Channel Estimation for Massive MIMO Systems

Mingduo Liao

PhD

University of York

Electronic Engineering

November 2022

Abstract

Massive multiple input multiple output (MIMO) systems can significantly improve the channel capacity by deploying multiple antennas at the transmitter and receiver. Massive MIMO is considered as one of key technologies of the next generation of wireless communication systems. However, with the increase of the number of antennas at the base station, a large number of unknown channel parameters need to be dealt with, which makes the channel estimation a challenging problem. Hence, the research on the channel estimation for massive MIMO is of great importance to the development of the next generation of communication systems. The wireless multipath channel exhibits sparse characteristics, but the traditional channel estimation techniques do not make use of the sparsity. The channel estimation based on compressive sensing (CS) can make full use of the channel sparsity, while use fewer pilot symbols. In this work, CS channel estimation methods are proposed for massive MIMO systems in complex environments operating in multipath channels with static and time-varying parameters. Firstly, a CS channel estimation algorithm for massive MIMO systems with Orthogonal Frequency Division Multiplexing (OFDM) is proposed. By exploiting the spatially common sparsity in the virtual angular domain of the massive MIMO channels, a dichotomous-coordinate-decent-joint-sparse-recovery (DCD-JSR) algorithm is proposed. More specifically, by considering the channel is static over several OFDM symbols and exhibits common sparsity in the virtual angular domain, the DCD-JSR algorithm can jointly estimate multiple sparse channels with low computational complexity. The simulation results have shown that, compared to existing channel estimation algorithms such as the distributed-sparsity-adaptive-matching-pursuit (DSAMP) algorithm, the proposed DCD-JSR algorithm has significantly lower computational complexity and better performance. Sec-

ondly, these results have been extended to the case of multipath channels with time-varying parameters. This has been achieved by employing the basis expansion model to approximate the time variation of the channel, thus the modified DCD-JSR algorithm can estimate the channel in a massive MIMO OFDM system operating over frequency selective and highly mobile wireless channels. Simulation results have shown that, compared to the DCD-JSR algorithm designed for time-invariant channels, the modified DCD-JSR algorithm provides significantly better estimation performance in fast time-varying channels.

Table of contents

List of tables

List of figures

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Dr. Yuriy Zakharov, for his guidance, support and encouragement during my Ph.D study.

I would like to express my sincere thanks to my family, without whom none of this work could have possibly happened.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Publications

Publications directly associated with thesis work

- M. Liao and Y. Zakharov, "DCD-based joint sparse channel estimation for OFDM in virtual angular domain," *IEEE Access*, vol. 9, pp. 102 081–102 090, 2021
- M. Liao and Y. Zakharov, "Estimation of time-varying channels in virtual angular domain for massive MIMO systems," *IEEE Access*, vol. 11, pp. 1923–1933, 2023

Mingduo Liao

November 2022

Chapter 1

Introduction

1.1 Research Background

In the past decades, with the development of wireless communication technology, multiple-input multiple-output (MIMO) has become the core technology of the fourth generation mobile communication system (4G) [7][8][9]. Compared with the third-generation mobile communication system, the 4G network can provide user terminals with greater bandwidth and longer coverage, higher data transfer rates, and system capacity [10] [11].With the rapid development of Internet technology, more data will be transmitted to user terminals through the wireless channel. However, with the increasing demand for data transmission rate by the users, and the increasingly prominent contradiction between limited spectrum resources and the rapidly growing number of wireless terminals, communication systems need to introduce more advanced wireless transmission technologies to provide reliable communication [12]. The commercial use of the 5th Generation Mobile Communication System (5G) has already been deployed. Compared with the 4G network, the 5G network can provide about 1000 times the capacity upgrade, higher peak data rate [13]. The 5G wireless communication networks are being deployed worldwide from 2020 and more capabilities are in the process of being standardized, such as mass connectivity, ultra-reliability, and guaranteed low latency. However, 5G will not meet all requirements of the future in 2030 and beyond, and sixth generation (6G) wireless communication networks are expected to provide global coverage, enhanced spectral/energy/cost efficiency, better intelligence level and security, etc [14]. To meet these requirements, 6G networks will rely on new enabling technologies, i.e., air interface and transmission technologies and novel network architecture, such as massive MIMO, waveform design, multiple access, channel coding schemes, multi-antenna technologies, network slicing, cell-free architecture, and cloud/fog/edge computing. In [15], it indicates that 6G will have four new advantages: first, to satisfy the requirement of global coverage, 6G will not be limited to terrestrial communication networks, which will need to be complemented with non-terrestrial networks such as satellite networks, thus achieving a space-airground-sea integrated communication network. Second, all spectra will be fully explored to further increase data rates and connection density, including the sub-6 GHz, millimeter wave (mmWave), terahertz (THz), and optical frequency bands. Third, facing the big datasets generated by the use of extremely heterogeneous networks, diverse communication scenarios, large numbers of antennas, wide bandwidths, and new service requirements, 6G networks will enable a new range of smart applications with the aid of artificial intelligence (AI) and big data technologies. Fourth, network security will have to be strengthened when developing 6G networks.

The core technology of the 4G wireless communication network system includes the MIMO technology and orthogonal frequency division multiplexing (OFDM) technology. The MIMO technology configures multiple antennas at the transmitting end and the receiving end (the number of antennas configured at the transmitting and receiving end generally does not exceed 8 [16]). Then, the terminal at the receiving end receives the signal and performs operations such as channel estimation, to recover the transmitted data accurately. The MIMO technology improves the spectral efficiency of the communication system, which lays a foundation for each user to perform high-data-rate communications. Furthermore, under the spatial multiplexing gain and spatial diversity gain provided by the MIMO technology, the use of the MIMO technology can not only improve the channel capacity in the system but also effectively improve the stability and reliability of the system [17] [18]. The OFDM technology is a multi-carrier modulation scheme used in wireless local area networks, digital audio broadcasting, and other fields, this technology decomposes the data stream into multiple sub-data streams to reduce the transmission rate, and then transmits them in parallel on multiple sub-carriers that are mutually orthogonal to each other, so as to achieve the purpose of combating the channel frequency selective fading or narrowband interference. By combining OFDM and the MIMO together, the MIMO-OFDM cannot only improve the transmission and spectral efficiency of wireless communication systems, but also improve the reliability of wireless systems and tolerance to multipath interference and environmental noise through diversity technology. The MIMO-OFDM technology is used in the LTE (Long Term Evolution, a standard for wireless broadband communication for mobile devices and data terminals, offers high-speed data communication, low latency, and advanced features such as multimedia streaming, high-quality voice and video calls, and mobile broadband internet access), WiMAX (a wireless communication technology that provides high-speed broadband connectivity over long distances. It is based on the IEEE 802.16 standard and operates on licensed or unlicensed frequencies.), and network protocol standards such as IEEE 802.15.4 (a standard that defines the physical and MAC layer specifications for low-rate wireless personal area networks) [19].

With the rapid growth of global mobile communication service demand, the traditional MIMO system has gradually reached its bottleneck in performance, due to the limitations of its theoretical basis and hardware configuration [20]. Additionally, the problems faced by 4G wireless communication systems have become increasingly prominent, including the difficulty to satisfy users' demands for data transmission rate, the scarcity of available spectrum resources, the saturation of communication system capacity, and the continuous increase in the power consumption of communication equipment [21].

According to current industry research, the improvement of the network service capabilities of 5G and future wireless mobile communication systems will be carried out simultaneously in three research directions: wireless transmission technology, wireless network technology and new spectrum resource mining. Among them, for wireless transmission technology, scholars have proposed Massive MIMO technology [22], filter bank-based multi-carrier modulation technology, full-duplex (Full-duplex) technology [23], Non-Orthogonal Multiple Access (NOMA) technology [24], etc. For wireless network technology, scholars have proposed new architectures, such as ultra-dense heterogeneous network technology [25] and end-to-end communication (Device-to-Device Communication, D2D) technology [26]. In terms of mining new spectrum resources, millimetre wave (mmWave) communication technology [27], visible light communication technology [28], etc., have become valuable transmission solutions in the field of wireless mobile communication in the future.

In view of the three types of key technologies for improving the capacity of 5G wireless communication network systems, the massive MIMO technology has become a bright spot among many alternative key technologies, attracting scholars and engineers. It is widely regarded as the most promising transmission technology for 5G communication networks [29].

In the traditional MIMO wireless communication system, the base station is usually configured with up to eight antennas. In 2010, Bell Labs professor Thomas L. Marzetta first proposed the concept of massive MIMO [30]. In the massive MIMO communication systems, by employing large number of antennas at the base station, and through the spatial multiplexing, the spectral efficiency of the wireless communication system can be greatly increased. Compared with the traditional MIMO system, the massive MIMO system has better communication performance, the main advantages of massive MIMO systems are [31] [32] [33]:

- The capacity of the communication system has greatly increased. In a massive MIMO system, as the number of antennas configured at the base station increases, the wireless transmission channel between each terminal and the antenna at the base station exhibits a progressive orthogonality characteristic. At this time, the interference of additive white Gaussian noise and small-scale fading (such as Rayleigh fading and Rice fading) in the the communication system can be ignored or effectively eliminated. According to Shannon's theorem, the channel capacity of the communication system will be greatly increased.

- The spectral efficiency of the communication system is greatly improved. Compared with the traditional MIMO system, the massive MIMO communication system's ability to analyse the spatial dimension is significantly enhanced. Therefore, massive MIMO technology can improve the utilization of spectrum resources among multiple terminals in the same cell.

In a wireless communication system, the acquisition of channel parameters (such as multipath gain and delay information, etc.) of the wireless channel between the transmitter and receiver is important, which relates to the performance of the signal detection and decoding. Channel parameters can be obtained through channel estimation techniques, therefore, the accuracy of channel estimation can have a significant impact on the performance of the massive MIMO-OFDM wireless communication systems.

However, the number of antennas in a massive MIMO system generally ranges from tens to hundreds, and the number of channels to be estimated between the base station and the user are proportional to the number of antennas at the base station. Therefore, as the number of antennas at the base station of the massive MIMO system increases, the computational complexity of the channel estimation process will increase accordingly, and the pilot overhead of the system will also increase. At present, massive MIMO communication systems generally operate in two modes: Time Division Duplex (TDD) and Frequency Division Duplex (FDD). Among them, due to the limited number of users in the TDD massive MIMO system, the uplink channel is easier to be estimated at the base station, and the channel information obtained by the uplink channel can be used to estimate the downlink channel, according to the channel reciprocity. Therefore, most researchers use TDD mode to avoid problems such as high computational complexity and large pilot overhead. However, in a massive MIMO system in TDD mode, there might be a calibration error between the uplink and the downlink channels, where the calibration error could be the result of the Doppler effect and signal attenuation, and this would lead to errors in the measurements of signal quality, such as signal-to-noise ratio, which can affect the performance of the system, furthermore, it is easy to produce pilot contamination, which would cause inaccurate channel estimation [34]. Because of its good ability to resist interference between different cells, FDD still occupies the main position in the wireless communication systems. For the base station, obtaining accurate downlink channel information in a massive MIMO system based on FDD mode is a very challenging problem, due to the computational complexity of downlink channel estimation, and the high pilot overhead in the signal feedback process [35]. Therefore, in order to further improve the communication quality of massive MIMO systems, research on high-performance and high-accuracy wireless channel estimation algorithms has become a hot research topic.

1.2 Existing Research and Challenges

In the MIMO wireless communication systems, channel estimation techniques are usually divided into blind channel estimation and non-blind channel estimation, according to whether pilot symbols are required. The blind channel estimation algorithm recovers the channel state information (CSI) by making full use of the statistical characteristics of the data at the receiver, without the need for pilot symbols [36]. For MIMO-OFDM systems, the literature proposes an effective blind channel estimation technique based on subspace method [37], this algorithm mainly uses the subspace of the correlation matrix of the received signal, and the orthogonality of the noise signal to estimate the coefficients of the unknown channel. Since in practical communication systems, the statistical properties of the received signal such as the correlation matrix are unknown, most existing blind channel estimation algorithms require enough received blocks to construct a reliable sample covariance matrix [38]. In [39], by collecting the covariance matrix of the signal from a set of selected subcarriers, and employing a low-order constellation diagram, a blind channel estimation algorithm for MIMO-OFDM systems that reduces the number of required receive blocks is proposed. Since the blind channel estimation algorithm does not need pilot information, it can make the communication system reduce the dependence on the pilot. However, this method still requires a long coherence time to obtain accurate CSI, and the algorithm is based on statistical analysis, which has high computational complexity. Generally, blind channel estimation algorithm is difficult to meet the needs of high-speed and high-quality communication in wireless channel scenarios with time-varying environments. Therefore, most wireless communication systems with high data rates and low delay adopt the non-blind channel estimation algorithm based on pilot symbols.

The non-blind channel estimation algorithm based on pilot symbols usually inserts an appropriate number of a pilot symbols into the transmitted signal in some way (such as block, comb, etc.), and the receiving end uses the known pilot symbols and the received signal information for channel estimation [40]. These kinds of channel estimation algorithm can be generally divided into pilot-based channel estimation algorithm based on time domain pilots and pilot-based channel estimation algorithm based on frequency domain pilots. Among them, the pilot-based channel estimation algorithm based on the time-domain pilot estimate the channel in the time-domain, and then applying the fast Fourier transform to get the channel information in frequency domain. The pilot-based channel estimation algorithm based on the frequency domain pilot first obtains the frequency domain state response of the wireless channel at each pilot symbol, and then obtain the frequency domain channel information at each data symbol through interpolation. Traditional linear channel estimation techniques based on pilot symbols include the Least Squares (LS) algorithm [41], the Minimum Mean-Square Error (MMSE) algorithm [42], and so on. Among them, the LS method adopts the method of pseudo-inverse solution, this algorithm is simple in operation but is sensitive to the influence of environmental noise, especially in the case of low signal-to-noise ratio (SNR), the accuracy of the obtained CSI is low. The MMSE method greatly improves the performance of channel estimation by using the second-order statistics of the channel, however, the computational complexity is high.

In recent years, many research works have indicated that, with the rapid increase in the channel bandwidth of the wireless communication system and the multiplication of the antennas at the base station, many wireless multipath channels related to the scattering environment show sparse characteristics in the time domain, frequency domain, and spatial domain [43] [44] [45]. Traditional algorithms such as LS and MMSE rely on the linear reconstruction scheme and do not exploit the inherent sparse nature of the channel, resulting in over-utilization of communication resources such as the number of pilot symbols. This problem has been solved by making full use of the sparse characteristics of wireless channels, by employing the compressive sensing (CS) theory to sparse channel estimation in wireless communication systems, this system can accurately estimate CSI with low pilot overhead, thereby improving the utilization of spectrum resources.

1.2.1 Sparse channel estimation for MIMO-OFDM systems

For sparse channel estimation in the MIMO-OFDM systems, a lot of research has been carried out by scholars. The work [46] introduced the concept of channel sparsity and proposed a new method for sparse channel estimation, namely the compressed channel sensing technique, based on the full exploitation of the sparse structure of the wireless channel and the compressive sensing theory; the application of this technique enables wireless transceivers to sense and adapt to the wireless environment in order to improve spectral efficiency and resist noise interference. Wireless communication scenarios are usually accompanied by multipath effects, which causes the frequency-selective fading of the channel. Based on the MIMO-OFDM systems, [47] converts the frequency-selective sparse channel estimation problem into a compressive sensing problem, and proposes a practical suboptimal solution method that reconstructs the channel coefficients, using the Orthogonal Matching Pursuit (OMP) algorithm, which provides a high channel reconstruction accuracy. In [48], an optimised OMP algorithm is proposed to further improve the accuracy of the channel estimation, however, the computational complexity is high. In order to reduce the computational complexity, the Compressive Sampling Matching Pursuit (CoSaMP) algorithm was proposed in [49], which effectively improves the spectrum resource utilization at the cost of a smaller computational complexity. However, all the above algorithms require the sparsity of channel state information as a known condition for the reconstruction process, so the practicality of these algorithms is limited. In [50], an Adaptive Step Size Sparsity Adaptive Matching Pursuit (AS-SaMP) based channel estimation algorithm was proposed. Compared to other sparse algorithms, this algorithm does not require a *priori* knowledge of channel sparsity and can adaptively adjust the step size to approximate the true sparsity. A frequency-selective channel estimation scheme based on Block Stagewise Orthogonal Matching Pursuit (Block StOMP) for the MIMO-OFDM systems is proposed in [51], this algorithm is able to reconstruct the CSI with high accuracy by exploiting the common support set of different sparse channel impulse responses. However, in the wireless communication environment of high-speed mobile scenarios, the channel is not only affected by the frequency-selective fading caused by the multipath effect, but also influenced by the time-selective fading caused by the Doppler effect. This complex wireless channel is often referred to as a time-frequency doubly selective fading channel. In this case, the coefficient of the channel impulse response is time-varying, which means that a large number of channel parameters need to be estimated at the receiving end. As a result, recovering a time-frequency doubly-selective fading channel requires more pilot symbols than a frequency-selective fading channel, which inevitably reduces spectrum resource utilisation and makes channel estimation more challenging [52] [53]. To address the complex time-frequency doubly-selective fading channel estimation problem, [3] exploits the temporal and spatial correlation of the channel: by exploiting the temporal correlation of the channel, the fact that the wireless channel varies slowly over time is utilized to estimate the channel, which makes channel estimation more accurate, the spatial correlation of the channel coefficients is used to reduce the number of channel parameters that need to be estimated, which leads to better estimation accuracy.

1.2.2 Sparse channel estimation for Massive MIMO systems

The work [54] utilizes the sparse structure characteristic of TDD mode massive MIMO channel in delay domain and spatial domain, by converting Multiple Measurement Vectors (MMV) into Single Measurement Vector (SMV), it proposed a channel estimation technique based on pilot sequence and compressive sensing, namely Adaptive Orthogonal Matching Pursuit (AOMP). Although this algorithm can reduce the pilot overhead and further improve the accuracy of channel estimation, the overall computational complexity is high, which is not conducive to be used in wireless communication network with high transmission rate of massive MIMO system.

Although the massive MIMO system in the TDD mode has the characteristics of reciprocity and provides a solution to the problem of high pilot overhead, many wireless communication systems still use the FDD mode at present. Therefore, it is of great importance to solve the channel estimation problem of massive MIMO systems in FDD mode. When the system adopts the FDD mode, the channel reciprocity feature of the communication system no longer exists, and the channels corresponding to the uplink and downlink wireless links are independent from each other. For the FDD mode, the downlink of the wireless communication system can directly use the pilot symbol to estimate the channel, and the base station needs to obtain the estimated channel information through the feedback from the uplink channel. For downlink channel estimation of massive MIMO systems in FDD mode, [55] uses compressive sensing technology to reduce the pilot overhead in the channel estimation and feedback process, and uses the joint sparsity of the channel to propose a compressive sensing channel estimation scheme, namely Joint Orthogonal Matching Pursuit (JOMP), so that the compressed signal is received at the user end, while channel estimation is performed at the base station end. Furthermore, [56] proposed a method for channel estimation using the local common support set of the channel, this method utilizes the temporal correlation in the acquired channel, while provides accurate channel estimation, and significantly reduces pilot overhead.

After channel estimation, the user can use various decoding techniques to recover the transmitted symbol from the received signal. In digital communication systems, the transmitted symbol is typically encoded and modulated before transmission. The receiver can use techniques such as demodulation and decoding to recover the original symbol from the received signal. The channel estimation information can be used to compensate for the channel effects on the received signal, thereby improving the accuracy of the decoding process.

1.2.3 Challenges

With current research work, although the sparse channel estimation technology of the MIMO-OFDM system and massive MIMO system has made some progress, it still faces the following challenges that need to be addressed:

1. The accuracy of sparse channel estimation in the MIMO-OFDM systems still needs to be further improved. In traditional channel estimation, the number of employed pilot symbols is determined by the Nyquist sampling theory, so the communication

system needs more pilots to estimate the channel accurately, and the rate of spectrum resource utilization is low. At present, although sparse channel estimation algorithms based on compressive sensing have been widely used in the MIMO-OFDM systems, which could reduce the number of pilot symbols to a certain extent. However, with the increasing complexity of wireless communication systems, they still cannot meet the requirements of the MIMO-OFDM systems for obtaining accurate CSI, and most of the current sparse channel estimation algorithms require the $priori$ knowledge of channel sparsity, which is not suitable for practical environment. Therefore, it is of great importance to develop channel estimation algorithms that do not depend on $priori$ knowledge of channel sparsity, and to find methods that can effectively improve the performance of sparse channel estimation.

2. There is "pilot contamination" in the uplink of TDD massive MIMO systems. In a TDD massive MIMO system, when multiple terminals in multiple adjacent cells use the same pilot symbol sequence and send pilots to the base station at the same time, the base station will not only receive the data sent by a certain terminal, but also receive the same pilot symbols sent by terminals from other cell at the same time. At this time, data interference may cause disorder in the channel estimation process, and this phenomenon is called "pilot contamination". With the continuous increase of the number of antenna at the base station of a massive MIMO system, the spectral efficiency and capacity of the communication system have been further improved, but the pilot contamination has become more serious, and has become a key factor affecting the performance of the massive MIMO systems. Therefore, it is of great significance to reduce pilot contamination by optimizing pilot design and channel estimation techniques.

3. The accuracy of downlink channel estimation of FDD massive MIMO system is low. At present, the downlink sparse channel estimation for massive MIMO systems in FDD mode mainly focuses on using the sparse characteristics of time, frequency and space domains. However, these methods usually adopt certain techniques for

channel estimation under specified sparse characteristics. For example, when the channel matrices of different users are sparse and share common support sets, this sparse feature can provide new ideas and methods for channel estimation. However, this sparse feature is difficult to keep consistent in different scenarios and environments. Therefore, wireless communication systems need to adopt different sparse channel estimation techniques to adapt to complex communication environment. and it is necessary to find more efficient downlink channel estimation algorithms for FDD massive MIMO systems.

4. The channel estimation efficiency of massive MIMO system is low. For FDD massive MIMO systems, sparse channel estimation technology based on compressive sensing has been widely studied and applied in recent years. However, most of these traditional greedy-based compressive sensing sparse channel estimation algorithms use an iterative optimization strategy, to solve the underdetermined optimization problem in the compressive sensing model. With the increase of the number of antennas at the base station, the scale of the channel of the wireless communication system also increases exponentially. The intensive computation of the iterative optimization and the inability to guarantee the global optimum, have become the bottleneck in the application of compressive sensing in channel estimation, thus limiting the compressive sensing technology to non-real-time application scenarios.

1.3 Conclusion

In this chapter, we have introduced the research background of this work, and discussed the current research in channel estimation techniques for MIMO and massive MIMO systems, and discussed main challenges of the channel estimation in the massive MIMO systems. Since main challenges of the channel estimation in the massive MIMO systems are the channel estimation accuracy, pilot contamination and channel estimation complexity, therefore, in this work, we focus on solving the channel estimation problem in FDD massive MIMO system, thus proposed channel estimation algorithms that have high accuracy and low computational complexity. The main content and research results of this thesis are as follows:

1. A CS channel estimation algorithm for massive MIMO systems with Orthogonal Frequency Division Multiplexing (OFDM) is proposed. By exploiting the spatially common sparsity in the virtual angular domain of the massive MIMO channels, a dichotomous-coordinate-decent-joint-sparse-recovery (DCD-JSR) algorithm is proposed. More specifically, by considering the channel is static over several OFDM symbols and exhibits common sparsity in the virtual angular domain, the DCD-JSR algorithm can jointly estimate multiple sparse channels with low computational complexity. The simulation results have shown that, compared to existing channel estimation algorithms such as the distributed-sparsity-adaptive-matching-pursuit (DSAMP) algorithm, the proposed DCD-JSR algorithm has significantly lower computational complexity and better performance.

2. These results have been extended to the case of multipath channels with time-varying parameters. This has been achieved by employing the basis expansion model to approximate the time variation of the channel, thus the modified DCD-JSR algorithm can estimate the channel in a massive MIMO OFDM system operating over frequency selective and highly mobile wireless channels. Simulation results have shown that, compared to the DCD-JSR algorithm designed for time-invariant channels, the modified DCD-JSR algorithm provides significantly better estimation performance in fast time-varying channels.

Chapter 2

Fundamental Techniques

2.1 Introduction

In this chapter, we will introduce the fundamental techniques used in this work. In Section 2.2, the MIMO system is introduced, after that, in Section 2.3, we introduce OFDM techniques, and both of them are key technologies for the next generation of wireless communication systems. Furthermore, in Section 2.4, we introduce the massive MIMO, which is an extension of the MIMO technology. We introduce the time varying channel model in Section 2.5. In Section 2.6, we discuss traditional channel estimation algorithms and pilot symbols, and then the sparse channel in Section 2.7.

In this work, capital and small bold fonts are used to denote matrices and vectors, respectively, $(\mathbf{x})_n$ denotes the nth element of the vector $\mathbf{x}$, $\mathbf{R}^{\mathrm{q}}$ denotes the qth column of the matrix $\mathbf{R}$, and $\mathbf{R}_{\mathrm{n}}$ denotes the nth row of the matrix $\mathbf{R}$, $\mathbf{R}_{\mathrm{m,n}}$ denotes an element of the matrix $\mathbf{R}$. The transpose operator is given by $(.)^{\mathrm{T}}$, $(.)^{*}$ denotes the conjugate operator, $(.)^{\dagger}$ denotes the Moore-Penrose inversion, and $(.)^{\mathrm{H}}$ denotes the Hermitian transpose operator. The ℓ_0-norm and ℓ_2-norm are represented by $||.\|_0$ and $||.\|_2$, respectively. We use I to denote a support, $|\mathrm{I}|$ is the cardinality of the support I, I^{c} is the complement of I, $\mathbf{R}_{\mathrm{I}}$ is a matrix obtained from $\mathbf{R}$, and which only contains rows corresponding to support I. $\mathbf{R}_{\mathrm{I,I}}$ is an $|\mathrm{I}| \times |\mathrm{I}|$ matrix obtained from $\mathbf{R}$ by collecting elements from columns and rows corresponding to I, and $\mathbf{x}_{\mathrm{I}}$ is the subset of $\mathbf{x}$ that includes non-zero elements from $\mathbf{x}$ corresponding to I. We

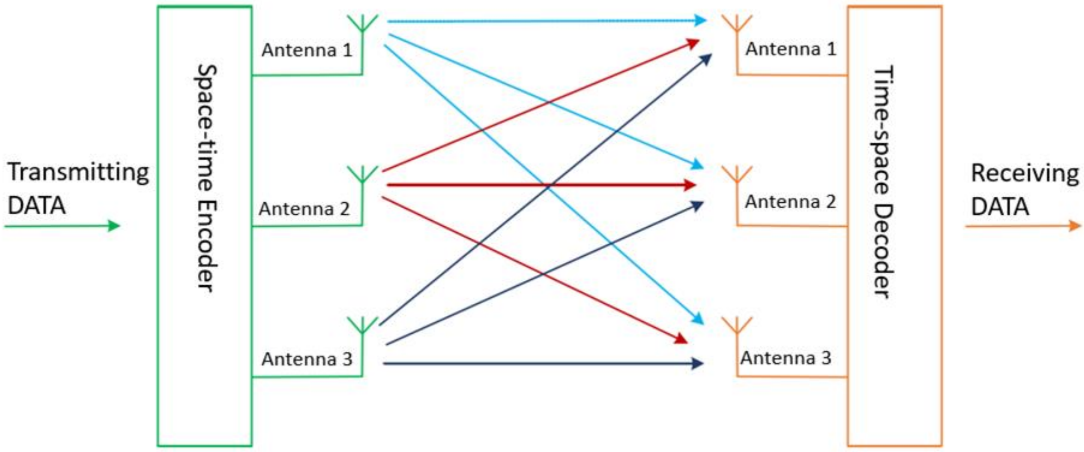

Fig. 2.1 Block diagram of the MIMO system (adapted from [1])

use $\mathbf{h}$ to denote a channel vector and $\tilde{\mathbf{h}}$ to denote the channel vector in the virtual angular domain, $\tilde{\mathbf{h}}_{\mathrm{n}}$ denotes the channel vector corresponding to the nth subcarrier. $\Re$ denotes the real part of a complex number, and $\mathrm{j} = \sqrt{-1}$.

2.2 MIMO systems

The MIMO technology is one of the key technologies for the next generation of wireless communication systems, the architecture of a typical the MIMO system is shown in Fig. 2.1. By configuring multiple antennas at the transmitter and receiver of the wireless link to transmit and receive data at the same time, the MIMO technology could reduce the channel fading. Furthermore, in a MIMO wireless communication system, by considering the impulse responses of the wireless channels between the antennas at the transmitting end and the antennas at the receiving end are independent of each other, we can consider that there are multiple parallel signal transmission channels in the communication environment. Therefore, for wireless communication systems, the MIMO technology can make full use of the effective resources in the space domain, and significantly enhance the performance in terms of data transmission rate and anti-interference, while improving the channel capacity of the system

[17] [18].Thus, the characteristics of the MIMO technology are mainly reflected in: diversity, space division multiplexing, beamforming, etc., which are respectively introduced as follows

The channel capacity is the maximum rate at which information can be transmitted over a channel subject to a certain level of noise and distortion. For a MIMO system with M transmit antennas and n_r receive antennas, the channel capacity is given by the following equation:

$$C = \log_2 \det \left(\mathbf{I}_{n_r} + \frac{P}{\sigma^2} \mathbf{H}\mathbf{H}^H \right) bps/Hz \tag{2.1}$$

where C is the channel capacity in bits per second (bps) per Hz , P is the transmit power, σ^2 is the noise power, $\mathbf{H}$ is the $n_r \times M$ channel matrix, and $\mathbf{I}_{n_r}$ is the $n_r \times n_r$ identity matrix. The matrix $\mathbf{H}$ represents the wireless channel between the transmitter and receiver, and each element of the matrix corresponds to the channel gain between a particular transmit antenna and receive antenna.

1. Diversity Technique

Diversity techniques are often embodied in two aspects: transmit diversity and receive diversity. The transmit diversity means the antenna array transmits multiple independent and uncorrelated data with the same information, thereby overcomes the selective fading of the channel, and improves the quality and reliability of the communication system. This type of diversity technique is commonly referred to as Space Time Block Code (STBC) [57]. The most important algorithm of receive diversity is the maximum ratio combining (MRC) [57], which recovers the transmitted signal by multiplying the different received signals by the corresponding fading coefficients, and then combining the corresponding signals together, thus mitigating the fading effect and obtaining diversity gain [58]. The diversity gain can be expressed as follows [57]:

$$G = \lim_{P_t \to \infty} \frac{C}{\log_2 P_t} \tag{2.2}$$

where G is the diversity gain, C is the channel capacity, and P_t is the transmit power. The diversity gain represents the improvement in system performance due to the use of multiple antennas, and it depends on the spatial correlation of the wireless channel.

2. Space Division Multiplexing

In a MIMO system, by employing antenna array, space division multiplexing technique enables the same frequency band of the communication system to be reused in different spaces, thus, different beams can be formed in the directions of different users. Therefore, space division multiplexing can improve the bandwidth utilization of the communication system. In addition, when each transmit antenna port transmits signal streams simultaneously and independently, the application of space division multiplexing technology can increase the data transmission rate of the communication system, in proportion to the number of transmit antennas [59]. In a wireless communication system with M transmit antennas and n_r receive antennas, the transmitted signal y can be expressed as follows [57]:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} \tag{2.3}$$

The transmitted signal $\mathbf{y}$ can be separated into multiple spatial channels using a linear transformation $\mathbf{V}$, such that:

$$\mathbf{y}' = \mathbf{V}^H \mathbf{y} \tag{2.4}$$

where $\mathbf{y}'$ is the vector of transmitted signals after separation. The matrix $\mathbf{V}$ is chosen such that the columns of $\mathbf{V}$ are orthonormal, and the matrix $\mathbf{H}$ is approximately diagonal when multiplied by $\mathbf{V}$:

$$\mathbf{HV} \approx \mathbf{\Lambda} \tag{2.5}$$

where $\mathbf{\Lambda}$ is a diagonal matrix containing the singular values of $\mathbf{HV}$. The diagonal elements of $\mathbf{\Lambda}$ represent the channel gains of the spatial channels. Each spatial channel can be used to transmit a separate data stream, which can be detected at the receiver using maximum likelihood detection or other techniques. The overall data rate of the system can be increased by using more transmit antennas, which increases the number of spatial channels.

3. Beamforming

Different from space division multiplexing, which is used to increase the capacity of wireless communication systems by spatially separating multiple data streams, beamforming is a technology utilized in wireless communication systems to enhance the dependability and quality of signal transmission by directing radio waves towards a specific direction [60]. It involves the use of multiple antennas to form a directional beam of radio waves that can be steered towards the receiver. The basic concept of beamforming is to regulate the phases and amplitudes of signals transmitted by each antenna to constructively combine the waves in a particular direction while canceling them out in other directions [57]. This can be achieved either through analog beamforming or digital beamforming [61]. Analog beamforming uses phase shifters and amplifiers to adjust the signal transmitted by each antenna and combines them with the help of passive components to produce a directional beam of radio waves. On the other hand, digital beamforming digitally adjusts the phase and amplitude of the signal transmitted by each antenna through signal processing algorithms and combines them to create a directional beam of radio waves. Beamforming can be applied to various wireless communication systems, such as Wi-Fi networks, satellite communication systems, and cellular networks, to enhance the signal-to-noise ratio, expand the range of the signal, and diminish interference from other sources [57]. Beamforming is a crucial technique for enhancing the performance of wireless communication systems and is expected to become increasingly important in the future with the rising demand for wireless data.

2.3 OFDM techniques

As a key technology of broadband high-speed communication systems, the OFDM transmission scheme has been widely used in WiMAX (World Interoperability for Microwave Access) wireless communication system, and LTE (Long Term Evolution) standard. The basic principle of OFDM technology is to decompose a high-speed serial data stream into multiple low-rate sub-data streams, then the data are converted into a parallel sequence and divided into multiple subcarriers using inverse fast Fourier transform (IFFT). [62].

The transmitter can usually insert a guard interval between OFDM symbols, and the length of the guard interval is greater than the maximum delay spread in the wireless channel, so as to resolve the Inter-Symbol Interference (ISI) as much as possible, thus, it is ensured that the transmitted symbols will not affect each other due to multipath components. Besides, the process of signal propagation of a wireless communication would affected by multipath effect and Doppler effect, which would cause Inter-Carrier Interference (ICI), that is, destroy the mutually orthogonal relationship between the sub-carriers, thereby causing data interference between the sub-carriers. In order to eliminate ICI, the transmitting end can insert a cyclic prefix (CP) in the guard interval of the OFDM symbol [63].

2.4 Massive MIMO

Massive MIMO technology was proposed by Professor Marzett at Bell Labs [30]. This technology is considered as one of the most important technologies in the 5G communication system, which is a further extension of the MIMO technology in the existing 4G communication system [33].

In massive MIMO systems, the number of antennas at the base station is typically much larger than the number of users being served. This creates a highly over-determined system, where the base station has more degrees of freedom than necessary to communicate with all the users in the cell [31].

One of the key advantages of massive MIMO is its ability to increase the spectral efficiency of wireless communication systems [31]. Spectral efficiency refers to the amount of data that can be transmitted over a given frequency band. With traditional MIMO systems, the spectral efficiency increases with the number of antennas, but only up to a certain point. With massive MIMO, the spectral efficiency continues to increase with the number of antennas, as long as there are enough users to communicate with.

Another key advantage of massive MIMO is it achieveS higher data rates, due to following reasons [31]: first, massive MIMO employs a large number of antennas, which enables the base station to transmit multiple independent data streams to the UE. The more antennas are used, the more independent data streams can be transmitted simultaneously, which increases the spectral efficiency of the system. Second, massive MIMO benefits from spatial multiplexing, which means that the system can transmit different data streams to different UE in the same time and frequency resources. This allows multiple UE to be served simultaneously, which increases the overall capacity of the system. Third, massive MIMO uses advanced signal processing algorithms, such as precoding and beamforming, to optimize the transmission of the data streams. These algorithms take advantage of the spatial channel information to enhance the signal-to-interference-plus-noise ratio (SINR), which further increases the data rates that can be achieved.

In summary, massive MIMO can outperforms the traditional MIMO systems since it can harvest more benefits from spatial multiplexing and large number of antennas.

2.5 Simulator of time-varying fading channels

The analysis of wireless mobile channel is one of the most important parts in the research of mobile communication systems. Among them, multipath propagation and the Doppler effect are the main characteristics of a multipath fading wireless channel, the transmitted signal is reflected and refracted, which makes the received signal consist of a superposition of several waves, these waves may cause fading of the received signal. Therefore, it is necessary to characterize the channel response, to ensure that systems could operate at acceptable performance levels during fading.

Since the orientation and material properties of obstacles between the transmitter and receiver are usually not known in advance, or may vary with time, the received signal is often characterized as stochastic. In the case of the sum-of-sinusoid (SOS) simulator, the received signal is the sum of randomly-phased-sinusoids. The idea that a received signal can be represented as a superposition of a finite number of waves has been existed for decades [64]. Bello [65], Gilbert [66] and Clarke [67] were the first to propose these multipath fading channel model. After Bello introduced his model [68], Jakes designed an SOS simulator, which is widely used in the modeling and simulation of Rayleigh fading channels in urban areas.

The Clarkes' model considers a frequency-nonselective fading channel comprised of propagation paths, the low-pass fading process is given by [67]

$$\mathrm{g(t)} = \mathrm{E}_0 \sum_{n=1}^{\mathrm{N}} \mathrm{C_n} \exp[\mathrm{j}\omega_\mathrm{d}\mathrm{t}\cos\alpha_\mathrm{n} + \phi_\mathrm{n}], \tag{2.6}$$

where $\mathrm{g(t)}$ is the impluse response, E_0 is a scaling constant, $\mathrm{C_n}$, α_n and ϕ_n are random path gain, angle of incoming wave, and initial phase associated with the nth propagation path, and ω_d is the maximum radian Doppler frequency when $\alpha_\mathrm{n} = 0$.

Based on Clarkes' reference model, and by selecting [69]

$$\mathrm{C_n} = \frac{1}{\sqrt{\mathrm{N}}}, \tag{2.7}$$

$$\alpha_\mathrm{n} = \frac{2\pi\mathrm{n}}{\mathrm{N}},\ \mathrm{n} = 1, 2, ...\mathrm{N}, \tag{2.8}$$

$$\phi_\mathrm{n} = 0,\ \mathrm{n} = 1, 2, ...\mathrm{N}, \tag{2.9}$$

Jakes derived his well known simulation model for Rayleigh fading channels. The simplification in Equation(2.7) makes this model deterministic and wide-sense non-stationary [69].

2.5.1 Sparse channel

A sparse channel model assumes that the wireless channel is characterized by only a few significant propagation paths, while the remaining paths can be considered as noise or interference. Such a model is applicable in scenarios where the wireless propagation environment is dominated by line-of-sight (LOS) or near-LOS components. The channel impulse response in a sparse channel can be represented as a sparse vector, where the non-zero elements correspond to the significant propagation paths.

In practice, the sparsity of the wireless channel can be exploited to reduce the complexity of the channel estimation and equalization algorithms in a communication system. For example, compressed sensing techniques can be used to estimate the sparse channel from a limited number of measurements, while sparse equalization techniques can be used to reconstruct the transmitted symbols from the received signal.

Sparse channel models are commonly used in millimeter-wave communication systems, where the propagation environment is highly directional and the number of signal paths is limited. They are also used in massive MIMO systems, where the number of antennas is large and the channel response is expected to be sparse in the spatial domain.

2.5.2 Non-sparse channel

A non-sparse channel refers to a communication channel in which the channel matrix has a significant number of non-zero entries or coefficients [57]. In other words, the channel matrix contains a large number of channel taps, which correspond to the different propagation paths between the transmitter and receiver. In non-sparse channels, the channel matrix cannot be represented by a sparse matrix, i.e., a matrix in which the majority of the entries are zero. Non-sparse channels are commonly found in indoor environments, such as buildings and houses, where multiple reflections and scattering cause multiple signal paths and a rich multipath environment.

2.5.3 Static channel

A static channel is a communication channel whose characteristics remain constant over time [57]. In other words, the channel does not change its behavior, such as attenuation, phase shift, and delay, over the course of a communication session or transmission. A static channel can be modeled as a time-invariant system that can be represented by a fixed channel matrix. Static channels are common in wired communication systems, where the physical channel characteristics do not change significantly over time. However, in wireless communication systems, the channel characteristics are typically time-varying due to factors such as fading, shadowing, and interference, and are therefore modeled as dynamic or time-varying channels.

2.5.4 Time-varying channel

A time-varying channel is a type of communication channel in which the channel response changes over time [57]. This means that the output signal is not simply a scaled version of the input signal, but is instead affected by time-varying factors such as fading, interference, and noise.

Mathematically, we can describe a time-varying channel as follows: let $x(t)$ be the input signal and $y(t)$ be the output signal. Then, the relationship between $x(t)$ and $y(t)$ can be written as:

$$y(t) = h(t)x(t) + w(t) \tag{2.10}$$

where $h(t)$ is the time-varying channel gain, which represents the effect of the channel on the input signal at time t, and $w(t)$ is the additive noise that is present in the channel.

The time-varying channel gain $h(t)$ is a function of time and can be represented as a complex-valued function that varies over time and frequency. The channel can be modeled as a linear time-varying (LTV) system, meaning that its response varies with time but is still linear with respect to the input signal.

In the frequency domain, the time-varying channel's frequency response can be represented as:

$$H(f,t) = |h(f,t)|e^{j\phi(f,t)} \tag{2.11}$$

where f is the frequency, t is the time, $|h(f,t)|$ is the magnitude of the frequency response, and $\phi(f,t)$ is the phase of the frequency response. The magnitude and phase of the channel's frequency response can vary over time and frequency, which can result in variations in the channel's gain and delay.

Time-varying channels are common in wireless communication systems, where the signal is transmitted through a medium that can introduce fading, interference, and noise. In order to overcome the effects of the time-varying channel, various techniques such as equalization, diversity, and coding can be used to improve the quality and reliability of the transmitted signal.

2.6 Traditional channel estimation algorithms based on pilot symbols

Since the signal is affected by such factors as multipath effect and Doppler effect during the transmission process, and multipath effect and Doppler effect would cause fading and time delay in the wireless propagation channel, which result in intersymbol interference in the communication system. In order to effectively overcome the interference, the receiver needs to obtain accurate channel information. Therefore, accurate and fast channel estimation is one of the core technologies for efficient wireless communication systems [70].

The traditional channel estimation algorithm based on pilot symbols usually inserts pilot symbols between the OFDM symbols, thus makes the receiver can estimate the channel according to the received sub-carrier information and the known pilot symbols. Therefore, both the pilot symbol and the channel estimation algorithm can directly affect the performance of the wireless communication system.

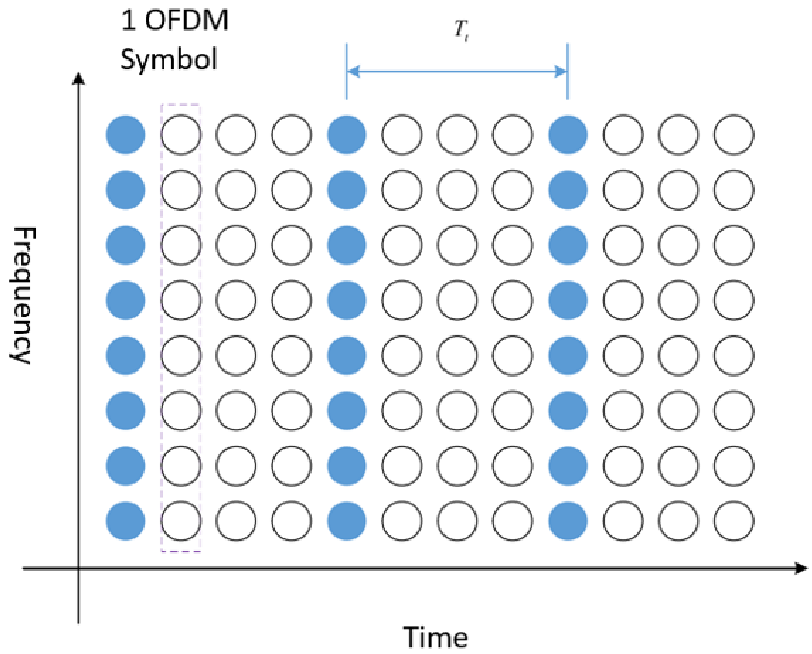

Fig. 2.2 Structure of block pilot (adapted from [2]

2.6.1 Pilot structure

In the channel estimation process based on pilot symbols, there can be the following structures of pilot symbols: block pilot, comb pilot and lattice pilot frequency

1. Block Pilot structure

Fig. 2.2 shows the structure of block pilots. For this kind of pilot structure, the communication system transmits pilot symbols along the time axis direction, the pilot symbol of each period is placed on all sub-carriers. The system can then use these pilot symbols to perform time-domain interpolation, thus reconstructing the complete channel information along the time axis. Assuming that the period of the pilot symbol is T_t, in order to effectively obtain the characteristics of complex time-varying channel, the period T_t of the pilot symbol is usually same as the coherence time of the channel. In addition, since the coherence time of the wireless communication channel is inversely proportional to the Doppler frequency shift f_d of the system, the period $\mathrm{T_t}$ of sending pilot symbols satisfies the following formula:

$$\mathrm{T_t} \leq \frac{1}{\mathrm{f_d}}. \tag{2.12}$$

For the block pilot, since the pilot symbol is inserted into the sub-carriers at a frequency of a certain fixed period of time, the pilot structure is suitable for a slow time-varying channel. However, for a fast time-varying fading channel, this kind method to obtain CSI would makes the system more complex, since it will need to insert more pilot symbols.

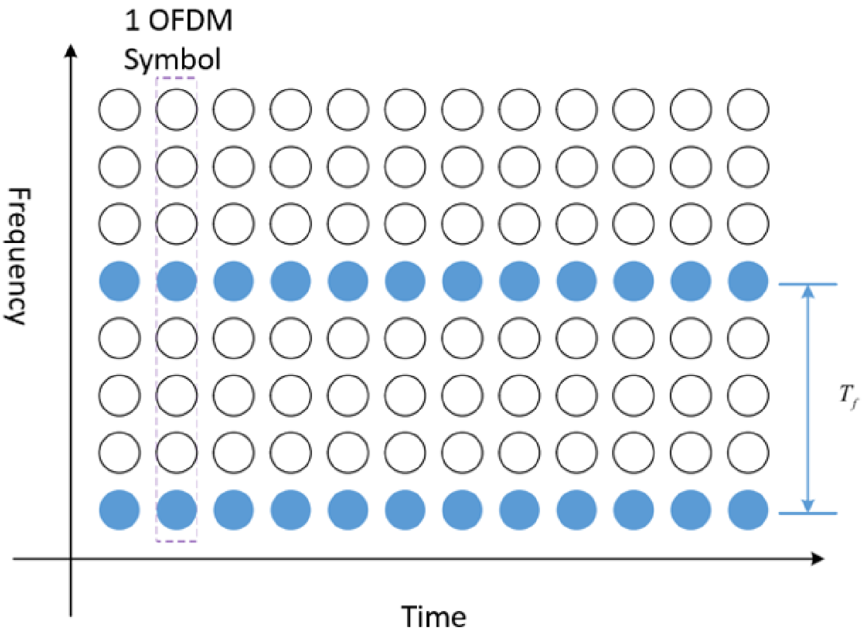

Fig. 2.3 Structure of comb pilot (adapted from [2]))

2. Comb pilot

Fig. 2.3 shows the structure of comb pilots. For this kind of pilot structure, the communication system transmits pilot data along the frequency axis direction. The system can use these pilot symbols to perform frequency domain interpolation, thus reconstructing the complete channel information along the frequency axis. Assuming that the frequency period of the pilot symbol is $\mathrm{T_f}$, in order to effectively obtain the

CSI, the period $\mathrm{T_f}$ of the pilot symbol sent by the transmitter must be consistent with the coherence bandwidth of the channel. However, because the coherence bandwidth of the wireless communication system channel is inversely proportional to the maximum delay extension τ_{max} of the channel, the time period $\mathrm{T_f}$ should satisfy

$$\mathrm{T_f} \leq \frac{1}{\tau_{\mathrm{max}}}. \tag{2.13}$$

Since the structure of comb pilot, since pilot symbol is inserted into some sub-carriers along the frequency axis at a fixed time period, it is suitable for fast time-varying fading channels rather than frequency selective channels.

3. Lattice Pilot

Fig. 2.4 shows the structure of the lattice pilot. For this kind of pilot structure, to

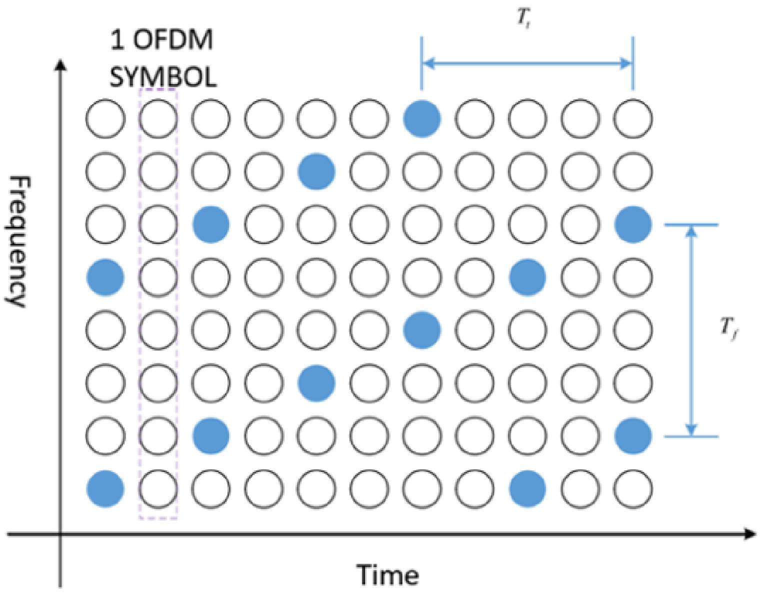

Fig. 2.4 Structure of lattice pilot (adapted from [3])

perform the channel estimation, the communication system transmits pilot data along the time axis and frequency axis with a fixed periods in time and frequency, while the pilot symbol is distributed in the sub-carriers of the OFDM symbols. The system can then use these pilot symbols to perform time/frequency domain interpolation, to obtain complete channel information. Generally, in order to effectively obtain the characteristics of time-varying and frequency-selective wireless channels, the period $\mathrm{T_t}$ and $\mathrm{T_f}$ of the pilot symbols sent by the transmitter, must be consistent with the coherence time and coherence bandwidth of the channel, which can be written as

$$\mathrm{T_t} \leq \frac{1}{\mathrm{f_d}} \text{ and } \mathrm{T_f} \leq \frac{1}{\tau_{\max}} \tag{2.14}$$

2.6.2 Least Square (LS) channel estimation

In communication engineering, the Least Squares (LS) algorithm is commonly used for channel estimation and equalization in digital communication systems. For slow time-varying frequency-selective channels, when the receiver estimates the channel through pilot symbols, LS is the most basic and widely used frequency-domain channel estimation algorithm [58].

For a single antenans user, we consider $\mathbf{x} \in \mathbf{C}^{M\times 1}$ as the transmitted signal vector, $\mathbf{y} \in \mathbf{C}^{J\times 1}$ as the received signal vector. Thus the received signal vector can be obtain:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} \tag{2.15}$$

where $\mathbf{H} \in \mathbf{C}^{J\times M}$ is the matrix that represents the channel response, $\mathbf{w} \in \mathbf{C}^{J\times 1}$ is the noise vector.

The estimated channel matrix $\hat{\mathbf{H}}$ is obtained by minimizing the cost function:

$$J(\hat{\mathbf{H}}) = ||\mathbf{y} - \hat{\mathbf{H}}\mathbf{x}||^2. \tag{2.16}$$

The solution to the least squares problem can be obtained by setting the derivative of the cost function with respect to $\mathbf{H}$ to zero, which leads to:

$$\hat{\mathbf{H}} = (\mathbf{x}^H\mathbf{x})^{-1}\mathbf{x}^H\mathbf{y} \tag{2.17}$$

In practice, the transmitted signal $\mathbf{x}$ may contain a known pilot signal, which is used to estimate the channel response vector $\mathbf{h}$. The pilot signal is inserted in the transmitted signal at regular intervals, and the corresponding received pilot signal is used to estimate the channel response vector $\mathbf{h}$. The estimated channel response vector is then used to equalize the received signal, which involves dividing the received signal by the estimated channel response vector, to remove the effect of the channel from the received signal.

Once the channel impulse response has been estimated using the LS algorithm, it can be used for equalization to remove the effects of the channel distortion from the received signal. The LS algorithm can also be extended to handle time-varying channels by using a sliding window approach, in which the channel impulse response is estimated using a limited window of received signal samples at each time instant.

2.6.3 MMSE algorithm

The Minimum Mean Square Error (MMSE) algorithm is a commonly used method for channel estimation in communication engineering. It uses the mean square error criterion to find the channel estimate, the cost function can be written as:

$$J(\hat{\mathbf{H}}) = \mathrm{E}\left[||\mathbf{y} - \mathbf{x}\hat{\mathbf{H}}\mathbf{h}||^2\right]. \tag{2.18}$$

The mean squared error criterion for MMSE channel estimation method is minimizing $J(\hat{\mathbf{h}})$. Taking the partial derivative of $\hat{\mathbf{h}}$ and setting it to zero yields the solution for MMSE:

$$\hat{\mathbf{H}}_{MMSE} = \mathbf{R}_{\mathbf{hy}}\mathbf{R}_{\mathbf{yy}}^{-1}\mathbf{y}, \tag{2.19}$$

where $\mathbf{R}_{\mathbf{Hy}}$ and $\mathbf{R}_{\mathbf{yy}}$ can be obtained by:

$$\mathbf{R}_{\mathbf{hy}} = E[\mathbf{h}\mathbf{Y}^H] = (\mathbf{R}_{HH})\mathbf{y}, \tag{2.20}$$

$$\mathbf{R}_{\mathbf{yy}} = E[\mathbf{y}\mathbf{y}^H] = \mathbf{x}(\mathbf{R}_{HH})\mathbf{x}^H + \sigma_w^2\mathrm{I}, \tag{2.21}$$

here, $\mathbf{R}_{\mathbf{hy}}$ is the cross-covariance matrix between the channel matrix $\mathbf{H}$ and the received data matrix $\mathbf{Y}$, $\mathbf{R}_{yy}$ is the autocovariance matrix of the received data matrix $\mathbf{Y}$, $\mathbf{R}_{HH}$ is the autocovariance matrix of the channel matrix $\mathbf{H}$, σ_w^2 is the noise variance of the wireless communication environment. Hence, by combining Equation (2.20) and (2.21), the solution for MMSE channel estimation can be obtained by:

$$\hat{\mathbf{H}}_{MMSE} = \mathbf{R}_{HH}[\mathbf{R}_{HH} + \sigma_w^2(\mathbf{x}\mathbf{x})^{-1}]^{-1}\hat{\mathbf{H}}_{LS}. \tag{2.22}$$

The MMSE method considers the influence of the system's AWGN signal, thus it is more accurate than traditional LS methods. However, this algorithm requires the calculation of the cross-covariance matrix between the channel information matrix and the received signal matrix , the autocovariance matrix of the received signal matrix , and the autocovariance matrix. Additionally, the algorithm also requires knowledge of the noise variance in the wireless communication environment. Therefore, this method has a high computational complexity. As the MMSE method involves matrix inversion, the computational complexity increases when the number of pilot subcarriers in a MIMO-OFDM system is large.

The MMSE algorithm can also be extended to handle time-varying channels by using a sliding window approach, in which the channel impulse response is estimated using a limited window of received signal samples at each time instant.

2.6.4 Compressive sensing

The Shannon/Nyquist sampling theorem indicates that in order to avoid losing relevant data information when sampling the signal, the sampling frequency of the system needs to satisfy twice or more than the original signal bandwidth [71]. In applications such as digital image and signal processing, if the Nyquist sampling frequency is too high, it will lead to the collection of too many samples and produce redundancy, which makes data compression a necessary condition for storage and transmission. In addition, in other applications such as radio receivers, medical imaging equipment, and high-speed analog-to-digital and digital-to-analog converters, further increasing the sampling rate of the device can be very expensive. Thus, a breakthrough sampling method has been proposed based on the related theory of compressive sensing [72]. This method can capture and represent a signal with sparse characteristics in a transformed space using a measurement matrix and project it to a lowdimensional space at a significantly lower sampling rate. It can be proved that this method can retain a sufficient amount of intrinsic structural features and data information of the original sampled signal by using linear projections, and then the original sampled signal can be reconstructed from these projections with high probability using an optimization process. Under this theoretical framework, the sampling frequency of compressive sensing technology will not depend on the bandwidth of the original data, but is determined by the sparse nature of the data and the Restricted Isometry Pmperty [73] of the measurement matrix.

Compared to traditional channel estimation techniques such as least squares (LS) or minimum mean squared error (MMSE) estimation, which requires a large number of measurements, thus limits the channel estimation efficiency in band-limited systems, compressive sensing can be employed to overcome this limitation by acquiring a compressed representation of the channel response, which can then be used to estimate the channel coefficients [71]. The compressed representation can be obtained by multiplying the transmitted signal with a designed matrix (such as discrete Fourier transform matrix), resulting in a small number of measurements.

By employing compressive sensing technique, the channel can be estimated by solving the optimization problem [74], one common optimization problem used in compressive sensing for channel estimation is the ℓ_1-norm minimization problem [75]:

$$\hat{\mathbf{h}} = \text{argmin}_{\mathbf{h}} ||\mathbf{y} - \mathbf{\Phi h}||_2^2 + \lambda ||\mathbf{h}||_1 \tag{2.23}$$

where $\mathbf{y}$ is the received signal, $\mathbf{\Phi}$ is the measurement matrix, $\hat{\mathbf{h}}$ is the estimated channel response, and λ is a parameter that controls the tradeoff between fidelity to the measured data and sparsity of the solution.

The compressed sensing approach to channel estimation can be further enhanced by exploiting the specific structures of the channel response. For example, the channel response may be sparse in a certain basis or transform domain, such as the discrete cosine transform or wavelet transform. By using these transforms, the channel response can be more efficiently represented and reconstructed from the compressed measurements [57].

According to the different estimation approaches, the current sparse channel estimation algorithms based on CS theory can be classified into three categories: basis pursuit algorithm, Bayesian algorithm, greedy algorithm and dichotomous coordinate descent algorithm. In addition, the related introduction and performance analysis of these three types of algorithms is as follows

1. The basis pursuit algorithm solves a convex optimization problem, such as interior point method, Bregman iterative algorithm, and iterative hard thresholding algorithm, etc. Although this kind of method can accurately reconstruct the original signal with high probability and can obtain the global optimal solution, in the real environment, in the process of processing high-dimensional signals, this kind of optimization algorithm needs to solve the equations with respect to multiple unknown variables, which results in high computational complexity.

2. Bayesian algorithms can be divided into Bayesian compressive sensing [76] and Variational Bayesian compressive sensing (VBCS) [77] algorithms. A sparse channel estimation algorithm based on the Bayesian algorithm requires a joint probability distribution of all unknown variables and measured signals, which makes the computational complexity of the algorithm high, and it cannot accurately estimate the sparsity of the signal. Therefore, Bayesian algorithm is rarely used in practice.

3. The greedy algorithm mainly adopts the strategy of sparse approximation to indirectly solve the problem of sparse signal reconstruction. Such methods mainly include OMP algorithm [78], generalized OMP algorithm [79], CoSaMP algorithm [80], SP algorithm [81], SAMP algorithm [82] and so on. The greedy algorithm is characterized by low algorithm complexity, high solution sparsity and high reconstruction accuracy. Therefore, greedy algorithms have received extensive attention in practical applications.

4. The dichotomous coordinate descent (DCD) algorithm is a multiplication-free and division free iterative technique, which guarantees convergence to the true solution under realistic assumptions [83]. In [74], based on solving either the $\ell_2\ell_1$ or $\ell_2\ell_0$

optimization problem using dichotomous coordinate descent (DCD) iterations, a family of greedy sparse algorithm based on the Homotopy and DCD algorithm was proposed for recovery of complex-value sparse signals. According to the simulation results in [74], the $\ell_2\ell_0$ DCD algorithm is most attractive due to its low computational complexity and high accuracy, which is used in this work.

2.7 Sparse wireless communication channels

In a MIMO system, the wireless channel is the link between the transmitter antenna and the receiver antenna. Signals can travel through many different paths from the transmitter to the receiver. For example, in some scenarios, there is no direct path between the transmitter and the receiver, however, there are various interacting bodies in the propagation environment of wireless signals, such as buildings, mountains, trees and walls in the outdoor environment. Therefore, the transmission process of the wireless signal from the transmitter to the receiver can also be realized by means of reflection, diffraction or scattering on the surfaces of these interacting bodies, resulting in multipath effects. The distance, time, phase and fading degree of the wireless signal from the transmitting end to the receiving end after transmission through multiple channels are different. Therefore, the state information of the wireless channel is composed of multiple taps with different time delays and amplitudes.

Since different sparse multipath have different propagation time, the phase for different sparse multipath are different, which would lead to the interference in the system. Typically, the coherence bandwidth is approximately equal to the inverse of the maximum multipath delay [49]. In the spectrum domain, if the coherence bandwidth is smaller than the bandwidth of the wireless channel, frequency selective fading will occur during the signal transmission process of the communication system, that is, the waveform of the received signal shows that the power of the signal is enhanced at some frequency points, and the power of the signal is weakened in other parts. The multipath effect will seriously affect the transmission quality of the signal in the wireless communication system, if the state information of the wireless channel can be obtained timely and accurately during the communication process, the MIMO wireless communication system can effectively reduce the influence of fading factors.

However, in practical the MIMO systems, the acquisition of CSI is performed through channel estimation algorithms. The traditional linear channel estimation algorithms such as LS and MMSE assumes that the taps of the wireless multipath channel are dense, that is, the taps corresponding to each path of the wireless channel have a large amplitude. However, in recent years, many scholars have shown through actual measurement results that wireless channels have sparse characteristics [43] [44]. That is, a small number of channel taps contain most of the energy, while most of the remaining taps have zero or close to zero energy. Because the wireless multipath channel satisfies the sparsity condition of compressive sensing, the communication system can employ an observation matrix with fewer pilot symbols, to obtain better channel estimation performance. As wireless communication systems become increasingly complex, pilot overhead increases dramatically, the sparse nature of wireless channels provides a realistic basis and necessary conditions for the wide application of compressive sensing in the field of channel estimation, and provides a new idea and approach for reducing pilot overhead, while further improves the accuracy of channel estimation.

2.8 Conclusion

Chapter 2 introduces and analyzes the basic principles of the MIMO system model and related technologies involved in the subject research. For the MIMO-OFDM wireless communication systems, this chapter first introduces the diversity, space division multiplexing, beamforming and channel capacity related to the MIMO technology, as well as the transmission model of the wireless channel of the MIMO system, and then introduces the OFDM system. Then, massive MIMO systems are introduced. Based on the basic theory and method of compressive sensing, this chapter indicates the practical application value of using the greedy algorithm with sparse approximation strategy to solve the sparse channel estimation problem.

Chapter 3

Sparse Channel Estimation for OFDM in Virtual Angular Domain

3.1 Introduction

Since in the MIMO-OFDM system, the signal is affected by the multipath effect during the transmission process, this causes the intersymbol interference in the received signal. In order to effectively overcome interference, the receiver needs to obtain accurate channel information. Therefore, reliable channel estimation is one of the key technologies for efficient wireless communication systems.

Massive MIMO has been proposed for next generations of communication systems, since it provides higher spectral efficiency [31], [84]. It can enhance the spectral efficiency by orders of magnitude by equipping the wireless transmitter with a large number of antennas and exploiting the increased degree of freedom in the spatial domain.

Pilot aided channel estimation is widely used in MIMO systems [85]. For channel estimation in a MIMO system with a small number of antennas, orthogonal pilots are often used [86], [87]. However, the pilot overhead increases with the number of antennas [88]. Employing orthogonal pilots for channel estimation would cause unacceptable pilot overhead because of the massive number of antennas at the base station (BS) [4]. In [4], a compressive sensing based channel feedback scheme was proposed, which can reduce the pilot overhead

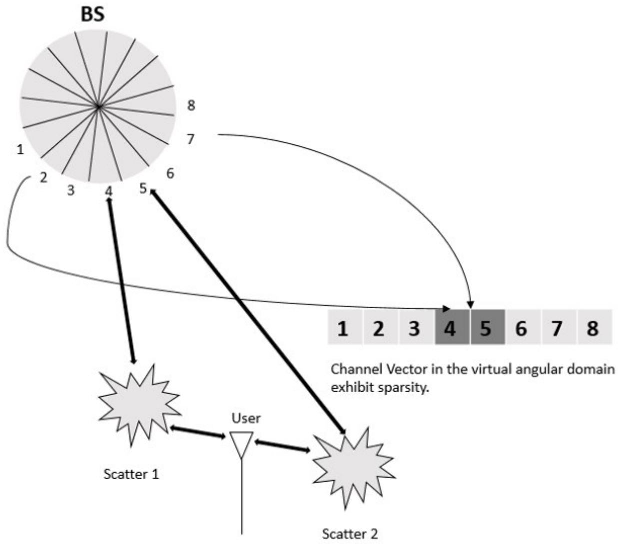

Fig. 3.1 Sparse channel in the virtual angular domain.

and achieve good channel state information (CSI) acquisition, where a feedback scheme is how the communication system perform the signal transmission process and collect CSI. In this chapter, we focus on the channel estimation in the feedback scheme.

Experiments and research have shown that due to the small angle spread seen from a BS between a user and BS, massive MIMO channels exhibit sparsity in the virtual angular domain, as shown in Fig. 3.1, the virtual angular domain is a concept used to describe the spatial domain in a wireless communication system, and refers to a virtual domain that is created by mapping the spatial domain onto a discrete set of virtual angles [89]. Furthermore, according to [88], [4], [57], when applying the orthogonal frequency division multiplexing (OFDM), because of the spatial propagation property of the wireless channel, such as the number of scatterers is nearly unchanged over the system bandwidth, the common sparsity is shared by different subcarriers, which is referred to as the spatially common sparsity over multiple subcarriers. Often, massive MIMO channels can be considered as quasi-static over a coherence time interval [57]. Furthermore, since the angle variation from the user to the BS is relatively slow, and can be often neglected, the support set of the channel in the virtual angular domain can be regarded as unchanged over several OFDM symbols, which is referred to as spatially common sparsity over multiple OFDM symbols [4] [57]. By exploiting the common sparsity in the virtual angular domain, we can jointly estimate the channel for multiple subcarriers.

Sparse recovery techniques are attractive for channel estimation [90],[91], [92]. There are two ways to find sparse representation, convex optimization and greedy methods [74]. Greedy methods typically have lower complexity [93], such as the orthogonal matching pursuit (OMP) [94], matching pursuit (MP) [93], compressive sampling matching pursuit (CoSAMP) [95]. However, they may provide limited performance when the signal is not very sparse or the noise is too high [96]. Convex optimization algorithms such as Your ALgorithms for ℓ_1 (YALL1) [97], which employs the alternating direction method, provide high accuracy, but the complexity is high [74], [98], [99]. For channel estimation, we usually deal with complex-valued problems [74]. The sparse recovery algorithm used in this chapter is for solving complex-valued problems.

The low-complexity coordinate descent (CD) search can be implemented to estimate the channel [83], [100]. In [74], algorithms applying dichotomous CD (DCD) iterations for solving $\ell_2\ell_0$ and $\ell_2\ell_1$ optimization problems have been proposed. By exploiting the DCD, the use of multiplications have been minimized, which significantly reduces the algorithm complexity and makes it well suited for real-time implementation [74]. Here we are interested in the DCD algorithm for the $\ell_2\ell_0$ optimization since it outperforms such greedy algorithms as MP and OMP [74].

The DCD algorithm for $\ell_2\ell_0$ optimization is a greedy algorithm [74], different from the CD algorithm [100], [101]. It does not optimize the step size for each iteration, but employs a set of step sizes defined by the fixed-point representation of the solution [74]. It has been indicated in [74] and [83], that the computational complexity of the algorithm is dominated by the computational complexity of a small number of successful iterations, while most of the operations of the DCD algorithm are additions and bit-shifts, which makes it suitable for implementation on real-time design platforms, such as digital signal processors and field-programmable gate arrays [102].

Since the DCD algorithm in [74] can only deal with single sparse channel at one time, by exploiting the spatially common sparsity in the virtual angular domain of the massive

MIMO channels, a DCD-Joint-Sparse-Recovery (DCD-JSR) algorithm is proposed here. The DCD-JSR algorithm can jointly estimate multiple sparse channels and provide accurate CSI acquisition with a low computational complexity. Simulation results show that the proposed algorithm has better mean square error (MSE) performance than the Distributed-Sparsity-Adaptive-Matching-Pursuit (DSAMP) algorithm proposed in [4] for solving the same problem.

This chapter is organized as follows. Section 3.2 describes the system model. Section 3.3 presents the proposed DCD-JSR algorithm. In Section 3.5, numerical examples are analysed and, finally, Section 3.6 presents the conclusion.

3.2 System Model And Problem Formulation

3.2.1 Channel Estimation Scheme

Generally, channel estimation scheme can be divided into two categories: uplink channel estimation and downlink channel estimation. In uplink channel estimation, the base station estimates the channel from the user. This is typically done using pilot signals, which are sent from the user to the base station. The pilot signals are known to both the user and the base station, and are used to estimate the channel between the two devices.

The uplink channel estimation process usually involves the following steps: firstly, the user sends pilot symbols to the base station. After that, the base station uses the pilot symbols to estimate the channel between the user and the base station. Then the base station sends feedback to the user, informing them of the estimated channel conditions. Finally, the user can then use this information to adjust their transmission parameters, such as power level and modulation scheme, to optimize their communication with the base station.

For donwlink channel estimation, the conventional method to acquire the CSI in frequency-division-duplexing (FDD) systems is as follows: the BS transmits pilot symbols to a user, so the user can estimate the downlink CSI locally and then feed it back to the BS via an uplink channel [103]. If we are employing conventional CSI estimation techniques (such as the minimum mean square error (MMSE) estimator), since the number of pilots required at the

BS has to scale linearly with the number of transmit antennas at the BS [55], it would cause prohibitively large overhead for both pilot training (downlink) and CSI feedback (uplink). Hence, to solve the overhead issues, as suggested in [4], the channel estimation is performed at the BS. The channel estimation scheme is summarized as follows.

1 In each OFDM symbol, every BS antenna broadcasts pilot symbols to users, the kth user receives the signal $\mathbf{y}_{\mathrm{k}}$ and feeds it back to the BS. The BS recovers the CSI for each user based on the feedback signals $\mathbf{y}_{\mathrm{k}}$, $\mathrm{k} = 1, ..., \mathrm{K}$. As shown in Fig.3.2 each OFDM symbol contains N subcarriers, while P subcarriers are used to transmit pilot symbols. The user feeds back the received signal to the BS without performing downlink channel estimation.

2 At the BS, a channel estimation algorithm is used to jointly estimate multiple sparse virtual angular domain channels, which are assumed to have the same support I. The least squares (LS) algorithm [104] is employed to acquire the CSI based on an estimate of the common support I.

3 The base station can then use this information to adjust their transmission parameters, such as power level and modulation scheme, to optimize their communication with the user.

3.2.2 Channel Model

In a typical FDD massive MIMO system, consider a coherence time interval consisting of J OFDM symbols. M antennas are employed at the BS to serve K single-antenna users simultaneously, where $\mathrm{M} \gg \mathrm{K}$. At the tth OFDM symbol, $1 \leq \mathrm{t} \leq \mathrm{J}$, for the nth subcarrier, $1 \leq \mathrm{n} \leq \mathrm{N}$, the received signal for the kth user, $1 \leq \mathrm{k} \leq \mathrm{K}$, is given by:

$$\mathrm{y}_{\mathrm{k,n}}^{\mathrm{t}} = \left(\mathbf{h}_{\mathrm{k,n}}^{\mathrm{t}}\right)^{\mathrm{T}} \mathbf{x}_{\mathrm{n}}^{\mathrm{t}} + \mathrm{w}_{\mathrm{k,n}}^{\mathrm{t}}, \tag{3.1}$$

where $\mathbf{h}_{\mathrm{k,n}}^{\mathrm{t}} \in \mathrm{C}^{\mathrm{M}\times 1}$ represents the downlink channel between the kth user and M antennas, $\mathbf{x}_{\mathrm{n}}^{\mathrm{t}} \in \mathrm{C}^{\mathrm{M}\times 1}$ is the vector of transmitted symbols (data or pilot symbols) and $\mathrm{w}_{\mathrm{k,n}}^{\mathrm{t}}$ is the

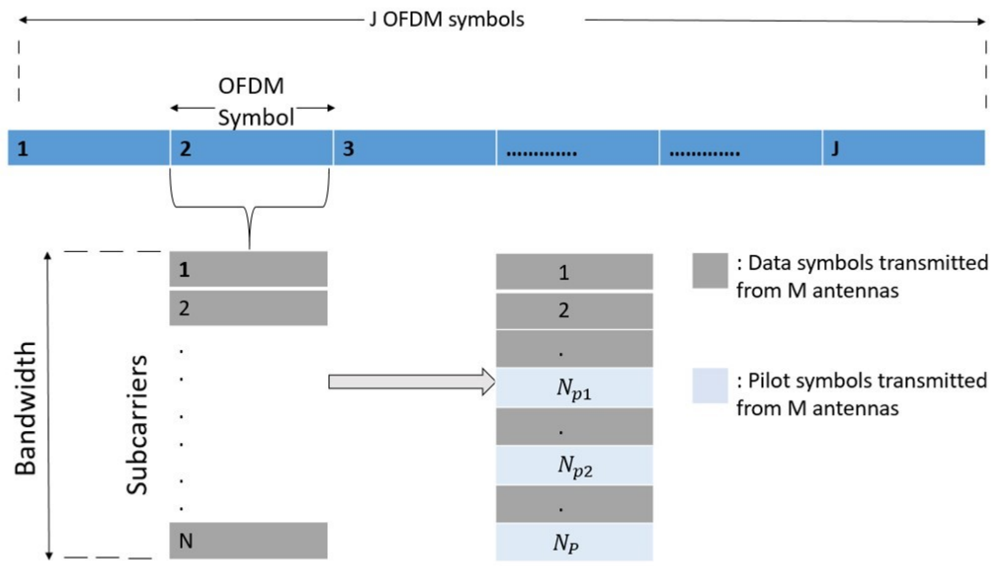

Fig. 3.2 Each OFDM symbol contains N subcarriers, while P subcarriers are used to transmit pilot symbols.

corresponding additive white Gaussian noise (AWGN). For a single user, we can drop the index k, thus we can write:

$$\mathrm{y}_{\mathrm{n}}^{\mathrm{t}} = \left(\mathbf{h}_{\mathrm{n}}^{\mathrm{t}}\right)^{\mathrm{T}} \mathbf{x}_{\mathrm{n}}^{\mathrm{t}} + \mathrm{w}_{\mathrm{n}}^{\mathrm{t}}. \tag{3.2}$$

Matrix $\mathbf{A}_{\mathrm{B}}$ is used to modify the channel vector $\mathbf{h}_{\mathrm{n}}^{\mathrm{t}}$ into a vector $\tilde{\mathbf{h}}_{\mathrm{n}}^{\mathrm{t}}$ in the virtual angular domain, and it is determined by the geometric structure of the antenna array. We consider a uniform linear array with the antenna spacing $\mathrm{d} = \lambda/2$, where λ is the wavelength of the carrier frequency, then $\mathbf{A}_{\mathrm{B}}$ becomes the discrete Fourier transform (DFT) matrix [4]. Thus we obtain:

$$\mathrm{y}_{\mathrm{n}}^{\mathrm{t}} = \left(\tilde{\mathbf{h}}_{\mathrm{n}}^{\mathrm{t}}\right)^{\mathrm{T}} \mathbf{A}_{\mathrm{B}}^{*} \mathbf{x}_{\mathrm{n}}^{\mathrm{t}} + \mathrm{w}_{\mathrm{n}}^{\mathrm{t}}, \tag{3.3}$$

where, $(\mathbf{h}_{\mathrm{n}}^{\mathrm{t}})^{\mathrm{T}} = \left(\tilde{\mathbf{h}}_{\mathrm{n}}^{\mathrm{t}}\right)^{\mathrm{T}} \mathbf{A}_{\mathrm{B}}^{*}$. and the channel vector in the angular domain divides the covering area of the BS into angular intervals. The mth element of $\tilde{\mathbf{h}}_{\mathrm{n}}^{\mathrm{t}}$ corresponds to the mth virtual angle, where $1 \leq \mathrm{m} \leq \mathrm{M}$.

According to experimental study [89] and analysis [55], in practical massive MIMO systems, the BS is usually at a high elevation with a limited number of scatterers (relative to the number of antennas), and the scatterers at the user side are relatively rich. In other words, the BS might only have few active transmit directions for the kth user, which means that the number of multipath arrivals dominating the majority of channel energy is small, and the channel vectors in the virtual angular domain exhibit sparsity. Thus, we have $|\mathrm{I}| \ll \mathrm{M}$, which means the channel exhibits sparsity in the virtual angular domain. Furthermore, as shown in Fig.3.3, according to [57] and [4], since the spatial propagation characteristics such as scatterers are almost unchanged over the system bandwidth, the subchannels associated with different subcarriers in the same OFDM symbol share common sparsity, namely:

$$\mathrm{I}_1^j = \mathrm{I}_2^j = ... = \mathrm{I}^j, 1 \leq j \leq J, \tag{3.4}$$

besides, in [105], it has been indicated that even in time-varying scenarios, the variation of the arrival angles is usually much slower than that of channel gains. This means the channel associated with J successive OFDM symbols shares common sparsity, thus we can obtain:

$$\mathrm{I}_n^1 = \mathrm{I}_n^2 = ... = \mathrm{I}_n^J \tag{3.5}$$

Hence, we can say that the support is same among J OFDM symbols and P pilot subcarriers. Moreover, Since the channel during J OFDM symbols is time invariant, the channel gain can be considered as unchanged during J OFDM symbols, which can be written as:

$$\tilde{\mathbf{h}}_{\mathrm{n}}^1 = \tilde{\mathbf{h}}_{\mathrm{n}}^2 = ... = \tilde{\mathbf{h}}_{\mathrm{n}}^{\mathrm{J}} = \tilde{\mathbf{h}}_{\mathrm{n}}. \tag{3.6}$$

In this chapter, we consider the pilot-aided channel estimation. The structure of the transmitted pilot symbols is shown in Fig.3.4. To provide accurate channel estimation with multiple pilot subcarriers, for the tth OFDM symbol, a part of subcarriers is used for transmitting pilot symbols $\mathbf{s}_{\mathrm{p}}^{\mathrm{t}} \in \mathrm{C}^{\mathrm{M}\times 1}$, and the received signal at the pilot subcarrier $\mathrm{n(p)}$ is given by:

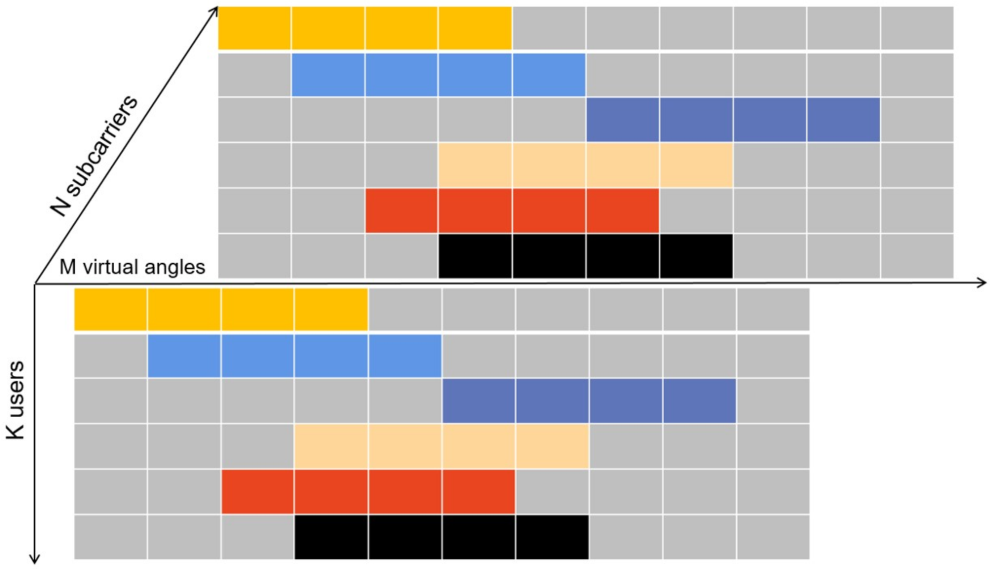

Fig. 3.3 The virtual angular domain channel vector exhibits common sparsity within the system bandwidth (adapted from [4]).

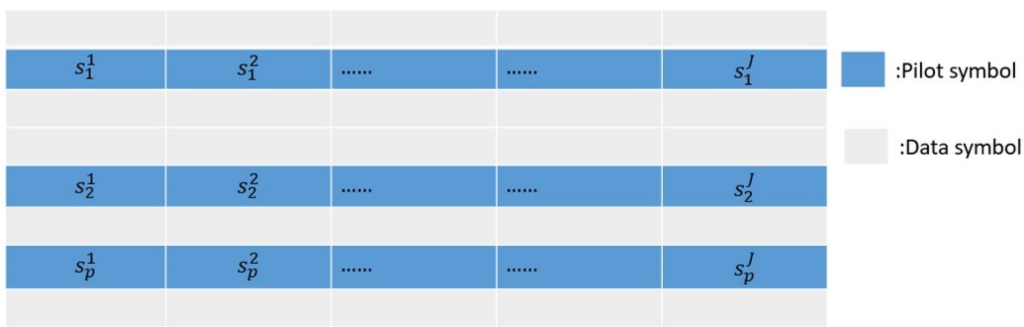

Fig. 3.4 Structure of the transmitted JP pilot symbols. Each pilot symbol corresponds to the pilot sequence transmitted from M antennas.

$$\mathrm{y}^{\mathrm{t}}_{\mathrm{n(p)}} = \left(\tilde{\mathbf{h}}_{\mathrm{n(p)}}\right)^{\mathrm{T}} \mathbf{A}^{*}_{\mathrm{B}}\mathbf{s}^{\mathrm{t}}_{\mathrm{p}} + \mathrm{w}^{\mathrm{t}}_{\mathrm{n(p)}}, \tag{3.7}$$

$$\left[\mathbf{s}^{\mathrm{t}}_{\mathrm{p}}\right]_{\mathrm{m}} = \mathrm{e}^{\mathrm{j}\theta_{\mathrm{t,m,p}}}, \tag{3.8}$$

$$1 \leq \mathrm{p} \leq \mathrm{P},\ 1 \leq \mathrm{m} \leq \mathrm{M},\ 1 \leq \mathrm{t} \leq \mathrm{J}$$

while $\theta_{\mathrm{t,m,p}}$ are independent random numbers uniformly distributed in $(0, 2\pi]$.

3.2.3 Problem Formulation

As described in Section 3.2.1, after receiving the signal from BS, the user will send the received signal back to the BS without performing the downlink channel estimation, where the feedback channel can be considered as an AWGN channel. [55] [106] [107]. Hence, for the tth OFDM symbol, at the pth pilot subcarrier, the signal received at the BS is given by:

$$\mathrm{r}^{\mathrm{t}}_{\mathrm{p}} = \left(\mathbf{s}^{\mathrm{t}}_{\mathrm{p}}\right)^{\mathrm{T}} \left(\mathbf{A}^{*}_{\mathrm{B}}\right)^{\mathrm{T}} \tilde{\mathbf{h}}_{\mathrm{n(p)}} + \mathrm{v}^{\mathrm{t}}_{\mathrm{p}} = \boldsymbol{\phi}^{\mathrm{t}}_{\mathrm{p}}\tilde{\mathbf{h}}_{\mathrm{n(p)}} + \mathrm{v}^{\mathrm{t}}_{\mathrm{p}},\ 1 \leq \mathrm{p} \leq \mathrm{P}. \tag{3.9}$$

Here,$\boldsymbol{\phi}^{\mathrm{t}}_{\mathrm{p}} = \left(\mathbf{s}^{\mathrm{t}}_{\mathrm{p}}\right)^{\mathrm{T}} \left(\mathbf{A}^{*}_{\mathrm{B}}\right)^{\mathrm{T}} \in \mathrm{C}^{1\times\mathrm{M}}$ is the sensing vector in the virtual angular domain for the pth pilot subcarrier at the tth OFDM symbol. $\tilde{\mathbf{h}}_{\mathrm{n(p)}} \in \mathrm{C}^{\mathrm{M}\times 1}$ is the sparse channel vector for the $\mathrm{n}\,(\mathrm{p})$th subcarrier, and $\mathrm{v}^{\mathrm{t}}_{\mathrm{p}}$ is the corresponding noise, which contains both downlink and uplink channel noise.

To provide an accurate channel estimation for the pth pilot subcarrier, the BS should jointly utilize the feedback signal over J successive OFDM symbols [4]. We collect the feedback signals $\mathrm{r}^{\mathrm{t}}_{\mathrm{p}}, 1 \leq \mathrm{t} \leq$ J, in a vector $\mathbf{r}_{\mathrm{p}} = \left[\mathrm{r}^{1}_{\mathrm{p}}, \mathrm{r}^{2}_{\mathrm{p}}, ..., \mathrm{r}^{\mathrm{J}}_{\mathrm{p}}\right]^{\mathrm{T}} \in \mathrm{C}^{\mathrm{J}\times 1}$, which contains the received signal for J OFDM symbols, then we have

$$\mathbf{r}_{\mathrm{p}} = \Phi_{\mathrm{p}}\tilde{\mathbf{h}}_{\mathrm{n(p)}} + \mathbf{v}_{\mathrm{p}},\ \ 1 \leq \mathrm{p} \leq \mathrm{P}, \tag{3.10}$$

where $\Phi_{\mathrm{p}} = \left[\mathbf{S}^{\mathrm{J}}_{\mathrm{p}} \left(\mathbf{A}^{*}_{\mathrm{B}}\right)^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathrm{C}^{\mathrm{J}\times\mathrm{M}}$ is the sensing matrix contains the sensing vector $\boldsymbol{\Phi}^{t}_{p}$ for J OFDM symbols, $\mathbf{S}_{\mathrm{p}} = \left[\mathbf{s}^{1}_{\mathrm{p}}, \mathbf{s}^{2}_{\mathrm{p}}, ..., \mathbf{s}^{\mathrm{J}}_{\mathrm{p}}\right]^{\mathrm{T}} \in \mathrm{C}^{\mathrm{J}\times\mathrm{M}}$ is the matrix contains J transmitted pilot symbols, and the noise vector $\mathbf{v}_{\mathrm{p}} = \left[\mathrm{v}_{\mathrm{p}}^{1}, \mathrm{v}_{\mathrm{p}}^{2}, ..., \mathrm{v}_{\mathrm{p}}^{\mathrm{J}}\right]^{\mathrm{T}} \in \mathrm{C}^{\mathrm{J}\times 1}$ contains both downlink and uplink noise for J OFDM symbols. Since the channels for all subcarriers exhibit common sparsity, we can jointly estimate the channels associated with multiple pilot subcarriers assuming the common support.

3.3 DCD-JSR Algorithm for the Channel Estimation in Virtual Angular Domain

In [4], the distributed sparsity adaptive matching pursuit (DSAMP) algorithm was proposed to jointly estimate multiple sparse channels by estimating the common support shared by different subcarriers in OFDM. However, simulation results show that it provides a limited performance when the number of OFDM symbols J used for the channel estimation is not high. In [74], the homotopy $\ell_2\ell_0$ DCD algorithm was proposed, which can be used to estimate the sparse channel, and it can provide accurate sparse estimation with low complexity. However, it was focused on a single sparse problem, and cannot jointly estimate multiple sparse channels. Therefore, based on [4] and [74], we propose the DCD-JSR algorithm, which can jointly estimate multiple sparse channels with a common support.

To simplify notation, we replace $\tilde{\mathbf{h}}_{\mathrm{n(p)}}$ with $\mathbf{h}_{\mathrm{p}} \in \mathbf{C}^{\mathrm{M}\times 1}$, which is the channel vector to be estimated. We denote $\tilde{\mathbf{h}}_{\mathrm{p}}$ as the final vector estimate. The DCD-JSR algorithm is summarized as follows.

1. For each pilot subcarrier, the $\ell_2\ell_0$ homotopy DCD algorithm is employed to acquire an estimate of $\mathbf{h}_{\mathrm{p}}$.

2. Based on the $\mathbf{h}_{\mathrm{p}}$ estimate, a common support $\tilde{\mathrm{I}}$ is found by analysing the distribution of the estimates.

3. Based on the common support $\tilde{\mathrm{I}}$, the final channel vector estimate $\tilde{\mathbf{h}}_{\mathrm{p}}$ is acquired by using the LS algorithm [104] on the support.

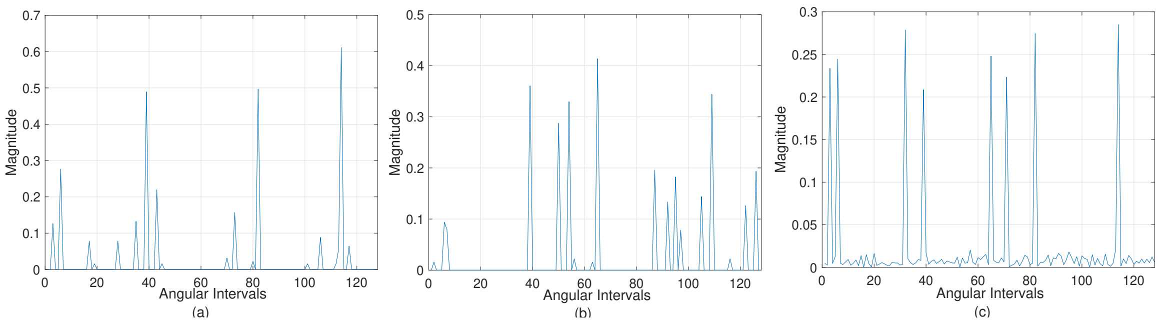

Fig. 3.5 Magnitudes of elements of vectors: (a) $\tilde{\mathbf{h}}_1$, (b) $\tilde{\mathbf{h}}_{64}$, (c) $\mathbf{q}$

921 3.3.1 Channel Estimation Using the $\ell_2\ell_0$ Homotopy DCD Algorithm

---

Algorithm 1 $\ell_2\ell_0$ homotopy DCD algorithm

Initialization:vector $\mathbf{h}_{\rm p} = \mathbf{0}$, $\mathrm{I_p} = \emptyset$, $\mathbf{b}_{\rm p} = \Phi_{\rm p}{}^{\rm H}\mathbf{r}_{\rm p}$,
$\mathbf{R}_{\rm p} = \mathbf{\Phi}_{\rm p}{}^{\rm H}\mathbf{\Phi}_{\rm p}$.

1: $\mathrm{g} = \arg\max \left|(\mathbf{b}_{\rm p})_{\rm k}\right|^2/(\mathbf{R}_{\rm p})_{\rm k,k}$,
$\tau_{\max} = (1/2)\max_{\rm k}\left|(\mathbf{b}_{\rm p})_{\rm k}\right|^2/(\mathbf{R}_{\rm p})_{\rm k,k}$,
$\tau = 0.5\left|(\mathbf{b}_{\rm p})_{\rm g}\right|^2/(\mathbf{R}_{\rm p})_{\rm g,g}$, $\mathrm{I_p} = \{\mathrm{g}\}$.

2: **Repeat** until the termination condition is met:

3: **If** the support $\mathrm{I_p}$ has been updated **then**
Solve $(\mathbf{R}_{\rm p})_{\rm I_p,I_p}(\mathbf{h}_{\rm p})_{\rm I_p} = \mathbf{f}_{\rm p}$,
where $\mathbf{f}_{\rm p} = (\Phi_{\rm p})^{\rm H}_{\rm I_p}\mathbf{r}_{\rm p}$
$\mathbf{c} \leftarrow \mathbf{b} - (\mathbf{R}_{\rm p})_{\rm I_p,I_p}(\mathbf{h}_{\rm p})_{\rm I_p}$

4: Update the regularization parameter : $\tau \leftarrow \gamma\tau$

5: Add the g-th element element into the support $\mathrm{I_p}$,
where $\mathrm{g} \in \mathrm{I_p^c}$,
and $\mathrm{g} = \arg\max\limits_{\mathrm{k}\in\mathrm{I_p^c}} \frac{\left|(\mathbf{c})_{\rm k}\right|^2}{(\mathbf{R}_{\rm p})_{\rm k,k}}$ s.t $\left|(\mathbf{c})_{\rm g}\right|^2 > 2\tau(\mathbf{R}_{\rm p})_{\rm g,g}$,
then assign to $(\mathbf{h}_{\rm p})_{\rm g}$ the value $(\mathbf{c})_{\rm g}/(\mathbf{R}_{\rm p})_{\rm g,g}$,
update $\mathbf{c} \leftarrow \mathbf{c} - (\mathbf{h}_{\rm p})_{\rm g}\mathbf{R}_{\rm p}^{\rm g}$.

6: Remove the gth element from the support $\mathrm{I_p}$,
where $\mathrm{g} \in \mathrm{I_p}$, and
$\mathrm{g} = \arg\min\limits_{\mathrm{k}\in\mathrm{I_p}}\left[\frac{1}{2}\left|(\mathbf{h}_{\rm p})_{\rm k}\right|^2(\mathbf{R}_{\rm p})_{\rm k,k} + \Re\left\{(\mathbf{h}_{\rm p})_{\rm k}^*(\mathbf{c})_{\rm k}\right\}\right]$,
s.t. $\frac{1}{2}\left|(\mathbf{h}_{\rm p})_{\rm g}\right|^2(\mathbf{R}_{\rm p})_{\rm g,g} + \Re\left\{(\mathbf{h}_{\rm p})_{\rm g}^*(\mathbf{c})_{\rm g}\right\} < \tau$
for every removed element,
update $\mathbf{c} \leftarrow \mathbf{c} + (\mathbf{h}_{\rm p})_{\rm g}\mathbf{R}_{\rm p}^{\rm g}$ and set $(\mathbf{h}_{\rm p})_{\rm g} = 0$.

To estimate the channel at the pth pilot subcarrier using the $\ell_2\ell_0$ homotopy DCD algorithm, we consider the signal model

$$\mathbf{r}_{\mathrm{p}} = \Phi_{\mathrm{p}}\mathbf{h}_{\mathrm{p}} + \mathbf{v}_{\mathrm{p}}. \tag{3.11}$$

It is worth to mention that since $\mathbf{h}_{\mathrm{p}}$ is sparse in the virtual angular domain, only $|\mathrm{I}|$ elements of the channel vector $\mathbf{h}_{\mathrm{p}}$ are non-zero. We consider that the observation matrix Φ_{p} is available and the support I is unknown.

Based on [74], we can find an estimate of $\mathbf{h}_{\mathrm{p}}$ by applying the homotopy DCD algorithm to the $\ell_2\ell_0$ optimization, considering the minimization of the cost function

$$\mathbf{J}_{\tau}(\mathbf{h}_{\mathrm{p}}) = \frac{1}{2}\left\|\mathbf{r}_{\mathrm{p}} - \Phi_{\mathrm{p}}\mathbf{h}_{\mathrm{p}}\right\|_2^2 + \tau\left\|\mathbf{h}_{\mathrm{p}}\right\|_0. \tag{3.12}$$

Here, $\tau \in [0,1)$ is a regularization parameter. The second term in (4.6) makes it non-convex problem and the solution of it is NP-hard. To solve the problem, we initially assign the support set $\mathrm{I}_{\mathrm{p}} = \emptyset$, and by adding new elements into the support or removing elements from the support in several iterations following the proposition in [74], we can find an estimate of $\mathbf{h}_{\mathrm{p}}$. Therefore we need to assign initially a high value to the regularization parameter $\tau = \tau_{\text{max}}$ which can dominate the cost function to provide an empty support $\mathrm{I}_{\mathrm{p}} = \emptyset$. In the homotopy iterations, by gradually reducing value of τ as $\tau \leftarrow \gamma\tau$, where $\gamma \in [0,1)$, new elements can be added to the support or removed from the support [74]. The algorithm stops when $\tau < \tau_{\text{min}}$, where $\tau_{\text{min}} = \mu_{\tau}\tau_{\text{max}}$ and $\mu_{\tau} \in [0,1)$ is a predefined parameter, and $(\mathbf{h}_{\mathrm{p}})_{\mathrm{g}}$ is the gth element of the pth estimated channel vector $\mathbf{h}_{\mathrm{p}}$. The structure of the employed $\ell_2\ell_0$ DCD homotopy algorithm is shown in Algorithm 4.1.

As shown in Algorithm 4.1, by solving the LS problem $(\mathbf{R}_{\mathrm{p}})_{\mathrm{I_p},\mathrm{I_p}}(\mathbf{h}_{\mathrm{p}})_{\mathrm{I_p}} = \mathbf{f}_{\mathrm{p}}$ at step 3, $\mathbf{h}_{\mathrm{p}}$ is estimated. According to [74], instead of using the matrix inversion to solve the LS problem, the DCD iterations [74], as shown in Algorithm 4.2, are employed at step 3 in Algorithm 4.1. When the DCD iterations start, an LS solution for the vector $\mathbf{h}_{\mathrm{p}}$ and the vector $\mathbf{c}$ found at the previous iteration are used as the initialization of the DCD algorithm, which results in reduction of the computational complexity. In the DCD iterations, $\mathrm{N_u}$ is the maximum number of successful iterations and a successful iteration means that the solution is updated in the iteration, $\mathrm{M_b}$ and H are predefined parameters.

Algorithm 2 DCD iterations for LS minimization

Input: $\mathbf{h}_\mathrm{p}$, $\mathbf{c}$, $\mathrm{I_p}$, $\mathbf{R}_\mathrm{p}$
Initialization: $\mathrm{s}=0$, $\delta=\mathrm{H}$
1: **for** $\mathrm{m}=1,...,\mathrm{M_b}$ **do until** $\mathrm{s}=\mathrm{N_u}$
2: $\delta=\delta/2$, $\boldsymbol{\alpha}=[\delta,-\delta,\mathrm{j}\delta,-\mathrm{j}\delta]$, State =0
3: **for** $\mathrm{n}=1,...,|\mathrm{I_p}|$ **do**: $\mathrm{v}=\mathrm{I_p}(\mathrm{n})$
4: **for** $\mathrm{k}=1,...,4$ **do**
5: **if** $\Re\{(\boldsymbol{\alpha})_\mathrm{k}(\mathbf{c})^*_\mathrm{v}\}>\left[(\mathbf{R}_\mathrm{p})_{\mathrm{v,v}}\right]\delta^2/2$ **then**
6: $(\mathbf{h}_\mathrm{p})_\mathrm{v}\leftarrow(\mathbf{h}_\mathrm{p})_\mathrm{v}+(\boldsymbol{\alpha})_\mathrm{k}$, $\mathbf{c}\leftarrow\mathbf{c}-(\boldsymbol{\alpha})_\mathrm{k}\mathbf{R}^\mathrm{v}_\mathrm{p}$
7: State=1, $\mathrm{s}\leftarrow\mathrm{s}+1$
8: **if** State=1, **go to** step 3

3.3.2 Common Support Acquisition and Joint Channel Estimation

In this section, the process of estimating the common support I is presented. As an example, we consider a scenario with $\mathrm{P}=64$ pilot subcarriers, $\mathrm{M}=128$ transmit antennas, signal to noise ratio $\mathrm{SNR}=20$ dB, $\mathrm{J}=20$ OFDM symbols and $|\mathrm{I}|=8$.

According to [4], among M coordinates of the channel vector $\mathbf{h}_\mathrm{p}$, the vast majority of the channel energy will concentrate on $|\mathrm{I}|$ coordinates, which are non-zero elements in $\mathbf{h}_\mathrm{p}$. Since we can estimate the channel at the pth pilot subcarrier using the $\ell_2\ell_0$ homotopy DCD algorithm, we can find an estimate of the common support $\tilde{\mathrm{I}}$ by jointly analysing estimates $\tilde{\mathbf{h}}_\mathrm{p}$ of vectors $\mathbf{h}_\mathrm{p}$ for all pilot subcarriers.

In Fig.3.5(a) and Fig.3.5(b), magnitudes of elements of vectors $\tilde{\mathbf{h}}_1$ and $\tilde{\mathbf{h}}_{64}$ are shown. For estimation of the joint support, we compute

$$\mathbf{q}=\left(\sum_{\mathrm{p}=1}^{\mathrm{P}}\left|\tilde{\mathbf{h}}_\mathrm{p}\right|\right)/\mathrm{P}. \tag{3.13}$$

An estimate $\tilde{\mathrm{I}}$ of the common support I is obtained using thresholding, as a set of elements in the vector $\mathbf{q}$, satisfying the condition

$$\tilde{\mathrm{I}} = \{\mathrm{k} : (\mathbf{q})_{\mathrm{k}} > \xi\}, \tag{3.14}$$

where ξ is a predefined threshold parameter.

Based on the estimate $\tilde{\mathrm{I}}$, the LS algorithm [104] is employed as follows:

$$(\mathbf{R}_{\mathrm{p}})_{\tilde{\mathrm{I}},\tilde{\mathrm{I}}} \left(\tilde{\mathbf{h}}_{\mathrm{p}}\right)_{\tilde{\mathrm{I}}} = \mathbf{f}_{\tilde{\mathrm{I}}}, \tag{3.15}$$

$$\mathbf{f}_{\tilde{\mathrm{I}}} = (\Phi_{\mathrm{p}})_{\tilde{\mathrm{I}}}^{\mathrm{H}} \mathbf{r}_{\mathrm{p}}. \tag{3.16}$$

Here, $\left(\tilde{\mathbf{h}}_{\mathrm{p}}\right)_{\tilde{\mathrm{I}}}$ is the final estimate of the channel vector $\mathbf{h}_{\mathrm{p}}$ on the support $\tilde{\mathrm{I}}$.

3.4 DSAMP algorithm

In [4], Zhen Gao proposes an algorithm for estimating the channel matrix in a frequency-division duplex massive multiple-input multiple-output (MIMO) system. The algorithm is based on exploiting the spatial sparsity of the channel matrix and assumes that the channel matrix has a common sparsity pattern across all antennas.

The proposed algorithm, called the distributed sparsity adaptive matching pursuit algorithm (DSAMP), utilizes training data consisting of received signal vectors and transmit signal matrices. The algorithm iteratively estimates the common support set and common support pattern of the channel matrix using the residual matrix obtained from the difference between the received and estimated transmit signals. The common support set refers to the indices of the non-zero entries that are common across all antennas of the channel matrix. Next, the algorithm estimates the common and non-common support matrices and coefficients by solving a sparse recovery problem subject to a constraint on the residual error. Finally, the estimated channel matrix is obtained by combining the estimated common and non-common support matrices.

Overall, the DSAMP algorithm [4], which was developed from the sparsity adaptive matching pursuit algorithm [82], can acquire multiple sparse channel vectors for different pilot subcarriers simultaneously. The DSAMP algorithm has been shown to provide a better channel estimation performance than the orthogonal matching pursuit, sparsity adaptive matching pursuit and subspace pursuit algorithms [4]. We use the DSAMP performance as a benchmark to assess the performance of the proposed DCD-JSR algorithm.

3.5 Simulation Results

3.5.1 MSE of the Channel Estimation

We will be assessing the algorithm performance using the mean square error (MSE) of the channel estimation. The MSE is given by

$$\mathrm{MSE} = \frac{\left|\left|\mathbf{h}_{\mathrm{p}} - \tilde{\mathbf{h}}_{\mathrm{p}}\right|\right|_2^2}{\left|\left|\mathbf{h}_{\mathrm{p}}\right|\right|_2^2}, \tag{3.17}$$

$$\left|\left|\tilde{\mathbf{h}}_{\mathrm{p}}\right|\right|_2 = \sqrt{\sum_{\mathrm{m}=1}^{\mathrm{M}} \left|\left(\tilde{\mathbf{h}}_{\mathrm{p}}\right)_{\mathrm{m}}\right|^2}, \tag{3.18}$$

where $\tilde{\mathbf{h}}_{\mathrm{p}}$ is the estimated channel vector and $\mathbf{h}_{\mathrm{p}}$ is the true channel vector. When analysing the performance of the estimators, we will also calculate the probability of the estimated support $\tilde{\mathrm{I}}$ to be exactly the same as the support I to be estimated.

3.5.2 Numerical Results

In this section, we consider simulation scenarios corresponding to a MIMO system with a uniform linear array. We compare the channel estimation performance of the DCD-JSR and DSAMP algorithms. The performance of the oracle LS algorithm [104] with known support is adopted as the performance bound. In most scenarios, we consider two cases, $\mathrm{SNR} = 10$ dB and $\mathrm{SNR} = 20$ dB.

To provide the best MSE performance, the threshold p_{th} for the DSAMP algorithm and ξ for the DCD-JSR algorithm need to be adjusted. As shown in Fig.3.6, for $\mathrm{SNR} = 20$ dB, the DCD-JSR algorithm has the best MSE performance when $\xi = 0.055$. In Fig.3.7, it can be seen that when $\mathrm{SNR} = 20$ dB and $\mathrm{p}_{\mathrm{th}} = 0.1$, the DSAMP algorithm achieves the best MSE

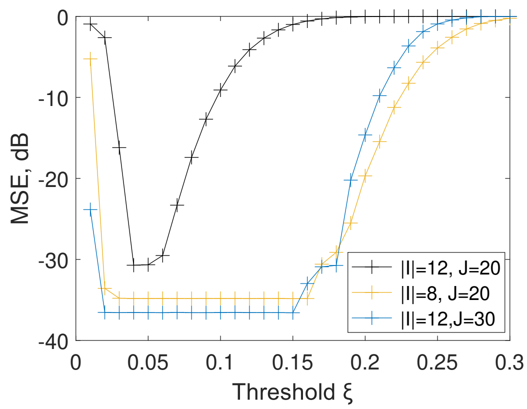

Fig. 3.6 MSE performance of the DCD-JSR algorithm against the threshold ξ, SNR=20 dB, the number of pilot subcarriers $\mathrm{P} = 64$, $\mathrm{M} = 128$.

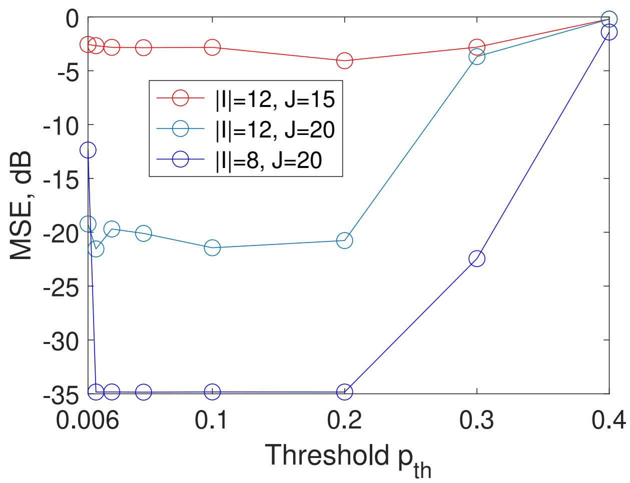

Fig. 3.7 MSE performance of the DSAMP algorithm against the threshold $\mathrm{p_{th}}$, SNR=20 dB, the number of pilot subcarriers $\mathrm{P} = 64$, $\mathrm{M} = 128$.

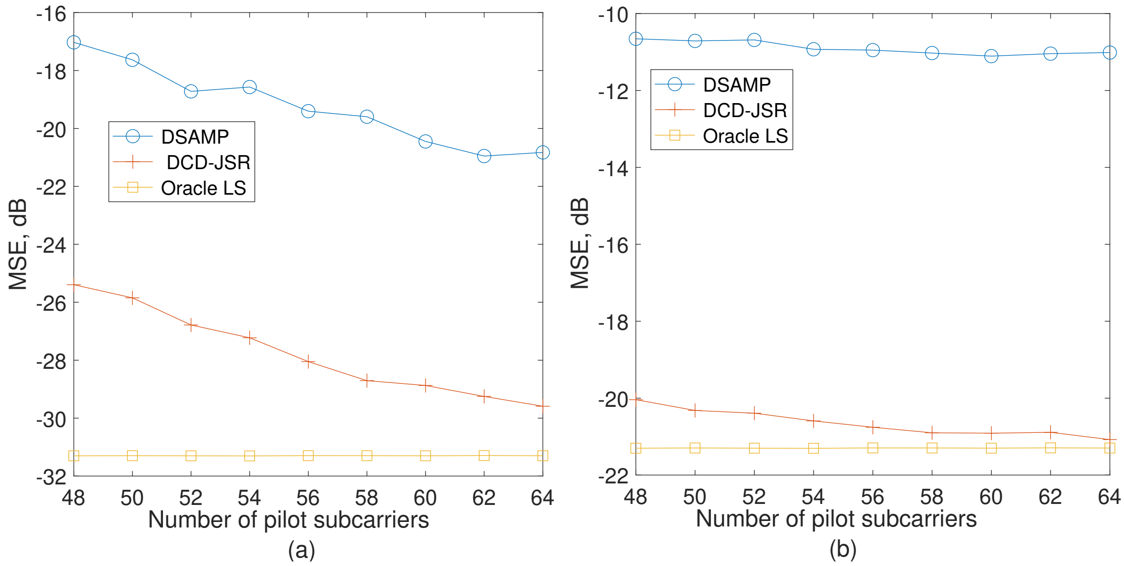

Fig. 3.8 MSE performance of Oracle LS, DSAMP, and DCD-JSR algorithms against the number of pilot subcarriers, $\mathrm{M} = 128$, $\mathrm{J} = 20$: (a) $\mathrm{SNR} = 20$ dB, (b) $\mathrm{SNR} = 10$ dB.

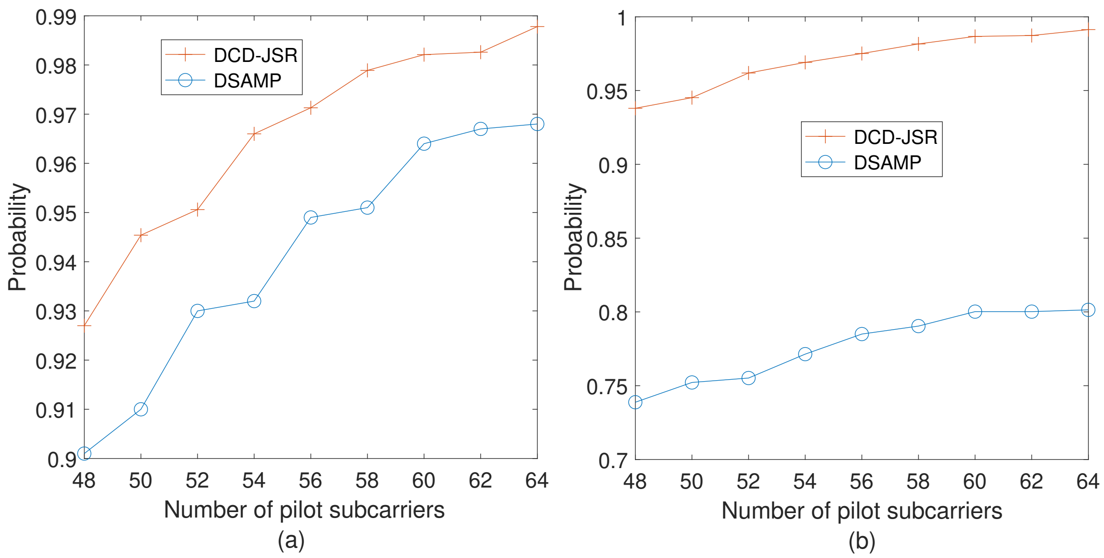

Fig. 3.9 Probability of perfect support estimation for DSAMP and DCD-JSR algorithms against the number of pilot subcarriers, $\mathrm{M} = 128$ $\mathrm{J} = 20$: (a) $\mathrm{SNR} = 20$ dB, (b) $\mathrm{SNR} = 10$ dB.

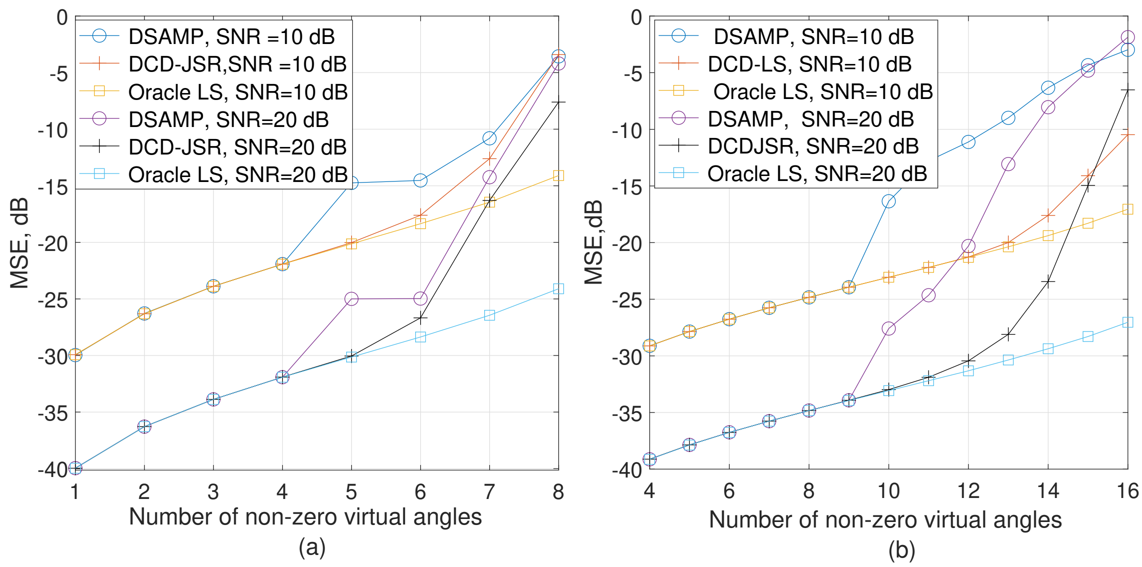

Fig. 3.10 MSE performance of Oracle LS, DSAMP, DCD-JSR algorithms against the number of non-zero virtual angles $\mathrm{M} = 128$, $\mathrm{P} = 64$: (a) J=10, (b) J=20.

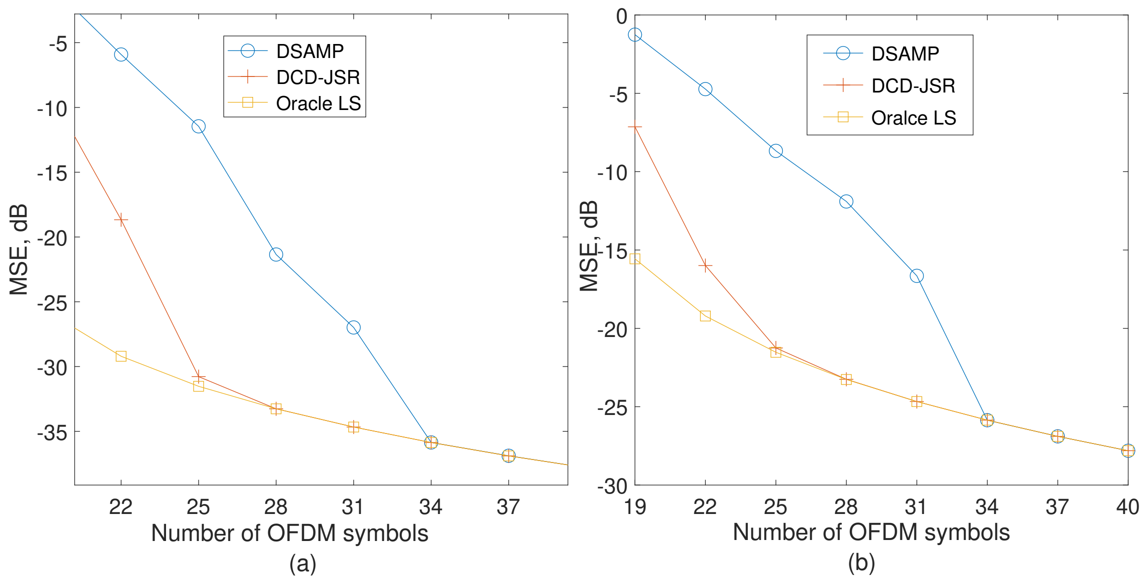

Fig. 3.11 MSE performance of Oracle LS, DSAMP, and DCD-JSR algorithms against the number of OFDM symbols $\mathrm{M} = 128$, $\mathrm{P} = 64$, $|\mathrm{I}| = 16$: (a) $\mathrm{SNR} = 20$ dB, (b) $\mathrm{SNR} = 10$ dB.

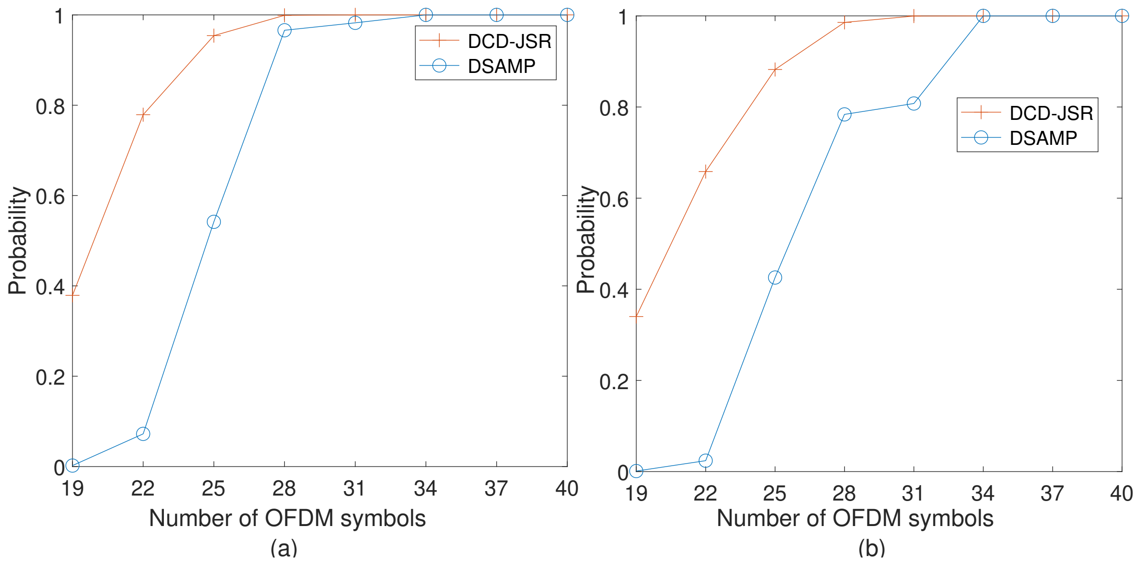

Fig. 3.12 Probability of perfect support estimation for DSAMP and DCD-JSR algorithms against the number of OFDM symbols, $M = 128$, $P = 64$, $|I| = 16$: (a) $SNR = 20$dB, (b) $SNR = 10$ dB.

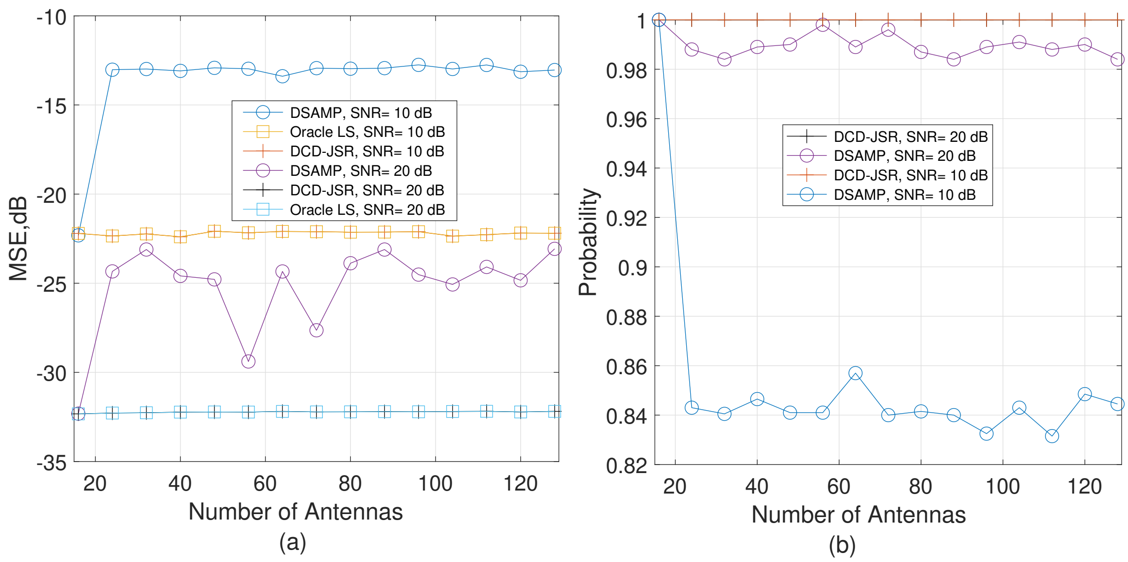

Fig. 3.13 Performance of Oracle LS, DSAMP, and DCD-JSR algorithms against the number of antennas, $J = 20$, $P = 64$ (a) MSE. (b) Probability of perfect support estimation.

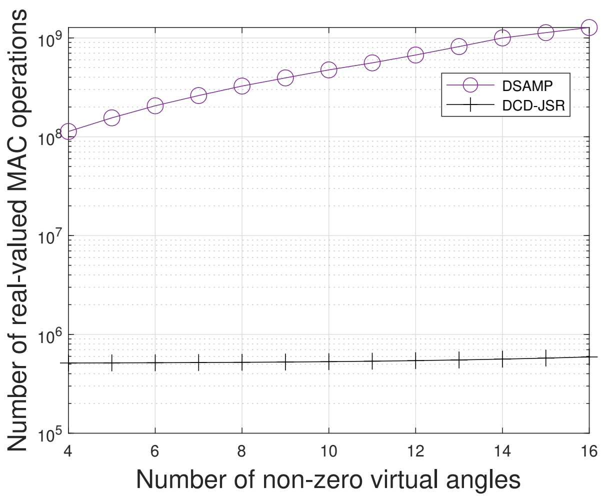

Fig. 3.14 Computational complexity of the DSAMP algorithm and the DCD-JSR algorithm, $\mathrm{M} = 128$, $\mathrm{J} = 20$, $\mathrm{P} = 64$, $\mathrm{SNR} = 20$ dB.

performance. Similarly, appropriate values of ξ and $\mathrm{p_{th}}$ for different SNR can be obtained. In this chapter, for the DCD-JSR algorithm, $\xi = 0.05$ is considered for both $\mathrm{SNR} = 20$ dB and $\mathrm{SNR} = 10$ dB; for the DSAMP algorithm, $\mathrm{p_{th}}$ is set to be 0.1 and 0.17 for $\mathrm{SNR} = 20$ dB and $\mathrm{SNR} = 10$ dB, respectively.

In Fig.3.8(a) and Fig.3.8(b), we consider scenarios with different number of pilot subcarriers. The number of pilot subcarriers varies from 48 to 64, and we set $\mathrm{M} = 128$, $|\mathrm{I}| = 12$, the number of simulation trials is $\mathrm{N_s} = 10000$. It can be seen that both the DSAMP and DCD-JSR algorithms benefit from the increasing number of pilot subcarriers, but a larger number of subcarriers results in lower spectral efficiency, since a smaller number of subcarriers are used for data transmission. However, the DCD-JSR algorithm shows significantly better MSE performance.

Fig.3.9(a) and Fig.3.9(b), for different number of pilot subcarriers and different SNR, show the probability of the perfect support estimation by the DSAMP and DCD-JSR algorithms, where the perfect support estimation means that the estimated support is exactly the same as the true support. In Fig.3.9, it can be seen that, compared to the DSAMP algorithm, the DCD-JSR algorithm provides a better probability of correct support estimation. This explains the better MSE performance of the DCD-JSR algorithm, as seen in Fig.7. Compared to the DSAMP algorithm, the DCD-JSR algorithm requires less pilot subcarriers to provide a specified probability of correct support estimation under same scenario.

In Fig.3.10(a) and Fig.3.10(b), we show the MSE performance for scenarios with $\mathrm{J} = 10$ and $\mathrm{J} = 20$ at different SNR. We set $\mathrm{M} = 128$, $\mathrm{P} = 64$, and the number of simulation trials $\mathrm{N_s} = 10000$. In Fig.3.10(a), for J=10, at $\mathrm{SNR} = 10$ dB, and $|\mathrm{I}| \leq 6$, the DCD-JSR algorithm approaches the performance of the oracle LS algorithm [104], while the DSAMP does it only for $|\mathrm{I}| \leq 4$. In Fig.3.10(b), for J=20, when $\mathrm{SNR} = 10$ dB, the DCD-JSR algorithm approaches the performance of the oracle LS algorithm [104] for $|\mathrm{I}| \leq 13$, whereas the DSAMP algorithm does not show the LS performance even for $|\mathrm{I}| = 10$. When $\mathrm{SNR} = 20$ dB, the DCD-JSR algorithm could approach the oracle performance until $|\mathrm{I}| = 13$, while the DSAMP does not. Hence, in these scenarios, the DCD-JSR algorithm outperforms the DSAMP algorithm.

Fig.3.11(a) and Fig.3.11(b) present results for different number of employed OFDM symbols J. The number of simulation trials is $\mathrm{N_s} = 10000$, $\mathrm{M} = 128$, $\mathrm{P} = 64$. It can be seen that the DCD-JSR algorithm outperforms the DSAMP algorithm for both $\mathrm{SNR} = 20$ dB and $\mathrm{SNR} = 10$ dB, and requires less OFDM symbols to approach the performance of the oracle LS channel estimator.

Fig.3.12(a) and Fig.3.12(b) compare the probability of perfect support estimation by the DSAMP and DCD-JSR channel estimators. It can be seen that the DCD-JSR channel estimator outperforms the DSAMP channel estimator: at $\mathrm{SNR} = 20$ dB, the DCD-JSR channel estimator needs $\mathrm{J} = 28$ to provide the perfect support estimation, while the DSAMP algorithm needs $\mathrm{J} = 34$, i.e., a lower number of OFDM symbols is required by the DCD-JSR algorithm. Thus, it is easy to see that, compared to the DSAMP channel estimator, the DCD-JSR channel estimator requires less OFDM symbols for an accurate support estimation.

In Fig.3.13, we consider the case where the massive MIMO system employs different number of antennas. The number of antenna varies from 16 to 128, the number of simulation trials is $\mathrm{N_s} = 10000$. We set the number of OFDM symbols $\mathrm{J} = 20$ and number of non-zero virtual angles $|\mathrm{I}| = 11$. In Fig.3.13(a), it can be seen that when $\mathrm{SNR} = 10$ dB, there exists a significant performance gap between the DSAMP algorithm and oracle LS algorithm, while the DCD-JSR algorithm approaches the oracle performance for any number of antennas. For the higher SNR, $\mathrm{SNR} = 20$ dB, the DCD-JSR channel estimator approaches the oracle performance for any number of antennas, while the DSAMP algorithm does not.

Fig.3.13(b) shows the probability of perfect support estimation in these scenarios. It can be seen that the DCD-JSR algorithm always provides perfect support estimation, while the DSAMP algorithm does not. Thus, we can see that with a large number of antennas, the DCD-JSR channel estimator provides a better MSE performance and more accurate support estimation than the DSAMP algorithm.

To estimate the computational complexity of the algorithms, we decided to update the computational complexity after each line of the algorithm code (both the algorithms have been implemented in Matlab) where an operation occurs. In the DCD-JSR algorithm, most of the operations are additions [74]; to simplify the comparison, we also count the pure additions as multiply-accumulate (MAC) operations.

Fig.3.14 shows the computational complexity against the number of non-zero virtual angles. We consider the $\mathrm{SNR} = 20$ dB, $J = 20$ and average the results over $\mathrm{N_s} = 10000$ simulation trials. It can be seen that the DCD-JSR algorithm has significantly lower complexity. Thus we can say that, compared to the DSAMP algorithm [4], the DCD-JSR algorithm exhibits lower computational complexity.

3.6 Conclusion

In this chapter, based on the original $\ell_2\ell_0$ DCD algorithm, a DCD-JSR algorithm has been proposed to jointly estimate the channel for multiple pilot subcarriers in the virtual angular domain in an FDD massive MIMO system. The DSAMP algorithm is used to compare the channel estimation performance with the DCD-JSR algorithm in different simulation scenarios. Simulation results have shown that the proposed DCD-JSR algorithm outperforms the DSAMP algorithm, and requires less OFDM symbols and employed pilot subcarriers for accurate channel estimation, whereas it also exhibits a significantly lower computational complexity.

Chapter 4

Estimation of time-varying channels in virtual angular domain for massive MIMO systems

4.1 Introduction

In [108] and [109], and many other publications, it has been shown that, the time-varying channel can be approximate accurately by employing the basis expansion model (BEM). Consequently, estimation of a realization of random process describing the time-varying channel is transformed into estimation of a few-time-invariant expansion coefficients [110]. In [111], the Karhunen-Loeve BEM has been proposed for estimating the time-varying channel, however, it is very sensitive to the variations of the channel statistics. In [112] and [113], algebraic polynomial BEMs have been employed to estimate the time-varying channel, where the channel vectors can be approximated as a linear combination of a set of polynomials. In [109], the experimental results have indicated that, by employing the Legendre polynomials as the BEM, the channel variation could be accurately approximated. Thus, in this chapter, we consider employing the Legendre polynomials to approximate the time-variation of the channel in the virtual angular domain..

In this chapter, by combining the DCD-JSR algorithm, proposed in Chapter 3 and the BEM, we have show that the modified DCD-JSR algorithm can estimate the channel in OFDM system operating over frequency selective and highly mobile wireless time-varying channels. Simulation results show that, compared to the original DCD-JSR algorithm, the modified DCD-JSR algorithm could provide better MSE performance when estimating time-varying channels.

The chapter is organized as follows. Section 4.2 describes the system model. In Section 4.3, channel model for time varying channel is introduced. In Section 4.4, the processes of employing the basis functions, and estimating the time-varying channel are described. In Section 4.5, numerical examples are analysed and, finally, Section 4.6 presents the conclusion.

4.2 System Model

4.2.1 Channel Estimation Approach

The channel estimation is performed at the BS. The channel estimation scheme is summarized as follows.

1. In each OFDM symbol, every BS antenna broadcasts pilot symbols to users, the kth user receives the signal $\mathbf{y}_{\mathrm{k}} \in \mathrm{C}^{\mathrm{M}\times 1}$ and feeds it back to the BS. The BS recovers the CSI for each user based on the feedback signals $\mathbf{y}_{\mathrm{k}}$, $\mathrm{k} = 1, ..., \mathrm{K}$. Each OFDM symbol contains N subcarriers, while P subcarriers are used to transmit pilot symbols. The user feeds back the received signal to the BS without performing downlink channel estimation.

2. At the BS, by employing the BEM to approximate the time variation of the channel, the DCD-JSR algorithm can jointly estimate the common support I for multiple sparse virtual angular domain channels. The least squares (LS) algorithm [104] is then employed to acquire the CSI based on an estimate of the common support I.

4.2.2 Received Signal

For the tth OFDM symbol, at the pth pilot subcarrier, the signal received at the BS is given by:

$$\mathrm{r}_{\mathrm{p}}^{\mathrm{t}} = \boldsymbol{\phi}_{\mathrm{p}}^{\mathrm{t}} \tilde{\mathbf{h}}_{\mathrm{n(p)}}^{\mathrm{t}} + \mathrm{v}_{\mathrm{p}}^{\mathrm{t}},\ 1 \leq \mathrm{p} \leq \mathrm{P},\ 1 \leq \mathrm{t} \leq \mathrm{J}. \tag{4.1}$$

Here, $\boldsymbol{\phi}_{\mathrm{p}}^{\mathrm{t}} = \left(\mathbf{s}_{\mathrm{p}}^{\mathrm{t}}\right)^{\mathrm{T}} (\mathbf{A}_{\mathrm{B}}^{*})^{\mathrm{T}} \in \mathrm{C}^{1\times \mathrm{M}}$ is the sensing vector defined by the DFT matrix and the pilot symbols. $\tilde{\mathbf{h}}_{n(p)}^{t} \in \mathrm{C}^{\mathrm{M}\times 1}$ is the sparse channel vector for the $\mathrm{n(p)}$th subcarrier, and $\mathrm{v}_{\mathrm{p}}^{\mathrm{t}}$ is the corresponding noise, which contains both downlink and uplink channel noise.

4.3 Time varying channel

In practice, due to the user mobility, the propagation of wireless signals would face the time-varying environment [114]. Due to the simple implementation of the Legendre polynomial matrix, using the Legendre polynomial matrix as basis functions with a period equal to the length of the investigated interval has been considered in the literature [109]. In this chapter, we consider that for the pth pilot subcarrier, the time-varying channel vector can be approximated by $\mathrm{N_b}$ basis functions:

$$\begin{aligned} &\hat{\mathbf{h}}_{\mathrm{n(p)}}^{\mathrm{t}} = \sum_{i=1}^{\mathrm{N_b}} \mathrm{b_i(t)} \mathbf{c}_{\mathrm{i,p}}, \\ &1 \leq \mathrm{t} \leq \mathrm{J}, \end{aligned} \tag{4.2}$$

where $\mathrm{b_i(t)}$ is the tth element of a vector $\mathbf{b}_{\mathrm{i}} \in \mathrm{C}^{\mathrm{J}\times 1}$ representing samples of the basis function $\mathrm{b_i(t)}$, $\mathbf{c}_{\mathrm{i,p}} \in \mathrm{C}^{\mathrm{M}\times 1}$ are expansion coefficients for the ith basis function at the pth pilot subcarrier. By employing basis functions to approximate the time variations of the channel, we decompose the channel variation into a set of linear combinations of basis functions. This can help to reduce the complexity of the problem by allowing us to focus on the linear behavior of the channel over small time intervals.

In this chapter, we consider employing the Legendre polynomials as the basis functions , this is because it has been indicated in [109], the Legendre polynomial can be employed to represent rapidly time-varying fading channel, while with low computational complexity. The Legendre polynomials are defined as:

$$\mathrm{b_i(t)} = \frac{1}{2^{\mathrm{i}-1}\mathrm{i}!}\frac{\mathrm{d}^{\mathrm{i}-1}}{\mathrm{dt}^{\mathrm{i}-1}}[(\mathrm{t}^2-1)^{\mathrm{i}-1}],\ \mathrm{i} \geq 1. \tag{4.3}$$

For the tth OFDM symbol at the pth pilot subcarrier, by substituting (4.2) into (4.1) , we can obtain:

$$\mathrm{r}_\mathrm{p}^\mathrm{t} = \boldsymbol{\phi}_\mathrm{p}^\mathrm{t}\sum_{\mathrm{i}=1}^{\mathrm{N_b}}\mathrm{b_i(t)}\mathbf{c}_{\mathrm{i,p}} + \mathrm{v}_\mathrm{p}^\mathrm{t},\ 1 \leq \mathrm{p} \leq \mathrm{P},\ 1 \leq \mathrm{t} \leq \mathrm{J}. \tag{4.4}$$

Since $\tilde{\mathbf{h}}_{\mathrm{n(p)}}^{\mathrm{t}} \in \mathrm{C}^{\mathrm{M}\times 1}$ exhibit common sparsity, the expansion coefficient vectors $\mathbf{c}_{\mathrm{i,p}} \in \mathrm{C}^{\mathrm{M}\times 1}$ also exhibit common sparsity. Thus, the task of estimating JM channel coefficients is transformed into estimating only $\mathrm{N_b|I|}$ expansion coefficients with usually $\mathrm{N_b} \ll \mathrm{J}$ and $|\mathrm{I}| \leq \mathrm{M}$.

We collect the received signal samples $\mathrm{r}_\mathrm{p}^\mathrm{t}, 1 \leq \mathrm{t} \leq\ \mathrm{J}$, in a vector $\mathbf{r}_\mathrm{p} = \left[\mathrm{r}_\mathrm{p}^1, \mathrm{r}_\mathrm{p}^2, ..., \mathrm{r}_\mathrm{p}^\mathrm{J}\right]^\mathrm{T} \in \mathrm{C}^{\mathrm{J}\times 1}$, then we have:

$$\mathbf{r}_\mathrm{p} = \sum_{i=0}^{\mathrm{N_b}}\mathbf{F}_{\mathrm{i,p}}\mathbf{c}_{\mathrm{i,p}} + \mathbf{v}_p,\ \ 1 \leq \mathrm{p} \leq \mathrm{P}, \tag{4.5}$$

where $\mathbf{F}_{\mathrm{i,p}}^\mathrm{t} = \boldsymbol{\phi}_\mathrm{p}^\mathrm{t}\mathrm{b_i(t)} \in \mathrm{C}^{1\times\mathrm{M}}$, $\mathbf{F}_{\mathrm{i,p}} = [\mathbf{F}_{\mathrm{i,p}}^1, \mathbf{F}_{\mathrm{i,p}}^2, ..., \mathbf{F}_{\mathrm{i,p}}^\mathrm{t}] \in C^{\mathrm{J}\times\mathrm{M}}$ is a matrix whose tth row is $\mathbf{F}_{\mathrm{i,p}}^\mathrm{t}$, and $\mathbf{v}_\mathrm{p} = \left[\mathrm{v}_\mathrm{p}^1, \mathrm{v}_\mathrm{p}^2, ..., \mathrm{v}_\mathrm{p}^\mathrm{J}\right]^\mathrm{T} \in \mathrm{C}^{\mathrm{J}\times 1}$ is the noise vector. Since the expansion coefficients exhibit common sparsity, we can firstly estimate the common support and then find the expansion coefficients.

4.4 DCD-JSR algorithm for time-varying channels

Here, the homotopy DCD algorithm [74] is used to estimate the support of the expansion coefficients, as shown in Table 4.1. First, we apply the homotopy DCD algorithm to the $\ell_2\ell_0$ optimization problem of minimizing:

$$\mathbf{J}_\tau(\tilde{\mathbf{c}}_{\mathrm{i,p}}) = \frac{1}{2}\left\|\mathbf{r}_\mathrm{p} - \mathbf{F}_{\mathrm{i,p}}\tilde{\mathbf{c}}_{\mathrm{i,p}}\right\|_2^2 + \tau\left\|\tilde{\mathbf{c}}_{\mathrm{i,p}}\right\|_0. \tag{4.6}$$

Table 4.1 $\ell_2\ell_0$ homotopy DCD algorithm

Initialization: For the ith expansion coefficient at the pth pilot subcarrier, vector $\mathbf{d} = \mathbf{0}$, $\mathrm{I_p} = \emptyset$, $\mathbf{b}_\mathrm{p} = \mathbf{F}_{\mathrm{i,p}}{}^{\mathrm{H}}\mathbf{r}_\mathrm{p}$, $\mathbf{R}_\mathrm{p} = \mathbf{F}_{\mathrm{i,p}}{}^{\mathrm{H}}\mathbf{F}_{\mathrm{i,p}}$.

1: $g = \arg\max_{\mathrm{k}\in\mathrm{I_p^c}} \left|(\mathbf{b}_\mathrm{p})_\mathrm{k}\right|^2 / (\mathbf{r}_\mathrm{p})_{\mathrm{k,k}}$,
$\tau_{\text{max}} = (1/2)\max_\mathrm{k} \left|(\mathbf{b}_\mathrm{p})_\mathrm{k}\right|^2 / (\mathbf{R}_\mathrm{p})_{\mathrm{k,k}}$,
$\tau = 0.5\left|(\mathbf{b}_\mathrm{p})_\mathrm{g}\right|^2 / (\mathbf{R}_\mathrm{p})_{\mathrm{g,g}}$, $\mathrm{I_p} = \{g\}$.

2: $\mathbf{Repeat}$ until the termination condition is met:

3: $\mathbf{If}$ the support $\mathrm{I_p}$ has been updated $\mathbf{then}$
Solve $(\mathbf{R}_\mathrm{p})_{\mathrm{I_p,I_p}} [\tilde{\mathbf{c}}_{\mathrm{i,p}}]_{\mathrm{I_p}} = \mathbf{f}_\mathrm{p}$,
where $\mathbf{f}_\mathrm{p} = [(\mathbf{F}_{\mathrm{i,p}}]_{\mathrm{I_p}}^H \mathbf{R}_\mathrm{p}$
$\mathbf{d} \leftarrow \mathbf{b} - (\mathbf{R}_\mathrm{p})_{\mathrm{I_p,I_p}} [\tilde{\mathbf{c}}_{\mathrm{i,p}}]_{\mathrm{I_p}}$

4: Update the regularization parameter : $\tau \leftarrow \gamma\tau$

5: Add the g-th element into the support $\mathrm{I_p}$,
where $g \in \mathrm{I_p}{}^c$,
and $g = \arg\max_{\mathrm{k}\in\mathrm{I_p}{}^c} \frac{\left|(\mathbf{c})_\mathrm{k}\right|^2}{(\mathbf{R}_\mathrm{p})_{\mathrm{k,k}}}$ s.t $\left|(\mathbf{c})_\mathrm{g}\right|^2 > 2\tau\,(\mathbf{R}_\mathrm{p})_{\mathrm{g,g}}$,
then assign to $[\tilde{\mathbf{c}}_{\mathrm{i,p}}]_\mathrm{g}$ the value $(\mathbf{d})_\mathrm{g} / (\mathbf{R}_\mathrm{p})_{\mathrm{g,g}}$,
update $\mathbf{d} \leftarrow \mathbf{d} - [\tilde{\mathbf{c}}_{\mathrm{i,p}}]_\mathrm{g}\,\mathbf{R}_p^\mathrm{g}$.

6: Remove the gth element from the support $\mathrm{I_p}$,
where $g \in \mathrm{I_p}$, and
$g = \arg\min_{\mathrm{k}\in\mathrm{I_p}} \left[\frac{1}{2}\left|[\tilde{\mathbf{c}}_{\mathrm{i,p}}]_\mathrm{k}\right|^2 (\mathbf{R}_\mathrm{p})_{\mathrm{k,k}} + \Re\left\{[\tilde{\mathbf{c}}_{\mathrm{i,p}}]_\mathrm{k}^* (\mathbf{c})_\mathrm{k}\right\}\right]$,
s.t. $\frac{1}{2}\left|(\mathbf{c}_{\mathrm{i,p}})_\mathrm{g}\right|^2 (\mathbf{R}_\mathrm{p})_{\mathrm{g,g}} + \Re\left\{[(\tilde{\mathbf{c}}_{\mathrm{i,p}}]_\mathrm{g}^* (\mathbf{d})_\mathrm{g}\right\} < \tau$
for every removed element,
update $\mathbf{d} \leftarrow \mathbf{d} + (\tilde{\mathbf{c}}_{\mathrm{i,p}})_\mathrm{g}\,\mathbf{R}_p^\mathrm{g}$ and set $[(\tilde{\mathbf{c}}_{\mathrm{i,p}}]_\mathrm{g} = 0$.

7: If $\tau < \tau_{\min}$, Stop.

Table 4.2 DCD iterations for LS minimization

Input: $\tilde{\mathrm{c}}_{\mathrm{i,p}}$, $\mathbf{d}$, $\mathrm{I_p}$, $\mathbf{R}_\mathrm{p}$
Initialization: $\mathrm{s} = 0$, $\delta = \mathrm{H}$
1: **for** $\mathrm{m} = 1, ..., \mathrm{M_b}$ **do until** $\mathrm{s} = \mathrm{N_u}$
2: $\delta = \delta/2$, $\boldsymbol{\alpha} = [\delta, -\delta, j\delta, -j\delta]$, State =0
3: **for** $\mathrm{n} = 1, ..., |\mathrm{I_p}|$ **do**: $\mathrm{e} = \mathrm{I_p}(\mathrm{n})$
4: **for** $\mathrm{k} = 1, ..., 4$ **do**
5: **if** $\Re\{(\alpha)_\mathrm{k}(d)_e^*\} > \left[(\mathrm{R_p})_{\mathrm{e,e}}\right]\delta^2/2$ **then**
6: $[\tilde{\mathrm{c}}_{\mathrm{i,p}}]_\mathrm{e} \leftarrow [\tilde{\mathrm{c}}_{\mathrm{i,p}}]_\mathrm{e} + (\alpha)_\mathrm{k}$, $\mathbf{d} \leftarrow \mathbf{d} - (\alpha)_\mathrm{k}\,\mathbf{R}_\mathrm{p}{}^{(e)}$
7: State=1, $\mathrm{s} \leftarrow \mathrm{s} + 1$
8: **if** State=1, **go to step 3**

Here, we solve the optimization problem for the pth pilot subcarrier of the ith expansion coefficient, and $\tau \in [0, 1)$ is a regularization parameter. The second term in (4.6) makes it a non-convex problem and the solution of it is NP-hard. To solve the problem, we initially assign the support set $\mathrm{I_p} = \emptyset$, and by following the proposition in [74] we can add new elements into the support or remove elements from the support in several iterations, thus, the estimated expansion coefficients $\tilde{\mathbf{c}}_{\mathrm{i,p}}$ can be obtained.

Therefore we need to assign initially a high value to the regularization parameter $\tau = \tau_{\text{max}}$, so that the second term in (4.6) dominates the cost function to provide an empty support $\mathrm{I_p} = \emptyset$. In the homotopy iterations, by gradually reducing value of τ as $\tau \leftarrow \gamma\tau$, where $\gamma \in [0, 1)$, new elements can be added to the support or removed from the support [74]. The algorithm stops when $\tau < \tau_{\text{min}}$, where $\tau_{\text{min}} = \mu_\tau \tau_{\text{max}}$ and $\mu_\tau \in [0, 1)$ is a predefined parameter.

To reduce the computational complexity [74], instead of solving the LS problem in Table 4.1, step 3, we employ the DCD iterations to solve the LS problem, as shown in Table 4.2, where $\mathrm{N_u}$ is the number of successful DCD iterations, and a successful DCD iteration means that the solution is updated.

Following is an example of how we estimate the common support for the expansion coefficients. For the simulation scenario, we consider a massive MIMO system, SNR=20 dB, $\mathrm{M} = 128, \mathrm{P} = 64, \mathrm{J} = 100$, the normalized Doppler frequency $\mathrm{f_d T} = 0.05$ and $\mathrm{N_b} = 3$, $|\mathrm{I}| = 8$.

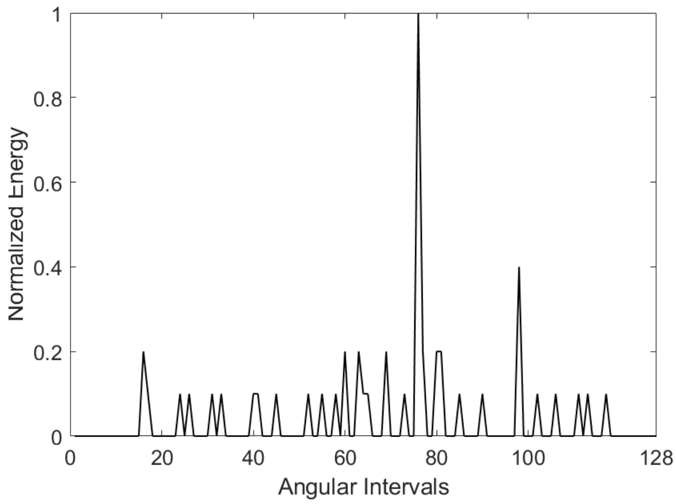

Fig. 4.1 Average of normalized energy of elements of the vector $\tilde{\mathbf{c}}_{1,1}$, for the expansion coefficient of the first basis function (zero-order Legendre polynomial) for the first pilot subcarrier against the angular intervals. $\mathrm{SNR} = 20$ dB, $\mathrm{J} = 100$, $|\mathrm{I}| = 8$.

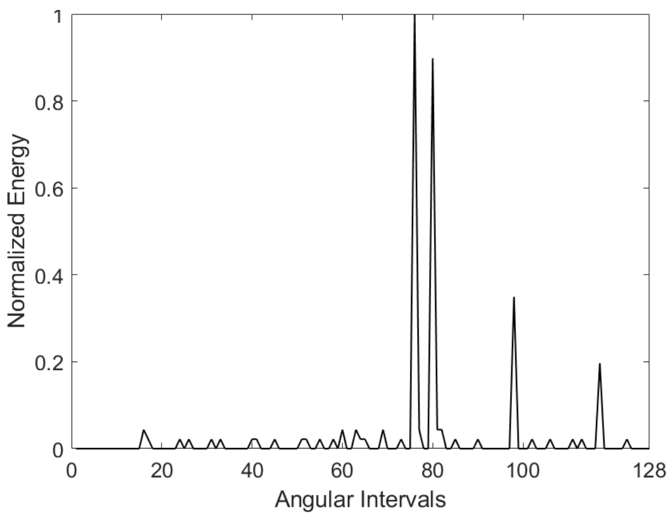

Fig. 4.2 Average of the normalized energy of elements of the vector $\mathbf{q}_1$, for the expansion coefficients of two basis functions (zero-order and first order Legendre polynomials) and all pilot subcarriers against the angular intervals. $\mathrm{SNR} = 20$ dB, $\mathrm{J} = 100$, $|\mathrm{I}| = 8$.

Fig. 4.1 shows the normalized energy of the elements of the estimated expansion coefficient $\mathbf{c}_{1,1}$, for the first basis function (zero-order Legendre polynomial) at the first pilot subcarrier in the angular intervals. It is easy to see that, due to the large variance of the energy of the elements in the angular intervals, we cannot clearly identify the support.

Since all expansion coefficients share a common support, for the pth pilot subcarrier, we can compute another vector with contribution from all the expansion coefficients:

$$\mathbf{q}_{\mathrm{p}} = \sum_{\mathrm{i}=1}^{\mathrm{N_b}} |\tilde{\mathbf{c}}_{\mathrm{i,p}}|^2, \tag{4.7}$$
$$\tilde{\mathbf{q}}_{\mathrm{p}} = \mathbf{q}_{\mathrm{p}}/(\max(\mathbf{q}_{\mathrm{p}})\mathrm{N_b}). \tag{4.8}$$

Here, as shown in Fig. 4.2, $\tilde{\mathbf{q}}_{\mathrm{p}}$ is a vector that contains normalized energy of elements for all expansion coefficients at the pth pilot subcarrier in the angular intervals. The new plot shows clearly 4 of 8 non-zero directions. However, the variance is still large and we cannot estimate reliably the support at this step.

As indicated in the previous section, since the expansion coefficients $\mathbf{c}_{\mathrm{i,p}}$ share the common support among P subcarriers, we can compute a new vector, which takes this into account:

$$\mathbf{q} = \sum_{\mathrm{i}=1}^{\mathrm{N_b}} \sum_{\mathrm{p}=1}^{\mathrm{P}} |\tilde{\mathbf{c}}_{\mathrm{i,p}}|^2, \tag{4.9}$$
$$\tilde{\mathbf{q}} = \mathbf{q}/(\max(\mathbf{q})\mathrm{N_b}\mathrm{P}). \tag{4.10}$$

As shown in Fig.4.3, here, $\tilde{\mathbf{q}} \in \mathrm{C}^{\mathrm{M}\times 1}$ is a sparse vector with elements averaging contribution from all pilot subcarriers and all expansion coefficients. We can acquire now the common support $\tilde{\mathrm{I}}$ by using the hard thresholding

$$\tilde{\mathrm{I}} = \{\mathrm{k} : [\mathbf{q}]_{\mathrm{k}} > \xi \max[\mathbf{q}]\}, \tag{4.11}$$

where ξ is a predefined thresholding parameter.

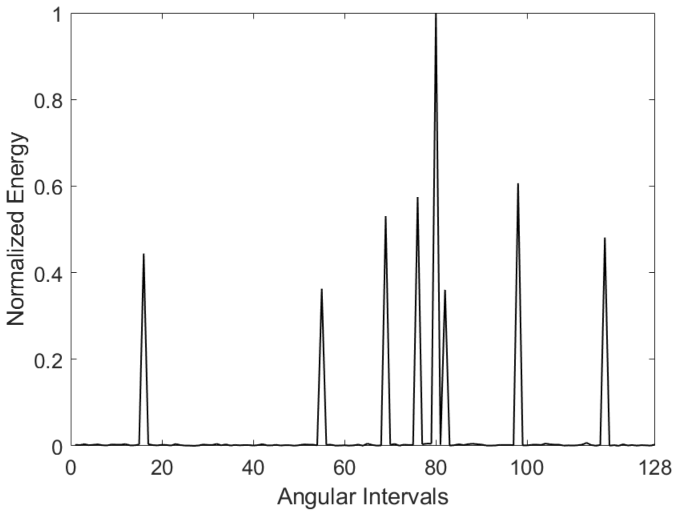

Fig. 4.3 Average of the normalized energy of elements of the vector $\mathbf{q}$, for all basis functions and all pilot subcarriers against the angular intervals. $\mathrm{SNR} = 20\mathrm{dB}$, $\mathrm{J} = 100$, $|\mathrm{I}| = 8$.

Based on the support estimate $\tilde{\mathrm{I}}$, the MMSE approach [104] is employed to estimate the expansion coefficients $\tilde{\mathbf{c}}_{\mathrm{p}}$:

$$\tilde{\mathbf{c}}_{\mathrm{p}} = \mathbf{R}_{\mathrm{pp}}(\mathbf{R}_{\mathrm{gg}} + \sigma^2|\mathbf{w}|)^{-1} \tag{4.12}$$

$$\mathbf{R}_{\mathrm{pp}} = \mathbf{F}'_{\mathrm{p}}\mathbf{r}_{\mathrm{p}} \tag{4.13}$$

Here, $\mathbf{w}$ is the identity matrix with size of $|\tilde{\mathrm{I}}| \times |\tilde{\mathrm{I}}|$, $\tilde{\mathbf{c}}_{\mathbf{p}} \in \mathrm{C}^{1\times \mathrm{MN_b}}$ is a vector containing $\tilde{\mathbf{c}}_{\mathrm{i,p}}$ for $\mathrm{N_b}$ basis functions, and $\mathbf{F}_{\mathrm{p}} \in \mathrm{C}^{\mathrm{J}\times \mathrm{MN_b}}$ is the matrix containing $\mathrm{N_b}$ vectors $\mathbf{F}_{\mathrm{i,p}}$, and σ^2 is the noise variance, which is assumed to be known.

4.5 Simulation Results

The MSE of the channel estimation will be used to asses the algorithm performance. First, we compute the MSE for the tth OFDM symbol at the pth subcarrier:

$$\mathrm{MSE}_{\mathrm{p}}^{\mathrm{t}} = \frac{||\hat{\mathbf{h}}_{\mathrm{p}}^{\mathrm{t}} - \tilde{\mathbf{h}}_{\mathrm{n(p)}}^{\mathrm{t}}||_2^2}{||\tilde{\mathbf{h}}_{\mathrm{n(p)}}^{\mathrm{t}}||_2^2}, \tag{4.14}$$

$$||\tilde{\mathbf{h}}^{\mathrm{t}}_{\mathrm{n(p)}}||_2^2 = (\tilde{\mathbf{h}}^{\mathrm{t}}_{\mathrm{n(p)}})^{\mathrm{H}}(\tilde{\mathbf{h}}^{\mathrm{t}}_{\mathrm{n(p)}}), \tag{4.15}$$

where $\hat{\mathbf{h}}^{\mathrm{t}}_{\mathrm{p}}$ is the estimated channel vector obtained from (4.2), and $\tilde{\mathbf{h}}^{\mathbf{t}}_{\mathbf{p}}$ is the true channel vector. Then the overall MSE for the channel estimation is computed by:

$$\mathrm{MSE} = \frac{1}{\mathrm{JP}}\sum_{\mathrm{p}=1}^{\mathrm{P}}\sum_{\mathrm{t}=1}^{\mathrm{J}}\mathrm{MSE}^{\mathrm{t}}_{\mathrm{p}}. \tag{4.16}$$

The MSE in (4.16) is further averaged over the simulation trials.

We consider simulation scenarios corresponding to a MIMO system with a uniform linear array. For massive MIMO systems, in most simulation scenarios, we consider the number of antennas $\mathrm{M} = 128$, the number of pilot subcarriers $\mathrm{P} = 64$, the sampling frequency $\mathrm{f_s} = 15.36$ MHz, the time interval for one OFDM symbol $\mathrm{T} = 66.7\ \mu\mathrm{s}$, the carrier frequency $\mathrm{f_c} = 2.5$ GHz and the number of simulation trials $\mathrm{N_s} = 500$. The performance of the oracle LS algorithm [27] with known support is adopted as the performance bound. In most scenarios, we consider two cases, $\mathrm{SNR} = 10$ dB and $\mathrm{SNR} = 20$ dB, and the user mobility with $\mathrm{v} = 120\frac{\mathrm{km}}{\mathrm{h}}$ and $\mathrm{v} = 300\frac{\mathrm{km}}{\mathrm{h}}$. The Doppler frequency $\mathrm{f_d}$ can be obtained by using:

$$\mathrm{f_d} = \mathrm{f_c}\frac{\mathrm{v}}{\mathrm{v_c}}, \tag{4.17}$$

where $\mathrm{v_c} = 3 \times 10^8$ m/s is the light speed.

The channel estimation performance is investigated in several ways. First, we investigate the MSE performance of the proposed algorithm with different number of employed OFDM symbols, then we compare the MSE performance against the number of basis functions, with normalized Doppler frequencies $\mathrm{f_dT} = 0.02$, $\mathrm{f_dT} = 0.05$, where the Doppler frequency is approximately $\mathrm{f_d} = 300$ Hz, $\mathrm{f_d} = 700$ Hz, respectively. After that, we compare the MSE performance for different SNR, considering the normalized Doppler frequency $\mathrm{f_dT} = 0.05$. The MSE performance against the number of DCD-iterations is investigated to show the convergence of the proposed algorithm, and the MSE performance against the number of employed antennas is also investigated. Furthermore, we compare the MSE performance of the DCD-JSR algorithm and the distributed sparsity adaptive matching pursuit (DSAMP) algorithm from [4] against the number of non-zero virtual angles, the oracle MMSE algorithm with known support is adopted as the performance bound [104]. At last, the computational complexity of the DCD-JSR algorithm is analyzed.

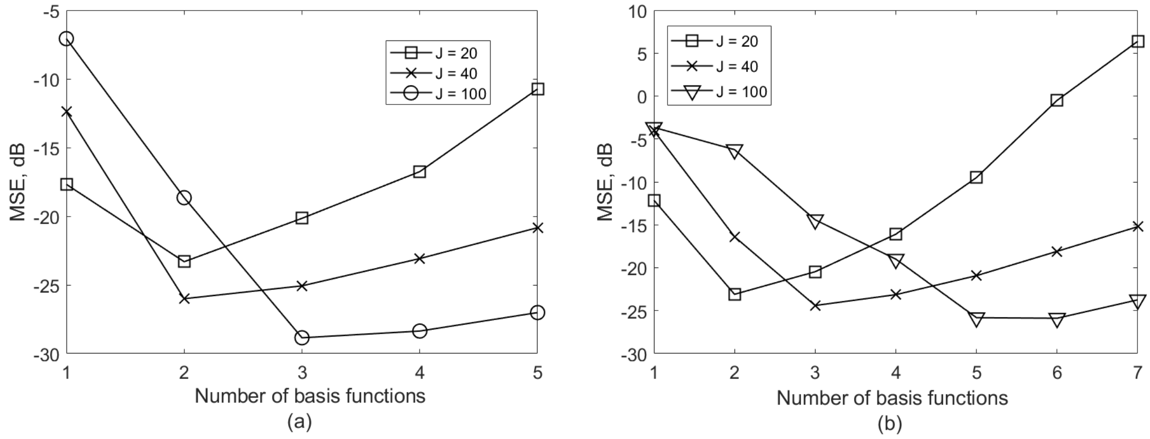

Fig. 4.4 MSE performance of the modified DCD-JSR algorithm against the number of employed basis functions; $\mathrm{SNR} = 20$ dB, $|\mathrm{I}| = 3$. (a) $\mathrm{f_d T} = 0.02$, (b) $\mathrm{f_d T} = 0.05$.

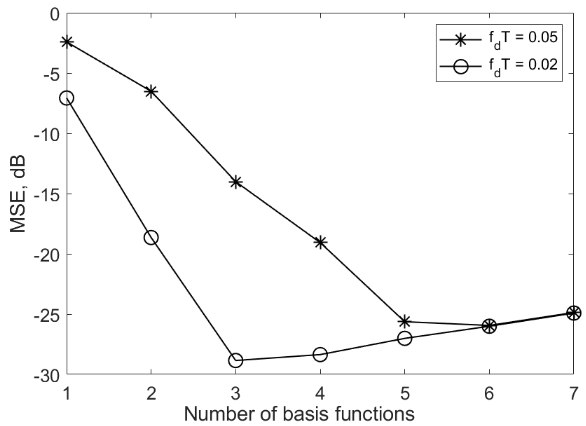

Fig. 4.5 MSE performance of the modified DCD-JSR algorithm against the number of employed basis functions; $\mathrm{SNR} = 20$ dB, $\mathrm{J} = 100$, $|\mathrm{I}| = 3$.

In Fig. 4.4, we show the MSE performance of the DCD-JSR algorithm for different number of employed OFDM symbols, for $\mathrm{SNR} = 20$ dB, $|\mathrm{I}| = 3$, $\mathrm{f_d T} = 0.02$ and $\mathrm{f_d T} = 0.05$. It can be seen that, as the number of OFDM symbols increases, the better MSE performance is provided by employing more basis functions. Thus we can conclude that, with longer data

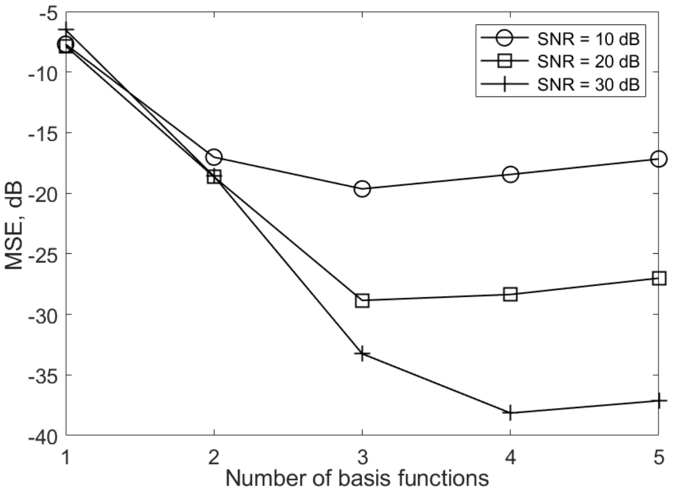

Fig. 4.6 MSE performance of the modified DCD-JSR algorithm against the number of employed basis functions; $\mathrm{f_d T} = 0.02$, $\mathrm{J} = 100$, $|\mathrm{I}| = 3$.

packets, more basis functions are required for the DCD-JSR algorithm to provide the best MSE performance, and the minimum MSE is also reduced. It can be seen that with increase in the number of basis function and OFDM symbols, better approximation accuracy can be achieved. This is explained by improving the model of time varying channel (channel with the Doppler effect) when using basis functions, compared to the static channel model.

In Fig. 4.5, we compare the MSE performance for different normalized Doppler frequencies, $\mathrm{f_d T} = 0.02$ and $\mathrm{f_d T} = 0.05$. The number of employed OFDM symbols is set to $\mathrm{J} = 100$, and the number of non-zero virtual angles $|\mathrm{I}| = 3$. It can be seen that, for the higher normalized Doppler frequency, we need to employ more basis functions to obtain the best MSE performance. In other words, to provide the best MSE performance, the number of basis functions to be employed increases with the normalized Doppler frequency, which means that with higher user mobility, the more basis functions is required to provide accurate channel estimation.

In Fig. 4.6, we investigate the number of basis functions required to provide the best MSE performance under different SNR scenarios, for the case $\mathrm{f_d T} = 0.02$, $\mathrm{J} = 100$, $|\mathrm{I}| = 3$. It can be seen that, as the SNR increases, the number of basis functions required to provide the best MSE performance is increased. This is because when the SNR is low, the main issue for channel estimation is the noise, a small number of basis functions is required to approximate

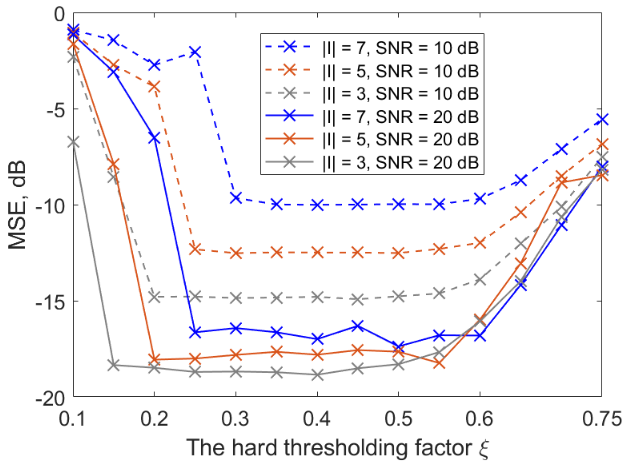

Fig. 4.7 MSE performance of the modified DCD-JSR algorithm against the hard thresholding factor ξ; $\mathrm{f_d T} = 0.05$, $\mathrm{N_b} = 3$, $\mathrm{J} = 40$.

the channel. When the SNR is high, the algorithm can focus more on the time variation of the channel, thus the number of basis functions required will be larger. Hence, we can say that, the number of basis functions required to approximate the time-varying channel should be higher for higher SNR.

To provide the best MSE performance, the thresholding factor ξ needs to be properly adjusted. In Fig. 4.7, we investigate the MSE performance of the DCD-JSR algorithm against the hard thresholding factor ξ, for the case $\mathrm{N_b} = 3$, $\mathrm{f_d T} = 0.05$, $\mathrm{J} = 40$. It is clear that, for both cases $\mathrm{SNR} = 10$ dB and $\mathrm{SNR} = 20$ dB, as the number of non-zero virtual angles $|\mathrm{I}|$ increases, the range for the thresholding factor ξ which can provide the best MSE performance decreases. However, in all the cases, the thresholding factor can be chosen in the interval $[0.30, 0.55]$ to provide the minimum MSE.

Fig. 4.8 shows the MSE performance of the DCD-JSR algorithm in scenarios with different number of non-zero virtual angles against the number of DCD iterations, for the case $\mathrm{N_b} = 2$, $\mathrm{f_d T} = 0.02$, $\mathrm{J} = 40$. It can be seen that, in all these scenarios, after a few DCD iterations, the algorithm converges to the best MSE. However, the smaller number of non-zero angles, the faster is the convergence. For $|\mathrm{I}| \leq 9$, a single DCD iteration is enough for the convergence.

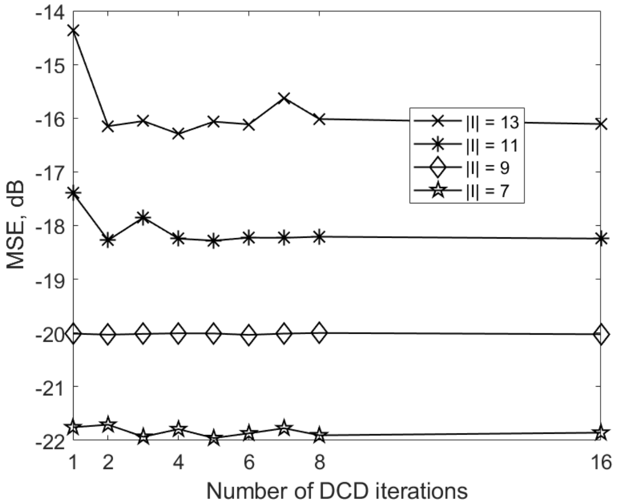

Fig. 4.8 MSE performance of the modified DCD-JSR algorithm against the number of DCD iterations; $\mathrm{f_d T} = 0.02$, $\mathrm{N_b} = 2$, $\mathrm{J} = 40$.

In Fig. 4.9(a), we show the MSE performance for different number of employed antennas, for the case $\mathrm{N_b} = 3$, $\mathrm{f_d T} = 0.05$, $\mathrm{J} = 20$, $|\mathrm{I}| = 4$. It can be seen that, with a small number of antennas, the MSE performance of the DCD-JSR algorithm is poor. For $\mathrm{SNR} = 10$ dB, it requires $\mathrm{M} = 56$ to approach the oracle performance, for $\mathrm{SNR} = 20\mathrm{dB}$, it requires at least $\mathrm{M} = 32$. In Fig. 4.9 (b), we show the probability of perfect support estimation against the number of employed antennas, where a perfect support estimation means the estimated support is exactly the same as the true support, for the case $|\mathrm{I}| = 4$, $\mathrm{f_d T} = 0.05$, $\mathrm{J} = 20$. It can be seen that, at $\mathrm{SNR} = 10$ dB, with a small number of employed antennas, we cannot estimate the support correctly until $\mathrm{M} = 64$; this explains why the MSE performance is poor with small number of antennas. We have run our simulations up to $\mathrm{M} = 512$ and observed that the MSE performance does not change.

In Fig. 4.10(a) and Fig. 4.10(b), for the DCD-JSR algorithm with $\mathrm{N_b} = 1$ and $\mathrm{N_b} = 2$, and the distributed sparsity adaptive matching pursuit algorithm (DSAMP) [4], we show the MSE performance for different number of non-zero virtual angles, and the probability of perfect support estimation, respectively; here, the DCD-JSR algorithm with $\mathrm{N_b} = 1$ corresponds to the version of the DCD-JSR algorithm previously proposed in Chapter 3 for time-invariant channels. For simulation scenario, we consider the normalized Doppler

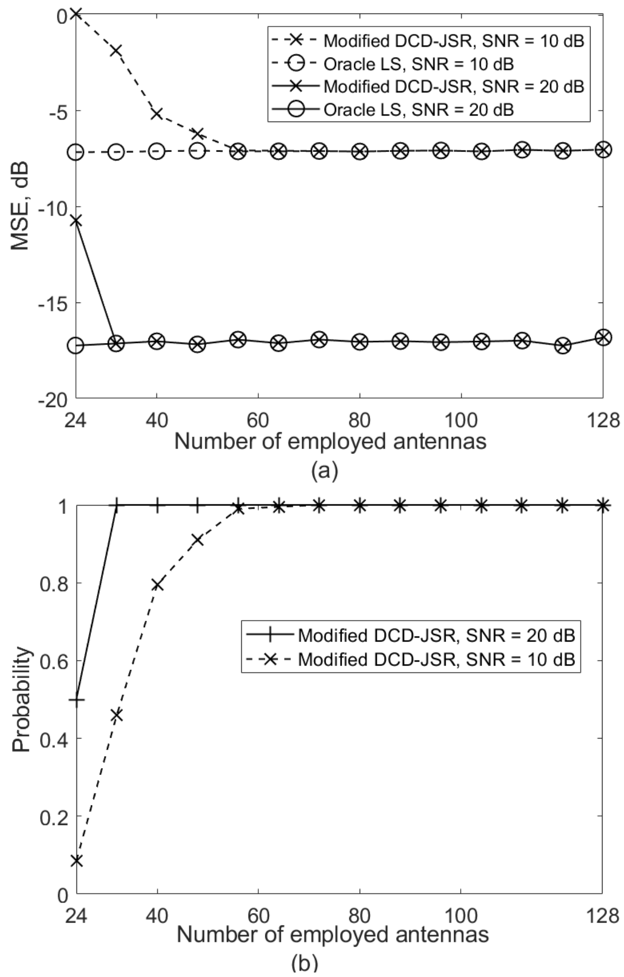

Fig. 4.9 (a) MSE performance of the modified DCD-JSR algorithm against the number of employed antennas; (b) Probability of perfect support estimation against the number of employed antennas. $\mathrm{f_d T} = 0.05$, $\mathrm{N_b} = 3$, $\mathrm{J} = 20$, $|\mathrm{I}| = 4$.

frequency $\mathrm{f_d T} = 0.02$, $\mathrm{J} = 40$, $\mathrm{SNR} = 20$ dB. It can be seen that, in Fig. 4.10(a), when $\mathrm{f_d T} = 0.02$, since the time variation of the channel is slow, we can estimate the channel quite well using only one basis function in the DCD-JSR algorithm or using the DSAMP algorithm, while both of them shows the MSE performance close to the oracle performance. The DCD-JSR algorithm with $\mathrm{N_b} = 2$ also shows close to the oralce MSE performance. In Fig. 4.10(b), it is seen that for the DCD-JSR algorithm with $\mathrm{N_b} = 2$, the support estimation is slightly better than that for the DCD-JSR algorithm with $\mathrm{N_b} = 1$ and DSAMP algorithm, which explains why the DCD-JSR algorithm with $\mathrm{N_b} = 2$ can provide a better MSE performance in this case.

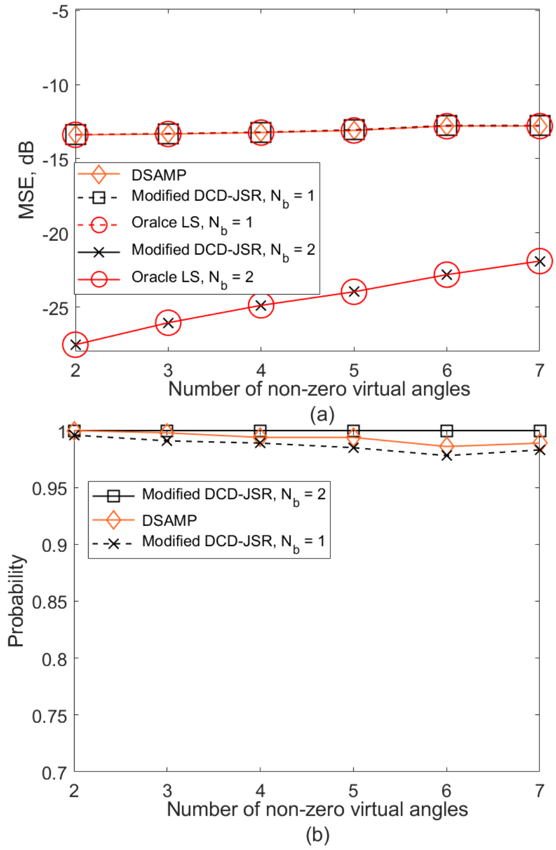

Fig. 4.10 (a) MSE performance against the number of non-zero virtual angles; (b) Probability of perfect support estimation against the number of non-zero virtual angles. $\mathrm{f_d T} = 0.02$, $\mathrm{J} = 40$, $\mathrm{SNR} = 20$ dB.

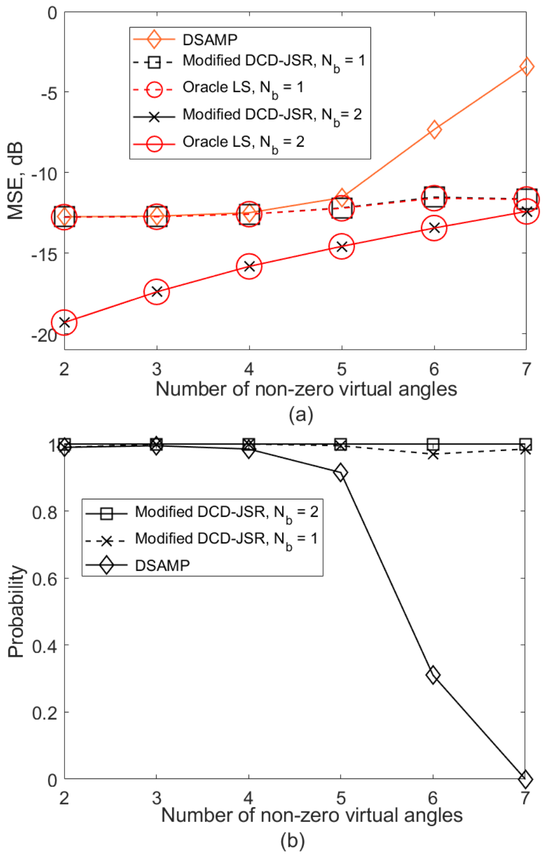

Fig. 4.11 (a) MSE performance against the number of non-zero virtual angles; (b) Probability of perfect support estimation against the number of non-zero virtual angles. $\mathrm{f_d T} = 0.02$, $\mathrm{J} = 40$, $\mathrm{SNR} = 10$ dB.

In Fig. 4.11, a lower SNR is considered compared to Fig. 4.10, $\mathrm{SNR} = 10$ dB. It can be seen that, when the noise level is higher, the MSE performance provided by the DCD-JSR algorithm with $\mathrm{N_b} = 2$ still shows close to the oracle performance and provides the perfect support estimation, whereas the DCD-JSR algorithm with a single basis function $(\mathrm{N_b} = 1)$ and the DSAMP algorithm, both developed for time-invariant channels, show inferior MSE performance and support estimation, although the DCD-JSR algorithm with $\mathrm{N_b} = 1$ still outperforms the DSAMP algorithm.

In Fig. 4.12(a) and Fig. 4.12(b), for the DCD-JSR algorithm, and the DSAMP algorithm, we show the MSE performance for different number of non-zero virtual angles, and the probability of perfect support estimation, respectively. For simulation scenario, we consider the normalized Doppler frequency $\mathrm{f_d T} = 0.05$, $\mathrm{J} = 40$, and $\mathrm{SNR} = 20$ dB. It can be seen

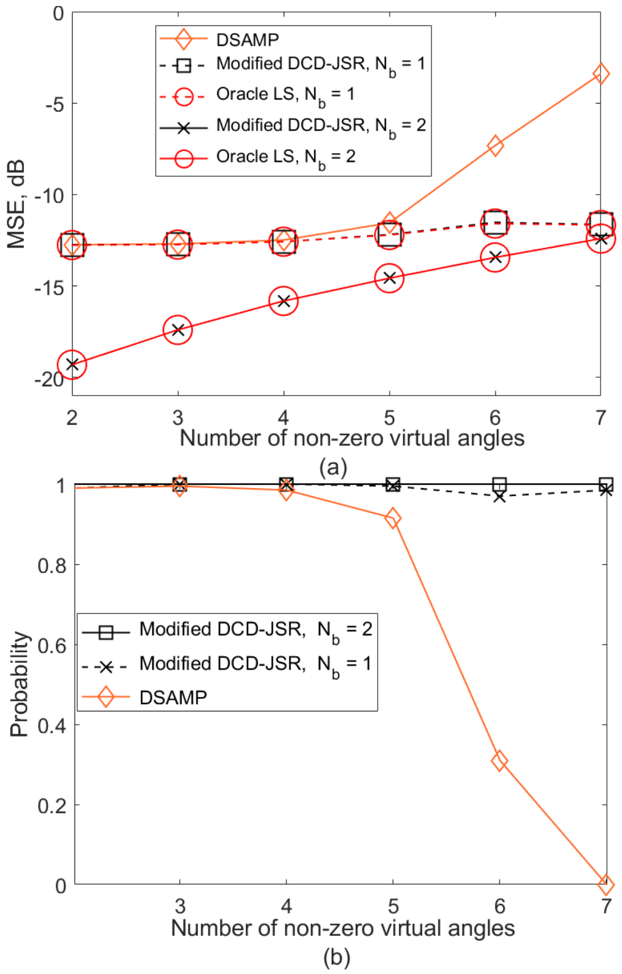

Fig. 4.12 (a) MSE performance against the number of non-zero virtual angles; (b) Probability of perfect support estimation against the number of non-zero virtual angles. $\mathrm{f_d T} = 0.05$, $\mathrm{J} = 40$, $\mathrm{SNR} = 20$ dB.

in Fig. 4.12(a) that the DCD-JSR algorithm shows close to the oracle MSE performance, while the DSAMP algorithm shows a poor performance. In Fig. 4.12(b), it is seen that the DCD-JSR algorithm with $\mathrm{N_b} = 3$ always provides the perfect support estimation, while the DCD-JSR algorithm with only one basis function $\mathrm{N_b} = 1$ shows inferior performance, and the DSAMP algorithm cannot estimate the support accurately. This is because the DSAMP algorithm is developed for static channel, and when $\mathrm{f_d T} = 0.05$, i.e. the time variation of the channel is fast, the algorithm is incapable of providing a high estimation performance.

In Fig. 4.13, results are shown for a higher noise level compared to Fig. 4.12, we set $\mathrm{SNR} = 10$ dB. It can be seen in Fig. 4.13(a), that the DSAMP algorithm has again a poor MSE performance, while the DCD-JSR algorithm with $\mathrm{N_b} = 1$ and $\mathrm{N_b} = 3$ shows close to the oracle performance. In Fig. 4.13(b), it is clear that the DSAMP algorithm cannot provide an accurate support estimation in this case. The probability of perfect support estimation provided by the DCD-JSR algorithm with $\mathrm{N_b} = 1$ decreses as the number of non-zero virtual angels increases, while the DCD-JSR algorithm with $\mathrm{N_b} = 3$ can always provides the perfect support estimation. This is because as the time variation of the channel becomes faster, more basis functions is required to accurately approximate the channel.

Hence, from Fig. 4.10 to Fig. 4.13, we can conclude that the DCD-JSR algorithm outperforms the DSAMP algorithm. The improvement in the performance provided by the DCD-JSR algorithm against the DSAMP algorithm is more significant in time-varying channels. For faster time varying channels, by employing more basis functions,we can significantly improve the performance of the DCD-JSR algorithm compared to the case $\mathrm{N_b} = 1$ previously developed in Chapter 3 for static channels.

Fig. 4.14 shows the computational complexity of the DCD-JSR algorithm and DSAMP algorithm against the number of non-zero vitual angles, obtained for the case $\mathrm{f_d T} = 0.02$, $\mathrm{J} = 20$, $\mathrm{SNR} = 20$ dB. It can be seen that the DCD-JSR algorithm has significantly lower computational complexity than the DSAMP algorithm. The DCD-JSR algorithm, when $\mathrm{N_b} = 2$, has slightly higher computational complexity than the DCD-JSR algorithm with $\mathrm{N_b} = 1$, while the increase in the number of basis functions provides a significantly better MSE performance.

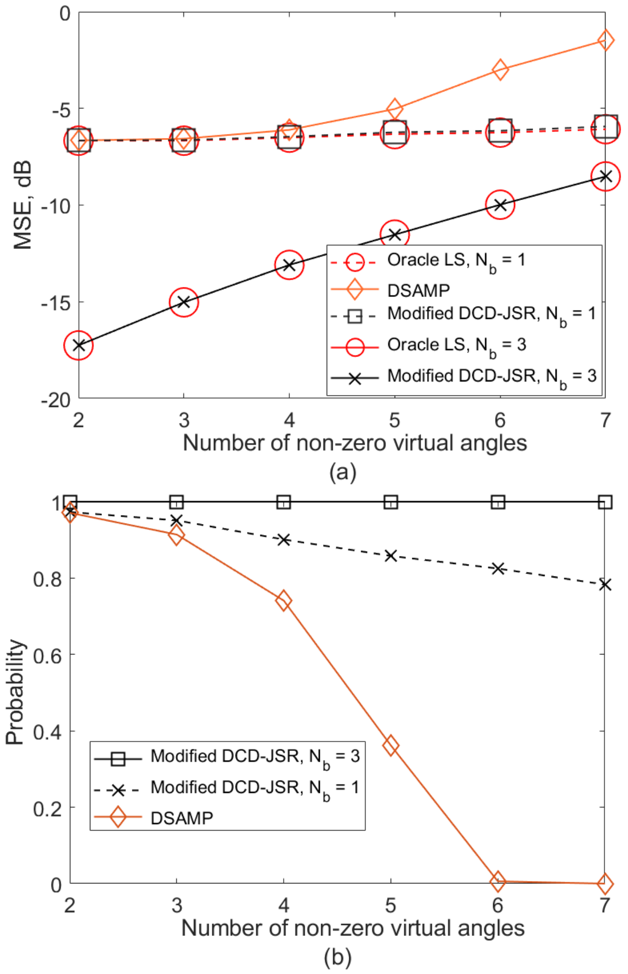

Fig. 4.13 (a) MSE performance against the number of non-zero virtual angles; (b) Probability of perfect support estimation against the number of non-zero virtual angles. $\mathrm{f_d T} = 0.05$, $\mathrm{J} = 40$, $\mathrm{SNR} = 10$ dB.

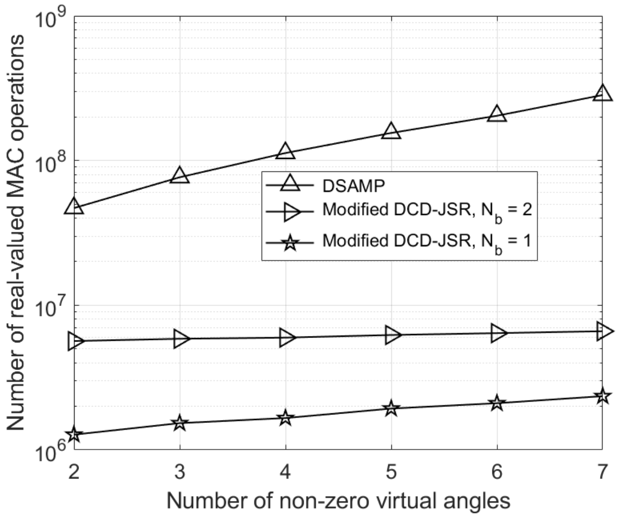

Fig. 4.14 Computational complexity of the modified DCD-JSR algorithm and the DSAMP algorithm; $\mathrm{f_d T} = 0.02$, $\mathrm{J} = 20$, $\mathrm{SNR} = 20$ dB.

4.6 Conclusion

In this chapter, by combining the BEM approach and the DCD-JSR algorithm, an efficient algorithm for estimation of the fast time-varying channels in virtual angular domain for massive MIMO systems is proposed. Simulation results have shown that compared to the previously proposed algorithms for channel estimation in virtual angular domain designed for time-invariant channels, the proposed DCD-JSR algorithm could provide significantly better MSE performance in time-varying channels.

Employing basis functions for channel estimation can provide better channel estimation performance for following reasons:

Firstly, basis functions can be chosen to capture the characteristics of the channel, including its time-varying effects due to the Doppler effect. This can lead to a more accurate representation of the channel and can improve the quality of the channel estimate.

Secondly, basis functions can be used to represent the channel as a linear combination of coefficients, which can be estimated using linear methods such as least squares or maximum likelihood estimation. This can be advantages in the presence of the Doppler effect because it can reduce the computational complexity.

Thirdly, basis functions can be adapted to match the statistics of the channel, which is important for accurately representing the channel variation over time. This can be achieved by adjusting the order of the basis functions or by using a weighted combination of different orders of basis functions. This can improve the accuracy of the channel estimate and can lead to better performance in terms of signal quality and system throughput.

Finally, by using basis functions that are well-suited for modeling time-varying signals, such as Legendre polynomials, it is possible to capture the rapid variations due to the Doppler effect and improve the accuracy of the channel estimate.

Chapter 5

Conclusion and Future work

With the rapid development of wireless communication technology, based on the MIMO-OFDM technology of 4G wireless communication system, massive MIMO technology has gradually become one of the most valuable enabling technologies in 5G communication system. The main principle of this technology is that by deploying a large-scale antenna array at the base station, combined with spatial multiplexing and beamforming technology, it can provide high-performance communication services for many user terminals in the cell on the same time-frequency resources. This technology can significantly improve the capacity and reliability of wireless communication systems. However, the performance of this technology is related to the performance of channel estimation in communication systems, so channel estimation is one of the core technologies of wireless communication systems. This work takes the massive MIMO system in a complex environment as the research background, uses compressive sensing and other methods, discusses and studies in the field of sparse channel estimation, and has achieved some scientific research results with certain value and application prospects. The main content and research results of this thesis are as follows:

1. A CS channel estimation algorithm for massive MIMO systems with Orthogonal Frequency Division Multiplexing (OFDM) is proposed. By exploiting the spatially common sparsity in the virtual angular domain of the massive MIMO channels, a dichotomous-coordinate-decent-joint-sparse-recovery (DCD-JSR) algorithm is proposed. More specifically, by considering the channel is static over several OFDM symbols and exhibits common

sparsity in the virtual angular domain, the DCD-JSR algorithm can jointly estimate multiple sparse channels with low computational complexity. The simulation results have shown that, compared to existing channel estimation algorithms such as the distributed-sparsity-adaptive-matching-pursuit (DSAMP) algorithm, the proposed DCD-JSR algorithm has significantly lower computational complexity and better performance.

2. These results have been extended to the case of multipath channels with time-varying parameters. This has been achieved by employing the basis expansion model to approximate the time variation of the channel, thus the modified DCD-JSR algorithm can estimate the channel in a massive MIMO OFDM system operating over frequency selective and highly mobile wireless channels. Simulation results have shown that, compared to the DCD-JSR algorithm designed for time-invariant channels, the modified DCD-JSR algorithm provides significantly better estimation performance in fast time-varying channels.

However, with the in-depth research work, the following problems need to be further improved and solved in future exploration and research:

1. Since in practical scenarios, user's mobility might not be static, which would cause the variation of the Doppler frequency, thus the normalized Doppler frequency $\mathrm{f_d T}$ would change with time, the channel estimation techniques proposed in this thesis might not be accurate, since they were designed for the static Doppler frequency. Therefore, it is necessary to propose channel estimation algorithms that can deal with the varying Doppler frequency.

2. In this work, we focus on estimating the channel with specified number of OFDM symbols, however, when the channel sparsity is far smaller than the number of OFDM symbols (which is the length of the measurement matrix), we do not need to employ such large number of OFDM symbols to send pilot symbols. Therefore, to improve the estimation and transmission efficiency, it is of great value to propose such channel estimation that can adaptively change the number of employed pilot symbols for channel estimation.

Acronyms

3D Three dimensions

4G 4th Generation mobile communication system

5G 5th Generation mobile communication system

AOMP Adaptive Orthogonal Matching Pursuit

As-SAMP Adaptive Step Size Sparsity Adaptive Matching Pursuit

BEM Basis Expansion Method

BS Base Station

Block StOMP Block Stagewise Orthogonal Matching Pursuit

CSI Channel State Information

CS Compressive Sensing

CoSAMP Compressive Sampling Matching Pursuit

D2D Device to Device

DCD Dichotomous Coordinate Decent

DFT Discrete Fourier Transform

DSAMP Distributed Sparsity Adaptive Matching Pursuit

FDD Frequency Division Duplex

FD Full Duplex

FFT Fast Fourier Transform

IFFT Inverse Fast Fourier Transform

JSR Joint Sparse Recovery

LS Least Squares

LTE Long Term Evolution

MIMO Multiple Input Multiple Output

MMSE Minimum Mean Square Error

MSE Mean Square Error

NOMA Non-Orthogonal Multiple Access

OFDM Orthogonal Frequency Division Multiplexing

OMP Orthogonal Matching Pursuit

RIP Restricted Isometry Pmperty

RLS Recursive Least Squares

SAMP Sparsity Adaptive Matching Pursuit

SMV Single Measurement Vector

SNR Signal to Noise Ratio

SOS Sum-Of-Sinusoid

SP Sampling Pursuit

TDD Time Division Duplex

VBCS Variational Bayesian compressive sensing

VL Visible Light

References

[1] P. Meerasri, P. Uthansakul, and M. Uthansakul, “Self-interference cancellation-based mutual-coupling model for full-duplex single-channel mimo systems,” *International Journal of Antennas and Propagation*, vol. 2014, pp. 1–10, 01 2014.

[2] S. K. Khadagade and N. Mittal, “Comparison of ber of ofdm system using qpsk and 16qam over multipath rayleigh fading channel using pilot-based channel estimation,” *International Journal of Engineering and Advanced Technology (IJEAT)*, pp. 2249–8958, 2013.

[3] Z. Gao, L. Dai, Z. Lu, C. Yuen, and Z. Wang, “Super-resolution sparse MIMO-OFDM channel estimation based on spatial and temporal correlations,” *IEEE Communications Letters*, vol. 18, no. 7, pp. 1266–1269, 2014.

[4] Z. Gao, L. Dai, Z. Wang, and S. Chen, “Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO,” *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6169–6183, 2015.

[5] M. Liao and Y. Zakharov, “DCD-based joint sparse channel estimation for OFDM in virtual angular domain,” *IEEE Access*, vol. 9, pp. 102 081–102 090, 2021.

[6] M. Liao and Y. Zakharov, “Estimation of time-varying channels in virtual angular domain for massive MIMO systems,” *IEEE Access*, vol. 11, pp. 1923–1933, 2023.

[7] S. Morsalin, K. Mahmud, and G. E. Town, “Scalability of vehicular M2M communications in a 4G cellular network,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 10, pp. 3113–3120, 2017.

[8] H. Liu, S. Hu, W. Zheng, Z. Xie, S. Wang, P. Hui, and T. Abdelzaher, “Efficient 3G budget utilization in mobile participatory sensing applications,” in *2013 Proceedings IEEE International Conference on Computer Communications*, pp. 1411–1419.

[9] S. A. Hoseinitabatabei, A. Mohamed, M. Hassanpour, and R. Tafazolli, “The power of mobility prediction in reducing idle-state signaling in cellular systems: A revisit to 4G mobility management,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3346–3360, 2020.

[10] D. Amzallag, R. Bar-Yehuda, D. Raz, and G. Scalosub, “Cell selection in 4G cellular networks,” *IEEE Transactions on Mobile Computing*, vol. 12, no. 7, pp. 1443–1455, 2012.

[11] I.-K. Fu, Y.-S. Chen, P. Cheng, Y. Yuk, R. Y. Kim, and J. S. Kwak, “Multicarrier technology for 4G wimax system WiMAX/LTE update,” *IEEE Communications Magazine*, vol. 48, no. 8, pp. 50–58, 2010.

[12] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, “5G: A tutorial overview of standards, trials, challenges, deployment, and practice,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, 2017.

[13] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, “What will 5G be?” *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.

[14] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, “What should 6G be?” *Nature Electronics*, vol. 3, no. 1, pp. 20–29, 2020.

[15] X. You, C.-X. Wang, J. Huang, X. Gao, Z. Zhang, M. Wang, Y. Huang, C. Zhang, Y. Jiang, J. Wang *et al.*, “Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts,” *Science China Information Sciences*, vol. 64, pp. 1–74, 2021.

[16] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, “Massive MIMO in real propagation environments: Do all antennas contribute equally?” *IEEE Transactions on Communications*, vol. 63, no. 11, pp. 3917–3928, 2015.

[17] H. A. Suraweera, I. Krikidis, G. Zheng, C. Yuen, and P. J. Smith, “Low-complexity end-to-end performance optimization in MIMO full-duplex relay systems,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 913–927, 2014.

[18] D. Wang, J. Wang, X. You, Y. Wang, M. Chen, and X. Hou, “Spectral efficiency of distributed MIMO systems,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 10, pp. 2112–2127, 2013.

[19] Q. Li, G. Li, W. Lee, M.-i. Lee, D. Mazzarese, B. Clerckx, and Z. Li, “MIMO techniques in WiMAX and LTE: a feature overview,” *IEEE Communications Magazine*, vol. 48, no. 5, pp. 86–92, 2010.

[20] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, “Multi-cell MIMO cooperative networks: A new look at interference,” *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1380–1408, 2010.

[21] A. Osseiran, V. Braun, T. Hidekazu, P. Marsch, H. Schotten, H. Tullberg, M. A. Uusitalo, and M. Schellman, “The foundation of the mobile and wireless communications

system for 2020 and beyond: Challenges, enablers and technology solutions," in *2013 IEEE 77th Vehicular Technology Conference*, pp. 1–5.

[22] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *arXiv preprint arXiv:1201.3210*, 2012.

[23] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, 2014.

[24] S. R. Islam, N. Avazov, O. A. Dobre, and K.-S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Communications Surveys and Tutorials*, vol. 19, no. 2, pp. 721–742, 2016.

[25] I. Hwang, B. Song, and S. S. Soliman, "A holistic view on hyper-dense heterogeneous and small cell networks," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 20–27, 2013.

[26] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 86–92, 2014.

[27] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, 2016.

[28] X. Liu, P. Tian, Z. Wei, S. Yi, Y. Huang, X. Zhou, Z.-J. Qiu, L. Hu, Z. Fang, C. Cong *et al.*, "Gbps long-distance real-time visible light communications using a high-bandwidth GaN-based micro-LED," *IEEE Photonics Journal*, vol. 9, no. 6, pp. 1–9, 2017.

[29] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, 2014.

[30] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[31] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An Overview Of Massive MIMO: Benefits And Challenges," *IEEE Journal of Selected Topics in Signal Processing.*, vol. 8, no. 5, pp. 742–758, 2014.

[32] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[33] K. Zheng, L. Zhao, J. Mei, B. Shao, W. Xiang, and L. Hanzo, “Survey of large-scale MIMO systems,” *IEEE Communications Surveys and Tutorials*, vol. 17, no. 3, pp. 1738–1760, 2015.

[34] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, “Pilot contamination and precoding in multi-cell TDD systems,” *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, 2011.

[35] Fang, Jun and Li, Xingjian and Li, Hongbin and Gao, Feifei, “Low-rank covariance-assisted downlink training and channel estimation for FDD massive MIMO systems,” *IEEE Wireless Communications*, vol. 16, no. 3, pp. 1935–1947, 2017.

[36] Q. Ling and T. Li, “Blind-channel estimation for MIMO systems with structured transmit delay scheme,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 55, no. 8, pp. 2344–2355, 2008.

[37] C. Shin, R. W. Heath, and E. J. Powers, “Blind channel estimation for MIMO-OFDM systems,” *IEEE Transactions on Vehicular Technology*, vol. 56, no. 2, pp. 670–685, 2007.

[38] S.-H. Fang, J.-Y. Chen, M.-D. Shieh, and J.-S. Lin, “Subspace-based blind channel estimation by separating real and imaginary symbols for cyclic-prefixed single-carrier systems,” *IEEE Transactions on Broadcasting*, vol. 59, no. 4, pp. 698–704, 2013.

[39] M. Hajjaj, W. Chainbi, and R. Bouallegue, “Low-rank channel estimation for MIMO MB-OFDM UWB system over spatially correlated channel,” *IEEE Wireless Communications Letters*, vol. 5, no. 1, pp. 48–51, 2015.

[40] P. Muneer and S. Sameer, “Pilot-aided joint estimation of doubly selective channel and carrier frequency offsets in OFDMA uplink with high-mobility users,” *IEEE Transactions on Vehicular Technology*, vol. 64, no. 1, pp. 411–417, 2014.

[41] M. K. Ozdemir and H. Arslan, “Channel estimation for wireless OFDM systems,” *IEEE Communications Surveys and Tutorials*, vol. 9, no. 2, pp. 18–48, 2007.

[42] J. N. Bae and J. Y. Kim, “Performance of multi-user MIMO OFDM channel estimation with LS and MMSE for 802.11 n systems,” in *2009 9th International Symposium on Communications and Information Technology*. IEEE, 2009, pp. 1–5.

[43] L. Dai, Z. Wang, and Z. Yang, “Spectrally efficient time-frequency training OFDM for mobile large-scale MIMO systems,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 251–263, 2013.

[44] W. Ding, F. Yang, W. Dai, and J. Song, “Time–frequency joint sparse channel estimation for MIMO-OFDM systems,” *IEEE Communications Letters*, vol. 19, no. 1, pp. 58–61, 2014.

[45] X. Ma, F. Yang, and J. Song, "Time-angle domain sparsity-based MIMO channel estimation approach in high mobility scenarios," *IEEE Communications Letters*, vol. 22, no. 9, pp. 1946–1949, 2018.

[46] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, 2010.

[47] Y. Peng, X. Yang, X. Zhang, W. Wang, and B. Wu, "Compressed MIMO-OFDM channel estimation," in *2010 IEEE 12th International Conference on Communication Technology*, pp. 1291–1294.

[48] D. J. R. Chisaguano, Y. Hou, T. Higashino, and M. Okada, "Low-complexity channel estimation and detection for MIMO-OFDM receiver with espar antenna," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 8297–8308, 2015.

[49] G. Gui, Q. Wan, W. Peng, and F. Adachi, "Sparse multipath channel estimation using compressive sampling matching pursuit algorithm," *arXiv preprint arXiv:1005.2270*, 2010.

[50] Y. Zhang, R. Venkatesan, O. A. Dobre, and C. Li, "Novel compressed sensing-based channel estimation algorithm and near-optimal pilot placement scheme," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2590–2603, 2015.

[51] D. Lee, "Mimo ofdm channel estimation via block stagewise orthogonal matching pursuit," *IEEE Communications Letters*, vol. 20, no. 10, pp. 2115–2118, 2016.

[52] K. Muralidhar and D. Sreedhar, "Pilot design for vector state-scalar observation Kalman channel estimators in doubly-selective MIMO-OFDM systems," *IEEE Wireless Communications Letters*, vol. 2, no. 2, pp. 147–150, 2012.

[53] A. N. Uwaechia and N. M. Mahyuddin, "Spectrum-efficient distributed compressed sensing based channel estimation for OFDM systems over doubly selective channels," *IEEE Access*, vol. 7, pp. 35 072–35 088, 2019.

[54] N. D. Lahbib, M. Cherif, M. Hizem, and R. Bouallegue, "Channel estimation for TDD uplink massive MIMO systems via compressed sensing," in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*. IEEE, pp. 1680–1684.

[55] X. Rao and V. K. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3261–3271, 2014.

[56] J. W. Choi, B. Shim, and S.-H. Chang, "Downlink pilot reduction for massive MIMO systems via compressed sensing," *IEEE Communications Letters*, vol. 19, no. 11, pp. 1889–1892, 2015.

[57] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[58] H. Zhao, Y. Tan, G. Pan, Y. Chen, and N. Yang, "Secrecy outage on transmit antenna selection/maximal ratio combining in MIMO cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 10 236–10 242, 2016.

[59] K. Zhang, C. Xiong, B. Chen, J. Wang, and J. Wei, "A maximum likelihood combining algorithm for spatial multiplexing MIMO amplify-and-forward relaying systems," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 12, pp. 5767–5774, 2015.

[60] S. Sun, T. S. Rappaport, R. W. Heath, A. Nix, and S. Rangan, "MIMO for millimeter-wave wireless communications: Beamforming, spatial multiplexing, or both?" *IEEE Communications Magazine*, vol. 52, no. 12, pp. 110–121, 2014.

[61] M. Sánchez-Fernández, S. Zazo, and R. Valenzuela, "Performance comparison between beamforming and spatial multiplexing for the downlink in wireless cellular systems," *IEEE transactions on wireless communications*, vol. 6, no. 7, pp. 2427–2431, 2007.

[62] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer in multiuser OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 4, pp. 2282–2294, 2014.

[63] H. Wu, J. Li, B. Dai, and Y. Liu, "Analysis of the impact of AGC on cyclic prefix length for OFDM systems," *IEEE Transactions on Communications*, vol. 66, no. 10, pp. 4783–4794, 2018.

[64] M. F. Pop and N. C. Beaulieu, "Limitations of sum-of-sinusoids fading channel simulators," *IEEE Transactions on Communications*, vol. 49, no. 4, pp. 699–708, 2001.

[65] P. Bello, "Characterization of randomly time-variant linear channels," *IEEE Transactions on Communications Systems*, vol. 11, no. 4, pp. 360–393, 1963.

[66] Gilbert, EN, "Energy reception for mobile radio," *Bell System Technical Journal*, vol. 44, no. 8, pp. 1779–1803, 1965.

[67] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell System Technical Journal*, vol. 47, no. 6, pp. 957–1000, 1968.

[68] W. C. Jakes and D. C. Cox, *Microwave mobile communications*. Wiley-IEEE press, 1994.

[69] P. Dent, G. E. Bottomley, and T. Croft, "Jakes fading model revisited," *Electronics Letters*, vol. 13, no. 29, pp. 1162–1163, 1993.

[70] C.-Y. Chen and W.-R. Wu, “Joint aod, aoa, and channel estimation for MIMO-OFDM systems,” *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 5806–5820, 2018.

[71] R. Baraniuk, M. A. Davenport, M. F. Duarte, and C. Hegde, “An introduction to compressive sensing,” 2014.

[72] E. J. Candès and M. B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[73] E. J. Candes, “The restricted isometry property and its implications for compressed sensing,” *Comptes Rendus Mathematique*, vol. 346, no. 9-10, pp. 589–592, 2008.

[74] Y. V. Zakharov, V. H. Nascimento, R. C. De Lamare, and F. G. D. A. Neto, “Low-complexity DCD-based sparse recovery algorithms,” *IEEE Access*, vol. 5, pp. 12 737–12 750, 2017.

[75] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok, “Introduction to compressed sensing.” 2012.

[76] S. Ji, Y. Xue, and L. Carin, “Bayesian compressive sensing,” *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2346–2356, 2008.

[77] Z. Yang, L. Xie, and C. Zhang, “Variational Bayesian algorithm for quantized compressed sensing,” *IEEE Transactions on Signal Processing*, vol. 61, no. 11, pp. 2815–2824, 2013.

[78] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, “Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition,” in *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*. IEEE, 1993, pp. 40–44.

[79] J. Wang, S. Kwon, and B. Shim, “Generalized orthogonal matching pursuit,” *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6202–6216, 2012.

[80] D. Needell and J. A. Tropp, “Cosamp: Iterative signal recovery from incomplete and inaccurate samples,” *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[81] W. Dai and O. Milenkovic, “Subspace pursuit for compressive sensing signal reconstruction,” *IEEE Transactions on Information Theory*, vol. 55, no. 5, pp. 2230–2249, 2009.

[82] T. T. Do, L. Gan, N. Nguyen, and T. D. Tran, “Sparsity adaptive matching pursuit algorithm for practical compressed sensing,” in *42nd Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA*. IEEE, 2008, pp. 581–587.

[83] Y. Zakharov and T. Tozer, "Multiplication-free iterative algorithm for LS problem," *Electron. Lett.*, vol. 40, no. 9, p. 1, 2004.

[84] Y. Xu, G. Yue, and S. Mao, "User grouping for massive MIMO in FDD systems: New design methods and analysis," *IEEE Access*, vol. 2, pp. 947–959, 2014.

[85] Y. Liu, Z. Tan, H. Hu, L. J. Cimini, and G. Y. Li, "Channel estimation for OFDM," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 4, pp. 1891–1908, 2014.

[86] D. Angelosante, E. Biglieri, and M. Lops, "Sequential estimation of multipath MIMO-OFDM channels," *IEEE Transaction on Signal Processing*, vol. 57, no. 8, pp. 3167–3181, 2009.

[87] M. Simko, P. S. Diniz, Q. Wang, and M. Rupp, "Adaptive pilot-symbol patterns for MIMO OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4705–4715, 2013.

[88] Z. Gao, L. Dai, W. Dai, B. Shim, and Z. Wang, "Structured compressive sensing-based spatio-temporal joint channel estimation for FDD massive MIMO," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 601–617, 2015.

[89] Y. Zhou, M. Herdin, A. M. Sayeed, and E. Bonek, "Experimental study of MIMO channel statistics and capacity via the virtual channel representation," *Univ. Wisconsin-Madison, Madison, WI, USA, Technical Report*, vol. 5, pp. 10–15, 2007.

[90] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization," *IEEE Transactions on Communications*, vol. 50, no. 3, pp. 374–377, 2002.

[91] W. Li and J. C. Preisig, "Estimation of rapidly time-varying sparse channels," *IEEE Journal of Oceanic Engineering*, vol. 32, no. 4, pp. 927–939, 2007.

[92] G. Z. Karabulut and A. Yongacoglu, "Sparse channel estimation using orthogonal matching pursuit algorithm," in *IEEE 60th Vehicular Technology Conference 2004-Fall*, vol. 6, pp. 3880–3884.

[93] P. Maechler, P. Greisen, N. Felber, and A. Burg, "Matching pursuit: Evaluation and implementatio for LTE channel estimation," in *2010 IEEE International Symposium on Circuits and Systems. Paris, France*, pp. 589–592.

[94] F. Ren, R. Dorrace, W. Xu, and D. Marković, "A single-precision compressive sensing signal reconstruction engine on FPGAs," in *IEEE 2013 23rd International Conference on Field programmable Logic and Applications. Porto, Portugal.*, pp. 1–4.

[95] J. Lu, H. Zhang, and H. Meng, "Novel hardware architecture of sparse recovery based on FPGAs," in *2nd International Conference on Signal Processing Systems. Dalian, China*, vol. 1, 2010, pp. V1–302.

[96] Y. C. Eldar, *Sampling theory: Beyond Bandlimited Systems*. Cambridge University Press, 2015.

[97] Y. Zhang, "User's Guide for YALL1: Your ALgorithms for ℓ_1 Optimization Online. Aavaliable: http://www.caam.rice.edu/optimization/," Tech. Rep., 2009.

[98] J. Huang, C. R. Berger, S. Zhou, and J. Huang, "Comparison of basis pursuit algorithms for sparse channel estimation in underwater acoustic OFDM," in *OCEANS'10 IEEE SYDNEY*, 2010, pp. 1–6.

[99] G. F. Edelmann and C. F. Gaumond, "Beamforming using compressive sensing," *Journal of the Acoustical Society of America*, vol. 130, no. 4, pp. EL232–EL237, 2011.

[100] J. Friedman, T. Hastie, H. Höfling, R. Tibshirani *et al.*, "Pathwise coordinate optimization," *The Annals of Applied Statistics*, vol. 1, no. 2, pp. 302–332, 2007.

[101] T. T. Wu, K. Lange *et al.*, "Coordinate descent algorithms for lasso penalized regression," *The Annals of Applied Statistics*, vol. 2, no. 1, pp. 224–244, 2008.

[102] J. Liu, Y. V. Zakharov, and B. Weaver, "Architecture and FPGA design of dichotomous coordinate descent algorithms," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 11, pp. 2425–2438, 2009.

[103] Q. Sun, D. C. Cox, H. C. Huang, and A. Lozano, "Estimation of continuous flat fading MIMO channels," in *IEEE Wireless Communications and Networking Conference*, vol. 1, 2002, pp. 189–193.

[104] J.-J. Van De Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On channel estimation in OFDM systems," in *IEEE Vehicular Technology Conference*, vol. 2, 1995, pp. 815–819.

[105] I. Telatar and D. Tse, "Capacity and mutual information of wideband multipath fading channels," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1384–1400, 2000.

[106] P. Cheng and Z. Chen, "Multidimensional compressive sensing based analog CSI feedback for massive MIMO-OFDM systems," in *IEEE Vehicular Technology Conference 2014-Fall, Vancouver*, pp. 1–6.

[107] P.-H. Kuo, H. Kung, and P.-A. Ting, "Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays," in *IEEE Wireless Communications and Networking Conference, Paris, France, April 1-5.*, 2012, pp. 492–497.

[108] Z. Tang, R. C. Cannizzaro, G. Leus, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2226–2238, 2007.

[109] H. Senol, E. Panayirci, and H. V. Poor, “Nondata-aided joint channel estimation and equalization for OFDM systems in very rapidly varying mobile channels,” *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4236–4253, 2012.

[110] Z. Tang and G. Leus, “Time-multiplexed training for time-selective channels,” *IEEE Signal Processing Letters*, vol. 14, no. 9, pp. 585–588, 2007.

[111] K. A. D. Teo and S. Ohno, “Optimal MMSE finite parameter model for doubly-selective channels,” in *GLOBECOM’05. IEEE Global Telecommunications Conference, 2005*, vol. 6, pp. 5–pp.

[112] S. Tomasin, A. Gorokhov, H. Yang, and J.-P. Linnartz, “Iterative interference cancellation and channel estimation for mobile OFDM,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 1, pp. 238–245, 2005.

[113] D. K. Borah and B. Hart, “Frequency-selective fading channel estimation with a polynomial time-varying channel model,” *IEEE Transactions on Communications*, vol. 47, no. 6, pp. 862–873, 1999.

[114] J. Zhao, H. Xie, F. Gao, W. Jia, S. Jin, and H. Lin, “Time varying channel tracking with spatial and temporal BEM for massive MIMO systems,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5653–5666, 2018.