_{k,v} = \text{bestindices}(\mathbf{z}_k, i)$

device k , through function 'bestindices', finds the set of indices corresponding to the i largest amplitude components of the final $\mathbf{z}_k$. These steps are captured in Algorithm 3. The weights a_{lk} may be updated assuming either of the two following adaptive weights methods similar to the one in [85].

5.7.2.1 Weight method based on support set as single entity

We introduce the instantaneous metric

$$b_{lk} \triangleq |\hat{\mathcal{S}}_{l,i} \setminus \hat{\mathcal{S}}_{k,i-1}| \quad (5.25)$$

that measures the number of different elements in the involved sets. We further smooth the quantity $b_{lk,i}$ through a first-order filter, say,

$$b_{lk,i} = (1 - v)b_{lk,i-1} + v|\hat{\mathcal{S}}_{l,i} \setminus \hat{\mathcal{S}}_{k,i-1}| \quad (5.26)$$

where $v \in (0, 1)$ is a small positive factor. Then, the combination weight assigned for each index j by device k to device l is approximated by

$$a_{lk,i}(j) \approx \frac{b_{lk,i}^{-1}}{\sum_{m \in \mathcal{L}_k^{in}} b_{mk,i}^{-1}}, l \in \mathcal{L}_k^{in}, k \in K, \forall j \in \hat{\mathcal{S}}_l. \quad (5.27)$$

At this point, it should be noted that the total number of common indices could be measured to determine the $b_{lk,i}$ s and the corresponding weight $a_{lk,i}$. In either case, all indices of a neighbour are graded with the same weight assigning the same significance to all indices even if some of them aren't included in the device's support set. This mechanism can efficiently identify devices' groups with similar sparsity patterns but

different noise levels (per group) and non-overlapping but common sparsity patterns inside the groups [92]. To manage the existence of overlapping sparsity profiles, like the one depicted in Figure 5.1, we assumed a per index weighting method in the majority voting mechanism.

5.7.2.2 Weight method based on, per-index, support set similarities

Here, our aim is to give an individual weight for each index of the support set (L_{AoD} in number) of all devices. To achieve this, each device k votes with 1 the support set indices that belong both to its neighbor l and its own support set. Then, a candidate index approved by the majority of neighboring devices wins. At each iteration i , each device k builds a matrix $\mathbf{Z}_k \in \mathbb{R}^{G_t \times |\mathcal{L}_k^{in}|}$ (initialized with zero values) which is updated by ones only in positions dictated by $\hat{\mathcal{S}}_{l,i-1} \cap \hat{\mathcal{S}}_{k,i} \neq \emptyset$ and fed as in input in (5.28). The weight that device k assigns to its neighbor l per support set index j at iteration i is estimated as follows

$$b_{lk,i}(j) = (1 - v)b_{lk,i-1}(j) + vZ_k(j, l) \quad (5.28)$$

$$a_{lk,i}(j) \approx \frac{b_{lk,i}(j)}{\sum_{m \in \mathcal{L}_k^{in}} b_{mk,i}(j)} \quad (5.29)$$

for $l \in \mathcal{L}_k^{in}$, $k \in K$ and all $j \in \hat{\mathcal{S}}_l \cap \hat{\mathcal{S}}_k$.

The elaborated weight methods are exploited in the voting mechanism in Algorithm 2 to enhance the channel estimation performance. In the following section, we will analytically estimate the error bounds of the proposed estimator.

5.8 Performance Evaluation

In this section, we analyze the performance of the proposed estimator from various aspects. More specifically, the analysis will concern i) the iterations of the algorithm, ii) the computational complexity, iii) the communication overhead of WDiSOMP compared to the rest considered schemes, iv) the channel estimation error in terms of the support set reconstruction quality and, finally, v) the impact of angular grid resolution assuming an overcomplete dictionary to minimize the channel quantization error.

5.8.1 Required iterations before termination and Computational Complexity

In Algorithm 2, the iterations of WDiSOMP are fixed, constant by construction, and equal to L_{AoD} . Also, they show how many times the devices exchange their support set indices (communication step) in the network. The communication step includes local processing, which is highly dependent on the available energy of the device, especially for devices with limited resources (e.g., battery-dependent) to run an energy harvesting process. In this study, we assume that the devices satisfy the minimum power requirements to solve the assigned tasks. SOMP is executed locally at each iteration of WDiSOMP, exploiting the obtained part of the support after voting to build an estimate of the full support set until its size is equal to L_{AoD} . The number of internal iterations of SOMP is exactly $L_{AoD} - i$ (since $|\hat{\mathcal{S}}_{k,voting}| = i$). As a result, in WDiSOMP running time, the total iterations of SOMP is $\sum_{i=1}^{L_{AoD}} (L_{AoD} - i) = \frac{L_{AoD}(L_{AoD}-1)}{2}$.

Moving on to the computational complexity of the proposed algorithm, we consider that operations like addition, subtraction, division and multiplication of any two numbers have computational complexity $O(1)$. Also, the complexity for two matrices multiplication of size $M \times N$ and $N \times P$ is $O(MPN)$.

Let us focus on communication round i . As we see in Figure 5.3, the WDiSOMP algorithm executed locally at each device k has two specific processing steps that actually contribute to the overall computational complexity. In particular, first, the complexity of SOMP is on the order of $O(N_r G_t T s_i)$ [94], where $s_i = L_{AoD} - i$ is the size of the support set that is estimated at iteration i . Second, the suggested voting mechanism contributes a small cost to the complexity of the algorithm due to the adaptive weights estimate. Both weighting methods should first compute the involved $b_{lk,i}$ s. In the first method, the cost mainly derives from the fact that each device k should measure the number of different indices between its support set and one of its neighbors ($|\hat{\mathcal{S}}_{l,i} \setminus \hat{\mathcal{S}}_{k,i-1}|$). In the second, per index method, the respective cost stems from i) the intersection operation ($\hat{\mathcal{S}}_{l,i} \cap \hat{\mathcal{S}}_{k,i-1}$) among the support set of device k and its neighbors, and ii) the term $vZ_k(j, l)$, $j = 1, 2, \dots, L_{AoD}$ whose activation depends on the previous step. In either case, the cost is in the order of the support set size or, else, sparsity level, i.e., $O(L_{AoD})$. After this sub-step, the cost of estimating the desired $a_{lk,i}$ for each one of its $K - 1$ neighbors (where $K - 1$ captures the $|\mathcal{L}_k^{in}|$) and itself involves simple operations, $K - 1$ sums (denominator) and 1 division, thus the cost is $O(K)$. In the former

method, the weights $a_{lk,i}$ are the same for all indices, while the latter estimates a weight per index.

Having available the weights $a_{lk,i}$, the computational complexity of the majority voting (Algorithm 3) is determined by: i) the number of additions to the zero-elements vector $\mathbf{z}_k$, which are equal to L_{AoD} (the support set size or sparsity level) for each of the $K - 1$ neighbors and the device itself, or, else, KL_{AoD} in total, and ii) the sorting method used in the function 'bestindices' to find the i most voted indices. In our case, the number of elements to be sorted is small and equal to the support set level L_{AoD} .

Finally, at Stage III of Algorithm 2, each device k uses an LS estimator to acquire the non-zero involved coefficients (i.e., the channel gains). This stage adds a cost which depends on the computation of the pseudo-inverse of matrix Φ , keeping only that columns dictated by $\hat{S}_k$, i.e. $O(L_{AoD}T^2 + L_{AoD}^3)$, and then by its multiplication with the measurements' matrix imposing an additional cost of $O(L_{AoD}TN_r)$.

It is noted here that, apart from the voting mechanism, the complexity of the other two local processing methods (namely, SOMP and LS) depends on the adopted algorithms. As this paper aims to study, for the first time, a fully-distributed channel estimation algorithm enhanced by carefully designed voting mechanisms that improve the cooperation among the devices, two well-known simple algorithms have been chosen. However, it should be clear that other candidates with lower complexity can also be selected.

Before moving on to the next section, let us compare the complexity of the proposed algorithm with the ones that will be compared in the simulations section. WDiSOMP has a slightly higher complexity than DiSOMP due to the employed voting mechanism. However, as demonstrated in the experiments, its performance is greatly improved, while there is no requirement for knowing beforehand the sparsity structure of the involved channels. As for the locally executed SOMP and the centralized SSOMP (namely, the one executed after all data are collected to a central point), the computational complexity is on the order of $O(N_rG_tTL_{AoD})$ and $O(KN_rG_tTL_{AoD})$, respectively. Again, as it will be demonstrated, their performance is worse.

5.8.2 Transmission Overhead

In this section, we compare the transmission overhead of the distributed schemes DiSOMP and WDiSOMP with the centralized scheme, SSOMP, in terms of the number of transmitted bits, denoted as $\mathcal{T}_o^d$ and $\mathcal{T}_o^c$, respectively. In the distributed schemes, $\mathcal{T}_o^d$ captures the amount of information exchanged between a device k and each neighbor, whereas in the centralized method, $\mathcal{T}_o^c$ stands for the communication overhead between BS and each device k . The transmitted bits are analytically evaluated for all algorithms, as follows.

In the case of the centralized solution SSOMP, we assume that each device k transmits to the BS separately the real and complex part of the measurements in $\bar{\mathbf{Y}}_k \in \mathbb{C}^{T \times N_r}$, assigning q bits per value, which corresponds to $2N_r T q$ bits from each device to BS. Then the BS, by employing SSOMP, utilizes the measurements from all devices to estimate the joint support set, whose indices are fed back to the devices delivering $L_{AoD} \lceil \log_2(G_t) \rceil$ bits. Hence, the total number of bits amounts to:

$$\mathcal{T}_o^c = K(2N_r T q + L_{AoD} \lceil \log_2(G_t) \rceil). \quad (5.30)$$

Moving on to the distributed algorithms, let us, first, assume a network topology in which each node k transmits to its $\mathcal{L}_k^{out}$ neighbors the L_{AoD} indices of the candidate support set derived by applying SOMP locally. In (W)DiSOMP, only L_{AoD} indices from $\{1, \dots, G_t\}$ (where G_t denotes the number of atoms) are transmitted, and each index is supposed to be encoded with $\lceil \log_2(G_t) \rceil$ bits, which entails $L_{AoD} \lceil \log_2(G_t) \rceil$ transmitted bits. Each device transmits in each iteration the full support set to its neighbors, and this process is repeated L_{AoD} times until the full sparsity level is achieved. Given that the proposed approach is fully distributed, $|\mathcal{L}_k^{out}| = K - 1$, the total transferred bits amount to:

$$\mathcal{T}_o^d = K(K - 1) \lceil \log_2(G_t) \rceil L_{AoD}^2. \quad (5.31)$$

From (5.31), it is shown that the transmission overhead of the distributed schemes is a quadratic function of the sparsity parameter L_{AoD} , considering the rest parameters constant.

In Table 5.2, we show the $\mathcal{T}_o^c$ in terms of the feedback quantization bits q , for $N_r = \{1, 2, 4, 8\}$. It is observed that the $\mathcal{T}_o^c$ cost is significantly impacted by the in-

	$N_r = 1$	$N_r = 2$	$N_r = 4$	$N_r = 8$
$q = 8$	5.34	10.14	19.74	38.94
$q = 16$	10.14	19.74	38.94	77.34
$q = 32$	19.74	38.94	77.34	154.14
$q = 64$	38.94	77.34	154.14	307.74

Table 5.2 Transmission overhead in Kbits of centralized SSOMP for $K = 10$, $G_t = 512$, $T = 30$, $|\mathcal{L}_k| = K - 1$, $L_{AoD} = 6$

G_t	256		512		1024	
	(W)DiSOMP	SSOMP	(W)DiSOMP	SSOMP	(W)DiSOMP	SSOMP
$q = 8$	25.92	10.08	29.16	10.14	32.40	10.20
$q = 16$	25.92	19.68	29.16	19.74	32.40	19.80
$q = 32$	25.92	38.88	29.16	38.94	32.40	39.00
$q = 64$	25.92	77.28	29.16	77.34	32.40	77.40

Table 5.3 Transmission overhead in Kbits of (W)DiSOMP and centralized SSOMP for $K = 10$, $N_r = 2$, $G_t = \{256, 512, 1024\}$, $T = 30$, $|\mathcal{L}_k^{out}| = K - 1$, $L_{AoD} = 6$

crease in receiving antennas and feedback quantization bits. In Figure 5.4, we capture the impact of training symbols T in the $\mathcal{T}_o^c$, $\mathcal{T}_o^d$ of SSOMP and WDiSOMP, respectively, assuming devices of single, two and four antennas. We have set the resolution of the AoDs dictionary (e.g., the number of atoms) equal to $G_t = 512 (\gg N_t = 128)$. The communication overhead of the distributed solutions is independent of the receiving antennas and thus kept at a constant level. Also, it is observed that the distributed solutions prevail for $N_r > 2$. These parameters affect only the complexity of the centralized solution, making the required communication cost prohibitive when the involved devices have many antenna elements.

Moreover, in Table 5.3, the centralized and distributed solutions are assessed in terms of the atoms G_t for various values of q . These outcomes show a subtle increase in the overhead of SSOMP and a higher increase in the one of (W)DiSOMP. In Table 5.4, the overhead of the distributed solutions is evaluated in terms of the sparsity level, selecting proper L_{AoD} compatible for mmWave systems [28], [27], [45] and setting fixed values at the rest parameters. The same table presents the performance of the centralized solution. Although the increase in the number of receiving antennas is accompanied by an essential increase in $\mathcal{T}_o^c$, this isn't verified with the G_t and L_{AoD} that are involved in the second term of (5.30).

From the above, we conclude that the distributed approaches, such as the one presented here, might be preferable in many cases encountered in practice (e.g., cases with

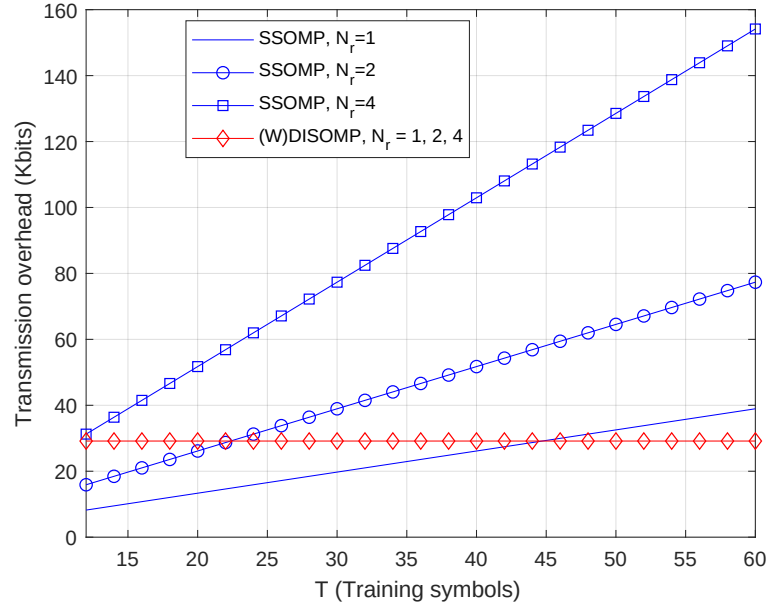


Fig. 5.4 Transmission overhead in Kbits of SSOMP, (W)DiSOMP for $K = 10$, $T = [12 : 60]$, $L_{AoD} = 6$, $N_r = \{1, 2, 4\}$, $G_t = 512$, $q = 32$ bits

L_{AoD}	5	6	7	8	9	10
(W)DiSOMP	20.25	29.16	39.69	51.84	65.61	81.00
SSOMP $N_r = 2$	38.85	38.94	39.03	39.12	39.21	39.30
SSOMP $N_r = 4$	77.25	77.34	77.43	77.52	77.61	77.70

Table 5.4 Transmission overhead in Kbits of SSOMP, (W)DiSOMP for $K = 10$, $T = [5 : 10]$, $L_{AoD} = [5 : 10]$, $N_r = \{2, 4\}$, $G_t = 512$, $q = 32$ bits

low sparsity and/or non-single antenna devices).

5.8.3 Impact of Support Set on Reconstruction Quality

The greedy approaches are based on the LS solution to estimate the complex values corresponding to the estimated support set. Their accuracy may deteriorate if part of the support set isn't correctly identified. To evaluate the proposed CS-based channel estimator and determine a lower error bound, we formulate the problem for the support set reconstruction assuming an oracle estimator, which considers perfect knowledge of the true AoDs.

First, we reform the measurements model assuming (5.8) to acquire the following

$$\bar{\mathbf{Y}}_k = \Phi_a \bar{\mathbf{H}}_a^k + \bar{\mathbf{N}}_k, \quad (5.32)$$

with $\bar{\mathbf{H}}_a^k = (\mathbf{A}_r \mathbf{H}_a^k)^H \in \mathbb{C}^{L_{AoD} \times N_r}$ and $\Phi_a = \mathbf{X}^H \mathbf{A}_t \in \mathbb{C}^{T \times L_{AoD}}$ the measurement matrix for the Least Squares (LS) problem. The problem in (5.32) is similar to the one in (5.20) if we keep the L_{AoD} known atoms (or, else, AoDs) involved in the channel, $\mathbf{D}_t$ is reduced to $\mathbf{A}_t$ (response matrix structure). Each column of (5.32) is captured by $\bar{\mathbf{y}}_{k,r}$ for $r = 1, 2, \dots, N_r$ and is written as

$$\bar{\mathbf{y}}_{k,r} = \Phi_a \bar{\mathbf{h}}_a^{k,r} + \bar{\mathbf{n}}_{k,r}, \quad (5.33)$$

with $\bar{\mathbf{h}}_a^{k,r} = \bar{\mathbf{H}}_a^k(:, r) \in \mathbb{C}^{L_{AoD} \times 1}$, $\bar{\mathbf{n}}_{k,r} = \bar{\mathbf{N}}_k(:, r) \in \mathbb{C}^{T \times 1}$. Given (5.33), the oracle LS approach estimates the L_{AoD} entries of $\bar{\mathbf{h}}_a^{k,r}$, requiring $\Phi_a^H \Phi_a$ be full rank, i.e., $T \geq L_{AoD}$ and assuming that AoDs are exactly known, by

$$\bar{\mathbf{h}}_a^{k,r,o} = (\Phi_a^H \Phi_a)^{-1} \Phi_a^H \bar{\mathbf{y}}_{k,r} \quad (5.34)$$

Then, the oracle estimate of $\mathbf{H}_k$, if in (5.19) we keep the L_{AoD} rows of $\bar{\mathbf{H}}_k^v$ and $\mathbf{D}_t$ is substituted by $\mathbf{A}_t$, is given by

$$\mathbf{H}_k^o = (\bar{\mathbf{H}}_a^{k,o})^H \mathbf{A}_t^H \quad (5.35)$$

with $\bar{\mathbf{H}}_a^{k,o} = [\bar{\mathbf{h}}_a^{k,o,1}, \bar{\mathbf{h}}_a^{k,o,2}, \dots, \bar{\mathbf{h}}_a^{k,o,N_r}]$.

An accurate recovery of the L_{AoD} -row matrix $\bar{\mathbf{H}}_a^k$ is equivalent to simultaneously recovering N_r vectors $\{\bar{\mathbf{h}}_a^{k,r}\}_{r=1}^{N_r}$. Considering the oracle estimator, which estimates the channel gain under the assumption that $\mathbf{A}_t$ is known, the MSE can be written as

$$\begin{aligned} \mathbb{E} \left[\|\mathbf{H}_k - \mathbf{H}_k^o\|_F^2 \right] &= \mathbb{E} \left[\sum_{r=1}^{N_r} \left\| \mathbf{h}_{k,r} - \mathbf{h}_{k,r}^o \right\|_2^2 \right] \\ &= \mathbb{E} \left[\sum_{r=1}^{N_r} \left\| \mathbf{A}_t \left(\bar{\mathbf{h}}_a^{k,r} - \bar{\mathbf{h}}_a^{k,o,r} \right) \right\|_2^2 \right] \\ &= \sum_{r=1}^{N_r} \mathbb{E} \left[\left\| \mathbf{A}_t \bar{\mathbf{h}}_e^{k,r} \right\|_2^2 \right], \end{aligned} \quad (5.36)$$

where $\bar{\mathbf{h}}_e^{k,r} \triangleq \bar{\mathbf{h}}_a^{k,r} - \bar{\mathbf{h}}_a^{k,o,r}$. Then, keeping the equivalent form of channel estimation

error in (5.36), the MSE can be lower bounded as

$$\mathbb{E} \left[\left\| \mathbf{A}_t \bar{\mathbf{h}}_e^{k,r} \right\|_2^2 \right] \geq \lambda_{\min} \left(\mathbf{A}_t^H \mathbf{A}_t \right) \mathbb{E} \left[\left\| \bar{\mathbf{h}}_e^{k,r} \right\|_2^2 \right] \quad (5.37)$$

The inequality holds due to the Rayleigh-Ritz ratio [41].

Combining (5.33) with (5.34) and assuming that $\mathbb{E} \left[\bar{\mathbf{n}}_{k,r} \bar{\mathbf{n}}_{k,r}^H \right] = \sigma^2 \mathbf{I}_T$, $\bar{\mathbf{h}}_a^{k,o,r} = \bar{\mathbf{h}}_a^{k,r} - \mathbf{Q} \bar{\mathbf{n}}_{k,r}$, the $\mathbb{E} \left[\left\| \bar{\mathbf{h}}_e^{k,r} \right\|_2^2 \right]$ can be written as

$$\begin{aligned} \mathbb{E} \left[\left\| \bar{\mathbf{h}}_e^{k,r} \right\|_2^2 \right] &= \mathbb{E} \left[\left\| \mathbf{Q} \bar{\mathbf{n}}_{k,r} \right\|_2^2 \right] \\ &= \mathbb{E} \left[\text{Tr} \{ \mathbf{Q} \bar{\mathbf{n}}_{k,r} \bar{\mathbf{n}}_{k,r}^H \mathbf{Q}^H \} \right] \\ &= \text{Tr} \{ \mathbf{Q} \mathbb{E} \left[\bar{\mathbf{n}}_{k,r} \bar{\mathbf{n}}_{k,r}^H \right] \mathbf{Q}^H \} \\ &= \sigma^2 \text{Tr} \{ \mathbf{Q} \mathbf{Q}^H \} \end{aligned} \quad (5.38)$$

with $\mathbf{Q} = (\Phi_a^H \Phi_a)^{-1} \Phi_a^H$. Since $\mathbf{Q} \mathbf{Q}^H = (\Phi_a^H \Phi_a)^{-1}$, the $\text{Tr} \{ \mathbf{Q} \mathbf{Q}^H \}$ is further processed, exploiting the arithmetic and harmonic means [16] and trace properties, as follows

$$\begin{aligned} \text{Tr} \{ \mathbf{Q} \mathbf{Q}^H \} &= \sum_{l=1}^{L_{AoD}} \lambda_l \left((\Phi_a^H \Phi_a)^{-1} \right) \\ &= L_{AoD} \left(\frac{1}{L_{AoD}} \sum_{l=1}^{L_{AoD}} \lambda_l \left((\Phi_a^H \Phi_a)^{-1} \right) \right) \\ &\geq \frac{(L_{AoD})^2}{\sum_{l=1}^{L_{AoD}} 1/\lambda_l \left((\Phi_a^H \Phi_a)^{-1} \right)} \\ &= \frac{(L_{AoD})^2}{\sum_{l=1}^{L_{AoD}} \lambda_l (\Phi_a^H \Phi_a)} \\ &= \frac{(L_{AoD})^2}{\text{Tr} \{ \Phi_a^H \Phi_a \}} \end{aligned} \quad (5.39)$$

From (5.36)-(5.39), we conclude that the MSE is lower-bounded as

$$\mathbb{E} \left[\left\| \mathbf{H}_k - \mathbf{H}_k^o \right\|_F^2 \right] \geq N_r \lambda_{\min} \left(\mathbf{A}_t^H \mathbf{A}_t \right) \frac{\sigma^2 L_{AoD}^2}{\text{Tr} \{ \Phi_a^H \Phi_a \}}. \quad (5.40)$$

Assuming an orthonormal basis, and specifically the DFT matrix, to construct the

AoDs dictionary $\mathbf{A}_t$ in Φ_a , it holds that $\mathbf{A}_t^H \mathbf{A}_t = \mathbf{I}_{L_{AoD}}$ and thus $\lambda_{\min}(\mathbf{A}_t^H \mathbf{A}_t) = 1$, simplifying (5.40) to

$$\mathbb{E} \left[\|\mathbf{H}_k - \mathbf{H}_k^o\|_F^2 \right] \geq N_r \frac{\sigma^2 L_{AoD}^2}{\text{Tr}\{\Phi_a^H \Phi_a\}}. \quad (5.41)$$

Denominator $\text{Tr}\{\Phi_a^H \Phi_a\}$ is equivalently written as $\sum_{i=1}^{L_{AoD}} \lambda_i(\Phi_a^H \Phi_a)$. To attain the lower bound, $\Phi_a^H \Phi_a$ should be a diagonal matrix with identical diagonal elements. Previous condition holds if and only if the eigenvalues λ_i 's are equal. In particular, if $\mathbf{X}\mathbf{X}^H = T\mathbf{I}$ and exploiting $\mathbf{A}_t^H \mathbf{A}_t = \mathbf{I}_{L_{AoD}}$, $\text{Tr}\{\Phi_a^H \Phi_a\} = L_{AoD}T$ and therefore

$$\mathbb{E} \left[\|\mathbf{H}_k - \mathbf{H}_k^o\|_F^2 \right] \geq N_r \sigma^2 \frac{L_{AoD}}{T}. \quad (5.42)$$

The bound in (5.42) increases linearly with the sparsity level L_{AoD} and the number of receive antennas N_r , while it is inversely proportional to the number of pilots T .

The oracle LS exploits the exact support-AoDs information; thus, under the previous conditions, its MSE will be the lower bound of WDiSOMP and every other estimator, which ignores the true AoDs and identifies them through a learning process under noisy conditions. Furthermore, when part of the estimated support set isn't correct, there will be a term in the error that captures the impact of the incorrect indices. This term is expected to become powerful in high SNR conditions adding a systematic deviation from the real support or channel and a respective floor. As part of erroneous indices is shrunk, the floor will gradually disappear, and the estimators' error will tend towards the lower bound. Such an approach could lead to designing more effective weighting methods than the ones elaborated here, suitable for multi-task joint sparsity to correct the occurred error (e.g., due to quantization). Although this study is out of scope here, it constitutes a challenging task which would deserve further investigation.

5.8.4 Impact of Dictionary Resolution

The aim of this subsection is to study the impact of the dictionary resolution, exploiting its mutual coherence, which will help us determine an upper bound for the error of the quantized channel matrix $\mathbf{H}$ in (5.11). The rationale behind the estimation of the upper-bound is to quantify it based on the quality of $\mathbf{D}_t$ and Φ , which depends on the number of grid points G_t and is measured by the mutual coherence. Assuming

an overcomplete dictionary $\mathbf{D}_t$ to minimize the quantization error, the symbols matrix should be properly designed to ensure the successful retrieval of the sparse channel matrix $\hat{\mathbf{H}}_k^v$ by the SOMP (locally).

For this purpose, we exploit (5.19) to write the error, in an equivalent form, as follows

$$\begin{aligned}
 \left\| \hat{\mathbf{H}}_k - \mathbf{H}_k \right\|_F^2 &= \left\| \left(\hat{\mathbf{H}}_k^v \right)^H \mathbf{D}_t^H - \left(\bar{\mathbf{H}}_k^v \right)^H \mathbf{D}_t^H \right\|_F^2 \\
 &= \left\| \left(\mathbf{D}_t \hat{\mathbf{H}}_k^v \right)^H - \left(\mathbf{D}_t \bar{\mathbf{H}}_k^v \right)^H \right\|_F^2 \\
 &= \left\| \mathbf{D}_t \left(\hat{\mathbf{H}}_k^v - \bar{\mathbf{H}}_k^v \right) \right\|_F^2 \\
 &= \sum_{r=1}^{N_r} \left\| \mathbf{D}_t \left(\hat{\mathbf{h}}_{k,r}^v - \bar{\mathbf{h}}_{k,r}^v \right) \right\|_2^2, \tag{5.43}
 \end{aligned}$$

where $\bar{\mathbf{H}}_k^v = [\bar{\mathbf{h}}_{k,1}^v, \bar{\mathbf{h}}_{k,2}^v, \dots, \bar{\mathbf{h}}_{k,N_r}^v]$.

An upper bound for the quantized channel estimation error is derived assuming that the supports of $\hat{\mathbf{h}}_{k,r}^v$ and $\bar{\mathbf{h}}_{k,r}^v$ are similar. Hence, the channel estimation error in (5.43), based on the Cauchy Schwarz inequality, satisfies

$$\begin{aligned}
 \left\| \hat{\mathbf{H}}_k - \mathbf{H}_k \right\|_F^2 &= \sum_{r=1}^{N_r} \left\| \left(\mathbf{D}_t \right)_{\mathcal{S}_k} \left(\left(\hat{\mathbf{h}}_{k,r}^v \right)_{\mathcal{S}_k} - \left(\bar{\mathbf{h}}_{k,r}^v \right)_{\mathcal{S}_k} \right) \right\|_2^2 \\
 &\leq \sum_{r=1}^{N_r} \left\| \left(\mathbf{D}_t \right)_{\mathcal{S}_k} \right\|_2^2 \left\| \left(\hat{\mathbf{h}}_{k,r}^v \right)_{\mathcal{S}_k} - \left(\bar{\mathbf{h}}_{k,r}^v \right)_{\mathcal{S}_k} \right\|_2^2. \tag{5.44}
 \end{aligned}$$

First, we will focus to find an upper bound for the term $\left\| \left(\mathbf{D}_t \right)_{\mathcal{S}_k} \right\|_2^2$. For this purpose, we define $\mathbf{G}_{\mathcal{S}_k} \triangleq \left(\mathbf{D}_t \right)_{\mathcal{S}_k}^H \left(\mathbf{D}_t \right)_{\mathcal{S}_k}$. Then, the squared spectral norm satisfies $\left\| \left(\mathbf{D}_t \right)_{\mathcal{S}_k} \right\|_2^2 = \lambda_{\max}(\mathbf{G}_{\mathcal{S}_k})$. Focusing on matrix $\mathbf{G}_{\mathcal{S}_k}$, otherwise, it is written as $\left(\mathbf{D}_t^H \mathbf{D}_t \right)_{\mathcal{S}_k}$ and given that $\mathbf{D}_t = \begin{bmatrix} \mathbf{a}_{t,1} & \dots & \mathbf{a}_{t,G_t} \end{bmatrix}$, $\mathbf{a}_{t,i} \in \mathbb{C}^{N_t \times 1}$ with exponential elements as in (5.9), is equal

to

$$\begin{aligned} \mathbf{G}_{\mathcal{S}_k} &= \begin{bmatrix} \mathbf{a}_{t,1}^H \mathbf{a}_{t,1} & \cdots & \mathbf{a}_{t,1}^H \mathbf{a}_{t,G_t} \\ \vdots & \cdots & \vdots \\ \mathbf{a}_{t,G_t}^H \mathbf{a}_{t,1} & \cdots & \mathbf{a}_{t,G_t}^H \mathbf{a}_{t,G_t} \end{bmatrix}_{\mathcal{S}_k} \\ &= \begin{bmatrix} 1 & \cdots & \mathbf{a}_{t,1}^H \mathbf{a}_{t,G_t} \\ \vdots & \cdots & \vdots \\ \mathbf{a}_{t,G_t}^H \mathbf{a}_{t,1} & \cdots & 1 \end{bmatrix}_{\mathcal{S}_k} \end{aligned} \quad (5.45)$$

Based on (5.45), $\mathbf{G}_{\mathcal{S}_k} = \mathbf{I} + \tilde{\mathbf{G}}_{\mathcal{S}_k}$ where

$$\tilde{\mathbf{G}}_{\mathcal{S}_k} = \begin{bmatrix} 0 & \cdots & \mathbf{a}_{t,1}^H \mathbf{a}_{t,G_t} \\ \vdots & \cdots & \vdots \\ \mathbf{a}_{t,G_t}^H \mathbf{a}_{t,1} & \cdots & 0 \end{bmatrix}_{\mathcal{S}_k}. \quad (5.46)$$

Exploiting the properties of the mutual coherence, $\mu(\mathbf{G}_{\mathcal{S}_k}) = \mu(\tilde{\mathbf{G}}_{\mathcal{S}_k})$, thus the off-diagonal entries of $\tilde{\mathbf{G}}_{\mathcal{S}_k}$ are upper-bounded by $\mu(\mathbf{D}_t)$. From this observation, all the off-diagonal entries of $\tilde{\mathbf{G}}_{\mathcal{S}_k}$ can be replaced by the $\mu(\mathbf{D}_t)$, constituting $\tilde{\mathbf{G}}'_{\mathcal{S}_k}$. Then, the Gershgorin circle theorem is utilized to find an upper-bound of the $\lambda_{\max}(\mathbf{I} + \tilde{\mathbf{G}}_{\mathcal{S}_k})$. In particular, it holds that $\lambda_{\max}(\mathbf{G}_{\mathcal{S}_k}) \leq \lambda_{\max}(\mathbf{I} + \tilde{\mathbf{G}}'_{\mathcal{S}_k}) \leq 1 + (L_{AoD} - 1)\mu(\mathbf{D}_t)$. From this inequality, (5.44) is rewritten as

$$\left\| \hat{\mathbf{H}}_k - \mathbf{H}_k \right\|_F^2 \leq \sum_{r=1}^{N_r} \bar{\mu} \cdot \left\| \hat{\mathbf{h}}_{k,r}^v - \bar{\mathbf{h}}_{k,r}^v \right\|_2^2, \quad (5.47)$$

where $\bar{\mu} = 1 + (L_{AoD} - 1)\mu(\mathbf{D}_t)$. In $\bar{\mu}$, the term $(L_{AoD} - 1)\mu(\mathbf{D}_t)$ denotes the imposed overhead in the sparse channel estimation error due to the discretization with resolution G_t higher than the transmitter antennas N_t . It depends on the mutual coherence of the dictionary $\mathbf{D}_t$, which, in turn, depends on the selected resolution G_t . When $G_t = N_t$, $\mu(\mathbf{D}_t) = 0$, and $\bar{\mu} = 1$. However, for $G_t \gg N_t$, $\mu(\mathbf{D}_t)$ attains values close to 1, making $\bar{\mu}$ and thus the upper bound of the channel error in (5.47) on the order of L_{AoD} .

In (5.47), $\left\| \hat{\mathbf{h}}_{k,r}^v - \bar{\mathbf{h}}_{k,r}^v \right\|_2^2$ represents the sparse channel vector estimation error of the SMV problem in (5.21). The proposed algorithm (WDiSOMP) aims to estimate the

MIMO channel by exploiting the commonalities in the support sets of the sparse channel matrices of nearby devices. All devices execute SOMP locally (using the measurements collected from all antenna elements), exchange the estimated support sets with their neighbors' and apply the weighted voting mechanism to estimate their support set. To ensure support recovery, the amplitudes of all nonzero elements must be large enough to overcome the noise. Hence, we define $\left| \bar{\mathbf{h}}_{k,r,\min}^v \right| = \min_{m \in \mathcal{S}_k} \left| \bar{\mathbf{h}}_{k,r}^v(m) \right|$ to derive a condition that the (S)OMP correctly identifies the support $\mathcal{S}_k$. Suppose that the event $\mathcal{B}_r = \left\{ \max_{1 \leq m \leq G_t} \left| \Phi^H(m) \bar{\mathbf{n}}_{k,r} \right| < \sigma \sqrt{(1 + \beta) \log G_t^2} \right\}$ (where constant $\beta > 0$) occurs simultaneously for all antenna elements and thus $\left| \bar{\mathbf{h}}_{k,r,\min}^v \right| (1 - (2L_{AoD} - 1)\mu(\Phi)) \geq 2 \left(\sigma \sqrt{(1 + \beta) \log G_t^2} \right)$ holds. An upper-bound of this estimation error is derived by considering the Lemma 3 for OMP performance guarantees in [11]. When the support $\mathcal{S}_k$ is successfully recovered, our estimator operates as the oracle from which the corresponding complex gains are estimated as $\left(\bar{\mathbf{h}}_{k,r}^{v,o} \right)_{\mathcal{S}_k} = \Phi_{\mathcal{S}_k}^\dagger \bar{\mathbf{y}}_{k,r}$ and its error is written as

$$\begin{aligned}
& \left\| \left(\bar{\mathbf{h}}_{k,r}^{v,o} \right)_{\mathcal{S}_k} - \left(\bar{\mathbf{h}}_{k,r}^v \right)_{\mathcal{S}_k} \right\|_2^2 = \\
& = \left\| \Phi_{\mathcal{S}_k}^\dagger \Phi_{\mathcal{S}_k} \left(\bar{\mathbf{h}}_{k,r}^v \right)_{\mathcal{S}_k} - \Phi_{\mathcal{S}_k}^\dagger \bar{\mathbf{n}}_{k,r} - \left(\bar{\mathbf{h}}_{k,r}^v \right)_{\mathcal{S}_k} \right\|_2^2 \\
& = \left\| \Phi_{\mathcal{S}_k}^\dagger \bar{\mathbf{n}}_{k,r} \right\|_2^2 \\
& = \left\| \left(\Phi_{\mathcal{S}_k}^H \Phi_{\mathcal{S}_k} \right)^{-1} \Phi_{\mathcal{S}_k}^H \bar{\mathbf{n}}_{k,r} \right\|_2^2 \\
& \leq \left\| \left(\Phi_{\mathcal{S}_k}^H \Phi_{\mathcal{S}_k} \right)^{-1} \right\|_2^2 \sum_{m \in \mathcal{S}_k} \left(\Phi^H(m) \bar{\mathbf{n}}_{k,r} \right)^2 \\
& \leq \frac{L_{AoD} ((1 + \beta) \sigma^2 \log G_t^2)}{(1 - (L_{AoD} - 1) \mu(\Phi))^2}, \tag{5.48}
\end{aligned}$$

where the last inequality is derived by using the definition of $\mathcal{B}_r$ and

$\left\| \left(\Phi_{\mathcal{S}_k}^H \Phi_{\mathcal{S}_k} \right)^{-1} \right\|_2^2 = \frac{1}{\lambda_{\min}(\Phi_{\mathcal{S}_k}^H \Phi_{\mathcal{S}_k})} \leq \frac{1}{(1 - (L_{AoD} - 1) \mu(\Phi))^2}$ by applying the Gershgorin circle theorem to $\lambda_{\min}(\Phi_{\mathcal{S}_k}^H \Phi_{\mathcal{S}_k})$. Combining (5.47) with (5.48), the upper bound is as

$$\begin{aligned}
\left\| \hat{\mathbf{H}}_k - \mathbf{H}_k \right\|_F^2 & \leq \frac{N_r L_{AoD} (1 + (L_{AoD} - 1) \mu(\mathbf{D}_t))}{(1 - (L_{AoD} - 1) \mu(\Phi))^2} \\
& \times ((1 + \beta) \sigma^2 \log G_t^2) \tag{5.49}
\end{aligned}$$

The upper bound of the squared error in (5.49) shows that the sparse channel recovery performance is affected by the mutual coherence of the measurement matrix Φ , for which $\mu(\Phi) < 1/(2L_{AoD} - 1)$ should hold in order to ensure reliable estimation of the support set in noisy conditions. The upper limit relates with the fact that some of the L_{AoD} involved AoDs in the channel matrix may occupy off-diagonal entries of the Gram matrix $\tilde{\mathbf{G}}$, with magnitudes, at most, equal to $\mu(\mathbf{D}_t)$. Given an overcomplete dictionary, the training symbols matrix should be properly designed to reduce the magnitudes of these off-diagonal entries.

Table 5.5 lists the right term in (5.49) (upper bound) of the instantaneous channel error in terms of the mutual coherence $\mu(\Phi)$ assuming an overcomplete DFT-based dictionary and $\beta = 0.5$, $N_t = 128$, $L_{AoD} = 6$. Assuming that the training pilots' design achieves $\mu(\Phi) = 0.05$, the obtained upper bound is higher in low SNR regimes and decreases linearly as the noise variance reduces (high SNR). Also, the upper bound increased by 48%, when the $G_t = 256$ with $\mu(\mathbf{D}_t) = 0.637$ increased to $G_t = 512$ with $\mu(\mathbf{D}_t) = 0.9$. This bound is maintained at low levels as low as possible, and close to zero is the $\mu(\Phi)$. To attain $\mu(\Phi)$ values close to zero, $\mathbf{X}$ should be appropriately constructed, which is out of scope in this work. In our study, the symbols are designed as random Gaussian variables, thus ensuring a “well-behaved” measurement matrix and hence correct channel reconstruction with high probability.

$G_t = 512, \mu(\mathbf{D}_t) = 0.900$				
$\mu(\Phi)$	0.1	0.05	0.0025	0.00125
$\sigma^2 = 0.01$	49.42	21.97	12.67	12.51
$\sigma^2 = 0.001$	4.94	2.19	1.26	1.25
$G_t = 256, \mu(\mathbf{D}_t) = 0.637$				
$\mu(\Phi)$	0.1	0.05	0.0025	0.00125
$\sigma^2 = 0.01$	33.40	14.85	8.56	8.46
$\sigma^2 = 0.001$	3.34	1.48	0.86	0.85

Table 5.5 Right term in (5.49) of the instantaneous channel error

5.9 Simulations

5.9.1 Multi-Antenna devices: simulation on real data

In this section, the elaborated algorithm is evaluated for the problem of channel estimation in a ULA mmWave massive MIMO system with fully-digital transmitters and

AoD Type	AoD Name	device 1	device 2	device 3	device 4	device 5	device 6	device 7	device 8	device 9	device 10
Global	AoD 1	144.34	144.34	144.34	144.34	144.34	144.34	144.34	144.34	144.34	144.34
Global	AoD 2	128.68	128.68	128.68	128.68	128.68	128.68	128.68	128.68	128.68	128.68
Common	AoD 3	112.02	112.02	112.02	112.02	112.02	112.02	93.58	93.58	93.58	93.58
Common	AoD 4	91.79	91.79	91.79	91.79	91.79	91.79	73.66	73.66	73.66	73.66
Common	AoD 5	64.05	64.05	64.05	64.05	64.05	53.57	53.57	53.57	53.57	53.57
Common	AoD 6	55.77	55.77	55.77	55.77	55.77	28.95	28.95	28.95	28.955	28.95

Table 5.6 Multitask device grouping according to AoDs sparsity patterns for $L_g = 2, L_c = 4$

receivers by employing SOMP locally.

To be more specific, the proposed distributed scheme is evaluated under two different voting mechanisms in a multi-tasking scenario. Moreover, its behavior will be compared to the individual SOMP (which is executed locally by each device), DiSOMP (which engages a simple majority voting mechanism) and the Structured SOMP (SSOMP), which is executed in a centralized manner at the BS. The performance measure, namely, Normalized Mean Squared Error (NMSE), is defined as

$$\text{NMSE} = \frac{1}{M} \sum_{m=1}^M \frac{\sum_{k=1}^K \|\mathbf{H}_k - \hat{\mathbf{H}}_k^m\|_F^2}{\sum_{k=1}^K \|\mathbf{H}_k\|_F^2} \quad (5.50)$$

where $\mathbf{H}_k, \hat{\mathbf{H}}_k^m$ are the correct and reconstructed channel vectors of device k at the m -th channel realization. In the simulations, $M = 50000$ captures the number of Monte Carlo iterations and $v = 0.1$ (which is utilized in the recursive equation of the adaptive weighting step in (5.26), (5.28)).

In the following, the NMSE behavior of these methods will be assessed i) in terms of the number of training symbols $T = [12, 40]$ for a specific Signal-to-Noise-Ratio (SNR) value, i.e., $\text{SNR} = 20\text{dB}$, and ii) for a fixed $T = 30$ in terms of SNR in the range $[0, 30]$.

We assume a network of $K = 10$ devices and a BS with $N_r = 2$ and $N_t = 128$ antenna elements, respectively. Following the relevant literature, e.g. [111], [95], a DFT-based overcomplete, predefined dictionary $\mathbf{D}_t$ is assumed (due to the similarity of the array response vectors with its columns) with resolution $G_t = 512$ atoms. Note that the AoDs and AoAs were selected from $[0, \pi]$, and the AoAs were randomly assigned. We assume that there are $L_g = 2$ global interest AoDs that contribute to all devices' channels and two different groups of $L_c = 4$ commonly shared AoDs (or channel paths) that contribute to devices' channels 1-5 and 7-10, respectively. Also, 1 out of 10 devices,

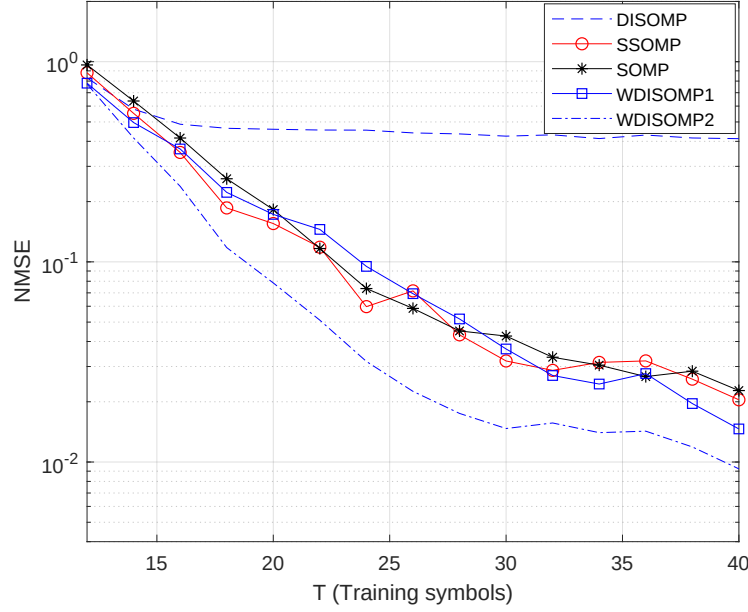


Fig. 5.5 Average NMSE over all nodes for $T = [12 : 2 : 40]$, $SNR = 20$ dB

and specifically device 6, partially shares its $L_c = 4$ AoDs with different subgroups. Table 5.6 presents the AoDs values that were considered to create the devices' channels and the corresponding grouping. The elements of the training matrix $\mathbf{X} \sim \mathcal{CN}(0, \frac{1}{T})$ are i.i.d. random variables [111]. Each device has no prior knowledge of the sparsity order apart from the total sparsity level ($L_{AoD} = 6$), which is similar to all devices.

Figure 5.5 depicts the average NMSE over all involved devices. It is noticed that WDiSOMP 1, which assumes voting on support set level, approaches the behavior of local SOMP. However, WDiSOMP 2 achieves superior performance than SOMP, and particularly DiSOMP, due to the employment of the voting method on the index level. For the latter, the floor in the performance is attained for low values of T and remains unchanged even when measurements T are high enough. This behavior is justified by the unknown sparsity structure, which WDiSOMP 1, 2 can learn via the suggested voting mechanisms. Also, WDiSOMP, in either case, attains the performance of SSOMP (the centralized solution), which, nonetheless, needs to know the number of global L_g and common L_c indices in the sparsity pattern division, as in [111].

Figure 5.6 demonstrates the channel estimation performance at the two subgroups and device 6, focusing on WDiSOMP 1, 2. It is reminded that these subgroups comprise devices 1-5 and 7-10, respectively, whose support set is not partially shared and device 6, which shares 4 out of 6 indices in its support set with the devices from different sub-

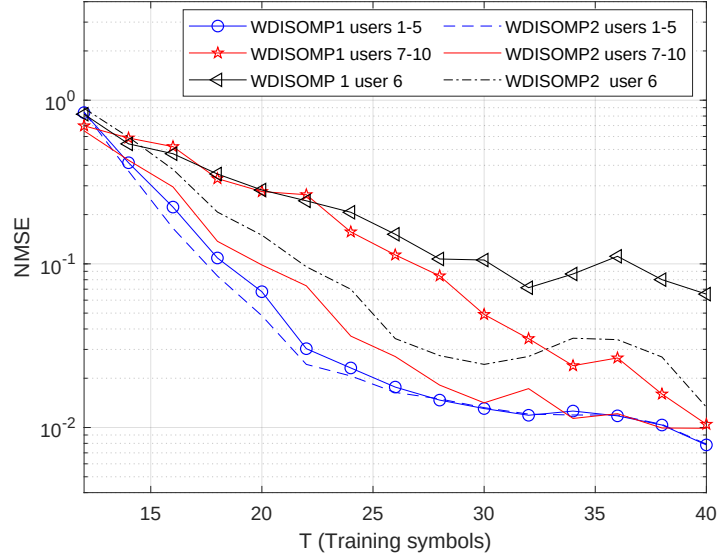


Fig. 5.6 NMSE of devices 1-5, devices 7-10 and device 6 for $T = [12 : 2 : 40]$, $SNR = 20$ dB

groups. It is observed that the behavior of the algorithm follows, as expected, the one observed in Figure 5.5, namely, follows the average NMSE. The performance of WDiSOMP2 in devices 1-5 enhanced slightly, while device 6, which captures the challenging multi-tasking condition, has considerably benefited from the new weighting mechanism that operates at the index level. The same holds for devices 7-10.

Now, we will focus on the second experiment of this study. In Figure 5.7, we show the average performance of all methods in terms of SNR for $T = 30$. Then, in Figure 5.8, the performance of WDiSOMP1 and WDiSOMP2 for the devices 1-5, 7-10 and 6 is captured. Both in this case, it is verified that the behavior of the proposed algorithm, under the per-index weighting method before voting, is superior to exploiting the support set as a global entity (i.e., not per index).

In Figure 5.9, we isolate the average behavior of WDiSOMP and local SOMP in terms of the noise level for $T = \{30, 60\}$. The theoretical lower error bound, presented in Section 5.8, is also plotted to validate the performance of the proposed scheme for the mmWave channel estimation. More specifically, the lower bound was derived based on (5.40). For $T = 30$, WDiSOMP1 has a similar performance to SOMP with a performance gap by WDiSOMP2 for high SNR. When the available measurements per device were doubled ($T = 60$), the two versions of WDiSOMP presented comparable behavior and, in high SNR, were better than SOMP (with a smaller gap in their performances than previously). Also, in high SNR, the floor in the performance of

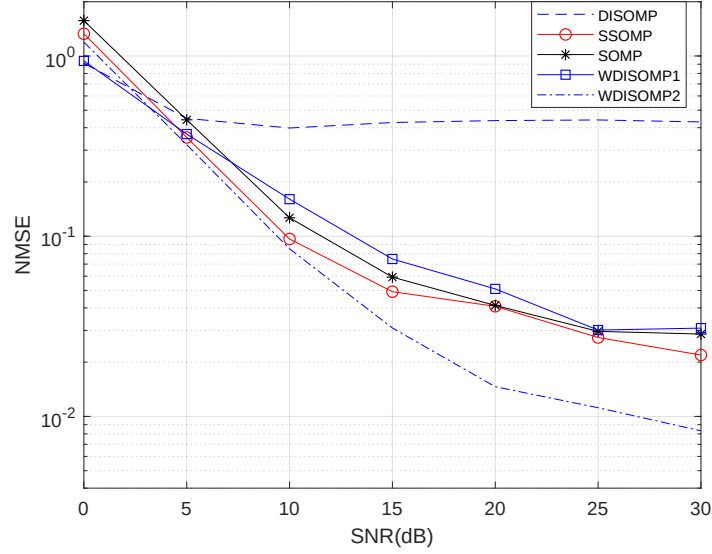


Fig. 5.7 Average NMSE of all devices for $\text{SNR} = [0 : 5 : 30]\text{dB}$, $T = 30$

WDiSOMP relates to the reconstruction error imposed by the mistaken support indices that the devices estimate (to find then the corresponding AoDs) compared to the oracle LS estimator, which assumes prior knowledge of them to estimate the channel. This experiment indicated that, in medium-to-high SNR, WDiSOMP benefited from the per index weighting method when the number of measurements was not enough. In low SNR, it was considerably reduced the distance between WDiSOMP1, WDiSOMP2 and the lower bound. However, they didn't approach the lower bound since the matrix $\mathbf{X}$ isn't properly designed, making the eigenvalues λ_i s of $\Phi_a^H \Phi_a$ different, even if it still holds that $\text{Tr}\{\Phi_a^H \Phi_a\} = \sum_{i=1}^{L_{AoD}} \lambda_i(\Phi_a^H \Phi_a) = L_{AoD}$. As mentioned in Section VI.C, the eigenvalues equality is a necessary condition to approach the lower bound, which in our case isn't satisfied.

From the above simulations, it is noted that WDiSOMP, with the use of either weighting method, can efficiently exploit, under different conditions, the inherent common sparsity structures and its performance is not impacted when various related device groupings coexist.

5.9.2 Single antenna devices: simulation on synthetic support sets

In this section, the performance of WDiOMP will be presented under three experiments. $K = 10$ single antenna devices are considered along with a BS with $N = 128$ antennas. An overcomplete, predefined dictionary $\mathbf{D}_t$ (using DFT) is employed with $G_t = 200$

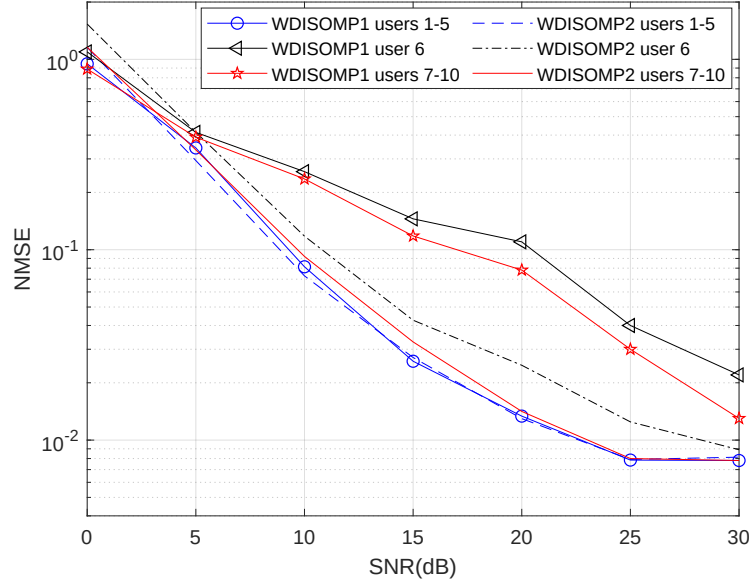


Fig. 5.8 Average NMSE of devices 1-5, 7-10 and device 6 for $\text{SNR} = [0 : 5 : 30]\text{dB}$, $T = 30$

atoms, which define a corresponding grid for $[-\pi, \pi]$, while $\phi_l \in [-\pi, \pi]$, and $T \in [10, 40]$. The elements of the training matrix $\mathbf{X}$ are i.i.d. random variables that follow $\mathcal{CN}(0, \frac{1}{T})$ (as in [111]). The L non-zero elements in $\mathbf{w}_k$ are also i.i.d. random variables following $\mathcal{N}(0, 1)$ (as in [103]). Each device has knowledge of only the total sparsity level L (i.e., not the individual parameters L_g, L_c). Finally, it is assumed that all devices are connected via a network and can share the positions of the non-zero elements in their $\mathbf{w}_k$'s, namely, their support sets. Thus, $\mathcal{L}_k^{\text{in}} = \mathcal{L}_k^{\text{out}} = \mathcal{V}$ holds.

The Average Support set Cardinality Error (ASCE) is considered for the evaluation. It takes values in the range $[0, 1]$ and is defined as follows

$$\text{ASCE} = 1 - \frac{1}{K} \frac{1}{M} \sum_{m=1}^M \sum_{k=1}^K \left\{ \frac{|\hat{\mathcal{S}}_k^m \cap \mathcal{S}_k|}{|\mathcal{S}_k|} \right\}. \quad (5.51)$$

For the results, $M = 1000$ Monte Carlo simulations were performed.

In the following, the performance of WDiOMP (using $v = 0.1$ for the calculation of the weights) will be compared with DiOMP (which employs a simple majority voting), the centralized algorithm SOMP [52] and the per device (locally) employed OMP algorithm. Three scenarios will be presented and the ASCE of the various approaches will be assessed versus the number of measurements T .

In the first experiment, depicted in Fig. 5.10 (left), ASCE is evaluated assuming

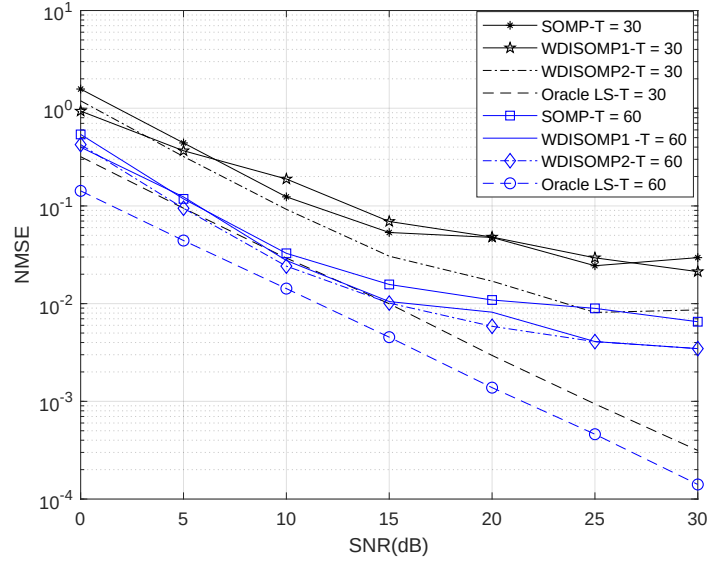


Fig. 5.9 Average NMSE of all devices for $\text{SNR} = [0 : 5 : 30]$, $T = \{30, 60\}$

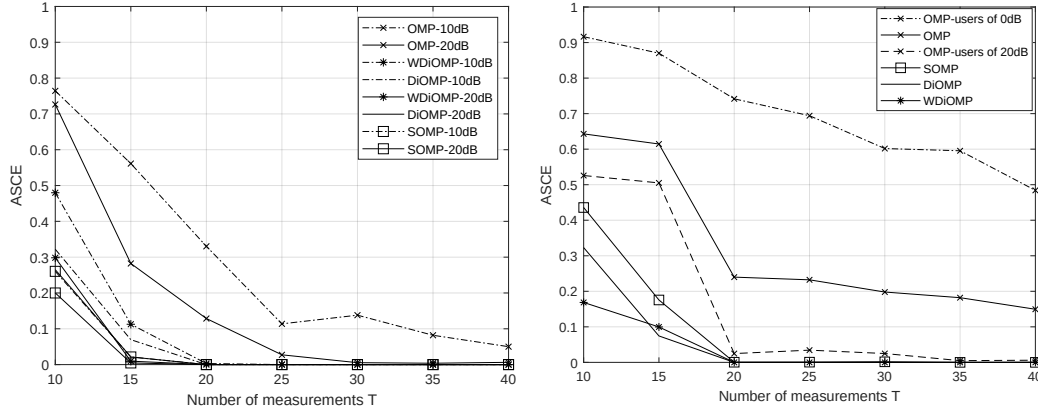


Fig. 5.10 (left) $L_g = 5$, $L_c = 0$, $\text{SNR} = \{10, 20\}$ dB, (right) Different SNR conditions per device. $L_g = 5$, $L_c = 0$.

that all devices share a global support set and considering two SNR values, namely, 10 and 20 dB. It is observed that both the DiOMP [90] and WDiOMP achieve a performance close to the one achieved by SOMP, which, being the point of reference for this experiment, was expected to demonstrate the best, though similar, performance. Additionally, the previous schemes perform considerably better than the per-device (locally) employed OMP algorithm, resulting, for instance, in about 40% reduction in required training measurements for achieving near zero ASCE, for $\text{SNR}=20$ dB. Moreover, for $\text{SNR}=10$ dB, the reduction is even more pronounced as the OMP approach does not reach a near zero ASCE for the T values considered. This experiment demonstrates the benefit of the devices when they participate in a collaborative channel estimation

procedure that exploits the adopted sparsity model, as opposed to the case in which each device operates in isolation. Finally, it should be mentioned that although DiOMP, WDiOMP and SOMP perform similarly, the first two do not utilize the BS during channel estimation.

In the second experiment, shown in Fig. 5.10(right), an “unbalanced” scenario is examined in which, although the devices share only a global support set, they operate in different SNR levels, namely, $\text{SNR}=20$ dB and $\text{SNR}=0$ dB for 7 and 3 devices, respectively. For the OMP case, in which the devices operate at isolation, apart from the overall mean performance (depicted by the “OMP” curve), the performances when focusing on the groups that face $\text{SNR}=20$ dB and $\text{SNR}=0$ dB, are also presented. It is observed that the collaboration of the devices (either centrally via SOMP or in a distributed fashion via DiOMP and WiOMP) improves considerably the estimation performance over the case in which the devices operate individually (via OMP) and, in particular, for devices facing both the better and (especially) the worse conditions. Finally, it is observed that WDiOMP is able to start from a lower ASCE which is attributed at its capability of weighting the contribution of each device, a behaviour that will be even more pronounced in the following experiment.

In the third experiment (see Fig. 5.11), another “unbalanced” scenario is examined. In more detail, the devices are divided into two groups that share a global support set (assuming $L_g = 5$), while the devices, of each group, share also a common support set of size $L_c = 3$. Additionally, all devices face $\text{SNR} = 20$ dB. Let us recall that only the total sparsity level ($L = 8$) is known (namely, L_g and L_c are considered unknown).

It is observed that WDiOMP performs similarly to the previous experiments and considerably better than SOMP and DiOMP. For the latter, the floor in their performance is a result of the unknown L_g and L_c . WDiOMP does not have this issue because it is able to infer (in a wide sense) the sparsity structure via the employed weighting voting mechanism. Also, it is noted that SOMP, here, is shown in order to assess the performance of DiOMP and should not be considered as the centralized version of the examined scenario. Furthermore, WDiOMP is better than OMP and requires about 40% less measurements for achieving a near-zero ASCE (in the case of OMP no floor appears as only L is relevant). Furthermore, it is mentioned that, although not shown here, the weights calculated by a particular device that employs WDiOMP, tend to similar

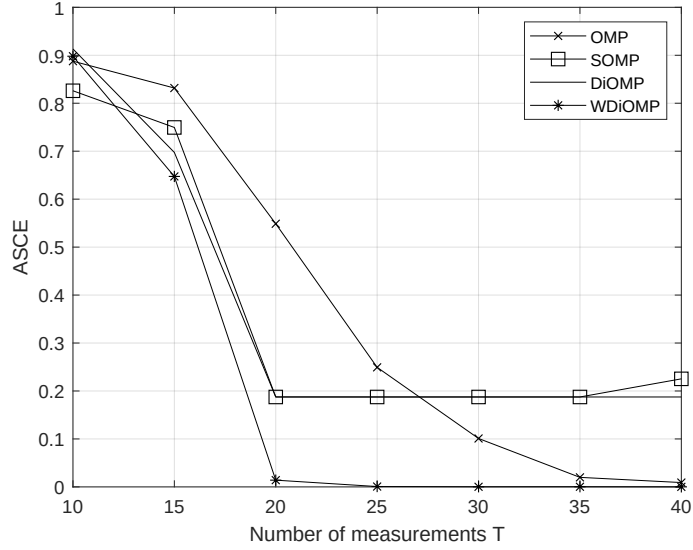


Fig. 5.11 Two groups of devices facing global and common support sets. $L_g = 5$, $L_c = 3$, SNR = 20dB.

(and larger) values for devices in the same group while the weights are smaller for the remaining devices, as hinted in Sec 5.7. As a final remark, it is noted that no NMSE curves for estimating $\mathbf{h}$ in (5.15) are presented due to space limitations. However, the conclusions are similar to the ones for ASCE. From the previous results, WDiOMP demonstrates a robust performance for all considered experimental scenarios via the proposed weighting voting mechanism.

CHAPTER 6

Conclusions and Future Works

In conclusion, channel estimation is a challenging task in millimeter wave MIMO systems and is necessary for efficient beamforming (data transmission). In the context of this dissertation, in Chapter 4, the mmWave MIMO channel estimation is addressed by combining a preprocessing scheme with power spectrum methods for DoA estimation (MUSIC for ULA and H-MUSIC for hybrid architectures), while in Chapter 5, is tackled as a sparse recovery problem with the aid of CS.

More specifically, we investigated the recoverability of a ULA snapshot applicable to sub- and fully-connected hybrid antenna arrays. The recommended scheme can recover the full snapshot as if we had a conventional antenna array either for time-varying or constant source signals. The phase shifters of the hybrid antenna array were designed to maximize the STNR of the restored snapshot. Typical simulation results have been presented for the AoA estimation problem, confirming the efficacy of the proposed approach under the optimal solution. Also, an advantage of this scheme is that, even if correlated source signals impinge on the antenna array, spatial smoothing techniques for classical ULAs can be applied before estimating the desired angles. At this point, it should be noted that, after the application of the preprocessing scheme, any high-resolution power spectrum methods could be utilized to estimate the desired AoAs, such as ESPRIT, apart from MUSIC.

Although the recoverability of the full snapshot was studied in hybrid antenna arrays using ULAs [104], there are other architectures that could be studied as future extensions of this work. Among different array geometries, fractal-wavelet modeling theory [44, 32, 114, 34, 33, 35], widely used in antenna theory, and its application to small antenna arrays and to the case of mmWave bands would be interesting. Additionally, other antenna array configurations could be studied like co-prime configurations

[115]. Finally, lens-antennas, as already have been applied in mmWave communication systems, would be an interesting direction [13].

Moreover, in the second part of this study, the channel estimation problem is addressed by a distributed approach that exploits the joint sparsity among proximal devices' sparse channels to estimate the desired parameters collaboratively. Initially, a fully distributed algorithm, WDiOMP, for the multi-device MISO downlink channel estimation was developed that exploited a general sparsity support model for the involved channel vectors. The improved performance of WDiOMP has been validated under synthetic sparse channels (namely, support sets) in various scenarios, even when the structure of the sparsity support is unknown. Then, the WDiOMP was extended and generalized to efficiently operate for multi-antenna devices. The measurements model allows the proposed solution to be directly applicable to scenarios with either single-antenna devices, following a Single Measurement Vector method, e.g., OMP, or multi-antenna devices, utilizing a Multiple Measurement Vector method, e.g., SOMP. Hence, the downlink multi-device MIMO channel estimation is solved via a fully distributed method, WDi(S)OMP, that exploits the same generalized model for the sparsity support of the involved MIMO channel matrices. The performance of WDiSOMP has been evaluated under the two proposed adaptive weighting methods, which are involved in the voting mechanism when the structure of the sparsity support is unknown and non-uniform under a multi-tasking framework.

WDi(S)OMP remains well-performing even when non-uniform overlapping sparsity groups occur in the network. It is achieved by incorporating a new voting mechanism which estimates an individual weight per support set index, not a global one as in our prior scheme [93], and further improves the channel estimation accuracy and reduces training time by the BS. Also, the transmission overhead of the proposed algorithm was investigated and compared to the corresponding cost of the centralized solution. The outcomes revealed the efficacy of the method under specific configurations of the involved parameters.

A promising future direction is to examine and broaden the proposed method in hybrid architectures [62], [51], since, in this work, fully-digital transmitters and receivers are considered. Furthermore, another goal is to enhance the performance of the proposed algorithm by focusing on the proper training matrix design. This analysis will

help attain the performance bound of the oracle LS estimator and ensure reliable channel recovery under the overcomplete dictionary (as already mentioned in the performance analysis). Another extension of the study could focus on alternative weighting methods in the voting process (quantifying the part of the error related to the deviation of the MSE from the lower bound) to optimize the mmWave channel estimation performance (under such a complicated joint sparsity framework). Also, we aim to investigate alternative methods, apart from (S)OMP, to solve the local problem in the suggested distributed methods WDi(S)OMP.

The increasing demands of new generations (5G and 6G) of wireless networks makes necessary the use of new cutting-edge technologies for their implementation. Artificial Intelligence and, especially, Machine Learning will play a key role by partially replacing classical algorithms, in 5G, mainly at the physical layer, while in 6G, fundamentally [65]. Machine learning tools are finding interesting applications in mmWave and massive MIMO systems. A future direction is to approach the mmWave channel estimation by designing Machine Learning-based methods that can learn some priors from the provided training dataset to provide high accuracy channel estimates with low training overhead. The use of deep learning techniques will allow the efficient use of mmWave systems even by users who move at high speeds due to the significantly better performance compared to traditional approaches, especially for systems consisting of large-scale antennas in the transmitter. Finally, a well-known challenging direction is to reformulate the suggested fully distributed cooperative approach under a CS-based federated learning framework to benefit from the advantages and capabilities (such as privacy) that this technique provides, exploiting the recent research effort in this field.

REFERENCES

- [1] I. Olkin A. W. Marshall. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, 1979.
- [2] Irfan Ahmed, Hedi Khammari, Adnan Shahid, Ahmed Musa, Kwang Soon Kim, Eli De Poorter, and Ingrid Moerman. A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives. *IEEE Communications Surveys and Tutorials*, 20(4):3060–3097, 2018.
- [3] Godfrey Anuga Akpakwu, Bruno J Silva, Gerhard P Hancke, and Adnan M Abu-Mahfouz. A survey on 5g networks for the internet of things: Communication technologies and challenges. *IEEE access*, 6:3619–3647, 2017.
- [4] Naser Al-Falahy and Omar YK Alani. Millimetre wave frequency band as a candidate spectrum for 5g network architecture: A survey. *Physical Communication*, 32:120–144, 2019.
- [5] Ahmed Alkhateeb, Omar El Ayach, Geert Leus, and Robert W Heath. Channel estimation and hybrid precoding for millimeter wave cellular systems. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):831–846, 2014.
- [6] Ahmed Alkhateeb, Jianhua Mo, Nuria Gonzalez-Prelcic, and Robert W Heath. MIMO precoding and combining solutions for millimeter-wave systems. *IEEE Communications Magazine*, 52(12):122–131, 2014.
- [7] Dimitris Ampeliotis, Christos Mavrokefalidis, and Kostas Berberidis. Distributed dictionary learning via projections onto convex sets. In *2017 25th European Signal Processing Conference (EUSIPCO)*, pages 2294–2298. IEEE, 2017.
- [8] Rafay Iqbal Ansari, Chrysostomos Chrysostomou, Syed Ali Hassan, Mohsen Guizani, Shahid Mumtaz, Jonathan Rodriguez, and Joel JPC Rodrigues. 5G D2D networks: Techniques, challenges, and future prospects. *IEEE Systems Journal*, 12(4):3970–3984, 2017.
- [9] Qazi Kamal Ud Din Arshad, Ahsan Ullah Kashif, and Ijaz Mansoor Quershi. A review on the evolution of cellular technologies. In *2019 16th International*

- Bhurban Conference on Applied Sciences and Technology (IBCAST)*, pages 989–993. IEEE, 2019.
- [10] Dror Baron, Michael B Wakin, Marco F Duarte, Shriram Sarvotham, and Richard G Baraniuk. Distributed compressed sensing. *Rice University, Depart. Electrical and Computer Engineering Technical Report TREE-0612*, 2006.
- [11] Zvika Ben-Haim, Yonina C Eldar, and Michael Elad. Coherence-based performance guarantees for estimating a sparse vector under random noise. *IEEE Transactions on Signal Processing*, 58(10):5030–5043, 2010.
- [12] Oussama Ben Smida, Slim Zaidi, Sofiène Affes, and Shahrokh Valaee. Robust distributed collaborative beamforming for wireless sensor networks with channel estimation impairments. *Sensors*, 19(5):1061, 2019.
- [13] Emil Björnson, Luca Sanguinetti, Henk Wymeersch, Jakob Hoydis, and Thomas L Marzetta. Massive mimo is a reality—what is next?: Five promising research directions for antenna arrays. *Digital Signal Processing*, 94:3–20, 2019.
- [14] Tadilo Endeshaw Bogale and Long Bao Le. Beamforming for multiuser massive mimo systems: Digital versus hybrid analog-digital. In *2014 IEEE Global Communications Conference*, pages 4066–4071. IEEE, 2014.
- [15] Nikola Bogdanović, Jorge Plata-Chaves, and Kostas Berberidis. Distributed incremental-based lms for node-specific adaptive parameter estimation. *IEEE Transactions on Signal Processing*, 62(20):5382–5397, 2014.
- [16] Peter S Bullen. *Handbook of means and their inequalities*, volume 560. Springer Science and Business Media, 2013.
- [17] Rafael Saraiva Campos. Evolution of positioning techniques in cellular networks, from 2g to 4g. *Wireless Communications and Mobile Computing*, 2017, 2017.
- [18] Robin Chataut and Robert Akl. Massive mimo systems for 5g and beyond networks—overview, recent trends, challenges, and future research direction. *Sensors*, 20(10):2753, 2020.
- [19] Lalit Chettri and Rabindranath Bera. A comprehensive survey on internet of things (iot) toward 5g wireless systems. *IEEE Internet of Things Journal*, 7(1):16–32, 2019.
- [20] Sheng-Fu Chuang, Wen-Rong Wu, and Yen-Ting Liu. High-resolution AoA estimation for hybrid antenna arrays. *IEEE Transactions on Antennas and Propagation*, 63(7):2955–2968, 2015.

- [21] Jisheng Dai, An Liu, and Vincent KN Lau. Joint channel estimation and user grouping for massive MIMO systems. *IEEE Transactions on Signal Processing*, 67(3):622–637, 2018.
- [22] Jisheng Dai, An Liu, and Hing Cheung So. Non-uniform burst-sparsity learning for massive MIMO channel estimation. *IEEE Transactions on Signal Processing*, 67(4):1075–1087, 2018.
- [23] Yacong Ding and Bhaskar D Rao. Dictionary learning-based sparse channel representation and estimation for FDD massive MIMO systems. *IEEE Transactions on Wireless Communications*, 17(8):5437–5451, 2018.
- [24] Rui Dinis and Alberto Palhau. A class of signal-processing schemes for reducing the envelope fluctuations of cdma signals. *IEEE transactions on communications*, 53(5):882–889, 2005.
- [25] Omar El Ayach, Sridhar Rajagopal, Shadi Abu-Surra, Zhouyue Pi, and Robert W Heath. Spatially sparse precoding in millimeter wave MIMO systems. *IEEE Transactions on Wireless Communications*, 13(3):1499–1513, 2014.
- [26] Dian Fan, Feifei Gao, Yuanwei Liu, Yansha Deng, Gongpu Wang, Zhangdui Zhong, and Arumugam Nallanathan. Angle domain channel estimation in hybrid millimeter wave massive mimo systems. *IEEE Transactions on Wireless Communications*, 17(12):8165–8179, 2018.
- [27] Zhen Gao, Linglong Dai, and Zhaocheng Wang. Channel estimation for mmwave massive mimo based access and backhaul in ultra-dense network. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2016.
- [28] Zhen Gao, Linglong Dai, Zhaocheng Wang, and Sheng Chen. Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO. *IEEE Transactions on Signal Processing*, 63(23):6169–6183, 2015.
- [29] Saswati Ghosh and Debarati Sen. An inclusive survey on array antenna design for millimeter-wave communications. *IEEE Access*, 7:83137–83161, 2019.
- [30] Andrea Goldsmith. *Wireless communications*. Cambridge university press, 2005.
- [31] Sotirios K Goudos, Panagiotis I Dallas, Stella Chatziefthymiou, and Sofoklis Kyriazakos. A survey of iot key enabling and future technologies: 5g, mobile iot, semantic web and applications. *Wireless Personal Communications*, 97(2):1645–1675, 2017.

- [32] Emanuel Guariglia. Spectral analysis of the weierstrass-mandelbrot function. In *2017 2nd International Multidisciplinary Conference on Computer and Energy Science (SpliTech)*, pages 1–6. IEEE, 2017.
- [33] Emanuel Guariglia. Harmonic sierpinski gasket and applications. *Entropy*, 20(9):714, 2018.
- [34] Emanuel Guariglia. Primality, fractality, and image analysis. *Entropy*, 21(3):304, 2019.
- [35] Emanuel Guariglia and Sergei Silvestrov. Fractional-wavelet analysis of positive definite distributions and wavelets on $'(\mathbb{C})$. In *Engineering Mathematics II*, pages 337–353. Springer, 2016.
- [36] Saeid Haghighatshoar and Giuseppe Caire. The beam alignment problem in mmwave wireless networks. In *2016 50th Asilomar Conference on Signals, Systems and Computers*, pages 741–745. IEEE, 2016.
- [37] Kais Hassan, Mohammad Masarra, Marie Zwingelstein, and Iyad Dayoub. Channel estimation techniques for millimeter-wave communication systems: Achievements and challenges. *IEEE Open Journal of the Communications Society*, 1:1336–1363, 2020.
- [38] Haibo He and He Jiang. Deep learning based energy efficiency optimization for distributed cooperative spectrum sensing. *IEEE Wireless Communications*, 26(3):32–39, 2019.
- [39] Shiwen He, Yan Zhang, Jiaheng Wang, Jian Zhang, Ju Ren, Yaoxue Zhang, Weihua Zhuang, and Xuemin Shen. A survey of millimeter-wave communication: Physical-layer technology specifications and enabling transmission technologies. *Proceedings of the IEEE*, 109(10):1666–1705, 2021.
- [40] Robert W Heath, Nuria Gonzalez-Prelcic, Sundeep Rangan, Wonil Roh, and Akbar M Sayeed. An overview of signal processing techniques for millimeter wave MIMO systems. *IEEE Journal of Selected Topics in Signal Processing*, 10(3):436–453, 2016.
- [41] Roger A Horn. Cr johnson matrix analysis. *Cambridge UP, New York*, 1985.
- [42] Jakob Hoydis, Kianoush Hosseini, Stephan Ten Brink, and Mérouane Debbah. Making smart use of excess antennas: Massive mimo, small cells, and tdd. *Bell Labs Technical Journal*, 18(2):5–21, 2013.

- [43] Wei Huang, Yongming Huang, Wei Xu, and Luxi Yang. Beam-blocked channel estimation for fdd massive mimo with compressed feedback. *IEEE Access*, 5:11791–11804, 2017.
- [44] Kuem C Hwang. A modified sierpinski fractal antenna for multiband application. *IEEE antennas and wireless propagation letters*, 6:357–360, 2007.
- [45] Jorge Iscar, Ismail Guevenc, Sener Dikmese, and Ahmed S Ibrahim. Optimal angular spread of the multipath clusters in mmwave systems under pilot contamination. In *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pages 1–6. IEEE, 2017.
- [46] Furqan Jameel, Zara Hamid, Farhana Jabeen, Sherali Zeadally, and Muhammad Awais Javed. A survey of device-to-device communications: Research issues and challenges. *IEEE Communications Surveys and Tutorials*, 20(3):2133–2168, 2018.
- [47] Suhanya Jayaprakasam, Sharul Kamal Abdul Rahim, and Chee Yen Leow. Distributed and collaborative beamforming in wireless sensor networks: Classifications, trends, and research directions. *IEEE Communications Surveys and Tutorials*, 19(4):2092–2116, 2017.
- [48] Imran Khan, Joel JPC Rodrigues, Jalal Al-Muhtadi, Muhammad Irfan Khattak, Yousaf Khan, Farhan Altaf, Seyed Sajad Mirjavadi, and Bong Jun Choi. A robust channel estimation scheme for 5g massive mimo systems. *Wireless Communications and Mobile Computing*, 2019, 2019.
- [49] Joonsuk Kim and Inkyu Lee. 802.11 wlan: history and new enabling mimo techniques for next generation standards. *IEEE Communications Magazine*, 53(3):134–140, 2015.
- [50] Farzana Kulsoom, Anna Vizziello, Hassan Nazeer Chaudhry, and Pietro Savazzi. Joint sparse channel recovery with quantized feedback for multi-user massive mimo systems. *IEEE Access*, 8:11046–11060, 2020.
- [51] Junho Lee, Gye-Tae Gil, and Yong H Lee. Channel estimation via orthogonal matching pursuit for hybrid mimo systems in millimeter wave communications. *IEEE Transactions on Communications*, 64(6):2370–2386, 2016.
- [52] Gang Li, Thakshila Wimalajeewa, and Pramod K Varshney. Decentralized and collaborative subspace pursuit: A communication-efficient algorithm for joint sparsity pattern recovery with sensor networks. *IEEE Trans. on Signal Processing*, 64(3):556–566, 2015.

- [53] Jing Li, DeChun Sun, and ZuJun Liu. A mimo channel prediction scheme based on multi-task learning. *Wireless Personal Communications*, 115(3):1869–1880, 2020.
- [54] Lin Li, Fangfang Chen, and Jisheng Dai. Separate DOD and DOA estimation for bistatic MIMO radar. *International Journal of Antennas and Propagation*, 2016, 2016.
- [55] Qinghua Li, Guangjie Li, Wookbong Lee, Moon-il Lee, David Mazzarese, Bruno Clerckx, and Zexian Li. Mimo techniques in wimax and lte: a feature overview. *IEEE Communications magazine*, 48(5):86–92, 2010.
- [56] Zhongjie Li, Wei Gao, Min Zhang, and Jiyuan Xiong. Multi-task deep learning based hybrid precoding for mmwave massive MIMO system. *China Communications*, 18(10):96–106, 2021.
- [57] Kaihua Luo, Xiaoping Zhou, Bin Wang, Jifeng Huang, and Haichao Liu. Sparse bayes tensor and DOA tracking inspired channel estimation for V2X millimeter wave massive mimo system. *Sensors*, 21(12):4021, 2021.
- [58] Nurul H Mahmood, Marta G Sarret, Gilberto Berardinelli, and Preben Mogenssen. Full duplex communications in 5g small cells. In *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1665–1670. IEEE, 2017.
- [59] Mohammad Majidzadeh, Aleksi Moilanen, Nuutti Tervo, Harri Pennanen, Antti Tolli, and Matti Latva-aho. Partially connected hybrid beamforming for large antenna arrays in multi-user MISO systems. In *28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1–6, 2017.
- [60] Elaine Crespo Marques, Nilson Maciel, Lirida Naviner, Hao Cai, and Jun Yang. A review of sparse recovery algorithms. *IEEE access*, 7:1300–1322, 2018.
- [61] Thomas L Marzetta. Massive mimo: an introduction. *Bell Labs Technical Journal*, 20:11–22, 2015.
- [62] Roi Méndez-Rial, Cristian Rusu, Nuria González-Prelcic, Ahmed Alkhateeb, and Robert W Heath. Hybrid mimo architectures for millimeter wave communications: Phase shifters or switches? *Ieee Access*, 4:247–267, 2016.
- [63] Rajarshi Middy, Nabajit Chakravarty, and Mrinal Kanti Naskar. Compressive sensing in wireless sensor networks: a survey. *IETE technical review*, 34(6):642–654, 2017.

- [64] Andreas F Molisch, Vishnu V Ratnam, Shengqian Han, Zheda Li, Sinh Le Hong Nguyen, Linsheng Li, and Katsuyuki Haneda. Hybrid beamforming for massive mimo: A survey. *IEEE Communications magazine*, 55(9):134–141, 2017.
- [65] Manuel Eugenio Morocho-Cayamcela, Haeyoung Lee, and Wansu Lim. Machine learning for 5g/b5g mobile and wireless communications: Potential, limitations, and future directions. *IEEE Access*, 7:137184–137206, 2019.
- [66] Sathiya Narayanan, Sujit Kumar Sahoo, and Anamitra Makur. Greedy pursuits assisted basis pursuit for reconstruction of joint-sparse signals. *Signal Processing*, 142:485–491, 2018.
- [67] Roula Nassif, Stefan Vlaski, Cédric Richard, Jie Chen, and Ali H Sayed. Multi-task learning over graphs: An approach for distributed, streaming machine learning. *IEEE Signal Processing Magazine*, 37(3):14–25, 2020.
- [68] M Nguyen. Massive mimo: A survey of benefits and challenges. *ICSES Trans. Comput. Hardw. Electr. Eng*, 4:1–4, 2018.
- [69] Zhitong Ni, J Andrew Zhang, Kai Yang, Fei Gao, and Jianping An. Estimation of multiple angle-of-arrivals with localized hybrid subarrays for millimeter wave systems. *IEEE Transactions on Communications*, 2019.
- [70] Yong Niu, Yong Li, Depeng Jin, Li Su, and Athanasios V Vasilakos. A survey of millimeter wave communications (mmwave) for 5g: opportunities and challenges. *Wireless networks*, 21(8):2657–2676, 2015.
- [71] Song Noh, Jiho Song, Junse Lee, and Heejung Yu. High-resolution and low-complexity direction of arrival estimation for hybrid array of subarrays. *IEEE Access*, 2022.
- [72] Jorge Plata-Chaves, Nikola Bogdanović, and Kostas Berberidis. Distributed diffusion-based lms for node-specific adaptive parameter estimation. *IEEE Transactions on Signal Processing*, 63(13):3448–3460, 2015.
- [73] Chuan Qin, J Andrew Zhang, Xiaojing Huang, Kai Wu, and Y Jay Guo. Fast angle-of-arrival estimation via virtual subarrays in analog antenna array. *IEEE Transactions on Wireless Communications*, 2020.
- [74] Zhijin Qin, Jiancun Fan, Yuanwei Liu, Yue Gao, and Geoffrey Ye Li. Sparse representation for wireless communications: A compressive sensing approach. *IEEE Signal Processing Magazine*, 35(3):40–58, 2018.

- [75] Anupama Rajoriya, Rohit Budhiraja, and Lajos Hanzo. Centralized and decentralized channel estimation in fdd multi-user massive mimo systems. *IEEE Transactions on Vehicular Technology*, 2022.
- [76] Sundeep Rangan, Theodore S Rappaport, and Elza Erkip. Millimeter-wave cellular wireless networks: Potentials and challenges. *Proceedings of the IEEE*, 102(3):366–385, 2014.
- [77] Xiongbiao Rao and Vincent KN Lau. Distributed compressive csit estimation and feedback for FDD multi-user massive MIMO systems. *IEEE Transactions on Signal Processing*, 62(12):3261–3271, 2014.
- [78] Theodore S Rappaport, Shu Sun, Rimma Mayzus, Hang Zhao, Yaniv Azar, Kevin Wang, George N Wong, Jocelyn K Schulz, Mathew Samimi, and Felix Gutierrez. Millimeter wave mobile communications for 5g cellular: It will work! *IEEE access*, 1:335–349, 2013.
- [79] Shiwei Ren, Xiaochuan Ma, Shifeng Yan, and Chengpeng Hao. 2-D unitary ESPRIT-like direction-of-arrival (DoA) estimation for coherent signals with a uniform rectangular array. *Sensors*, 13(4):4272–4288, 2013.
- [80] Abdelhamid Riadi, Mohamed Boulouird, and MM Hassani. An overview of massive-mimo in 5g wireless communications. In *Colloque international TELECOM*, pages 10–12, 2017.
- [81] Wonil Roh, Ji-Yun Seol, Jeongho Park, Byunghwan Lee, Jaekon Lee, Yungsoo Kim, Jaeweon Cho, Kyungwhoon Cheun, and Farshid Aryanfar. Millimeter-wave beamforming as an enabling technology for 5g cellular communications: Theoretical feasibility and prototype results. *IEEE communications magazine*, 52(2):106–113, 2014.
- [82] Athanasios A Rontogiannis and Kostas Berberidis. Efficient decision feedback equalization for sparse wireless channels. *IEEE Transactions on Wireless Communications*, 2(3):570–581, 2003.
- [83] Richard Roy and Thomas Kailath. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on acoustics, speech, and signal processing*, 37(7):984–995, 1989.
- [84] Kei Sakaguchi, Thomas Haustein, Sergio Barbarossa, Emilio Calvanese Strinati, Antonio Clemente, Giuseppe Destino, Aarno Pärssinen, Ilgyu Kim, Heesang Chung, Junhyeong Kim, et al. Where, when, and how mmwave is used in 5g and beyond. *IEICE Transactions on Electronics*, 100(10):790–808, 2017.

- [85] Ali H Sayed, Sheng-Yuan Tu, Jianshu Chen, Xiaochuan Zhao, and Zaid J Towfic. Diffusion strategies for adaptation and learning over networks: an examination of distributed strategies and network behavior. *IEEE Signal Processing Magazine*, 30(3):155–171, 2013.
- [86] Huan-Tang Sheng and Wen-Rong Wu. Blind aoa estimation with hybrid antenna arrays. *IEEE Transactions on Wireless Communications*, 20(8):5058–5070, 2021.
- [87] Feng Shu, Yaolu Qin, Tingting Liu, Linqing Gui, Yijin Zhang, Jun Li, and Zhu Han. Low-complexity and high-resolution DOA estimation for hybrid analog and digital massive MIMO receive array. *IEEE Transactions on Communications*, 66(6):2487–2501, 2018.
- [88] Foad Sahrabi and Wei Yu. Hybrid digital and analog beamforming design for large-scale antenna arrays. *IEEE Journal of Selected Topics in Signal Processing*, 10(3):501–513, 2016.
- [89] Shu Sun and Theodore S Rappaport. Millimeter wave mimo channel estimation based on adaptive compressed sensing. In *2017 IEEE international conference on communications workshops (ICC Workshops)*, pages 47–53. IEEE, 2017.
- [90] Dennis Sundman, Saikat Chatterjee, and Mikael Skoglund. Distributed greedy pursuit algorithms. *Signal Processing*, 105:298–315, 2014.
- [91] Maria Trigka and Elias Dritsas. An efficient distributed approach for cooperative spectrum sensing in varying interests cognitive radio networks. *Sensors*, 22(17):6692, 2022.
- [92] Maria Trigka, Christos Mavrokefalidis, and Kostas Berberidis. A distributed sparse channel estimation technique for mmwave massive MIMO systems. In *29th European Signal Processing Conference (EUSIPCO)*, 2021.
- [93] Maria Trigka, Christos Mavrokefalidis, and Kostas Berberidis. An efficient decentralized approach for mmwave MIMO channel estimation. In *4th International Balkan Conference on Communications and Networking (BalkanCom)*, 2021.
- [94] Joel A Tropp and Anna C Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on information theory*, 53(12):4655–4666, 2007.

- [95] Cheng-Rung Tsai, Yu-Hsin Liu, and An-Yeu Wu. Efficient compressive channel estimation for millimeter-wave large-scale antenna systems. *IEEE Transactions on Signal Processing*, 66(9):2414–2428, 2018.
- [96] David Tse and Pramod Viswanath. *Fundamentals of wireless communication*. Cambridge university press, 2005.
- [97] Georgios I Tsiropoulos, Animesh Yadav, Ming Zeng, and Octavia A Dobre. Co-operation in 5G hetnets: Advanced spectrum access and D2D assisted communications. *IEEE Wireless Communications*, 24(5):110–117, 2017.
- [98] Anthony Ngozichukwuka Uwaechia and Nor Muzlifah Mahyuddin. A comprehensive survey on millimeter wave communications for fifth-generation wireless networks: Feasibility and challenges. *IEEE Access*, 8:62367–62414, 2020.
- [99] Evangelos Vlachos, Aris S Lalos, and Kostas Berberidis. Stochastic gradient pursuit for adaptive equalization of sparse multipath channels. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 2(3):413–423, 2012.
- [100] Cheng-Xiang Wang, Ji Bian, Jian Sun, Wensheng Zhang, and Minggao Zhang. A survey of 5g channel measurements and models. *IEEE Communications Surveys and Tutorials*, 20(4):3142–3168, 2018.
- [101] Meng-Jie Wang, Jheng-Liang Cai, Fan-Shuo Tseng, and Chao-Yuan Hsu. A low-complexity 2-d angle of arrival estimation in massive mimo systems. In *2016 International Computer Symposium (ICS)*, pages 710–713. IEEE, 2016.
- [102] Peng Wang, Yonghui Li, Lingyang Song, and Branka Vucetic. Multi-gigabit millimeter wave wireless communications for 5g: From fixed access to cellular networks. *IEEE Communications Magazine*, 53(1):168–178, 2015.
- [103] Thakshila Wimalajeewa and Pramod K Varshney. Omp based joint sparsity pattern recovery under communication constraints. *IEEE Transactions on Signal Processing*, 62(19):5059–5072, 2014.
- [104] Kai Wu, Wei Ni, Tao Su, Ren Ping Liu, and Y Jay Guo. Recent breakthroughs on angle-of-arrival estimation for millimeter-wave high-speed railway communication. *IEEE Communications Magazine*, 57(9):57–63, 2019.
- [105] Ming Xiao, Shahid Mumtaz, Yongming Huang, Linglong Dai, Yonghui Li, Michail Matthaiou, George K Karagiannidis, Emil Björnson, Kai Yang, I Chih-Lin, et al. Millimeter wave communications for future mobile networks. *IEEE Journal on Selected Areas in Communications*, 35(9):1909–1935, 2017.

- [106] Wenwu Xie, Jian Xiao, Peng Zhu, and Chao Yu. Multi-task learning-based channel estimation for RIS assisted multi-user communication systems. *IEEE Communications Letters*, 2021.
- [107] Shujun Ye, Guanhao Liu, Yan Zhou, and Bo Dang. The enhanced beam sweeping algorithm for doa estimation in the hybrid analog-digital structure with nested array. *Scientific Reports*, 12(1):1–7, 2022.
- [108] Xianghao Yu, Jun Zhang, and Khaled B Letaief. Partially-connected hybrid precoding in mm-wave systems with dynamic phase shifter networks. *arXiv preprint arXiv:1705.00859*, 2017.
- [109] Jian A Zhang, Xiaojing Huang, Val Dyadyuk, and Y Jay Guo. Massive hybrid antenna array for millimeter-wave cellular communications. *IEEE Wireless Communications*, 22(1):79–87, 2015.
- [110] Ruoqiao Zhang, Jianyi Zhou, Ji Lan, Binqi Yang, and Zhiqiang Yu. A high-precision hybrid analog and digital beamforming transceiver system for 5g millimeter-wave communication. *IEEE Access*, 7:83012–83023, 2019.
- [111] Ruoyu Zhang, Honglin Zhao, and Jiayan Zhang. Distributed compressed sensing aided sparse channel estimation in FDD massive MIMO system. *IEEE Access*, 6:18383–18397, 2018.
- [112] Zheng Zhang, Yong Xu, Jian Yang, Xuelong Li, and David Zhang. A survey of sparse representation: algorithms and applications. *IEEE access*, 3:490–530, 2015.
- [113] Kan Zheng, Long Zhao, Jie Mei, Bin Shao, Wei Xiang, and Lajos Hanzo. Survey of large-scale mimo systems. *IEEE Communications Surveys and Tutorials*, 17(3):1738–1760, 2015.
- [114] Xianwei Zheng, Yuan Yan Tang, and Jiantao Zhou. A framework of adaptive multiscale wavelet decomposition for signals on undirected graphs. *IEEE Transactions on Signal Processing*, 67(7):1696–1711, 2019.
- [115] Chengwei Zhou, Yujie Gu, Shibo He, and Zhiguo Shi. A robust and efficient algorithm for coprime array adaptive beamforming. *IEEE Transactions on Vehicular Technology*, 67(2):1099–1112, 2017.

LIST OF PUBLICATIONS

International Scientific Conferences

M. Trigka, C. Mavrokefalidis and K. Berberidis, "An Effective Preprocessing Scheme for DoA Estimation in Hybrid Antenna Arrays," 2018 25th International Conference on Telecommunications (ICT), 2018, pp. 127-131, doi: 10.1109/ICT.2018.8464911.

Abstract

The aim of this work is to study the direction-of-arrival estimation problem in hybrid antenna arrays. To this end, we recommend an effective preprocessing scheme with the aim to recover the received signal as if a conventional (non-hybrid) antenna array was employed. Simulation results are also presented that illustrate the performance of the proposed scheme when compared to conventional MUSIC algorithm and its hybrid version.

M. Trigka, C. Mavrokefalidis and K. Berberidis, "AoA Estimation Scheme for Fully-Connected Hybrid Architecture Antenna Arrays," 2019 13th European Conference on Antennas and Propagation (EuCAP), 2019, pp. 1-5.

Abstract

In the context of this research work, we study the Angle-of-Arrival estimation problem in a fully-connected hybrid antenna array structure. The estimation problem is considered at the receiver side where the hybrid architecture exploits in a novel way all antenna elements of a uniform linear array. To this end, a preprocessing scheme proposed in [9] is properly extended so as to be applicable to the fully-connected architecture. The scheme is studied for the cases of non-constant and constant envelope signal sources. The new scheme enables recovery of the snapshots that would have been obtained if a conventional (non-hybrid), uniform, linear antenna array was employed. Simulation results are also presented that illustrate the performance of the proposed scheme, when compared to the conventional MUSIC algorithm and a recently proposed hybrid version of MUSIC.

M. Trigka, C. Mavrokefalidis and K. Berberidis, "A Distributed Sparse Channel Estimation Technique for mm Wave Massive MIMO Systems," 2021 29th European Signal Processing Conference (EUSIPCO), 2021, pp. 2154-2158, doi: 10.23919/EUSIPCO54536.2021.9616333.

Abstract

In this paper, we study the problem of sparse channel estimation via a collaborative and fully distributed approach. The estimation problem is formulated in the angular domain by exploiting the spatially common sparsity structure of the involved channels in a multi-user scenario. The sparse channel estimation problem is solved via an efficient distributed approach in which the participating users collaboratively estimate their channel sparsity support sets, before locally estimate the channel values, under the assumption that global and common support subsets are present. The performance of the proposed algorithm, named WDiOMP, is compared to DiOMP, local OMP and a centralized solution based on SOMP, in terms of the support set recovery error under various experimental scenarios. The efficacy of WDiOMP is demonstrated even in the case in which the underlining sparsity structure is unknown.

M. Trigka, C. Mavrokefalidis and K. Berberidis, "An efficient decentralized approach for mmWave MIMO Channel Estimation," 2021 International Balkan Conference on Communications and Networking (BalkanCom), 2021, pp. 31-35, doi: 10.1109/BalkanCom53780.2021.9593124.

Abstract

In the physical propagation environment, the channel matrices of neighboring users exhibit a joint sparsity structure due to the shared scatterers at the Base Station (BS) side. Based on this observation, we consider a multi-user Multiple-Input Multiple-Output (MIMO) system where the sparse channel estimation problem is tackled via an efficient fully distributed approach based on compressive sensing (CS). The involved users cooperatively estimate the sparsity support sets of the involved channels before individually estimate the channel coefficients, assuming that global and common sparsity support subsets exist. The performance of the proposed algorithm, named Weighted Distributed Simultaneous Orthogonal Matching Pursuit (WDiSOMP), is compared to Distributed Simultaneous Orthogonal Matching Pursuit (DiSOMP), local Simultaneous Orthogonal Matching Pursuit (SOMP) and a centralized solution based on SOMP in

terms of the channel estimation under a multi-tasking scenario.

International Scientific Journals

M. Trigka, C. Mavrokefalidis and K. Berberidis: Full snapshot reconstruction in hybrid architecture antenna arrays. *EURASIP J. Wirel. Commun. Netw.* 2020(1): 243 (2020)

Abstract

In the context of this research work, we study the so-called problem of full snapshot reconstruction in hybrid antenna array structures that are utilized in mmWave communication systems. It enables the recovery of the snapshots that would have been obtained if a conventional (non-hybrid) uniform linear antenna array was employed. The problem is considered at the receiver side where the hybrid architecture exploits in a novel way the antenna elements of a uniform linear array. To this end, the recommended scheme is properly designed so as to be applicable to overlapping and non-overlapping architectures. Moreover, the full snapshot recoverability is addressed for two cases, namely for time-varying and constant signal sources. Simulation results are also presented to illustrate the consistency between the theoretically predicted behaviors and the simulated results, and the performance of the proposed scheme in terms angle-of-arrival estimation, when compared to the conventional MUSIC algorithm and a recently proposed hybrid version of MUSIC (H-MUSIC).

M. Trigka, C. Mavrokefalidis and K. Berberidis: Efficient Distributed Multi-task Schemes for mmWave MIMO Channel Estimation. *IEEE Access*, 2022.

Abstract In this study, the problem of sparse channel estimation is investigated with the employment of a fully distributed approach. We exploit the spatially joint sparsity structure of the involved channels to formulate the channel estimation problem in the angular domain. The devices collaboratively estimate their channel sparsity support sets before the local estimation of the channel values, assuming the existence of global and common support subsets. The combination of the proposed distributed scheme with the classical Simultaneous Orthogonal Matching Pursuit (SOMP) algorithm is called Weighted Distributed Simultaneous Orthogonal Matching Pursuit (WDiSOMP). The performance of WDiSOMP is assessed under a multitask scheme considering a new weighted voting method. The new mechanism is applied at each support set index and enhances its estimation accuracy and, thus, the eventual estimation of the overall

channel. The mean squared error (MSE) is utilized to derive the performance bounds and assess the efficacy of the WDiSOMP estimator. Finally, the performance and the theoretical findings are evaluated via the comparison of WDiSOMP with Distributed Simultaneous Orthogonal Matching Pursuit (DiSOMP), local SOMP, and a centralized approach based on Structured SOMP (SSOMP) in terms of the MSE.

