UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Engineering and Information Technology

Fast and Accurate Estimation of Angle-of-arrival in Millimetre-wave Large-scale Hybrid Arrays

by

Kai Wu

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

Sydney, Australia

2019

Certificate of Authorship/Originality

I, Kai Wu declare that this thesis, is submitted in fulfilment of the requirements for the award of doctor of philosophy, in the Faculty of Engineering and Information Technology at the University of Technology Sydney. This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree except as fully acknowledged within the text. This thesis is the result of a research candidature jointly delivered with Xidian University as part of a Collaborative Doctoral Research Degree. This research is supported by the Australian Government Research Training Program.

Production Note:

Signature: Signature removed prior to publication.

Date: 18 Jan 2020

ABSTRACT

Fast and Accurate Estimation of Angle-of-arrival in Millimetre-wave Large-scale Hybrid Arrays

by

Kai Wu

Hybrid array is able to leverage array gains, transceiver sizes and costs for massive multiple-input-multiple-output (MIMO) systems in millimetre wave frequencies. Challenges arise from estimation of angle-of-arrival (AoA) in hybrid arrays, due to the array structure and the resultant estimation ambiguities and susceptibility to noises. In this thesis, we study the unambiguous and non-iterative AoA estimation in two types of hybrid arrays — Localized Hybrid Array of Phased SubArrays (LHA-PSAs) and Lens Antenna Arrays (LAAs). For each type, two AoA estimation approaches are proposed for narrowband and wideband, respectively. The main innovation of the approaches include:

- The deterministic sign rules and patterns in LHA-PSAs are unprecedentedly discovered, and exploited to eliminate the estimation ambiguities;
- The optimal trade-off between different error sources is achieved, minimising the wideband AoA estimation error in LHA-PSAs;
- A new wide beam synthesis method is developed for LAAs, which substantially improves the AoA estimation efficiency in LAAs;
- New spatial-frequency patterns are unveiled exploiting the spatial-wideband effect, leading to fast and accurate wideband AoA estimation in LAAs.

Performance analysis is provided for all the approaches with closed-form estimation (lower) bounds derived. Corroborated by simulations, our approaches are able to dramatically improve AoA estimation accuracy while reducing complexity and the number of training symbols, as compared to the state of the art. The estimation errors of our methods asymptotically approach the (lower) bounds.

Acknowledgements

First, I would like to thank my principal supervisor, Prof. Ren Ping Liu, for his persistent effort earning me this opportunity to study at UTS, and for his guidance, encouragement, enthusiasm, and technical comments through the whole PhD candidature. Prof. Liu has been devoting his precious time into organising group meetings, seminars and workshops frequently, which provides valuable chances for us to learn the latest technology trend in communications. I would also like to express my deep gratitude to the co-supervisor, Dr. Wei Ni, for his consistent trust, encouragement, guidance and enormous effort he paid for each of our research works.

I would like to express my special thanks to Prof. Jay Guo, the founding Director of the Global Big Data Centre (GBDTC), University of Technology Sydney (UTS), for proposing the concept and developing the prototype of the hybrid array of phased subarrays, early in 2009. This hybrid array, along with the related studies carried out by Prof. Xiaojing Huang and Prof. Andrew Zhang from GBDTC, motivate the researches in Chapters 2 and 3. In later 2018, Prof. Guo generously shared the idea of reconfigurable multi-beam antennas, which motivates the work in Chapter 4 and some ongoing researches.

I would like to acknowledge GBDTC, Faculty of Engineering and Information Technology (FEIT) and Graduate Research School (GRS) of UTS, and Xidian University, for the scholarships and other (financial) supports during this research.

To my friends, thank you all for the joy and sorrow we shared. To my families, I would not be able to march this far without your wholehearted supporting and sacrifice. To my partner, thank you for always being there.

Kai Wu

Sydney, Australia, 2019

List of Publications

Journal Papers

J-1 . **K. Wu, W. Ni, T. Su, R. P. Liu and Y. J. Guo**, *"Exploiting Spatial-Wideband Effect for Fast AoA Estimation at Lens Antenna Array,"* ***IEEE Journal of Selected Topics in Signal Processing***, accepted, 2019.

J-2 . **K. Wu, W. Ni, T. Su, R. P. Liu and Y. J. Guo**, *"Large-scale Hybrid Antenna Array for Millimeter-Wave/Terahertz High-Speed Railway Communication,"* ***IEEE Communications Magazine***, accepted, 2019.

J-3 . **K. Wu, W. Ni, T. Su, R. P. Liu and Y. J. Guo**, *"Expeditious Estimation of Angle-of-Arrival for Hybrid Butler Matrix Arrays,"* ***in IEEE Transaction on Wireless Communications (TWC)***, vol. 18, no. 4, pp. 2170-2185, April 2019.

J-4 . **Q. Li, T. Su and K. Wu**, *"Accurate DOA Estimation for Large-Scale Uniform Circular Array Using a Single Snapshot,"* ***IEEE Communications Letters***, vol. 23, no. 2, pp. 302-305, Feb. 2019.

J-5 . **K. Wu, W. Ni, T. Su, R. P. Liu and Y. J. Guo**, *"Efficient Angle-of-Arrival Estimation of Lens Antenna Arrays for Wireless Information and Power Transfer,"* ***in IEEE Journal on Selected Areas in Communications (JSAC)***, vol. 37, no. 2, pp. 116-130, Sep. 2018.

J-6 . **K. Wu, W. Ni, T. Su, R. P. Liu and Y. J. Guo**, *"Robust Unambiguous Estimation of Angle-of-Arrival in Hybrid Array With Localized Analog Subarrays,"* ***in IEEE TWC***, vol. 17, no. 5, pp. 2987-3002, May 2018.

J-7 . **K. Wu, W. Ni, T. Su, R. P. Liu and Y. J. Guo**, *"Fast and Accurate Estimation of Angle-of-Arrival for Satellite-Borne Wideband Communication System,"* ***in IEEE JSAC***, vol. 36, no. 2, pp. 314-326, Feb. 2018.

Contents

List of Figures

Abbreviation

Analog-to-Digital and Digital-to-Analog Converters (ADC/DAC)

Angle-of-Arrivals (AoAs)

Auxiliary Beam Pair (ABP)

Additive White Gaussian Noise (AWGN)

Base Stations (BSs)

Cyclic Prefix (CP)

Cramér-Rao Lower Bound (CRLB)

Coherent Signal-subspace Method (CSM)

Designs-of-Freedom (DoF)

Discrete Antenna Arrays (DAAs)

Discrete Fourier Transform (DFT)

Differential DFT Beams (DDBs)

Differential Beam Search (DBS)

Double Cross-Correlation (DCC)

DFT Beam Differences (DBDs)

Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT)

fifth-Generation (5G)

Gigabits per second (Gbps)

Geosynchronous Earth Orbit (GEO)

High-Speed Railway (HSR)

High-Speed Trains (HSTs)

Inverse DFT (IDFT)

Incoherent Signal sub-space Method (ISM)

Lens Antenna Arrays (LAAs)

Localized Hybrid Array of Phased SubArrays (LHA-PSA)

Low Earth Orbit (LEO)

Line-of-Sight (LoS)

Least Squared (LS)

Multiple-Input-Multiple-Output (MIMO)

MUltiple SIgnal Classification (MUSIC)

Mean Square Error (MSE)

MSE Lower Bound (MSELB)

millimetre-Wave (mmWave)

Non-Line-of-Sight (NLoS)

Orthogonal Frequency-Division Multiplexing (OFDM)

Parallel Interference Cancellation (PIC)

Right-Hand Side (RHS)

Radio Frequency (RF)

Signal-to-Noise Ratio (SNR)

Successive Interference Cancellation (SIC)

Singular Value Decomposition (SVD)

Simultaneous WIPT (SWIPT)

Wireless Information and Power Transfer (WIPT)

Wireless Power Transfer (WPT)

Wireless Information Transfer (WIT)

Chapter 1

Introduction

1.1 Background

Massive multiple-input-multiple-output (MIMO) has been widely accepted as an enabling technology for the emerging fifth-generation (5G) and beyond communication [9–11]. One reason is because the scarcity of communication spectrum pushes 5G towards millimetre-wave (mmWave) frequencies [12]. Another reason is that the short wavelength of mmWave allows for integrations of large numbers, i.e., up to hundreds, of miniaturized antennas in limited space, exploiting array gain to compensate for poor radio propagations of mmWave [13]. On the other hand, the physical sizes of radio frequency (RF) chains, consisting of analog-to-digital and digital-to-analog converters (ADC/DAC), power amplifiers, and filters, do not shrink, as the frequency grows higher [14]. To this end, a hybrid antenna array is of particular interest to massive MIMO [1, 3, 4, 10, 15–22], which consists of an adequate number of analog subarrays with phase controllable antenna elements. The analog subarrays are typically arranged localized to facilitate wiring and schematic design [1, 15]. In the localized hybrid arrays, antennas of a subarray are co-located, which is illustrated in Fig. 1.1. The number of RF chains required is equal to that of analog subarrays. The most studied hybrid arrays to date are shown in Fig. 1.1, as will be detailed in Section 1.2.

Massive MIMO is also a promising technique for future wideband multibeam Terabit satellite communications [23, 24], e.g., for extending network connectivities to remote or disjoint regional areas, as well as to aircrafts and vessels typically beyond the coverage of terrestrial wireless networks [25]. This is illustrated in Fig. 1.2. The large number of antennas in massive MIMO can increase substantially antenna gain-to-noise temperature and frequency reuse. It can also enable spatial

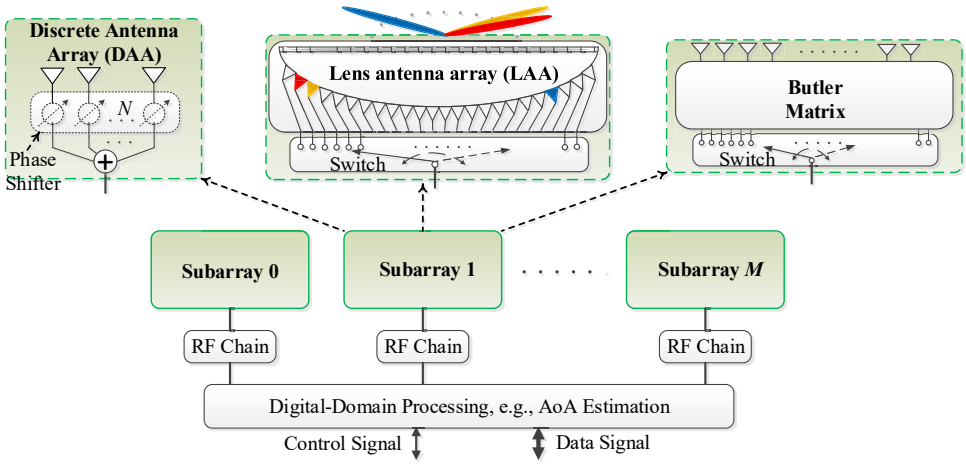

Figure 1.1 : Schematic diagram of a one-dimensional localized hybrid array, where the subarrays can be implemented by either generic discrete antenna elements, or highly integrated antennas such as lens and Butler matrix.

multiplexing, improving spectral efficiency of satellite communications [24]. On the other hand, massive MIMO is uniquely suited for future Terabit multibeam satellite communication systems, due to the fact that the communication spectrum becomes increasingly scarce [26] and future satellite communications are very likely to operate in higher Q/V band [27]. The short wavelength of Q/V frequency band allows for integrations of large numbers, i.e., up to hundreds or even thousands, of miniaturized antennas in a limited space. The high beamforming gains of massive MIMO are also able to compensate for poor radio propagations [13, 28, 29].

Massive MIMO is also extensively adopted in the studies of wireless information and power transfer (WIPT), due to the enormous designs-of-freedom (DoF). RF radiation has been extensively used for wireless communications. Recently, it has been increasingly considered to provide energy sources to power electronic devices. Wireless power transfer (WPT) is critical to self-sustained electronic devices and networks, especially in human-unfriendly environments where battery replacement and recharging are difficult or even impossible [30]. As a matter of course, WIPT has the great potential to develop environment-friendly and self-sustainable communication networks [31]. In general, there are two types of WIPT designs, namely, (a) simultaneous WIPT (SWIPT) which modulates energy signals to serve the purpose

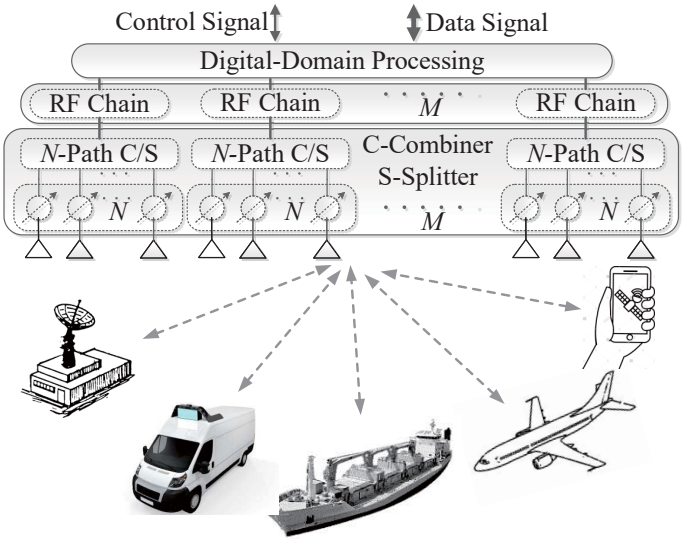

Figure 1.2 : Illustration of a one-dimensional linear localized hybrid antenna array composed of M subarrays. Each subarray has N antenna elements, each connected to an analog phase shifter.

of wireless information transfer (WIT) [29, 32, 33]; and (b) concurrent WPT and WIT, where the transmissions of information and energy signals take place at different transmitters [34–36]. Power transfer efficiency is the key measure of WIPT, due to the fast attenuation of wirelessly transfered power (with the growth of distance).

One of the most critical challenges in WIPT is energy efficiency. To achieve a high energy efficiency, it is desirable to use antennas with narrow beams and therefore high directivities. When these antennas are employed in practice, it is crucial to point the beams accurately in the directions of the transmitter. In the case of SWIPT, focusing on the strongest paths is to save energy for power transfer and improve the receive signal strength for information transmission [37, 38]. In the case of concurrent WPT and WIT, accurate estimation of the angle-of-arrivals (AoAs) of the strongest paths is still critical for focusing the power transfer on the paths, thereby improving power transfer efficiency [37, 38]. Accurate estimation of the AoAs of the strongest WPT paths is also crucial for precisely canceling the paths and hence accurately estimating the AoAs of the far weaker WIT paths. It can be particularly important to alleviate the error propagation in successive interference cancellation (SIC) and, in turn, to improve the data rate for WIT. To this end,

accurate AoA estimation and consequently accurate beamforming with focus on the strongest paths can substantially enlarge the rate-energy region of WIPT, and impact the trade-off between WPT and WIT [31]. Note that in this thesis, we focus on the estimation of AoAs rather than channel vector, which can be more efficient in mmWave massive MIMO channels with sparse nature.

With the energy constraint of the receivers, one would need to resort to high-accuracy estimation of the AoAs of radio signals without using feedback signal loops. We notice that in WIPT applications, the receiver is generally restricted in its own power consumption in order to purse a high energy harvesting efficiency. Therefore, the power that can be consumed for channel/AoA estimation can be very limited, which therefore requires an AoA estimation methods without feedback loops. Furthermore, the cost of the hardware and signal processing required for AoA estimation should be kept as low as possible. Currently, there is a strong interest in lens antenna arrays (LAAs) due to its high gain, and low complexity and cost, as compared to conventional discrete antenna arrays (DAAs) [39–43]. *One reason is that an LAA is more energy-efficient than a DAA [41].* An N-dimensional LAA can readily produce N discrete Fourier transform (DFT) beams [39], while an N-dimensional linear DAA requires N^2 number of $\log_2^N$-bit phase shifters to generate N DFT beams. A DAA would require more antennas than an LAA to achieve the same spatial resolution [42]. *Another reason is that LAAs have strong abilities of spatial interference suppression [39, 41].* LAAs can separate signals with different AoAs by exploiting different mainlobes of DFT beams. This is particularly interesting to concurrent WPT and WIT [35, 36], where different DFT beams can be generated to pinpoint WPT and WIT paths, and suppress interferences between paths. To the best of our knowledge, however, no effective AoA estimation technique using LAA has been reported.

Another critical application of mmWave massive MIMO is the high-speed railway (HSR) communication. Being an efficient and sustainable land transport method, future HSR is envisaged to be safer, greener and more convenient [44, 45]. Given increased passenger capacity of high-speed trains (HSTs), there are expected to be

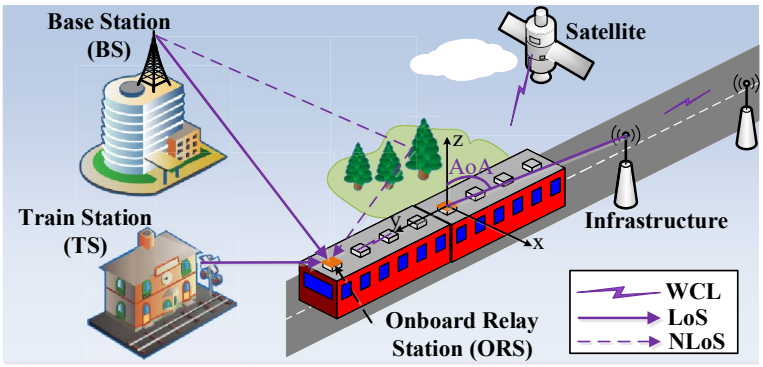

Figure 1.3 : Schematic diagram of the "smart" railway communication system with the interconnections between the HST and BS, TS, infrastructures, as well as satellite. WCL refers to wireless communication link.

a substantially increased number of wireless connections between passengers and the Internet. As illustrated in Fig. 1.3, a practical scenario is to have a train-top antenna array to act as the proxy to relay in-cabin (WiFi, cellular or Ethernet) traffic to track-side base stations (BSs) or communication satellites. Given the scarcity of frequencies and the demand for high data rates of tens of gigabits per second (Gbps) [44, 46], high carrier frequencies with broad bandwidths, such as mmWave and THz, are likely to be used. The mmWave frequency range of 24.25 – 52.6 GHz has been specified in 5G new radio, and considered in HST scenario [47]. Feasibility studies, including extensive measurement campaigns, have been conducted to confirm the validity of the consideration. For example, a 30 GHz large-scale hybrid array has been tested over an HST link between an onboard relay and track-side infrastructure [47]. A significant improvement of spectral efficiency compared with legacy LTE-configured HST has been demonstrated at the train speed of 500 km/h [47, Fig. 5]. THz antenna arrays still have issues in compact and efficient design [44], but have started to show a good prospect. For example, a recent work [48] has reported a 400-GHz THz antenna array with a measured gain of 33.66 dBi.

1.2 Large-scale Hybrid Antenna Array

Large-scale antenna arrays have a range of advantages, e.g., in the mmWave

frequency band. The arrays are able to achieve tunable/steerable narrow beams with high gains to combat the severe attenuation the mmWave bands. The narrow beams can also help suppress interference and improve the effective signal-to-interference-plus-noise ratio of intended signals. For example, the beamwidth of a $1,000$-element uniform linear array is about 0.0126 rad, which is only two percent of the beamwidth of a 20-element counterpart. The short wavelength of the mmWave frequency makes the integration of hundreds to thousands of antenna elements per array possible [49]. A linear aperture size of 150 cm can accommodate up to 300 elements with antenna spacing of half wavelength at 30 GHz, while it can only accommodate 20 elements at 2 GHz.

A hybrid antenna array becomes a preferable and cost-effective design of large-scale antenna arrays, where a large number of antennas are grouped and connected into a much smaller number of analog subarrays. Each antenna has an individual configurable phase shifter. Each subarray is connected to a single RF chain. This is because the physical sizes of RF chains, consisting of analog-to-digital and digital-to-analog converters, power amplifiers, and filters, do not shrink in the mmWave frequency band, as compared to the lower frequencies. Hardware impairments, like different delays in RF components, can be calibrated and suppressed before baseband digital signal processing. However, we notice that introducing analog arrays can incur more power consumption. The typical power consumption of a 4-bit mmWave phase shifter is 30 mW [49]. To produce a 16-dimensional DFT beam, a total of 16 phase shifters are required, consuming $30 \times 16 = 480$ mW power. To this end, when determining the size of the analog arrays, the power consumption and also the parameter estimation performance should be taken into account.

A one-dimensional linear large-scale hybrid array is shown in Fig. 1.1. The subarrays are typically arranged in a localized fashion to facilitate wiring and schematic design [50]. The subarrays can consist of discrete antenna elements. The resulting array is referred to as DAA, which is typically a one-dimensional uniform linear array. In a recent 3GPP HST evaluation [47], a 30-GHz large-scale hybrid array with 256 antennas was installed at an onboard relay. The array is rectangular with

four 8×8 analog subarrays and four RF chains. The antenna spacing is half of a wavelength. With considerations on energy and cost efficiency, and integration level, a subarray can be replaced with a lens antenna or a Butler matrix, as also shown in Fig. 1.2.

In the case of lens antenna array, each subarray is a lens antenna array [49, 51]. The lens can focus the microwave signals onto its parabola focal surface. By meticulously placing the antennas on the focal surface of the lens, a complete set of DFT beams can be activated with sinc-shape beam patterns and pointing directions evenly spaced within $[0, 2\pi)$. Each antenna, referred to as a *beam port*, can be energized to activate a DFT beam. A switch can be used to select a beam port, and connect the selected beam port to the RF chain hardwired to the switch [49].

In the case of Butler matrix array, each subarray is a Butler matrix which is a passive analog beamforming circuit. The Butler matrix also generates DFT beams. Each beam port can activate an individual DFT beam, like the lens antenna. Also like the lens antenna array, a switch can be used to select and connect a beam port to an RF chain in the Butler matrix arrays. The beam pattern of a DFT beam is a sinc function within $[0, 2\pi)$. The beam width of the mainlobe depends on the number of antennas. The mainlobe can be shifted (or rotated) by changing the phases of the antennas [51, eq. 7].

The Lens antenna arrays and Butler matrix arrays can be much more energy-efficient than the phase shifter based DAAs due to their high integration level. The lens and Bulter matrix are passive beamformers, while the phase shifters consume non-negligible powers. The typical power consumption of a 4-bit mmWave phase shifter is 30 mW [49]. To produce a 16-dimensional DFT beam, a total of 16 phase shifters are required, consuming $30 \times 16 = 480$ mW power. In contrast, the power consumption of an equivalent lens antenna array is only 4 mW [49].

The one-dimensional Butler matrices and lens antennas can both be readily extended to two-dimensional uniform planar arrays, due to the fact that the DFT beams can be decoupled losslessly between the azimuth and elevation [3]. The DAAs can also be readily extended to two-dimensional uniform planar arrays if

DFT beams are adopted. There are also other forms of hybrid antenna arrays, such as fully-connected array and uniform circular array. Overall, they are less tractable due to too many cross-points (in fully connected hybrid arrays) or weak directivity (in uniform circular arrays). Therefore, they are far less popular in practice.

1.3 AoA Estimation — Importance, Challenges and State of the art

The AoA estimation is indispensable for efficient utilization of mmWave hybrid arrays in future communications. It allows for accurately configuring the arrays to quickly capture the impinging signal, form narrow and strong beams, and achieve high signal-to-noise ratio (SNR) and throughput. The fast and accurate AoA estimation can also avoid the round-trip delay of the typical channel estimation and feedback, hence speeding up beamforming and guaranteeing seamless and reliable connectivity. Conventional channel estimation techniques involving channel sounding, estimation, and feedback would become inadequate. This is because the number of RF chains is much smaller than that of antennas (beams) in mmWave frequency bands given the increasingly compact design of antennas. In contrast, conventional channel estimation approaches were typically designed for digital arrays with the equal numbers of antennas and RF chains in lower frequencies. Important channel information which was accessible to digital arrays, such as the phase difference between antennas, is not readily accessible in the mmWave hybrid antenna arrays. In this section, the importance and challenges of the AoA estimations in the popular hybrid arrays (shown in Fig. 1.1) are elaborated on. The state-of-the-art AoA estimation methods in different narrowband/wideband hybrid arrays are also introduced and summarized. In this thesis, we regard an antenna array with a negligible spatial-wideband effect as a narrowband array; whereas in contrast, when the effect is prominent we refer to an antenna array as a wideband array.

1.3.1 Narrowband Hybrid Array of Phased Subarrays

A critical issue to be addressed in localized hybrid arrays is accurate estimation of AoA, which is critical to beamforming, interference suppression or localization [18, 20]. The hybrid nature and the sheer scale of massive MIMO hinder the use of conventional estimation techniques [1, 3, 15, 16]. Cross-correlation has been increasingly used for AoA estimation in localized hybrid arrays [1, 3, 4, 15, 17]. Cross-correlations were first taken to estimate the phase difference between consecutive analog subarrays, i.e., Nu, using the same phase shift at all subarrays. N is the number of antennas per subarray. $u = \frac{2\pi d \sin\phi}{\lambda}$, where ϕ is the azimuth AoA at a uniform linear array with antenna spacing d and wavelength λ. However, estimating u from the estimate of Nu exhibits inherent ambiguity with N possible results [1, 3, 15, 17]. The ambiguities would further deteriorate, coupled with a zigzag effect which can misjudge an estimate of Nu around $(2k \pm 1)\pi$ with an error around 2π, $k = 0, 1, \cdots, N-1$, in the presence of non-negligible noises [15, 17].

In a recent work [4], subarray-specific time-invariant phase shifts were designed. As we proved in [4] for special localized hybrid arrays with N being a multiple of the number of subarrays M, the complex gains of subarrays in the direction of ϕ form a Fourier series with a consistent phase difference u between coefficients, provided the subarrays adopt evenly spaced phase shifts between $[0, 2\pi)$ per symbol. Given that the gain of the m-th subarray and e^{jmNu} are coupled multiplicatively in the received signal of the subarray, one could first estimate Nu and suppress e^{jmNu} in the received signal, then take inverse discrete Fourier transform (IDFT) and correlate the Fourier coefficients to unambiguously estimate u and hence ϕ. However, the estimation of Nu is non-trivial and can incur ambiguities, due to the use of subarray-specific phase shifts [4]. The complex gains of subarrays can have different phases in the direction of ϕ. The estimate of Nu, obtained from the cross-correlation of received signals, can be corrupted by the different phases of the subarray gains, compromising the estimation accuracy of Nu and u. There are also undetectable angles in the case that the AoA happens to be the nulls of all subarrays. Moreover, the phase shifts are limited to a special case with N being a multiple of M. The results are inapplicable

to general localized hybrid arrays.

1.3.2 Wideband Hybrid Array of Phased Subarrays

Accurate and fast estimation of AoA is the key enabling technology underlying wideband multi-beam Terabit satellite-borne communication system [52–54]. In the case of widely deployed geosynchronous Earth orbit (GEO) satellites, the a-priori knowledge on the locations of fixed ground stations can help the satellites to steer beams. However, it is still possible that there is a misalignment due to the inclination of the satellites and the orbit perturbations [55, 56]. Let alone aircrafts and vessels, of which the a-priori location knowledge can be unavailable or inaccurate at the satellites [57]. On the other hand, low Earth orbit (LEO) satellites have been increasingly deployed for Terabit satellite communications. Unlike GEO satellites, the coverage areas of LEO satellites dynamically change [55]. In addition, the instantaneous information on the trajectory and orbit of an LEO satellite can be inaccurate, or even unknown [58]. In both cases of GEO and LEO satellites, the accurate and fast estimation of AoA is vital for improving the beamforming gain and radio link quality, as well as increasing the data rate, especially in low SNR regions.

Unfortunately, the on-board accurate and efficient wideband AoA estimation in satellite-borne localized hybrid arrays is yet to be addressed. With the hybrid nature and the sheer scale of the arrays, conventional wideband AoA estimation techniques – classical signal subspace based algorithms, e.g., incoherent signal subspace method (ISM) [59], coherent signal-subspace method (CSM) [60] and the subsequent improvement algorithm [61, 62] are inapplicable or unsuitable, since they would require a reliable initial AoA estimation, and iterative refinements for a fine result [63]. The estimation delay would be lengthy, and the estimation accuracy can be poor, both subject to the initial accuracy.

Revisiting the recent work [4], subarray-specific time-invariant phase shifts were designed, such that the receive signals of the subarrays, with phases calibrated by Nu, can form a series of Fourier coefficients with a phase difference of u. Cross-

correlations can be taken between the coefficients to estimate u unambiguously. However, the phase shifts would result in a new ambiguity in estimating Nu with two possible estimates π apart. Combining the cross-correlation between all pairs adjacent subarrays can be destructive due to undetermined signs of the amplitude parts of the cross-correlations which leads to the ambiguous estimation of Nu in the first place. In [4], the signs were heuristically adjusted before the cross-correlations were added up. Weak cross-correlations would be likely corrupted by receive noises, exhibit incorrect signs, and hence be incorrectly adjusted. The combination of the cross-correlations could remain destructive.

We note that these algorithms have been typically developed for narrow-band systems. The only exception is the algorithm in [3], which was later applied to wideband systems in [1], where narrow-band operations, i.e., cross-correlation between subarrays, were independently conducted at each sub-carrier followed by cross-correlations between sub-carriers. Unfortunately, the issue of error propagation remains [1]. Moreover, the cross-correlation between sub-carriers can substantially reduce the estimation SNR, as the frequency interval (in terms of the number of sub-carriers), over which the cross-correlation is taken, appears as the coefficient of Nu and needs to be used to divide the phase of the cross-correlation to estimate Nu. In this sense, the interval can amplify the estimation error of Nu.

1.3.3 Narrowband Lens Antenna Arrays

Accurate knowledge of the AoA is the key to the effective applications of LAAs and massive MIMO in general, because the AoA allows for correct selection of the DFT beam pointing at the dominant path [51, 64, 65]. Moreover, channel estimation can be substantially simplified by first estimating the AoA and then the path responses in the consecutive transmission blocks [66, 67]. This is because the AoAs of mmWave/THz channels have a much larger coherent time (or in other words, stay unchanged longer) than the path responses [44, 68].

Due to the distinct array physics and structure of the LAA, existing AoA estimation techniques become either inapplicable [2, 6, 7, 21, 69–73], or incur inefficiency

and/or inaccuracy [74, 75]. In [74, 75], the beamspace (spanned by DFT beams) channel estimation techniques were developed for LAAs by enumerating all DFT beams and selecting those with the outputs above a predefined threshold. The least square (LS) approximation was performed on the selected beams for channel estimations in [74]. In [43], compressive sensing was employed for LAAs and the beamspace channel structure was used to improve the estimation accuracy of each channel. The estimation errors of these approaches can be as large as half the width of a DFT beam of an LAA [43, 74, 75].

Most existing works have been focused on the channel estimation of DAAs, and are not directly applicable to the AoA estimation of LAAs. Channel estimation has been extensively studied for WPT [69–72] or SWIPT [76, 77]. It was performed by taking the LS approximation [69, 70, 76, 77], by combining the Kalman filter with the LS approximation [71], or by using the maximum likelihood estimation [72]. In [78, 79], the transmitter was designed to receive one-bit feedback from the receiver and the channel estimation was refined recursively by adjusting the transmit beamforming at the cost of significant delays. In these works, a single-path WPT was assumed at a receiver. In [2, 6, 21], codebook-based channel estimation techniques were developed for millimeter wave massive multiple-input-multiple-output. The techniques were based on DAAs, and are not directly applicable to LAAs. The estimation errors of these approaches can be up to the beamwidth based resolution of the codebooks [2, 6, 21].

In [5], the state-of-the-art AoA estimation algorithm, referred to as ABP, was designed for DAAs, where DFT beamforming was employed. Specifically, the two DFT beams with the two strongest outputs were identified; and the AoA was estimated from the ratio of the power difference of the DFT beams to the total power of the beams. ABP was able to achieve better estimation accuracy than the other existing methods in [2, 6, 21, 43, 74, 75], through the exploitation of the deterministic amplitudes of DFT beams. However, for the AoA estimation of a path, ABP enumerated the DFT beams to search for the two adjacent beams with the two strongest outputs, incurring long estimation delays.

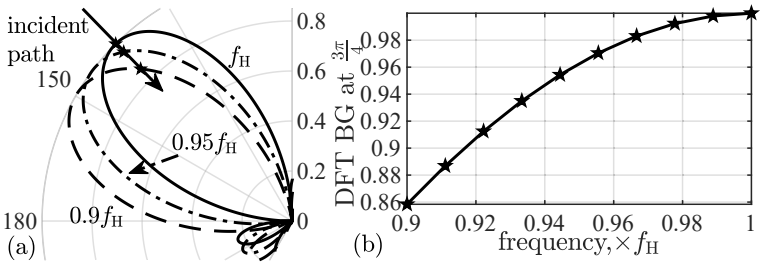

Figure 1.4 : Illustration of the spatial-wideband effect, where an 8-element linear array is considered. The spatial responses of the 4th DFT beam at different frequencies are plotted in Fig. 1.4(a). The beamforming gain (BG) of the second DFT beam at $\theta = \frac{3\pi}{4}$ across frequencies is plotted in Fig. 1.4(b).

1.3.4 Wideband Lens Antenna Arrays

It is non-trivial to estimate the AoA in wideband LAAs. As revealed in [80–82], a prominent reason is that a spatial-wideband effect, also known as a beam squint effect [82], is non-negligible in mmWave/THz wideband massive MIMO, and leads to the frequency-dependent spatial responses, as depicted in Fig. 1.4.* Consider a typical linear mmWave massive MIMO with 128 antennas and a bandwidth of 1 GHz at the 60 GHz carrier frequency, the maximum delay caused by across-aperture propagation can be as large as 1.058 times of a sampling interval [66, 80]. The spatial-wideband effect has been hindering the effective AoA estimation and beamforming, particularly for passive, analog arrays (such as LAA) where beams cannot be adjusted independently for different frequencies.

Another reason behind the difficulty in the AoA estimation of LAAs is that the parameter for the estimation of the AoA, i.e., the phase difference between adjacent antenna elements, is obscure, due to the DFT beamforming (or in other words, the high integration level of the LAAs). Moreover, only a subset of (combined) DFT beams can be selected to feed into the RF chains at any instant, since the number

*The spatial-wideband effect refers to that phenomenon that the beam pointing direction and the beamwidth of a wideband array vary with the frequency. Weighting a wideband large-scale array by the conventional steering vector at θ, i.e., $\left[1, e^{\frac{2\pi d \sin\theta}{\lambda}}, \cdots, e^{\frac{2\pi(N-1)d\sin\theta}{\lambda}}\right]^{\mathrm{T}}$, the beam points at θ if $f = f_{\mathrm{H}}$. N denotes the total number of antenna elements. As f decreases, the beam pointing direction becomes larger than θ and the beamwidth expands; see Fig. 1.4(a). d is antenna spacing, λ is the wavelength of the largest frequency f_{H}. Due to the effect, a DFT beam can have different spatial responses across frequencies; see Fig. 1.4(b).

of RF chains is typically much smaller than that of the DFT beams [43, 75].

Only a small number of studies [8, 43, 51, 74, 75] have been carried out on the channel/AoA estimation of LAAs. Most of the studies [43, 51, 74, 75] were focused on the narrowband channel/AoA estimation, and typically required a large number of training symbols in search of the dominant path [43, 74, 75]. In [51], the accurate AoA estimation was achieved in LAAs; however, the method was targeted for narrowband systems. In [8], compressive sensing was applied for the channel estimation in LAAs, where the spatial-wideband effect was accounted for by identifying the different path supports across frequency sub-carriers. The accuracy achieved by [8] was shown to be much better than that of the previous methods (developed for DAAs) [83, 84] in which the spatial-wideband effect was overlooked. However, the AoA estimation was not pursued in [8, 83, 84]. An accurate estimation of the AoA for wideband signals was achieved in [7], which, despite being developed based on DAAs, performed DFT beamforming. However, the N DFT beams were enumerated at the cost of a long training delay, and the spatial-wideband effect was overlooked. It is worth pointing out that the few existing methods [8, 66], that did take the spatial-wideband effect into account, only take countermeasures to offset the effect passively. This would prolong the estimation delay or incur a high computational complexity.

1.3.5 Fast Angular Estimation and Tracking of HST

The train-top array and its counterparts at the track-side BSs or satellites are expected to be large-scale to produce narrow beams with high gains, thereby combating severe attenuation at mmWave frequencies. The arrays are also anticipated to be hybrid with localized architectures in consideration of cost-efficiency and size. Omni-directional transmission could be possible at lower frequencies, but the limited bandwidths available at those frequencies are not wide enough to support Gbps [44]. To this end, accurate estimation and reliable tracking of the AoA from an HST to the track-side BSs or satellites, or the other way around, is critical to the implementation of the mmWave HSR communication systems.

The estimation and feedback could become excessively frequent and likely to be outdated in the HSR application where the trains can travel at very high speeds of up to 500 km/h. It is important that a train-top mmWave antenna array and its counterpart at the track-side BSs (or satellites) can estimate the AoAs passively and instantly based on the impinging signals from each other, steer and lock their beams, and keep tracking the changes of AoAs to adapt the beams.

Seen from the previous elaboration, to estimate the AoA in the large-scale hybrid antenna arrays is non-trivial with several key challenges to be properly addressed. None of existing techniques are able to address the challenges holistically. Below we highlight the challenges of the AoA estimation in hybrid arrays, particularly combining the HST applications.

1. There is a long-standing estimation ambiguity originating from the architecture of localized hybrid antenna arrays where the received signals at the different antenna elements of a subarray are mixed (or added up) before being sent into an RF chain. The phase difference between adjacent antenna elements, which gives the estimate of the AoA, becomes obscure. Only the phase difference between adjacent subarrays is available, which is a multiple of the inter-element phase difference and needs to be divided by the number of antenna elements in each subarray. Unfortunately, a division of an angle (or phase) gives ambiguous results due to the 2π periodicity of angles (or phases).

2. In most cases, HSTs have unobstructed line-of-sight (LoS) to the track-side BSs. When penetrating buildings and propagating through foliage, mmWave signals attenuate significantly. For these reasons, the mmWave HSR channels are typically dominated by strong LoS paths [44, 46, 85], and can be modeled as Rician channels [86]. Ray-tracing techniques have been employed to simulate the Rician factor – the ratio between the power in the LoS path and the power in the scattered paths – in typical outdoor scenarios. As reported in [85, Tab. II], the power in the LoS path is at least 11 dB stronger than the total power in the other paths in the outdoor 30 GHz channels.

3. The HSR channels can also exhibit strong temporal variations, due to the speed of up to 500 km/h of the HSTs [87]. The conventional omnidirectional broadcast is not suitable for the mmWave frequencies, because of severe attenuations at the frequencies. The received SNR can be very low. For example, the received SNR is around -20 dB at each antenna, when the carry frequency is 100 GHz, the bandwidth is 1 GHz, the transmit power is 20 dBm, the transmitter-receiver distance is 100 m, and both the transmit and receive antennas have unit gains [44]. It is necessary to exploit the strong directivity and gains of efficient mmWave antennas to track and uninterruptedly serve the HSTs.

1.4 Literature Review

The state-of-the-art AoA estimation methods in mmWave hybrid arrays, to be used as benchmarks, are summarized in this section.

The two best-known AoA estimation techniques are multiple signal classification (MUSIC) [19] and estimation of signal parameters via rotational invariance techniques (ESPRIT) [19]. Developed originally for full-digital arrays [19], these techniques exploit the orthogonality of the signal and noise subspaces to estimate the signal subspace which is the span of the array response vectors in the directions of the AoAs. MUSIC takes the autocorrelation of received array signals, and estimates the signal subspace as the eigenvectors associated with the large, meaningful eigenvalues of the autocorrelation matrix. ESPRIT divides an array into two subarrays, calculates the transformation matrix between the signal subspaces of the two subarrays, and estimates the AoA from the eigenvalues of the matrix. Computationally expensive singular value decomposition (SVD) is required in both techniques, and could hinder the scalability of the techniques. MUSIC and ESPRIT cannot be directly applied to hybrid antenna arrays, due to the RF combining at analog subarrays and the resultant obscurity of the phase offset information on individual antennas.

Some earlier works were focused on extending the results of analog or digital

arrays to hybrid arrays. In [19], N N-dimensional orthonormal bases were used to be the phase shifts of a subarray over N symbols. The steering matrix of the localized hybrid array became an $NM \times NM$ matrix, where every M columns account for a symbol, and each of the columns consists of the N-entry orthonormal steering vector of the corresponding subarray, and zeros at all the other entries. The received signal vector can be constructed accordingly over N symbols. The covariance of the steering matrix is the identity matrix, and does not distort the signal subspace of the covariance of the received signals. This allows the classical MUSIC and ESPRIT [88] to be applied, which, however, require singular value decomposition. The complexity would grow cubically with the total number of antennas NM, and become prohibitive since the number of antennas is typically in hundreds. In [21], the received signal of a hybrid array was restructured to be a linear function of a sparse vector collecting path gains in quantized directions. The AoA can be estimated through the non-zero element in the vector by using compressive sensing techniques. The coefficient of the vector, depending on a multi-resolution beamforming codebook, can be adjusted to scan different directions with different beamwidths. This algorithm was later extended to use a non-orthogonal codebook, where the beams, generated by different codewords, overlap [22]. However, these algorithms require the receiver to repeatedly select the codewords for the transmitter to narrow down the beam. The estimation delay would be lengthy, and the estimation accuracy would be subject to the resolutions of the codebooks.

More recent approaches have taken cross-correlations among subarrays for AoA estimation [1, 3, 17]. In [3], the same phase shifts were deployed across subarrays at every symbol to scan one direction. By accumulating the cross-correlations between the received signals of adjacent subarrays, the propagation phase offset between adjacent subarrays, Nu, was estimated to be the phase of the accumulated cross-correlation, denoted by $\widehat{Nu}$. The accumulation is constructive, since all the subarrays have the same complex gain in the direction of the AoA. u was estimated to be $\hat{u}_n = \frac{\widehat{Nu}+2k\pi}{N}$ $(k = -\lfloor \frac{N}{2} \rfloor, -\lfloor \frac{N}{2} \rfloor + 1, \cdots, \lfloor \frac{N}{2} \rfloor)$ with an odd number of N or $(N+1)$ possible estimates due to flooring $\lfloor \cdot \rfloor$. By adjusting the phase shifts, the N

or $(N+1)$ estimates were sequentially scanned in the following N or $(N+1)$ symbols. The estimate with the highest receive power was recursively scanned, and updated through the accumulated cross-correlations between the received signals of adjacent subarrays to recalibrate the scanning direction, till convergence. Unfortunately, a long estimation delay would occur, and the approach would also suffer from error propagations resulting from an incorrectly identified initial direction. The algorithm was later applied to wideband systems, whereas the issue of error propagation remained [1]. Authors in [17] endeavored to improve the estimation accuracy in the early symbols of [3] to suppress the error propagation. However, the results, though improved, were still susceptible.

In a recent work [4], a special linear localized hybrid array was considered, where N is a multiple of M. The phase shift was set to $\frac{2\pi m}{M}$ at the m-th subarray, $m = 0, 1, \cdots, M-1$. It was discovered that, given the phase shifts, the complex gains of the subarrays form a Fourier series with a consistent phase offset u in the coefficients. The Fourier coefficients could be recovered by suppressing e^{jmNu} in the the received signal of the m-th subarray, and then taking IDFT on the received signals. u could be estimated through the cross-correlations between consecutive Fourier coefficients. To suppress e^{jmNu}, Nu was estimated by evaluating the phase of the accumulated cross-correlations between the received signals of adjacent subarrays in the same way as it was in [3]. Unfortunately, the cross-correlations between the complex gains of adjacent subarrays did not have consistent signs due to the use of different phase shifts, as opposed to [3]. The accumulation of the cross-correlations of the received signals were destructive, even after the signs of the cross-correlations were heuristically calibrated [12]. Moreover, the sign of the accumulated cross-correlation was still undetermined, leading to estimation ambiguities in Nu and subsequently u.

For lens antenna arrays, earlier works [74, 75] enumerated the complete set of DFT beams to detect the pilot signals, selected the beams with the strong received power, and performed the least squared (LS) approximation to estimate the principle components of the beamspace channel. Note that in LAAs, the physical channel is

transformed by the DFT beamforming, and the resulting channel is referred to as the beamspace channel [75]. These works could be time-consuming, since they require a large number of training symbols, i.e., $\mathcal{O}(\frac{N}{K})$.

In [43], compressive sensing was adopted to reduce the number of training symbols. The random Bernoulli matrix (with ± 1 elements) was adopted for beam selection, which acts as the random sampling matrix in sparse sampling. By detecting the support of the strongest path, the principle components of the path were estimated through sparse recovery. The beamspace channel structure was exploited to improve sparse recovery performance. By canceling the estimated components, the next strongest path could be estimated.

In [51], the accurate AoA estimation was achieved in narrowband LAAs by establishing the one-to-one mapping between the AoA and the received power in the first two strongest differential DFT beams (DDBs). To identify the DDBs, a new beam synthesis method was designed in [51] by exploiting solely the analog beam selection network. As a result, a fast noise-resilient K-ary search of the strongest DDBs was developed, reducing the number of training symbols to $\mathcal{O}(\log_K(N))$. However, an extension of [7] to a wideband system is non-trivial. [7] cannot provide a benchmark for the wideband AoA estimation technique developed in this thesis.

The above methods [43, 51, 74, 75] were developed for narrowband systems. The wideband channel/AoA estimation were widely studied in DAA-based massive MIMO systems [7, 83, 84, 89, 90]. For instance, in [7], an accurate AoA estimation of wideband signals was achieved at hybrid DAAs by establishing a one-to-one mapping between the AoA and the selected DFT beams. However, these works are either inapplicable to LAAs, due to the optimized analog beamforming network which therein was based on phase shifters; or not tailored according to the properties of LAAs, such as the simultaneous multi-beam DFT beamforming and the beam selection-based transceiving. Moreover, these works [7, 83, 84, 89, 90] fail to address the spatial-wideband effect [82].

It was revealed in [8, 15, 66] that overlooking the spatial-wideband effect in mmWave massive MIMO systems can degrade both the channel estimation perfor-

mance and the system achievable rate. The channel estimation method [43] was extended to wideband LAAs in [8]. Unlike the previous works [7, 83, 84, 89, 90], the spatial-wideband effect was taken into account in [8] by adopting the sub-carrier-dependent path supports. To out best knowledge, [8] is the only work accounting for the spatial-wideband effect in the *channel estimation* for LAAs; and the *AoA estimation* in wideband LAA-based massive MIMOs has yet to be studied properly.

1.5 Thesis Organization

In this thesis, we present four new approaches to address the challenges for fast and accurate AoA estimations in hybrid arrays of phased subarrays and lens antenna arrays. The first two approaches are designed for the hybrid array of phased subarrays, one for narrowband and the other for wideband. The last two approaches are designed for lens antenna arrays, also one for narrowband and the other for wideband. The approaches all have their own respective innovations. In specific, the thesis is organised as follows:

1. Chapter 1 introduces the background of the AoA estimation in mmWave hybrid arrays, discusses the critical challenges, and reviews the state of the art.

2. Chapter 2 proposes a new AoA estimation in narrowband hybrid array of phased subarrays, where the estimation ambiguities are solved for the first time.

3. Chapter 3 develops a new wideband AoA estimation method in wideband hybrid array of phased subarrays, achieving the optimal trade-off between different errors/gains and substantially improving the estimation accuracy.

4. Chapter 4 proposes a new AoA estimation method for lens antenna array, which is accomplished by a new wide beam synthesis method and a one-to-one mapping established between judiciously selected DFT beams and AoA.

5. Chapter 5 develops a wideband AoA estimation methods in lens antenna array, specifically to exploit the spatial-wideband effect to improve estimation efficiency.

6. Chapter 6 presents the conclusions drawn from the results in earlier chapters of the thesis, and discusses the limitations and future research directions of this study.

7. Chapter 7 provides the proofs, analyses and derivations for chapters 2 to 5.

Chapter 2

Robust Unambiguous Estimation of Angle-of-Arrival in Hybrid Array with Localized Analog Subarrays

2.1 Introduction

Hybrid array is able to leverage array gains, transceiver sizes and costs for massive MIMO systems in millimeter wave frequencies. Challenges arise from estimation of AoA in localized hybrid arrays, due to the array structure and the resultant estimation ambiguities and susceptibility to noises. This chapter eliminates the ambiguities and enhances the tolerance to the noises based on our new discoveries. Particularly, by designing new subarray-specific time-varying phase shifts, we discover that the cross-correlations between the gains of consecutive subarrays have consistent signs except the strongest. This enables the cross-correlations to be deterministically calibrated and constructively combined for the noise-tolerant estimation of the propagation phase offset between adjacent subarrays. Given the phase offset, the AoA can be estimated unambiguously with few training symbols. We also derive a closed-form lower bound for the mean square error of AoA estimation. Corroborated by simulations, our approach is able to dramatically improve estimation accuracy by orders of magnitude while reducing complexity and training symbols, as compared to the state of the art. With the ambiguities eliminated, the estimation errors of our method asymptotically approach the lower bound, as training symbols increase.

In this chapter, we propose an unambiguous AoA estimator for general localized hybrid arrays where N does not need to be a multiple of M. By judiciously designing the phase shifts, we discover that the cross-correlations between the complex gains of adjacent subarrays have consistent signs except the strongest. The discovery

enables the cross-correlations between the received signals of adjacent subarrays to be correctly calibrated in sign and constructively combined for the estimation of Nu, since the strongest cross-correlation is the least likely to be corrupted by noises. With the phase shift design, we also lift the restriction of [12] to special arrays and inherit in general counterparts that u can be uniquely recovered from the Fourier coefficients of the complex gains of subarrays. As a result, the complex gains of the subarrays can be retained while the phase differences mNu ($m = 0, 1, \cdots, M-1$) are suppressed in the received signals, recovering the Fourier coefficients and hence the unambiguous estimation of u.

The key contributions of this chapter can be summarised in the following:

1. We discover that the cross-correlations of the complex gains between adjacent subarrays have consistent signs except the strongest. In the absence of explicit knowledge on the AoA, the discovery enables the cross-correlations to be correctly calibrated and coherently accumulated for unambiguous estimation of Nu;

2. We extrapolate the phase shift design of [12] for general localized hybrid arrays, and prove that u can be uniquely estimated from the Fourier coefficients of the received signals, after the propagation phases of the signals mNu ($m = 0, 1, \cdots, M-1$) are estimated and suppressed;

3. By properly configuring the generalized phase shifts, we eliminate undetectable AoAs, which were overlooked in [12];

4. We alleviate the zigzag effect of the AoA estimation, which often happens in the case that u is around $\pm\pi$ and misjudged by up to 2π due to non-negligible receive noises.

An analysis of estimation accuracy is carried out, and an asymptotic lower bound is derived for the mean square error (MSE) of the estimation. Evident from extensive simulations, our approach is able to dramatically reduce the MSE by orders of magnitudes with a fraction of complexity and training symbols, as compared to the

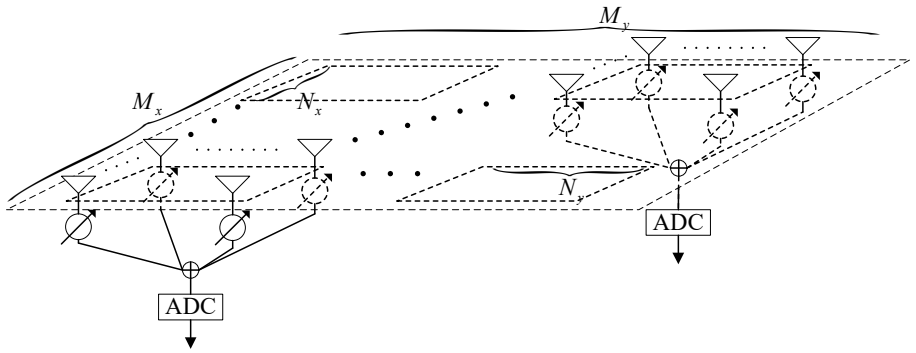

Figure 2.1 : A two-dimensional planar localized hybrid antenna array composed of $M_x \times M_y$ subarrays. Each subarray has $N_x \times N_y$ antenna elements, each connected to an analog phase shifter. The outputs at analog subarrays are further processed to generate digital samples. The RF generation and down-converters are suppressed.

state of the art. With the ambiguities eliminated, our approach is noise-tolerant. Its estimation errors asymptotically approach the lower bound, as the number of training symbols increases.

The rest of this chapter is organized as follows. In Section 2.2, the system model is described. In Section IV, we design the new subarray-specific time-varying phase shifts, discover the deterministic rule on the signs of the cross-correlations between the gains of consecutive subarrays, and develop unambiguous estimators for both Nu and u. In Section V, we derive the MSE lower bound (MSELB) of the estimations. Simulation and numerical results are provided to demonstrate the superior accuracy of the proposed approach in Section 2.5, followed by conclusions in Section 2.6.

2.2 System Model

Fig. 2.1 illustrates a rectangular localized hybrid array with $M_xN_x \times M_yN_y$ evenly spaced antenna elements in two dimensions. The array is divided into $M_x \times M_y$ localized rectangular analog subarrays, each comprised of $N_x \times N_y$ antenna elements. These antennas are connected to adjustable phase shifters, one for each, and aggregated at an ADC. (For clarity, the RF generation and down-converters are suppressed in the figure.)

The mmWave frequencies are dominated by directive LoS radio links with in-

significant multi-path interferences [1, 3, 91, 92]. Consider a single-path narrow-band received signal $\tilde{s}(t)$ with wavelength λ_c, elevation angle θ and azimuth angle ϕ, the received signal at the (m_x, m_y)-th analog subarray ($m_x = 0, 1, \cdots, M_x - 1$; $m_y = 0, 1, \cdots, M_y - 1$) can be written as [3, 4, 15, 16]

$$s_{m_x,m_y}(t) = \tilde{s}(t)e^{j2\pi f_D t} P^t_{m_x,m_y}(u_x, u_y)e^{j(m_x N_x u_x + m_y N_y u_y)} + \xi_{m_x,m_y}(t), \tag{2.1}$$

where f_D is the Doppler frequency due to the relative movement of the transmitter and receiver; $\xi_{m_x,m_y}(t)$ is the zero-mean additive white Gaussian noise (AWGN) at the (m_x, m_y)-th subarray; $P^t_{m_x,m_y}(u_x, u_y)$ is the radiation pattern of the (m_x, m_y)-th analog subarray at the t-th symbol, and

$$P^t_{m_x,m_y}(u_x, u_y) = \sum_{n_x=0}^{N_x-1} \sum_{n_y=0}^{N_y-1} P^{n_x,n_y}_{m_x,m_y}(u_x, u_y) \times e^{j(n_x u_x + n_y u_y + \alpha^t_{m_x,m_y}(n_x,n_y))}, \tag{2.2}$$

where $P^{n_x,n_y}_{m_x,m_y}(u_x, u_y)$ is the individual radiation pattern of the (n_x, n_y)-th antenna element ($n_x = 0, 1, \cdots, N_x - 1$, $n_y = 0, 1, \cdots, N_y - 1$) at the (m_x, m_y)-th analog subarray and $\alpha^t_{m_x,m_y}(n_x, n_y)$ is the phase shift at the antenna element and t-th symbol. n_x and n_y are the indexes for antennas along the x- and y-axes of the subarray, respectively. As in other works [1, 3, 4, 17, 19, 93], we assume here that $P^{n_x,n_y}_{m_x,m_y}(u_x, u_y) = 1$. u_x and u_y are beam-domain AoAs in x- and y-axis, respectively. They are defined as

$$u_x = 2\pi d_x \sin\theta \cos\phi/\lambda_c, \tag{2.3}$$

$$u_y = 2\pi d_y \sin\theta \sin\phi/\lambda_c, \tag{2.4}$$

where d_x and d_y are the distances between two adjacent antenna elements along the x-axis and y-axis, respectively.

The estimation of AoA, or more specifically θ and ϕ, is essentially to estimate u_x and u_y, since $\theta = \text{sign}\{u_x\} \sin^{-1}\left(\frac{\lambda_c\sqrt{u_x^2+u_y^2}}{2\pi\sqrt{d_x+d_y}}\right)$ and $\phi = \tan^{-1}(\frac{u_y d_x}{u_x d_y})$. $\text{sign}\{u_x\}$ takes the sign of u_x.

Take cross-correlation between the received signals of two consecutive analog

subarrays along the x-axis, as given by

$$\begin{aligned}\rho_{m_x}(t) &= s^*_{m_x,m_y}(t)s_{m_x+1,m_y}(t) \\ &= |\tilde{s}(t)|^2 P^{t*}_{m_x,m_y}(u_x,u_y)P^{t}_{m_x+1,m_y}(u_x,u_y)e^{jN_xu_x} + \tilde{\xi}_{m_x}(t),\end{aligned} \tag{2.5}$$

where the superscript "$*$" takes complex conjugate, and $\tilde{\xi}_{m_x}(t)$ is a zero-mean AWGN that can be given by

$$\begin{aligned}\tilde{\xi}_{m_x}(t) =& \xi^*_{m_x,m_y}(t)\xi_{m_x+1,m_y}(t) \\ &+\tilde{s}(t)^* e^{-j2\pi f_D t}P^{t*}_{m_x,m_y}e^{-j(m_xN_xu_x+m_yN_yu_y)}\xi_{m_x+1,m_y}(t) \\ &+\tilde{s}(t)e^{-j2\pi f_D t}P^{t}_{m_x+1,m_y}e^{-j((m_x+1)N_xu_x+m_yN_yu_y)}\xi_{m_x,m_y}(t).\end{aligned} \tag{2.6}$$

f_D is suppressed by cross-correlation. Likewise, we can also obtain $r_{m_y}(t)$ and $\tilde{\xi}_{m_y}(t)$.

Note from (2.5) that the estimation of u_x can be decoupled from that of u_y. For illustration convenience, we consider a one-dimensional linear localized hybrid array with M linear analog subarrays and N antenna elements per subarray. The subscripts "$_x$" and "$_y$" are suppressed in the rest of this chapter. (2.3) can be written as

$$u = \frac{2\pi d \sin\theta}{\lambda_c} \tag{2.7}$$

Nevertheless, the results of this chapter can be readily applied to two-dimensional planar localized hybrid arrays.

2.3 Proposed AoA Estimation for general Localized Hybrid Arrays

Considering a general localized hybrid array, we design new phase shifts and address the aforementioned estimation ambiguity of $\widehat{Nu}$. We also prove that u can be unambiguously estimated through the Fourier coefficients of $s_m(t)e^{-jm\widehat{Nu}}$, $m = 0, 1, \cdots, K-1$, provided $\widehat{Nu}$ is accurate.

2.3.1 Phase Shifts under General Localized Hybrid Arrays

We design new phase shift values for generalized hybrid arrays with N unnecessarily being a multiple of M, as given by

$$\alpha_m^t(n) = n\alpha_m^t = \frac{2\pi n(\text{mod}\{m, K\}P + t)}{L}, \tag{2.8}$$

where m, t and n are the indexes for analog subarrays, training symbols and antennas, respectively ($0 \leq m \leq M-1;\ 0 \leq t \leq P-1;\ 0 \leq n \leq N-1$); K is the number of different phase shifts adopted per symbol; P is the number of training symbols; L is the total number of different phase shifts in use, and $L = PK$. α_m^t is the slope of the phase shifts with respect to the antenna index in subarray m, and hence provides the beamforming direction of the subarray. Since N may not be a multiple of M, K is designed to be a factor of N satisfying:

$$2<K \leq M,\ N = QK,\ K, Q \in \mathbb{Z}^+, \tag{2.9}$$

where $\mathbb{Z}^+$ stands for positive integers.

Different from [4], the phase shifts in (2.8) are designed to accommodate general localized hybrid arrays where N may not be a multiple of M. Specifically, we discovered in [4] that, with linearly increased phase shifts over consecutive subarrays, the outputs of the subarrays can form a Fourier series with u unambiguously captured in the Fourier coefficients. This must be under the condition that N is an integral multiple of the number of subarrays involved. Our new design of $N = QK$ ($K \leq M$) partitions the total M subarrays by every K subarrays; and obtains $(Q-1)$ non-overlapping groups of K consecutive subarrays, and the Q-th group of the last K consecutive subarrays sharing $(QK - M)$ subarrays with the $(Q-1)$-th group. Each of the K-subarray groups can meet the aforementioned condition and unambiguously estimate u, even in the general cases where N may not be an integral multiple of M. Further, K evenly spaced phase shifts between 0 and 2π are applied to all the Q groups of K consecutive subarrays at a symbol. Between successive symbols, all the K phase shifts rotate clockwise by $\frac{2\pi}{L}$. This allows L evenly spaced

directions within $[0, 2\pi)$ to be scanned during P symbols, ensuring that the AoA is captured by one of the L beams with strong gain. Coherent accumulations can be conducted among the groups and symbols to improve resistance to noises.

Note that it is unnecessary to go through all P symbols to get an estimate of u, denoted by $\hat{u}$. As a matter of fact, we can only use a single symbol, e.g., $t = 0$. In the rest of this section, we will prove that the cross-correlations of the subarray gains $P_m^{t*}(u)P_{m+1}^t(u)$ in (2.5) have deterministic signs unaffected by the receive noise. The cross-correlations of the received signals $\rho_m(t)$ can be coherently combined across subarrays, within a single symbol or across multiple symbols, to increase the SNR of Nu estimation. We will also prove that, given $\widehat{Nu}$, u can be unambiguously estimated by taking the cross-correlations of the Fourier coefficients of the received signals within a symbol after e^{jmNu} $(m = 0, 1, \cdots, M-1)$ are suppressed in the signals. Both $\widehat{Nu}$ and $\hat{u}$ can be obtained within a single symbol.

Also note that the accuracy of $\hat{u}$ is not restrained by the resolution of phase shift, i.e., $\frac{2\pi}{L}$. As mentioned earlier, cross-correlations are taken to estimate u through the phase differences between subarrays, rather than scanning at the resolutions. Moreover, the L phase shifts defined in (2.8) specify L predefined directions. This is distinctively different from the scanning process of [3], where multiple ambiguous estimates of u are recursively scanned.

The proposed phase shift design in (2.8) can be readily extended from one-dimensional linear localized hybrid arrays to two-dimensional rectangular localized hybrid arrays. At the t-th symbol, the phase shift of the (n_x, n_y)-th antenna of the (m_x, m_y)-th subarray can be designed as

$$\alpha_{m_x,m_y}^t(n_x, n_y) = \frac{2\pi n_x(\mathrm{mod}\{m_x, K_x\}P_x + t)}{L_x} + \frac{2\pi n_y(\mathrm{mod}\{m_y, K_y\}P_y + t)}{L_y}, \quad (2.10)$$

where K_x and K_y are the extensions of K to the two dimensions along the x- and y-axes, respectively; P_x and P_y are the extensions of P; and $L_x = P_xK_x$ and $L_y = P_yK_y$ are the extensions of $L = PK$. The AoA estimation can be decoupled between the two dimensions of a rectangular localized hybrid array, since the

radiation pattern $P^t_{m_x,m_y}(u_x, u_y)$ of a two-dimensional rectangular antenna array is the product of individual radiation patterns along x- and y-axes; see (2.2), and the cross-correlations taken between the received signals of adjacent subarrays along one dimension can completely suppress the impact of the phase shifts designed for the other dimension; see (2.5). Specifically, given m_y, both the cross-correlations between the received signals of adjacent subarrays along the x-axis to estimate Nu_x and the cross-correlations of the Fourier coefficients to estimate u_x can eliminate $\frac{2\pi n_y(\text{mod}\{m_y,K_y\}P_y+t)}{L_y}$ which stays unchanged in the phases of the signals; and vice versa. u_x and u_y can be estimated in the same way as u, as will be elaborated on in Section 2.3.3.

2.3.2 Estimation of Nu

For illustration convenience and clarity, our discussions are focused on the first K subarrays using the K phase shifts. The index for the K subarrays is $m = 0, 1, \cdots, K-1$. The conclusion obtained in the rest of this section can be readily applied to the rest $(M-K)$ subarrays.

As per the t-th symbol, the phase shift at the n-th antenna of the m-th subarray is $n\alpha^t_m$, according to (2.8). By substituting (2.8) into (2.2), the complex gain of the m-th subarray $P^t_m(u)$ can be given by

$$P^t_m(u) = \sum_{n=0}^{N-1} e^{jnu} e^{-jn\alpha^t_m} \tag{2.11a}$$

$$= \sum_{n=0}^{N-1} e^{jnu} e^{-j2\pi \frac{\text{mod}\{m,K\}P+t}{L} n} \tag{2.11b}$$

$$= \frac{1 - e^{j2N\left(\frac{u}{2} - \pi \frac{\text{mod}\{m,K\}P+t}{L}\right)}}{1 - e^{j2\left(\frac{u}{2} - \pi \frac{\text{mod}\{m,K\}P+t}{L}\right)}} = e^{j(N-1)\omega^t_m} \frac{\sin(N\omega^t_m)}{\sin(\omega^t_m)}, \tag{2.11c}$$

where $\omega^t_m = \frac{u}{2} - \frac{\alpha^t_m}{2} = \frac{u}{2} - \pi\left(\frac{t}{L} + \frac{m}{K}\right)$ for $m = 0, 1, \cdots, K-1$, because $\alpha^t_m = \frac{2\pi(mP+t)}{L}$ for $m = 0, 1, \cdots, K-1$ and can be rewritten as $\alpha^t_m = \frac{2\pi m}{K} + \frac{2\pi t}{L}$ by plugging $L = PK$. (2.11a) adds up the complex gains of individual antennas of the subarray, nu is the propagation phase offset of antenna n with respect to antenna 0, and $n\alpha^t_m$ is the aforementioned phase shift deployed at antenna n; (2.11b) is achieved

by substituting (7) into (2.11a); and (2.11c) is obtained by plugging $L = PK$ into (2.11b). We note that $P_m^t(u)$ is a periodic function of m with the cycle of K.

By substituting (2.11) into (2.5), the cross-correlation of the received signals between the m-th and the $(m+1)$-th analog subarrays at symbol t can be given by

$$\rho_m(t) = s_m^*(t)s_{m+1}(t) = |\tilde{s}(t)|^2 G_m^t(u)e^{jNu} + \tilde{\xi}_m(t), \tag{2.12}$$

where $m = 0, 1, \cdots, K-1$ (for $K < M$. In the case of $K = M$, $s_K(t)$ is replaced by $s_0(t)$, exploiting the periodicity of the phase shifts, as will be discussed later), and $G_m^t(u) = P_m^{t*}(u)P_{m+1}^t(u)$ and can be written as

$$G_m^t(u) = e^{\frac{-j(N-1)\pi}{K}} \frac{\sin(N\omega_m^t)}{\sin(\omega_m^t)} \frac{\sin(N\omega_{m+1}^t)}{\sin(\omega_{m+1}^t)}. \tag{2.13}$$

Since $\xi_m(t)$ is typically zero-mean additive white Gaussian and independent and identically distributed (i.i.d), $\tilde{\xi}_m(t)$ can be approximated as a zero-mean AWGN, as given by

$$\begin{aligned}\tilde{\xi}_m(t) =& \xi_m^*(t)\xi_{m+1}(t) + \tilde{s}(t)P_{m+1}^t(u)e^{j(m+1)Nu}\xi_m^*(t) \\ &+ \tilde{s}(t)^* P_m^{t*}(u)e^{-jmNu}\xi_{m+1}(t).\end{aligned} \tag{2.14}$$

We attempt to coherently combine $\rho_m(t)$ for $m = 0, 1, \cdots, M-1$, and $t = 0, 1, \cdots, P-1$, to improve the robustness of $\widehat{Nu}$ against the noise and subsequently the estimation accuracy. This requires the coefficients $|\tilde{s}(t)|^2 G_m^t(u)$, $\forall m, t$ in (2.12), to be consistent in sign. The reason for taking constructive combination is because the complex gains of the subarrays $P_m^t(u)$, depending on u to be estimated, are unknown in prior, and so is the cross-correlation of the gains $G_m^t(u)$. Other combining techniques, such as maximal ratio combining, would require explicit knowledge on $G_m^t(u)$, $\forall m, t$, and therefore are inapplicable.

Theorem 2.1. *At a symbol t, only $G_{m'}^t(u) \neq 0$, with the largest amplitude, has the opposite sign to all the rest of $G_m^t(u)$, where m, $m' \in [0, K-1]$ and $m \neq m'$.*

Proof. In the case of $K < M$, $\sin(N\omega_{m+1}^t) = \sin\left(N\omega_m^t - \frac{N\pi}{K}\right) = (-1)^Q \sin(N\omega_m^t)$

for $m = 0, 1, \cdots, K-1$, since $\omega_{m+1}^t = \frac{u}{2} - \pi\big(\frac{t}{L} + \frac{m+1}{K}\big) = \omega_m^t - \frac{\pi}{K}$, and $Q = \frac{N}{K}$, as specified in (2.9). Therefore, $G_m^t(t)$ can be rewritten as

$$G_m^t(u) = e^{\frac{-j(N-1)\pi}{K}} \frac{(-1)^Q \sin^2(N\omega_m^t)}{\sin(\omega_m^t)\sin(\omega_{m+1}^t)}, m \in [0, K-1]. \tag{2.15}$$

Given Q, the sign of $G_m^t(u)$ only depends on that of $\sin(\omega_m^t)\sin(\omega_{m+1}^t)$.

For any u, $\sin(\omega_m^t) = \sin\big(\frac{u}{2} - \frac{\pi t}{L} - \frac{\pi m}{K}\big)$, $m \in [0, K-1]$, spans half a cycle of the sine function with the initial phase $(\frac{u}{2} - \frac{\pi t}{L})$. $\sin(\omega_{m+1}^t)$ is the shifted version of $\sin(\omega_m^t)$.

Lemma 2.1. *There exists a unique integer* $m' \in [0, K-1]$ *satisfying*

$$\sin(\omega_{m'}^t)\sin(\omega_{m'+1}^t) < 0.$$

Proof. See Appendix 7.1.1. □

Lemma 2.2. *Given* m'*, specified in Lemma 2.1, we have*

$$|\sin(\omega_{m'}^t)\sin(\omega_{m'+1}^t)| < |\sin(\omega_m^t)\sin(\omega_{m+1}^t)|,$$

for any $m \in [0, K-1]$, $m \neq m'$

Proof. See Appendix 7.1.2 □

Given m', as specified in Lemma 2.2, we have $\sin(N\omega_m^t) = \sin\left(N\omega_{m'}^t + \frac{\pi N(m'-m)}{K}\right)$ $= (-1)^{Q(m'-m)}\sin(N\omega_{m'}^t)$, since $\omega_m^t = \omega_{m'}^t + \frac{\pi(m'-m)}{K}$. Therefore,

$$\begin{aligned}
&\sin(N\omega_m^t)\sin(N\omega_{m+1}^t) \\
&= (-1)^{Q(m'-m)}\sin(N\omega_{m'}^t) \times (-1)^{Q(m'-m-1)}\sin(N\omega_{m'}^t) \\
&= (-1)^Q \sin^2(N\omega_{m'}^t).
\end{aligned} \tag{2.16}$$

From Lemma 2.2, the denominator of $G_m^t(u)$, $m = 0, 1, \cdots, K-1$, takes the

minimum at m', i.e.,

$$|G^t_{m'}(u)| = \left| \frac{(-1)^Q \sin^2(N\omega^t_{m'})}{\sin(\omega^t_{m'}) \sin(\omega^t_{m'+1})} \right| \tag{2.17a}$$

$$> \left| \frac{(-1)^Q \sin^2(N\omega^t_{m'})}{\sin(\omega^t_{m}) \sin(\omega^t_{m+1})} \right| \tag{2.17b}$$

$$= \left| \frac{\sin(N\omega^t_{m}) \sin(N\omega^t_{m+1})}{\sin(\omega^t_{m}) \sin(\omega^t_{m+1})} \right| \tag{2.17c}$$

$$= |G^t_m(u)|, m = 0, 1, \cdots, K-1; m \neq m', \tag{2.17d}$$

where (2.17a) and (2.17c) are obtained by plugging (2.16); (2.17b) is based on Lemma 2.2; and (2.17d) is based on (2.13).

By combining Lemma 2.1 and (2.17), the sign of $G^t_m(u)$ only changes at m', which, according to Lemma 2.2, provides the largest of $|G^t_{m'}(u)|$ for $m = 0, 1, \cdots, K-1$.

In the case of $K = M$, we can have $G^t_{K-1}(u) = P^{t*}_{K-1} P^t_0(u)$, i.e., the cross-correlation between the first and the last subarrays. For $m = 0, 1, \cdots, K-2$, $G^t_m(u)$ remains the same as (2.13). The discussion on the case of $K < M$ still applies, because

$$P^t_K(u) = e^{j(N-1)\left(\frac{u}{2} - \frac{\pi t}{L} - \frac{\pi K}{K}\right)} \frac{\sin\left(N\left(\frac{u}{2} - \frac{\pi t}{L} - \frac{\pi K}{K}\right)\right)}{\sin\left(\left(\frac{u}{2} - \frac{\pi t}{L} - \frac{\pi K}{K}\right)\right)}$$

$$= e^{j(N-1)\left(\frac{u}{2} - \frac{\pi t}{L}\right)} \frac{\sin\left(N\left(\frac{u}{2} - \frac{\pi t}{L}\right)\right)}{\sin\left(\frac{u}{2} - \frac{\pi t}{L}\right)} = P^t_0(u). \tag{2.18}$$

This concludes the proof. □

Following Theorem 2.1, we propose to align the signs of $\rho_m(t)$, before conducting constructive combination, as follows.

$$\tilde{\rho}_m(t) = \begin{cases} (-1)^Q \rho_m(t) & \text{if } m \neq m'; \\ (-1)^{Q+1} \rho_{m'}(t) & \text{if } m = m', \end{cases} \tag{2.19}$$

where $m = 0, 1, \cdots, K-1$; $t = 0, 1, \cdots, P-1$; and m' is the index to the cross-

Algorithm 1 Estimate Nu under generalized hybrid arrays

Initialize: Given N and M, select K ($\in (2, M]$, $\in \mathbb{Z}^+$),
  $N/K \in \mathbb{Z}^+$.
for $t \leftarrow 0, 1, \cdots, P-1$ **do**
  $s_m(t) \leftarrow$Analogue beamforming with α_m^t in (2.8);
  for all $m \in [0, M-2]$ **do**
    $\rho_m \leftarrow s_m^*(t) s_{m+1}(t)$;
  end for
  if $K = M$ **then**
    $\rho_{K-1} \leftarrow s_{K-1}^*(t) s_0(t)$;
  end if
  Find $\rho_{m'}(t)$ having the largest amplitude among
    $\rho_m(t)$ $(0 \le m \le K-1)$;
  for all $m \in [0, M-2]$ **do**
    $\tilde{\rho}_m(t) \leftarrow (-1)^Q \rho_m(t)$;
  end for
  for $m \leftarrow m' : K : M-2$ **do**
    $\tilde{\rho}_m(t) \leftarrow -\tilde{\rho}_m(t)$;
  end for
end for
$\bar{\rho} \leftarrow \frac{1}{P(M-1)} \sum_{t=0}^{P-1} \sum_{m=0}^{M-2} \tilde{\rho}_m(t)$;
return $\widehat{Nu} \leftarrow \text{ang}\left\{ e^{\frac{j\pi(N-1)}{K}} \bar{\rho} \right\}$.

correlation with the largest amplitude, as specified in Lemma 2.2.

The generalization of (2.19) to the case of $0 \le m \le M-1$ is straightforward. From (2.8), $\rho_m(t) = \rho_{\text{mod}\{m,K\}}(t) = \rho_m(t)$, $0 \le m \le M-1$. Out of the total M cross-correlations, there can be $\lfloor \frac{M}{K} \rfloor$ or $\lceil \frac{M}{K} \rceil$ with the opposite sign to the rest, i.e., one every K. These $\lceil \frac{M}{K} \rceil$ or $\lfloor \frac{M}{K} \rfloor$ cross-correlations are the strongest, and therefore the least susceptible to the noise and most unlikely to be misjudged in terms of sign, as compared to the rest. We can reliably calibrate their signs, as done in (2.19), combine $\tilde{\rho}_m(t)$ over $m \in [0, M-2]$ and $t \in [0, P-1]$, and estimate Nu with improved estimation accuracy, as summarized in Algorithm 1. The algorithm is noise-tolerant.

Algorithm 2 Estimate u under generalized hybrid arrays

Initialize: Given $\widehat{Nu}$ obtained from Algorithm 1.
for $t \leftarrow 0, 1, \cdots, P-1$ **do**
  for all $m \in [0, M-1]$ **do**
    $\tilde{s}_m(t) \leftarrow s_m(t) e^{-jm\widehat{Nu}}$;
  end for
  for $l \leftarrow 0, 1, \cdots, \lfloor M/K \rfloor - 1$ **do**
    $m \leftarrow lK$;
    $\mathbf{s}^l(t) \leftarrow \{\tilde{s}_m(t), \tilde{s}_{m+1}(t), \cdots, \tilde{s}_{m+K-1}(t)\}$;
    $\boldsymbol{a}^l(t) \leftarrow \mathrm{IDFT}\{\mathbf{s}^l(t)\}$;
    $\boldsymbol{r}^l_{n'}(t) \leftarrow \boldsymbol{a}^{l*}_{n'} \boldsymbol{a}^l_{n'+1}, n' = 0, 1, \cdots, K-2$;
  end for
  $\mathbf{s}^{\lfloor M/K \rfloor}(t) \leftarrow \{s_{\lfloor M/K \rfloor K}(t), \cdots, s_{M-1}(t),$
    $s_{M-\lfloor M/K \rfloor K}(t), \cdots s_{K-1}(t)\}$;
  $\boldsymbol{a}^{\lfloor M/K \rfloor}(t) \leftarrow \mathrm{IDFT}\big\{\mathbf{s}^{\lfloor M/K \rfloor}(t)\big\}$;
  $\boldsymbol{r}^{\lfloor M/K \rfloor}_{n'}(t) \leftarrow \boldsymbol{a}^{\lfloor M/K \rfloor *}_{n'} \boldsymbol{a}^{\lfloor M/K \rfloor}_{n'+1}, n' = [0, K-2]$;
end for
$\bar{r} \leftarrow \frac{1}{P\lceil M/K \rceil (K-1)} \sum_{t=0}^{P-1} e^{j\frac{2\pi t}{L}} \left[\sum_{l=0}^{\lfloor M/K \rfloor} \sum_{n'=0}^{K-2} \boldsymbol{r}^l_{n'}(t) \right]$;
return $\hat{u} \leftarrow \mathrm{ang}\{\bar{r}\}$.

2.3.3 Estimation of u under general localized hybrid arrays

Given $\widehat{Nu}$, we can estimate $P^t_m(u)$ by multiplying $e^{-jm\widehat{Nu}}$ to both sides of (2.1), i.e.,

$$s_m(t) e^{-jm\widehat{Nu}} = \tilde{s}(t) P^t_m(u) e^{j2\pi f_D t} e^{jm(Nu-\widehat{Nu})} + \xi_m(t) e^{-jm\widehat{Nu}}. \tag{2.20}$$

Suppose that e^{jmNu} can be precisely suppressed. u can be unambiguously estimated, as dictated in the following theorem.

Theorem 2.2. *Provided that $e^{jm(Nu-\widehat{Nu})} \to 1$, u can be estimated unambiguously by using (2.8) in general localized hybrid arrays.*

Proof. Let $n = n' + qK$ $(0 \le n' \le K-1,\ 0 \le q \le Q-1)$. (2.11) can be written as

$$\begin{aligned} P_m^t(u) &= \sum_{n'=0}^{K-1} \sum_{q=0}^{Q-1} e^{j(n'+qK)u} e^{-j2\pi \frac{mP+t}{L}(n'+qK)} \\ &= \sum_{n'=0}^{K-1} g_{n'}^t(u) e^{-j\frac{2\pi n' m}{K}}, m = 0, 1, \cdots, K-1, \end{aligned} \tag{2.21}$$

where

$$g_{n'}^t(u) = e^{jn'(u-\frac{2\pi t}{L})} e^{j(N-K)(\frac{u}{2}-\frac{\pi t}{L})} \frac{\sin(\frac{Nu}{2}-\frac{N\pi t}{L})}{\sin(\frac{Ku}{2}-\frac{K\pi t}{L})}, n' = 0, 1, \cdots, K-1, \tag{2.22}$$

gives the Fourier coefficients of $P_m^t(u)$.

Taking the IDFT of $s_m(t)e^{-jm\widehat{Nu}}$, $m = 0, 1, \cdots, K-1$, we obtain $a_{n'}(t) = \tilde{s}(t)e^{2\pi f_D t} g_{n'}^t(u) + z_{n'}(t)$, $n' = 0, 1, \cdots, K-1$, with $z_{n'}(t) = \text{IDFT}\{\xi_m(t)e^{-jm\widehat{Nu}}\}$.

A cross-correlation can be taken between any two consecutive IDFT points, i.e., $a_{n'}(t)$ and $a_{n'+1}(t)$, to achieve an unambiguous estimate of u, since the cross-correlation, denoted by $r_{n'}(t)$, can be given by

$$\begin{aligned} r_{n'}(t) &= a_{n'}^*(t) a_{n'+1}(t), n' = 0, 1, \cdots, K-2 \\ &= e^{j(u-\frac{2\pi t}{L})} \left| \frac{\sin(\frac{Nu}{2}-\frac{N\pi t}{L})}{\sin(\frac{Ku}{2}-\frac{K\pi t}{L})} \right|^2 |\tilde{s}(t)|^2 + \tilde{z}_{n'}(t), \end{aligned} \tag{2.23}$$

where

$$\tilde{z}_{n'}(t) = z_{n'}^*(t) z_{n'+1}(t) + \tilde{s}(t) C(u) e^{j(n'+1)(u-\frac{2\pi t}{L})} z_{n'}^*(t) + \tilde{s}(t)^* C(u)^* e^{-jn'(u-\frac{2\pi t}{L})} z_{n'+1}(t), \tag{2.24}$$

where $C(u) = e^{j(N-K)(\frac{u}{2}-\frac{\pi t}{L})} \frac{\sin(\frac{Nu}{2}-\frac{N\pi t}{L})}{\sin(\frac{Ku}{2}-\frac{K\pi t}{L})}$. □

Following Theorem 2.2, Algorithm 2 is proposed to estimate u. Consider the periodicity of the phase shifts in (2.8) over subarrays. The outputs of every K consecutive analog subarrays outputs can provide an estimate of u. There are a total of $\lfloor M/K \rfloor$ non-overlapping groups of K consecutive analog subarrays, including the

last group concatenated between the last $(M - \lfloor M/K \rfloor K)$ subarrays and the last $(K - M + \lfloor M/K \rfloor K)$ subarrays of the first cycle, exploiting the periodicity of the phase shifts designed in (2.8); see Lines 12 to 15. The step in Line 17 captures that constructively combining (2.23) for $n' = 0, 1, \cdots, K-2$, $l = 0, 1, \cdots, \lceil M/K \rceil$ and $t = 0, 1, \cdots, P-1$ is carried out to improve the tolerance of the estimation to noises and hence the estimation accuracy, where l indicates the l-th group of K consecutive subarrays.

In the case that $N \neq QM$ ($Q \in \mathbb{Z}^+$), the phase shifts in (2.8) enables us to find $K < M$ satisfying $N/K \in \mathbb{Z}^+$. Note that IDFT and cross-correlation were also used to estimate u from $P_m^t(u)$ in [4]. However, the phase shifts designed in [4] were limited to the case that N is a multiple of M. In contrast, Algorithm 2 proposed here does not require N to be a multiple of M. The phase shifts of [4] is a special case of (2.8) with $K = M$ and $P = Q$.

2.3.4 Suppression of the zigzag effect

There can be a zigzag effect on the estimation error of u, especially in the case that u is close to $\pm\pi$. Specifically, $\hat{u}$ can be misjudged up to 2π away from u due to non-negligible receive noises. For instance, $u = 0.98\pi$ can be misjudged to be $\hat{u} = -0.98\pi$ in the presence of noises, provided the maximum error is as small as 0.04π (and $\hat{u}$ is supposed to be within $[0.94\pi, 1.02\pi]$). The zigzag effect on $\hat{u}$ can substantially compromise the estimation accuracy of $\phi = \sin^{-1}(\lambda_c u / 2\pi d)$.

By comparing the amplitudes between all the outputs of K consecutive analog subarrays across P symbols, we can obtain the range of u, denoted by $\tilde{\beta}_u$. Suppose that the largest amplitude of all the outputs is at the m_1-th subarray and t_1-th symbol, and the second largest at the m_2-th subarray and t_2-th symbol. As noted, α_m^t divides $[0, 2\pi)$ into L evenly spaced segments. In the case that the SNR is high, u is most likely to be close to $\alpha_{m_1}^{t_1}$, within the segment specified by

$$\beta_u = \begin{cases} \left[\alpha_{m_1}^{t_1}, \alpha_{m_2}^{t_2} - \frac{\pi}{L}\right), & \text{if } \alpha_{m_1}^{t_1} = \alpha_{m_2}^{t_2} - \frac{2\pi}{L}; \\ \left(\alpha_{m_2}^{t_2} + \frac{\pi}{L}, \alpha_{m_1}^{t_1}\right], & \text{if } \alpha_{m_1}^{t_1} = \alpha_{m_2}^{t_2} + \frac{2\pi}{L}, \end{cases} \tag{2.25}$$

where $\frac{\pi}{L}$ specifies the half of the segment closest to $\alpha_{m_1}^{t_1}$. Since α_k^t is defined on $[0, 2\pi)$ while ang$\{\cdot\}$ is defined on $[-\pi, \pi)$, we project β_u onto $[-\pi, \pi)$ by

$$\tilde{\beta}_u = \begin{cases} \beta_u, & \text{if } \alpha_{m_1}^{t_1} < \pi; \\ \beta_u - 2\pi, & \text{if } \alpha_{m_1}^{t_1} \geq \pi. \end{cases} \tag{2.26}$$

If $\hat{u}$ and $\tilde{\beta}_u$ take opposite signs, the zigzag effect is likely to occur. To alleviate the effect, if $\hat{u}$ is within $-\tilde{\beta}_u$, we change the sign of $\hat{u}$. We note that the segments are likely to be correctly identified, because a SNR gain of around $10\log_{10}N$ dB* is typically achieved at the m_1-th subarray and symbol t_1.

2.3.5 Discussions on Undetectable Angles

There can be the special case where $G_m^t(u) = 0$, $m \in [0, K-1]$, and u can be undetectable. Theorem 2.1 does not apply and Algorithm 1 cannot achieve $\widehat{Nu}$ with accuracy. Nevertheless, we show that the accuracy of Algorithm 2 is unaffected by Algorithm 1 in this special case. We also prove that the use of as few as two training symbols can eliminate any undetectable angles using the new phase shifts in general localized hybrid arrays.

Lemma 2.3. *At any symbol t, if there exists $\tilde{m} \in [0, K-1]$ that $G_{\tilde{m}}^t(u) = 0$, then $G_m^t(u) = 0$ $\forall m \in [0, K-1]$.*

Proof. If $G_{\tilde{m}}^t(u) = 0$, then $\sin(N\omega_{\tilde{m}}^t) = 0$ or $\sin(N\omega_{\tilde{m}+1}^t) = 0$. In the case that $\sin(N\omega_{\tilde{m}}^t) = 0$, we have $\sin(N\omega_{\tilde{m}+1}^t) = \sin(N\omega_{\tilde{m}}^t - \frac{N\pi}{K}) = (-1)^Q \sin(N\omega_{\tilde{m}}^t) = 0$. In the case that $\sin(N\omega_{\tilde{m}+1}^t) = 0$, we have $\sin(N\omega_{\tilde{m}}^t) = \sin(N\omega_{\tilde{m}+1}^t + \frac{N\pi}{K}) = (-1)^Q \sin(N\omega_{\tilde{m}+1}^t) = 0$. We also have

$$\sin(N\omega_{\tilde{m}+r}^t) = \sin(N\omega_{\tilde{m}}^t - rQ\pi) = (-1)^{rQ} \sin(N\omega_{\tilde{m}}^t) = 0, r \in \mathbb{N}.$$

As a result, $\sin(N\omega_m^t) = 0$ and hence $G_m^t(u) = 0$ for $m = 0, 1, \cdots, K-1$. □

*$10\log_{10}N$ is the maximum achievable SNR improvement for beamforming with the proposed phase shifts when $u = \alpha_{m_1}^{t_1}$. While the maximum difference between u and $\alpha_{m_1}^{t_1}$ is π/L, it means with a considerably large L in massive localized hybrid arrays, the practical SNR improvement after beamforming can approximate or even be equal to $10\log_{10}N$ in dB.

From Lemma 2.3, we know that $\sin(N\omega_m^t) = 0$ for $m = 0, 1, \cdots, K-1$, if $G_{\tilde{m}}^t(u) = 0$ for any $\tilde{m} \in [0, K-1]$. Then, $\sin(N\omega_m^t)\sin(N\omega_{m+1}^t) = 0$. There are two scenarios in terms of the value of $\sin(\omega_m^t)$ or $\sin(\omega_{m+1}^t)$.

Scenario 1: $\sin(\omega_m^t) = 0$ or $\sin(\omega_{m+1}^t) = 0$, and $\sin(\omega_m^t) \neq \sin(\omega_{m+1}^t)$, where $0 \leq m \leq K-1$. Without loss of generality, we suppose $\sin(\omega_{\breve{m}}^t) = 0$. Therefore, $\omega_{\breve{m}}^t \in \{\pm a\pi, a \in \mathbb{N}\}$. Given $m \in [0, K-1]$, i.e., $\omega_m^t \in [\frac{u}{2} + \frac{\pi t}{L} - \pi + \frac{\pi}{K}, \frac{u}{2} + \frac{\pi t}{L}]$ (less then half a cycle), $\breve{m}$ is unique and $\sin(\omega_m^t) \neq 0$ for $m \neq \breve{m}$. Therefore, $P_m^t(u) = 0$ for $m \in [0, K-1]$ and $m \neq \breve{m}$; and

$$P_{\breve{m}}^t(u) = \lim_{\omega_{\breve{m}}^t \to \pm a\pi} e^{j(N-1)\omega_{\breve{m}}^t} \frac{\sin(N\omega_{\breve{m}}^t)}{\sin(\omega_{\breve{m}}^t)} \tag{2.27a}$$

$$= (-1)^{\pm a(N-1)} \lim_{\omega_{\breve{m}}^t \to \pm a\pi} \frac{N\cos(N\omega_{\breve{m}}^t)}{\cos(\omega_{\breve{m}}^t)} = N, \tag{2.27b}$$

where (2.27b) is based on the *L'Hospital's rule* [94]. (2.20) can be rewritten as

$$s_m(t)e^{-jm\widehat{Nu}} = \begin{cases} N\tilde{s}(t)e^{2\pi f_D t}e^{j\breve{m}(Nu-\widehat{Nu})} + \xi_{\breve{m}}(t)e^{-j\breve{m}\widehat{Nu}}, & m = \breve{m}; \\ \xi_m(t)e^{-jm\widehat{Nu}}, & m \neq \breve{m}, \end{cases} \tag{2.28}$$

The IDFT of $s_m(t)e^{-jm\widehat{Nu}}$, $m = 0, 1, \cdots, K-1$, can be given by

$$a_{n'}(t) = \tilde{s}(t)e^{2\pi f_D t}Qe^{j\breve{m}(Nu-\widehat{Nu})}e^{jn'(u-\frac{2\pi t}{L})} + z_{n'}(t), n' = 0, 1, \cdots, K-1, \tag{2.29}$$

where $g_{n'}^t(u)$ in (2.22) becomes† $Qe^{jn'(u-\frac{2\pi t}{L})}$, $z_{n'}(t) = \text{IDFT}\{\xi_m(t)e^{-jm\widehat{Nu}}\}$, and $e^{j\breve{m}(Nu-\widehat{Nu})}$ is constant for $n' = 0, 1, \cdots, K-1$, due to the poor accuracy of $\widehat{Nu}$ in this special case (c.f., Theorem 2.1, where $e^{j(Nu-\widehat{Nu})} \to 1$).

Here, u can be unambiguously estimated by evaluating the phase of the cross-

†Given $\omega_{\breve{m}} \in \{\pm a\pi, a \in \mathbb{N}\}$, we have $\frac{u}{2} - \frac{\pi t}{L} \in \{\frac{\breve{m}\pi}{K} \pm a\pi, a \in \mathbb{N}\}$. Both the denominator and numerator of the RHS of (2.22) are equal to zero. By using the *L'Hospital's rule* [94], we have

$$\lim_{u\to\breve{u}} g_{n'}^t(u) = \lim_{u\to\breve{u}} e^{jn'(u-\frac{2\pi t}{L})}e^{j(N-K)(\frac{u}{2}-\frac{\pi t}{L})}\frac{\sin\left(\frac{Nu}{2}-\frac{N\pi t}{L}\right)}{\sin\left(\frac{Ku}{2}-\frac{K\pi t}{L}\right)}$$

$$= (-1)^{\breve{m}(Q-1)\pm a(N-K)}Q\frac{\cos(\breve{m}Q\pi \pm aN\pi)}{\cos(\breve{m}\pi \pm aK\pi)}e^{jn'(u-\frac{2\pi t}{L})} = Qe^{jn'(u-\frac{2\pi t}{L})},$$

where $\breve{u} = \frac{2\breve{m}\pi}{K} + \frac{2\pi t}{L} \pm 2a\pi$.

correlation of two consecutive IDFT points, i.e., $a_{n'}(t)$ and $a_{n'+1}(t)$ in (2.29), as given by

$$r_{n'}(t) = a^*_{n'}(t)a_{n'+1}(t) = Q^2|\tilde{s}(t)|^2 e^{j(u-\frac{2\pi t}{L})} + \tilde{z}_{n'}(t). \tag{2.30}$$

where $\tilde{z}_{n'}(t)$ is the composite noise term in a similar form to (2.24). Eq. (2.30) is independent of $\widehat{Nu}$ which is suppressed through the cross-correlation. In other words, Algorithm 1 does not invalidate Algorithm 2 in the first scenario of the special case.

Scenario 2: $\sin(\omega_m^t) \neq 0$ for $m = 0, 1, \cdots, K-1$. The first term of the right-hand side (RHS) of (2.20) becomes $s_m(t)e^{-jm\widehat{Nu}} = 0$ for any $m = 0, 1, \cdots, K-1$, at the current symbol t. This is the worst-case scenario where the AoA is undetectable at the current symbol. Nevertheless, we can prove that by using another training symbol, the undetectable angle can be unambiguously detected, as dictated in Corollary 2.1.

Corollary 2.1. *Under the new subarray-specific time-varying phase shifts, no AoA remains undetectable for P training symbols, as long as $P \geq 2$ and $\gcd\{P, Q\} = 1$, where $\gcd\{P, Q\}$ takes the greatest common divisor of P and Q.*

Proof. This corollary can be proved by hypothesizing t_1 and $t_2 \in [0, P-1]$ $(t_1 \neq t_2)$ that $\sin(N\omega_m^{t_1}) = \sin(N\omega_m^{t_2}) = 0$, $\sin(\omega_m^{t_1}) \neq 0$, and $\sin(\omega_m^{t_2}) \neq 0\ \forall m$. $|t_1 - t_2| \leq P-1$. The following equalities would hold:

$$\begin{cases} \sin\left(N(\frac{u_n}{2} - \frac{\pi m_1}{K} - \frac{\pi t_1}{L})\right) = 0, \\ \sin\left(N(\frac{u_n}{2} - \frac{\pi m_2}{K} - \frac{\pi t_2}{L})\right) = 0, \end{cases} m_1, m_2 \in [0, K-1],$$

where m_1 and m_2 are the indexes for subarrays at symbol t_1 and t_2, respectively. As a result,

$$\begin{cases} \frac{u_n}{2} - \frac{\pi m_1}{K} - \frac{\pi t_1}{L} = \frac{a_1\pi}{N}, \\ \frac{u_n}{2} - \frac{\pi m_2}{K} - \frac{\pi t_2}{L} = \frac{a_2\pi}{N}, \end{cases} a_1, a_2 \in \{\mathbb{Z}^-, 0, \mathbb{Z}^+\}, \tag{2.31}$$

where $\mathbb{Z}^-$ stands for the set of negative integers. Given $L = PK$ and $N = QK$,

(2.31) can be written as

$$\frac{(t_1 - t_2)Q}{P} + Q(m_1 - m_2) = a_2 - a_1 \in \{\mathbb{Z}^-, 0, \mathbb{Z}^+\}. \tag{2.32}$$

However, $|t_1 - t_2| \leq P - 1$, and therefore, (2.32) cannot hold for $P \geq 2$ and $\gcd\{P, Q\} = 1$, contradicting with the hypothesis. As a result, no AoA remains undetectable across P symbols, if $P \geq 2$ and $\gcd\{P, Q\} = 1$. The corollary is proved. □

2.4 Analysis of AoA estimation accuracy

In this section, we analyze the MSEs of $e^{j\widehat{Nu}}$ and $\hat{u}$ and derive closed-form MSELBs.

2.4.1 MSELB of $e^{j\widehat{Nu}}$

The accuracy of $e^{j\widehat{Nu}}$ is of particular interest, which is the key to recover the Fourier series underlying the received signals of the subarrays. It depends on the SNR of $\rho_m(t)$ in (2.12), where $\tilde{\xi}_m(t)$ is the zero-mean AWGN with power $\sigma_{\tilde{\xi}}^2 = |\tilde{s}(t)|^2(|P_m^t(u)|^2 + |P_{m+1}^t(u)|^2)\sigma_\xi^2$. The signal power of $\rho_m(t)$ is $|\tilde{s}(t)|^4|G_m^t(u)|^2 = |\tilde{s}(t)|^4|P_m^t(u)|^2|P_{m+1}^t(u)|^2$. Therefore, the estimation SNR of $e^{j\widehat{Nu}}$, before the constructive accumulation, can be given by

$$\gamma_{\tilde{s},u}^{Nu} = \frac{|\tilde{s}(t)|^4|P_m^t(u)|^2|P_{m+1}^t(u)|^2}{|\tilde{s}(t)|^2(|P_m^t(u)|^2 + |P_{m+1}^t(u)|^2)\sigma_\xi^2} \leq \frac{|P_m^t(u)|^2 + |P_{m+1}^t(u)|^2}{4}\frac{|\tilde{s}(t)|^2}{\sigma_\xi^2}, \tag{2.33}$$

which is based on the identity inequality that a geometric mean is no greater than the corresponding quadratic mean.

To obtain the unconditional SNR of $\rho_m(t)$, we take the expectation of $\gamma_{\tilde{s},u}^{Nu}$ over $\tilde{s}$ and u, as given by

$$\gamma_{\tilde{s},u} = \mathbb{E}_{\tilde{s},u}\{\gamma_{\tilde{s},u}^{Nu}\} = \mathbb{E}_{\tilde{s}}\left\{\frac{|\tilde{s}(t)|^2}{\sigma_\xi^2}\right\}\mathbb{E}_u\left\{\frac{|P_m^t(u)|^2 + |P_{m+1}^t(u)|^2}{4}\right\} = \frac{1}{2}\gamma_0, \tag{2.34}$$

where $\mathbb{E}_{\tilde{s}}\left\{\frac{|\tilde{s}(t)|^2}{\sigma_\xi^2}\right\} = \frac{1}{N}\gamma_0$, $\gamma_0 (= \mathbb{E}_{\tilde{s}}\left\{\frac{|\tilde{s}(t)|^2}{\sigma_0^2}\right\})$ is the average receive SNR per antenna, $\sigma_0^2 = \frac{\sigma_\xi^2}{N}$ is the noise power at antennas, and $\mathbb{E}_u\left\{\frac{|P_m^t(u)|^2+|P_{m+1}^t(u)|^2}{4}\right\} = \frac{N}{2}$ by exploiting the property of discrete-time Fourier transform (DTFT) that $\mathbb{E}_u\{|P_m^t(u)|^2\} = \mathbb{E}_u\{|P_{m+1}^t(u)|^2\} = N$; see Appendix 7.1.4.

Considering the coherent combination of the cross-correlations in (2.12), the unconditional SNR can be given by $\bar{\gamma}^{Nu} = \frac{(M-1)P}{2}\gamma_0$. From (2.12), we can see that the variance of $e^{j\widehat{Nu}}$ is inversely proportional to the SNR of $\rho_m(t)$. Given the upper bound for the estimation SNR, $\bar{\gamma}^{Nu}$, we finally have

$$\text{MSELB}\{e^{j\widehat{Nu}}\} = \frac{1}{\bar{\gamma}^{Nu}}. \tag{2.35}$$

2.4.2 MSELB of $\hat{u}$

Assuming that $Nu = \widehat{Nu}$ and hence $e^{jm(Nu-\widehat{Nu})} \to 1$ in (2.20), we proceed to analyze the MSELB of $\hat{u}$. $\hat{u}$ is achieved by evaluating the phase of $r_{n'}(t)$ in (2.23). The accuracy of $\hat{u}$ depends on the SNR of $r_{n'}(t)$, where $\tilde{z}_{n'}(t)$ is a zero-mean AWGN with power $\sigma_{\tilde{z}}^2 = \text{Var}\{\tilde{z}_{n'}(t)\} = 2|\tilde{s}(t)|^2|C(u)|^2\sigma_z^2$. $C(u) = e^{j(N-K)(\frac{u}{2}-\frac{\pi t}{L})}\frac{\sin(\frac{Nu}{2}-\frac{N\pi t}{L})}{\sin(\frac{Ku}{2}-\frac{K\pi t}{L})}$. σ_z^2 is the noise power of $z_{n'}(t) = \text{IDFT}\{\xi_m(t)e^{-jm\widehat{Nu}}\}, m = 0, 1, \cdots, K-1, n' = 0, 1, \cdots, K-1$, satisfying that $\sigma_z^2 = \frac{\sigma_\xi^2}{K}$. σ_ξ^2 is the power of $\xi_m(t)$, i.e., the AWGN at the m-th subarray and the t-th symbol. Typically, $\xi_m(t)$ is i.i.d across different subarrays and symbols.

From (2.23) and (2.24), the conditional SNR for the estimation of u on $\tilde{s}(t)$ and u can be given by

$$\gamma_{\tilde{s},u} = \frac{m_{r_{n'}(t)}^2}{\sigma_{\tilde{z}}^2} = \frac{|\tilde{s}(t)|^2|C(u)|^2}{2\sigma_z^2}, \tag{2.36}$$

where $m_{r_{n'}(t)} = |\tilde{s}(t)|^2|C(u)|^2 e^{j(u-\frac{2\pi t}{L})}$ according to (2.23).

Taking the expectation of $\gamma_{\tilde{s},u}$, we can obtain the unconditional SNR, as given

by

$$\begin{aligned}\gamma = \mathbb{E}\{\gamma_{\tilde{s},u}\} &= \mathbb{E}\left\{\frac{|\tilde{s}(t)|^2}{2\sigma_z^2}\right\}\mathbb{E}_u\left\{|C(u)|^2\right\} \\ &= \frac{K}{2}\gamma_\xi \mathbb{E}_u\left\{|C(u)|^2\right\}, \end{aligned} \tag{2.37}$$

where $\gamma_\xi = \frac{\mathbb{E}\{|\tilde{s}(t)|^2\}}{\sigma_\xi^2}$, and $\mathbb{E}_u\{\cdot\}$ takes expectation over u.

Despite the proposed algorithm is general and does not rely on any specific distribution of the AoA, i.e., in (2.3), we assume a uniform distribution of u over $[-\pi, \pi)$ to analyze the estimation accuracy of the AoA, as extensively assumed in the literature [1, 3, 4, 17]. As a result,

$$\begin{aligned}\mathbb{E}_u\left\{|C(u)|^2\right\} &= \frac{1}{2\pi}\int_{-\pi}^{\pi}|C(u)|^2\,\mathrm{d}u \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin(\frac{Nu}{2}-\frac{N\pi t}{L})}{\sin(\frac{Ku}{2}-\frac{K\pi t}{L})}\right|^2\mathrm{d}u. \end{aligned} \tag{2.38}$$

The integral of the RHS of (2.38) can be calculated through the property of the Fourier transform, see Appendix 7.1.4. Different from Appendix 7.1.4, we use the K-decimation of $r_N(n)$, denoted by $r_N(mK)$. $r_N(mK) = 1$, if $m = 0, 1, \cdots, N/K - 1$; or $r_N(mK) = 0$, otherwise [95]. The DTFT of $r_N(mK)$, termed $f(u)$, can be given by $f(u) = \sum_{m=0}^{N/K-1} r_N(mK)e^{-juKm} = e^{j\frac{N-K}{2}u}\frac{\sin(Nu/2)}{\sin(Ku/2)}$. Similarly, we have $\mathbb{E}_u\{|C(u)|^2\} = N/K$. Substituting (2.38) into (2.37) , we obtain

$$\gamma = \frac{N}{2}\gamma_\xi. \tag{2.39}$$

Consider the coherent accumulation of the cross-correlations in (2.23) over $n' \in [0, K-2]$, $l \in \left[0, \lceil M/K\rceil - 1\right]$, and $t \in [0, P-1]$. Since $\sigma_\xi^2 = N\sigma_0^2$, the unconditional SNR for the estimation of u is given by

$$\bar{\gamma} = \frac{P\lceil M/K\rceil(K-1)N}{2}\gamma_\xi = \frac{P\lceil M/K\rceil(K-1)}{2}\gamma_0, \tag{2.40}$$

where $\lceil\cdot\rceil$ stands for ceiling.

In high SNR regions of flat-fading Rayleigh fading channels, the probability density function (pdf) of $\hat{u}$ is given by [3, 93, 96]

$$\tilde{f}(\hat{u},\bar{\gamma}) \approx \frac{\sqrt{\bar{\gamma}\pi^2+1}}{2\pi(1+\bar{\gamma}\hat{u}^2)^{3/2}}, \quad -\pi \leq u \leq \pi. \tag{2.41}$$

The MSE of $\hat{u}$ now can be calculated by

$$\begin{aligned}\sigma_{\hat{u}}^2 &= \int_{-\pi}^{\pi} (\hat{u})^2 \tilde{f}(\hat{u},\bar{\gamma})\, \mathrm{d}\hat{u} \\ &= \frac{\sqrt{\bar{\gamma}\pi^2+1}}{\pi\bar{\gamma}^{3/2}} \sinh^{-1}\left(\sqrt{\bar{\gamma}}\pi\right) - \frac{1}{\bar{\gamma}}.\end{aligned} \tag{2.42}$$

Note that (2.42) is under the assumption that $Nu = \widehat{Nu}$. Moreover, $(\lceil M/K \rceil K - M)$ from the outputs of the first K subarrays are reused to form the last K-point IDFT with the outputs of the last $M - \lfloor M/K \rfloor K$ subarrays. $\tilde{z}_{n'}(t)$ for $l = \lceil M/K \rceil$ can be correlated to those for $l = 1$, and therefore, (2.40) provides an upper bound for the unconditional SNR. As a result, (2.42) provides an MSELB for $\hat{u}$. However, it is clear that the MSELB is asymptotically tight, in other words, it gives the asymptotic expression for the MSE, as M and P become large. This is because $\widehat{Nu} \to Nu$ and the contribution of the last K-point IDFT per symbol to the estimation becomes increasingly negligible.

There is an opportunity to design the parameters P and K by using (2.42). Let σ_e^2 denote the target MSE of AoA estimation to be achieved. The required unconditional SNR, $\bar{\gamma}$, can be calculated by (2.42). From (2.40), $\bar{\gamma}$ is a function of P, K and γ_0, as N and M typically depend on the design technologies of mmWave antenna front-end, cost, and system complexity [97–99], and they are fixed a priori. γ_0 can be estimated based on system specifications, such as the maximum detectable range $R_{\max}$, and the corresponding smallest detectable signal power $P_{\min}$[16, 100]. From (2.40), the MSELB decreases with P, while K can take only a limited number of values, since $N/K \in \mathbb{Z}^+$ and $2 < K \leq M$. We can examine all possible values of P and K in (2.42), and choose those meeting the target. There would be a gap

between the MSELB and the actually achievable MSE given P and K, especially when the SNR is low. However, the gap is consistent, as revealed by extensive simulations in Section 2.5, and therefore can be compensated off-line.

2.5 Numerical and Simulation Results

In this section, simulation results are provided to validate the proposed approach and analysis, where uniform linear localized hybrid arrays are considered. The array configurations and other simulation parameters are provided in Table I; unless otherwise specified. We also simulate the state of the art, referred to as differential beam search (DBS) [3] and double cross-correlation (DCC) [4].

The comparisons between the proposed approach and the state of the art are fair; or otherwise, in favor of existing approaches, such as DBS [3], by allowing more training symbols. This is because DBS needs more training symbols to iteratively scan multiple possible estimates resulting from estimation ambiguities. The proposed approach is simulated typically with far fewer training symbols than DBS. Other simulation parameters are set to be identical between the proposed approach and the algorithms in [3] and [4]. For example, the same average receive SNR per antenna, γ_0, is taken for fair comparisons between different approaches under different array/symbol configurations.

Note that the proposed approach and DCC can accommodate any number of training symbols. Without loss of generality, for the proposed approach, we set $P = 4$ (unless otherwise specified), i.e., the phase shifts repeat every 4 symbols till the end of training symbols. Even if the number of training symbols is unequal to P, the symbols can be constructively combined and the AoA can be estimated in the same way as described in Section 2.3. The estimation accuracy can be analyzed in the same way as described in Section 2.4 (we just need to replace P with the number of training symbols). In contrast, DBS sequentially scans the N possible estimates of u, i.e., $\hat{u} = \text{mod}\left\{\frac{\widehat{Nu}+2n\pi}{N}, 2\pi\right\} - \pi, n = 0, 1, \cdots, N-1$, given $\widehat{Nu}$, due to the phase ambiguity. This limits the feasible number of training symbols to multiples of N. For fair comparisons between the proposed approach and the two benchmarks, we

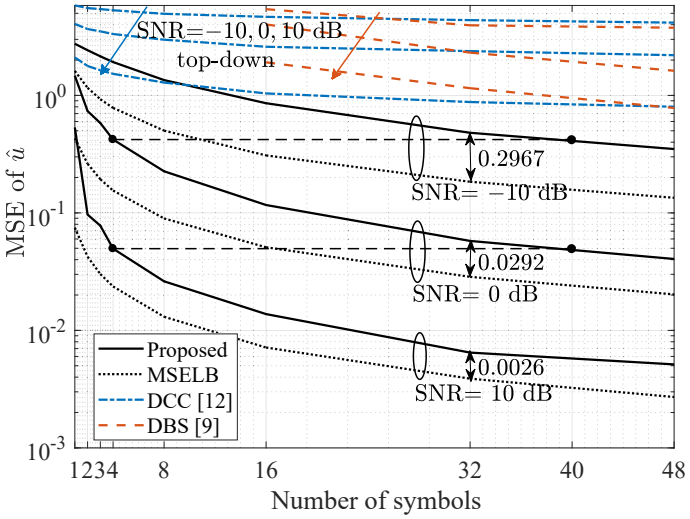

Figure 2.2 : MSE of $\hat{u}$ versus the number of symbols (also iterations for DBS [3]), where $M = 8$ and $N = 16$. For DBS [3], the AoA estimation can only be updated every 16 symbols, with result output at $P = 16, 32$ and 48.

always keep the number of training symbols consistent across the three algorithms.

Table 2.1 : Configurations of the localized hybrid array and other parameters

Parameter	Value	Parameter	Value
M	8	K $(\leq M)$	8
N	16	Q $(= N/K)$	2
N_{trials}*	4×10^4	P	4
AoA	$U_{[-\pi,\pi]}$**	SNR***	$[-10, 0, 10]$ dB

* N_{trials} is the number of independent trials for each SNR. The curves in all figures are the average of 4×10^4 independent trials.

** $U_{[-\pi,\pi]}$ represents the uniform distribution on $[-\pi, \pi]$.

*** The SNR is antenna-wise, i.e. γ_0 in (2.40).

Fig. 2.2 plots the MSEs of the proposed approach and the state of the art, as the number of training symbols increases. We show that the proposed algorithm is able to dramatically and increasingly outperform the state of the art with the increasing number of training symbols. The proposed approach can reduce the MSE by orders of magnitude. Particularly, the MSEs of the proposed approach keeps declining, due to the increasingly improved SNR. In contrast, the MSE of DCC flats out and DBS is far from convergence, as the consequence of the aforementioned phase

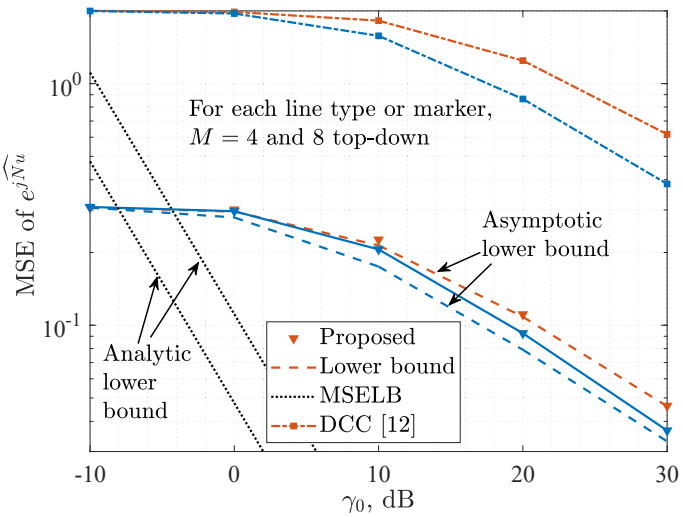

Figure 2.3 : MSE of $e^{j\widehat{Nu}}$ as γ_0 increases, where $N = 24$, and $P = 6$.

ambiguities in both methods. We also see that the MSELB becomes increasingly tight to the proposed approach, as the SNR grows[‡]. This validates our analysis of the MSELB, and demonstrates the superiority of the proposed approach in estimation accuracy. It is worth pointing out that DBS requires recursive search necessitating multiple training symbols to suppress the phase ambiguity, and therefore provides no estimation results until the 16-th symbol.

By increasing the training symbols, the estimation accuracy improves consistently with the estimation SNR in Fig. 2. For instance, by increasing the number of symbols P from 4 to 40, the estimation SNR increases by 10 dB. The case of $P = 4$ and $\gamma_0 = 0$ dB has the same estimation SNR as the case of $P = 40$ and $\gamma_0 = -10$ dB. γ_0 is the average receive SNR per antenna element. The MSEs of the two cases are 0.4211 and 0.4152, close to each other, as shown in Fig. 2. The case of $P = 4$ and $\gamma_0 = 10$ dB has the same estimation SNR as the case of $P = 40$ and $\gamma_0 = 0$ dB. The MSEs of the two cases are 0.0491 and 0.04962, even closer.

Fig. 2.3 plots the MSE of $e^{j\widehat{Nu}}$ as the receive SNR per antenna increases. Apart from the closed-form MSELB and simulation results, we also plot the MSE under the assumption that the strongest cross-correlation between the complex gains of

[‡]DBS, as an iterative searching approach, requires a large number of iterations to converge, e.g., more than 400 symbols [3, Fig. 10].

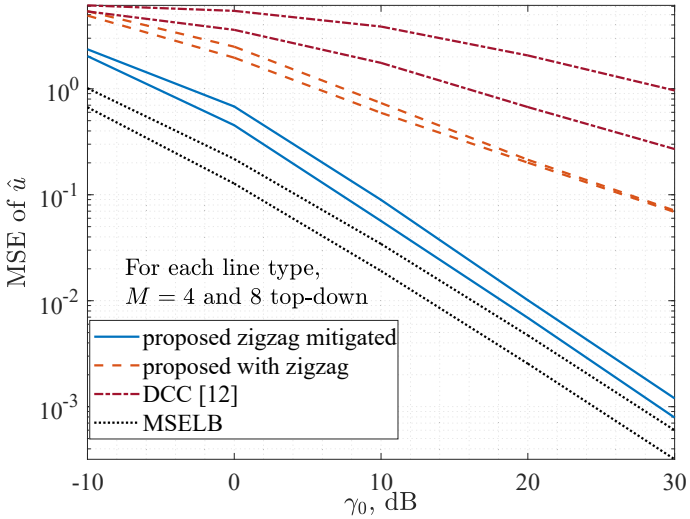

Figure 2.4 : MSE of $\hat{u}$ versus γ_0, where $N = 24$, $M = 4, 8$ and the number of training symbolsis 6. Both results with and without the proposed zigzag mitigation are provided, along with that of DCC [4] for comparison.

adjacent subarrays is always correctly identified and calibrated in sign. It provides the asymptotic MSELB of $e^{j\widehat{Nu}}$. By comparing the asymptotic MSELB and simulation results, we see that our discovery in Theorem 1 can help effectively identify and calibrate the strongest cross-correlation, and constructively accumulate the cross-correlations to approach the asymptotic MSELB. We also see that the MSELB is fairly loose due to the relaxation based on the identity inequality that a geometric mean is no greater than the corresponding quadratic mean; see (2.33). We further show that the proposed approach is able to dramatically improve the accuracy of $e^{j\widehat{Nu}}$, as compared to the state of the art. This contributes significantly to the improved accuracy of $\hat{u}$ in the proposed algorithm, as will be shown in Fig. 2.4.

Fig. 2.4 demonstrates the effectiveness of the proposed zigzag mitigation method described in (2.25) and (2.26), where, for comparison purpose, we also plot the MSEs of DCC, the MSELB and the estimation accuracy of the proposed algorithm without mitigating the zigzag effect. We can see that the zigzag effect can substantially compromise the estimation of u, even in the case of high SNRs. As a consequence, the estimation accuracy of the azimuth ϕ degrades, since u and ϕ are one-to-one mapping. We also see that the proposed mitigation of the zigzag effect in (2.25) and

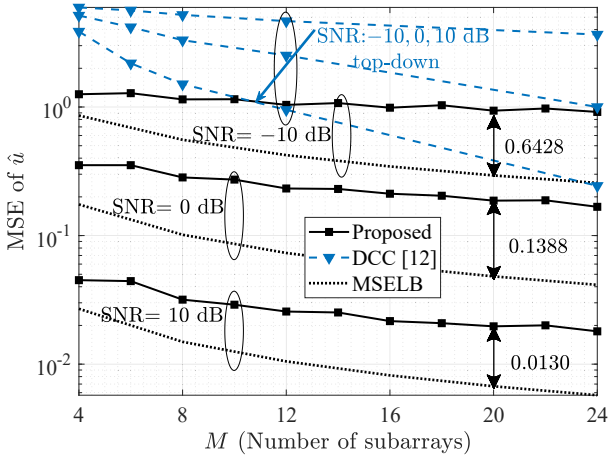

Figure 2.5 : MSE of $\hat{u}$ versus the number of subarrays M, where $N = 24$ and the number of training symbols is 8. For the proposed approach, M ranges from 4 to 24. While for [4], only five values can be taken, i.e. $4, 6, 8, 12$ and 24, to keep $N/M \in \mathbb{Z}^+$.

(2.26) becomes increasingly effective as the SNR increases. This is important for the proposed algorithm to approach the MSELB, as the SNR increases, as shown in Fig. 2.2.

It is worth pointing out that though mitigating the zigzag appears to dominate in the improvement of estimation accuracy in Fig. 2.4, yet in fact, the sign alignment based on our discovery in Theorem 2.1 has the dominating effort on the improvement. The sign alignment is able to reduce the MSE dominantly in absolute value, and lay the foundation for the further reduction through the proposed mitigation of the zigzag effect.

Fig. 2.5 compares the proposed approach and DCC [4] with the increasing number of subarrays M. We can confirm that the proposed approach is able to operate in general localized hybrid arrays where N is unnecessarily a multiple of M. Particularly, M can take any integer that is larger than K. In contrast, DCC only necessitates N to be a multiple of M. Given $N = 24$, M can only take the values of $4, 6, 8, 12$ and 24. As also corroborated in Fig. 2.5, the proposed approach is able to significantly and consistently outperform the state of the art. This is consistent with Fig. 2.2. Moreover, Fig. 2.5 shows that both algorithms increasingly improve their estimation accuracy with the growing number of subarrays. The conclusion

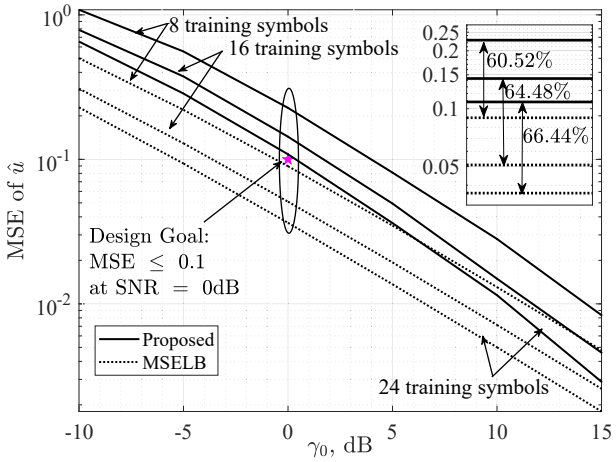

Figure 2.6 : MSE of $\hat{u}$ and MSELB versus γ_0, where $N = 16$ and $M = 8$.

drawn is that the increases of digital capability and the number of phase shifts per symbol help improve the accuracy of AoA.

Fig. 2.6 plots MSEs of $\hat{u}$ using the proposed algorithm with the growth of SNR, where $N = 16$, $M = 8$, and the number of training symbols is set to be $8, 16$ and 24. We observe that the gap between the actually achievable MSEs and the corresponding MSELBs keeps consistent, under different numbers of training symbols. Consider the SNR of 0 dB. The gap is 60.52% in the case of 8 training symbols, and the corresponding gap is 64.48% and 66.44% in the cases of 16 and 24 training symbols, respectively.

In light of this observation, we can design the number of training symbols to achieve a given requirement of MSE, through MSELB developed in (2.42). Suppose that the MSE requirement of AoA estimation is 0.1 at 0 dB, where $N = 16$, $M = 8$, and $K = 8$. Substituting the MSELB of 0.1 and the values of M and K into (2.42), we need more than 7 training symbols. Take 8 training symbols for simulations. The achievable MSE is 0.2272 at 0 dB, $55.99\%\big(= \frac{0.2272-0.1}{0.2272} \times 100\%\big)$ higher than the MSE requirement. Observed in Fig. 2.6, the MSELB corresponding to the achievable MSE with 8 training symbols is 0.08969. As a result, the MSELB, which provides the actual MSE of 0.1, can be predicted to be around $0.0395 (= 0.08969 \times (1 - 55.99\%))$. Substituting this in (2.42), we can see more than 22 training symbols are required.

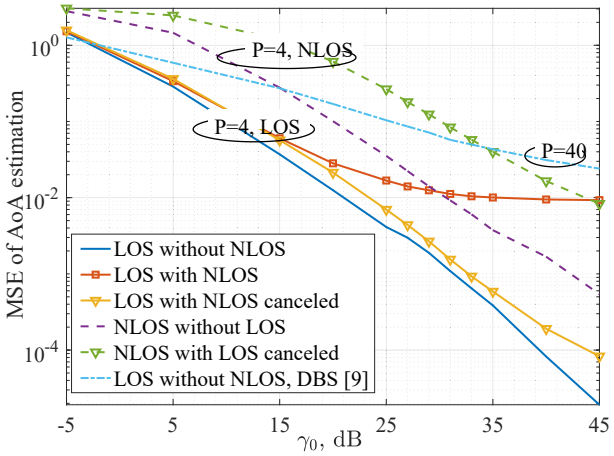

Figure 2.7 : MSE of $\hat{u}$ in the presence of a LoS path and a NLoS path, where $M = 4$, $N = 8$, and $P = 4$ for the proposed algorithm and $P = 40$ for DBS [3]. The received signal strength of the LoS path is set to be 5 dB stronger than that of the NLoS path, the LoS path is uniformly distributed within $[\frac{\pi}{4}, \frac{\pi}{3}]$ and the NLoS path is uniformly distributed within $[-\frac{\pi}{3}, -\frac{\pi}{4}]$.

Simulations confirm that 24 training symbols can provide an actually achievable MSE of 0.1084, which meets the MSE requirement.

Fig. 2.7 plots the estimation accuracy in the presence of a LoS path and a non-line-of-sight (NLoS) path, where the received signal strength of the LoS path is set to be 5 dB stronger than that of the NLoS path. Without loss of generality, the LoS path is uniformly distributed within $[\frac{\pi}{4}, \frac{\pi}{3}]$, and the NLoS path is uniformly distributed within $[-\frac{\pi}{3}, -\frac{\pi}{4}]$. SIC is carried out in coupling with the proposed approach to iteratively improve the estimation accuracies of both the LoS and NLoS paths. We also plot the proposed approach and DBS [9] in the absence of NLoS for references. We can see that, coupled with SIC, the proposed approach is able to substantially improve the estimation accuracy of both the LoS and NLoS paths, especially in the high SNR region. The improved accuracy of the LoS path is close to the lower bound performance achieved by the proposed approach in the absence of NLoS. We also see that the estimation accuracy of the LoS path by using the proposed approach can even be much lower in the presence of NLoS than that by using DBS in the absence of NLoS. It is indicated that estimation ambiguities can be as a significant source of inaccuracy as NLoS. The proposed approach eliminating

the ambiguities helps compensate for inaccuracy resulting from NLoS.

It is worth noting that the comparison study is in favor of DBS in Fig. 2.7, where $P = 4$ for the proposed approach while $P = 40$ for DBS. This is due to the fact that DBS has to repeatedly scan multiple estimates of u by using more training symbols. It is also interesting to point out that the curve of "LoS with NLoS", with the NLoS path deemed as interference, starts to saturate at $\gamma_0 \approx 23$dB, despite the per-antenna SIR is 5dB. This is because, by running Algorithm 2, the interference can be decoupled to a randomness-bearing part and a deterministic part with a deterministic phase and gain. Our analysis reveals that the per-antenna SIR and γ_0 are 5dB and 23.5354dB, respectively, when the random part of interference and the noise have the same power after coherent accumulation. The random part of interference can increasingly dominate over the noise, as γ_0 grows further. The deterministic part of the interference can lead to a noise-independent bias upon the AoA estimation of the LoS path. It can affect the saturated estimation accuracy; but would not affect the threshold of γ_0, beyond which the estimation accuracy saturates. Details are provided in Appendix 7.1.5.

2.6 Conclusion

In this chapter, we propose high-accuracy AoA estimation with new subarray-specific time-varying phase shifts in general localized hybrid arrays. We discover that the signs of the cross-correlations between consecutive subarrays are deterministic, and only the strongest cross-correlation takes a different sign from the rest. We propose to align the cross-correlations, achieving constructive combinations and improving tolerance to noises. Evident from extensive simulations, the estimation accuracy can be substantially improved by orders of magnitude through our design of phase shifts and our discovery, and asymptotically approach the MSELB. In our future work, we will extrapolate our discovery to more sophisticated wideband and multi-path environments.

Chapter 3

Fast and Accurate Estimation of Angle-of-arrival for Wideband Large-scale Hybrid Arrays

3.1 Introduction

Accurate estimation of AoA is critical to wideband satellite communications, but is susceptible to receive noises and can be ambiguous due to space/cost-effective hybrid antenna array designs with localized analog phased subarrays. As a matter of fact, there has yet to be an unambiguous estimator even for narrow-band systems. This chapter presents a novel and efficient wideband AoA estimator for unambiguously and accurately estimating the AoA of wideband impinging signals at satellite-borne localized hybrid antenna arrays. Embarking on new discoveries, the chapter addresses the critical challenge of estimation ambiguity which has yet to be addressed in the literature, and achieves superior estimation accuracy and efficiency in comparison to the state of the art, e.g., [1]. The contributions of the chapter can be summarized as follows.

1. We reveal that, by partitioning a hybrid array into groups of K consecutive subarrays (K is a factor of the number of antennas per subarray, N) and applying linearly increasing phase shifts per group, the received signals of each group, after a phase correction of Nu, can form a Fourier series at every sub-carrier with the AoA unambiguously captured in coefficients. To this end, the AoA can be efficiently estimated at every sub-carrier, by correcting phases of received signals, running IDFT, correlating the Fourier coefficients for each K-subarray group and accumulating coherently among groups.

2. We discover unprecedentedly that the cross-correlations of the received signals between adjacent subarrays, before corrupted by noises, have deterministically

changed phases at each sub-carrier, and can be reliably augmented and coherently accumulated to achieve a noise-resistant estimate of the phase to be corrected, i.e., Nu, in 1).

3. We extrapolate the sub-carrier-wise operations of 1) and 2) optimally to wideband signals and improve estimation accuracy substantially by exploiting coherence across sub-carriers. Specifically, the subarray-wise cross-correlations are correlated between sub-carriers and accumulated coherently to enhance the phase correction. We identify the optimal gap between sub-carriers to be correlated, balancing between the decreasing accumulation gain and the alleviating error amplifying effect of an increasing gap. We also prove that, after the phase correction, the optimal coherent accumulation gain can be readily achieved for the AoA estimation by adding up the cross-correlations of the Fourier coefficients across all sub-carriers.

Corroborated by extensive simulations and analysis, the proposed approach is able to significantly improve the estimation accuracy and reduce the estimation delay by up to 99.949% and 70% for the dominant path, respectively, even in low SNR regions, as compared to the state of the art. The proposed approach is also demonstrated to asymptotically approach the theoretical limit of the error variance exhibited by the estimates. To this end, the approach is uniquely suited for satellite-borne wideband multi-antenna communication systems which typically require high accuracy of AoA and operate under low SNRs.

The rest of the chapter is organized as follows. In Section 3.2, the system model and the problem statement are described. In Section 3.3, the new phase shifts are designed and the deterministic nature of the cross-correlations between subarrays is discovered. In Section 3.4, we optimize the frequency interval, followed by an unambiguous estimation of u in Section 3.5. In Section 3.6, the analysis and extensions of our approach are provided. The superiority of the proposed approach is demonstrated through simulations in Section 3.7, followed by conclusions in Section 3.8.

3.2 System Model and Problem Statement

For illustration convenience, we consider a one-dimensional linear localized hybrid array with M analog subarrays and N antenna elements per subarray, as shown in Fig. 1.2.* The antennas in the same subarray are connected to adjustable phase shifters, one for each, and then aggregated and connected to a RF chain. As a receiver, the RF chain consists of an amplifier, a down-converter and filters and down-converts RF signals into intermediate or baseband frequencies. The baseband signals are digitized by an ADC and transformed into the frequency domain.

The wideband satellite transmissions in Q/V band are dominated by directive LoS radio links with insignificant multi-path components [27]. Let $\tilde{s}(t)$ denote the complex incoming time-domain signal in the baseband. Assume that the frequency range of $\tilde{s}(t)$, is $[f_l, f_h]$, and $B = f_h - f_l$ is the frequency bandwidth. The spacing between adjacent antenna elements, denoted by d, is half the wavelength of the highest frequency, i.e. $d = \frac{\lambda_l}{2}$, where $\lambda_l = \frac{c}{f_h}$ and c is the light speed.

Every wideband symbol is digitalized, generating I samples. At symbol t, the i-th sample $\tilde{s}(t)$ is taken at $t_i = \frac{tI+i}{f_s}$, where f_s is the sampling frequency, t is the index for symbols ($t \geq 0$) and i is the index for samples within a symbol and hence the index for sub-carriers ($0 \leq i \leq I-1$). Suppose that the incoming direction of $\tilde{s}(t_i)$ is θ. The time-domain output of the m-th subarray can be written as [1]

$$s_m(t_i) = \sum_{n=0}^{N-1} P_m^n(\theta)\tilde{s}(t_i - \tau_n^m(\theta))e^{jn\alpha_m^t} + z_m(t_i), \tag{3.1}$$

where $P_m^n(\theta)$ is the radiation pattern of the n-th antenna element at the m-th subarray ($0 \leq n \leq N-1, 0 \leq m \leq M-1$); $z_m(t_i)$ is the zero-mean AWGN at the m-th subarray and is typically independent and identically distributed (i.i.d); $\tau_m^n(\theta) = \frac{(nd+mNd)\sin\theta}{c}$ is the propagation delay of the n-th antenna element at the m-th subarray with reference to the left-most antenna element; and α_m^t is the phase shift value of the m-th subarray at symbol t. The phase shifts are assumed to be

*Note that the algorithms proposed in this chapter can be readily applied in two-dimension rectangular hybrid arrays.

frequency flat [1, 101], i.e. the same phase shifts for all I samples of symbol t.

For the t-th wideband symbol, the I-point DFT of $s_m(t_i)$, denoted by $S_m(f_i), 0 \leq i \leq I-1$, can be given by

$$\begin{aligned} S_m(f_i) &= \sum_{n=0}^{N-1} P_m^n(\theta)\tilde{S}(f_i)e^{j2\pi f_i \tau_m^n(\theta)}e^{jn\alpha_m^t} + Z_m(f_i) \\ &= \tilde{S}(f_i)P_m(\rho_i, u)e^{jm\rho_i Nu} + Z_m(f_i), \end{aligned} \tag{3.2}$$

where $\tilde{S}(f_i) = \text{DFT}\{\tilde{s}(t_i)\}$; $Z_m(f_i) = \text{DFT}\{z_m(t_i)\}$; $u = \frac{2\pi d \sin\theta}{\lambda_l} = \pi \sin\theta$, since $d = \frac{\lambda_l}{2}$; $P_m(\rho_i, u) = \sum_{n=0}^{N-1} P_m^n(\theta)e^{jn\rho_i u}e^{jn\alpha_m^t}$ is the beamforming gain of the m-th subarray under α_m^t; and ρ_i is the normalized frequency:

$$\rho_i = \frac{f_i + f_l}{f_h} \in \left[\frac{f_l}{f_h}, 1\right], \tag{3.3}$$

where $f_i = \frac{iB}{I}, 0 \leq i \leq I-1$, is the centroid frequency of the i-th sub-carrier in the baseband. Different from narrowband scenarios, $P_m(\rho_i, u)$ depends on both u and f_i through ρ_i. The estimation of AoA, or more specifically, θ, is essentially to estimate u, since u and θ are one-to-one mapping, i.e., $\theta = \sin^{-1}(\frac{u\lambda_l}{2\pi d})$.

Take the cross-correlation of the frequency-domain outputs between two consecutive analog subarrays at the i-th sub-carrier ($0 \leq i \leq I-1$), as given by

$$R_m(f_i) = S_m^*(f_i)S_{m+1}(f_i) = |\tilde{S}(f_i)|^2 G_m(\rho_i, u)e^{j\rho_i Nu} + \tilde{Z}_m(f_i), 0 \leq m \leq M-2, \tag{3.4}$$

where $G_m(\rho_i, u) = P_m^*(\rho_i, u)P_{m+1}(\rho_i, u)$ and the composite noise component $\tilde{Z}_m(f_i)$ can be given by

$$\begin{aligned} \tilde{Z}_m(f_i) =& \tilde{S}^*(f_i)P_m^*(\rho_i, u)e^{-jm\rho_i Nu}Z_{m+1}(f_i) + \tilde{S}(f_i)P_{m+1}(\rho_i, u)e^{j(m+1)\rho_i Nu}Z_m^*(f_i) \\ &+ Z_m^*(f_i)Z_{m+1}(f_i), \end{aligned} \tag{3.5}$$

where $\tilde{Z}_m(f_i)$ can be modeled as a zero-mean AWGN, since $Z_m(f_i), 0 \leq m \leq M-1, 0 \leq i \leq I-1$, are i.i.d AWGNs. Provided the sign of $G_m(\rho_i, u)$ is known in (3.4),

$\rho_i Nu$ could be estimated by evaluating the phase of $R_m(f_i)$, but an ambiguity would arise from estimating u.

A second cross-correlation of $R_m(f_i)$ over f_i can be taken to eliminate the ambiguity [1], as given by

$$\begin{aligned}\tilde{R}_m(f_i) = R_m^*(f_i)R_m(f_{i+\triangle_i}) = |\tilde{S}(f_i)|^2|\tilde{S}(f_{i+\triangle_i})|^2 G_m^*(\rho_i, u)G_m(\rho_{i+\triangle_i}, u)e^{j\delta_{\rho_i}Nu} \\ + \breve{Z}_m(f_i), \quad 0 \le i \le I - \triangle_i - 1 \end{aligned} \tag{3.6}$$

where $\triangle_i$ is the difference of indexes for sub-carriers and $\delta_{\rho_i} = \rho_{i+\triangle_i} - \rho_i = \frac{\triangle_i B}{I f_h}$. $\breve{Z}_m(f_i)$, like $\tilde{Z}_m(f_i)$, is also a zero-mean AWGN and can be written as

$$\begin{aligned}\breve{Z}_m(f_i) =& |\tilde{S}(f_i)|^2 G_m^*(\rho_i, u)e^{-j\rho_i Nu}\tilde{Z}_m(f_{i+\triangle_i}) \\ &+ |\tilde{S}(f_{i+\triangle_i})|^2 G_m(\rho_{i+\triangle_i}, u)e^{j\rho_{i+\triangle_i}Nu}\tilde{Z}_m^*(f_i) \\ &+ \tilde{Z}_m^*(f_i)\tilde{Z}_m(f_{i+\triangle_i}).\end{aligned} \tag{3.7}$$

Provided $\text{sgn}\{G_m^*(\rho_i, u)G_m(\rho_{i+\triangle_i}, u)\} > 0$, we can carefully choose $\delta_{\rho_i}N \le 1$ so that $\triangle_i \le \frac{I f_h}{NB}$ and u can be estimated unambiguously by evaluating the phase of $\tilde{R}_m(f_i)$ [1]. The estimate of u, denoted by $\hat{u}$, is $\hat{u} = \frac{1}{\delta_{\rho_i}N}\arg\{\tilde{R}_m(f_i)\}$, which, unfortunately, has the estimation error amplified by $\frac{1}{\delta_{\rho_i}N} \ge 1$. $\text{sgn}\{\cdot\}$ takes sign and $\arg\{\cdot\}$ takes angle. From $\tilde{R}_m(f_i)$ in (3.6), $\text{sgn}\{G_m^*(\rho_i, u)G_m(\rho_{i+\triangle_i}, u)\} > 0$, $\forall m$ allows for coherent accumulation across subarrays and sub-carriers, hence improving noise resistence and estimation accuracy. In [1], $\text{sgn}\{G_m^*(\rho_i, u)G_m(\rho_{i+\triangle_i}, u)\} > 0$ was achieved by applying identical phase shifts for all subarrays, but led to iteratively scanning for initialization with a large number of training symbols and long estimation delays. Additionally, the convergent performance, i.e., estimation accuracy, would be affected by the selection of the initial phase shifts and error propagation. Other relevant algorithms, such as [4, 17], cannot guarantee consistent signs at all.

3.3 New Design of Phase Shifts

In this section, we propose to use different phase shifts across subarrays and symbols. With few symbols (even a single symbol at high receive SNR regions), the AoA can be estimated unambiguously with high accuracy. Particularly, the new phase shifts are able to address the aforementioned issue of inconsistent signs of $G_m(\rho_i, u)$.

With M subarrays and N antennas per subarray, the new phase shift values are given by

$$\alpha_m^t = \frac{2\pi(\text{mod}\{m, K\}P + \text{mod}\{t, P\})}{L}, \tag{3.8}$$

where m and t are the indexes for analog subarrays and training symbols, respectively ($0 \leq m \leq M-1,\ t \geq 0$); K is the number of different phase shifts adopted per symbol; P is the number of training symbols, and $L = PK$ is the total number of phase shifts. The L phase shifts are evenly spaced across $[0, 2\pi)$ with an interval of $\frac{2\pi}{L}$. The modulo operations, $\text{mod}\{\cdot\}$, indicate that α_m^t has the cycles of K and P for subarrays and symbols, respectively.

Here, K is designed to satisfy:

$$2<K \leq M,\ N = QK,\ K, Q \in \mathbb{Z}^+, \tag{3.9}$$

where $\mathbb{Z}^+$ is the set of positive integers, and Q is an auxiliary integer variable. For every symbol, K different phase shifts repeat every K subarrays.

By substituting (3.8) into (3.2), the beam pattern of the m-th subarray at any symbol t and sub-carrier i ($0 \leq m \leq K-1$, $0 \leq t \leq P-1$, and $0 \leq i \leq I-1$), i.e., $P_m(\rho_i, u)$, can be given by

$$\begin{aligned} P_m(\rho_i, u) &= \sum_{n=0}^{N-1} e^{jn\rho_i u} e^{-jn\alpha_m^t} = \sum_{n=0}^{N-1} e^{jn\rho_i u} e^{-j2\pi \frac{\text{mod}\{m,K\}P + \text{mod}\{t,P\}}{L} n} \\ &= e^{j(N-1)\omega_m^i} \frac{\sin(N\omega_m^i)}{\sin(\omega_m^i)}, \end{aligned} \tag{3.10}$$

where $\omega_m^i = \frac{\rho_i u}{2} - \pi\left(\frac{t}{L} + \frac{m}{K}\right)$. As extensively assumed in [1, 3, 4, 17], we set $P_m^n(\theta) = 1$, i.e., the antenna elements are omni-directional with unitary antenna gains, and can be suppressed.

From (3.10), we can see that $P_m(\rho_i, u)$ is periodic over m and t, with cycles of K and P, respectively. For illustration convenience, we take $m = 0, 1, \cdots, K-1$ and $t = 0, 1, \cdots, P-1$ in the following. The conclusion drawn can be readily applied to the rest $(M - K)$ subarrays within the P symbols, as will be discussed later.

As a result, $G_m(\rho_i, u)$ in (3.4) can be rewritten as

$$G_m(\rho_i, u) = e^{\frac{-j(N-1)\pi}{K}} \frac{\sin(N\omega_m^i)}{\sin(\omega_m^i)} \frac{\sin(N\omega_{m+1}^i)}{\sin(\omega_{m+1}^i)}, \tag{3.11}$$

which has deterministic and predictable signs, as dictated in Theorem 3.1. Two lemmas are provided before the theorem, as they play important roles in the proof of the theorem.

Lemma 3.1. *There exists a unique integer* $m' \in [0, K-1]$ *satisfying*

$$\sin(\omega_{m'}^i)\sin(\omega_{m'+1}^i) < 0.$$

Proof. See Appendix 7.2.1. □

Lemma 3.2. *Given* m' *specified in Lemma 3.1, we have*

$$|\sin(\omega_{m'}^i)\sin(\omega_{m'+1}^i)| < |\sin(\omega_m^i)\sin(\omega_{m+1}^i)|,$$

for any $m \in [0, K-1]$, $m \neq m'$

Proof. See Appendix 7.2.2. □

Theorem 3.1. *At any symbol* t *and sub-carrier* i *(*$i = 0, 1, \cdots, I-1$*), only* $G_{m'}(\rho_i, u)$ *with the largest non-zero amplitude has the opposite sign to all* $G_m(\rho_i, u)$, $m \neq m'$, *where* m, $m' \in [0, K-1]$.

Proof. In the case of $K < M$, $\sin(N\omega_{m+1}^i) = \sin\left(N\omega_m^i - \frac{N\pi}{K}\right) = (-1)^Q \sin(N\omega_m^i)$

for $m = 0, 1, \cdots, K-1$, since $\omega_{m+1}^i = \frac{\rho_i u}{2} - \pi\left(\frac{t}{L} + \frac{m+1}{K}\right) = \omega_m^i - \frac{\pi}{K}$, and $Q = \frac{N}{K}$, as specified in (3.9). Therefore, $G_m(\rho_i, u)$ can be rewritten as

$$G_m(\rho_i, u) = e^{\frac{-j(N-1)\pi}{K}} \frac{(-1)^Q \sin^2(N\omega_m^i)}{\sin(\omega_m^i)\sin(\omega_{m+1}^i)}, m \in [0, K-1]. \tag{3.12}$$

Given Q, the sign of $G_m(\rho_i, u)$ only depends on that of $\sin(\omega_m^i)\sin(\omega_{m+1}^i)$.

For any u and frequency point ρ_i, $\sin(\omega_m^i) = \sin\left(\frac{\rho_i u}{2} - \frac{\pi t}{L} - \frac{\pi m}{K}\right)$, $m \in [0, K-1]$, spans half a cycle of the sine function with the initial phase $(\frac{\rho_i u}{2} - \frac{\pi t}{L})$. $\sin(\omega_{m+1}^i)$ is the shifted version of $\sin(\omega_m^i)$.

Given m' specified in Lemma 3.1, we have $\sin(N\omega_m^i) = \sin\left(N\omega_{m'}^i + \frac{\pi N(m'-m)}{K}\right) = (-1)^{Q(m'-m)}\sin(N\omega_{m'}^i)$, since $\omega_m^i = \omega_{m'}^i + \frac{\pi(m'-m)}{K}$. Therefore,

$$\begin{aligned}\sin(N\omega_m^i)\sin(N\omega_{m+1}^i) =& (-1)^{Q(m'-m)}\sin(N\omega_{m'}^i) \times (-1)^{Q(m'-m-1)}\sin(N\omega_{m'}^i) \\ =& (-1)^Q \sin^2(N\omega_{m'}^i).\end{aligned} \tag{3.13}$$

From Lemma 3.2, the denominator of $G_m(\rho_i, u)$, $m = 0, 1, \cdots, K-1$, takes the minimum at m', i.e.,

$$|G_{m'}(\rho_i, u)| = \left|\frac{(-1)^Q \sin^2(N\omega_{m'}^i)}{\sin(\omega_{m'}^i)\sin(\omega_{m'+1}^i)}\right| \tag{3.14a}$$

$$> \left|\frac{(-1)^Q \sin^2(N\omega_{m'}^i)}{\sin(\omega_m^i)\sin(\omega_{m+1}^i)}\right| \tag{3.14b}$$

$$= \left|\frac{\sin(N\omega_m^i)\sin(N\omega_{m+1}^i)}{\sin(\omega_m^i)\sin(\omega_{m+1}^i)}\right| \tag{3.14c}$$

$$= |G_m(\rho_i, u)|, m = 0, 1, \cdots, K-1; m \neq m', \tag{3.14d}$$

where (3.14a) and (3.14c) are obtained by plugging (3.13); (3.14b) is based on Lemma 3.2; and (3.14d) is based on (3.11).

By combining Lemma 3.1 and (3.14), the sign of $G_m(\rho_i, u)$ only changes at m', which, according to Lemma 3.2, corresponds to the largest of $|G_{m'}(\rho_i, u)|$ for $m = 0, 1, \cdots, K-1$.

In the case of $K = M$, we can have $G_{K-1}(\rho_i, u) = P_{K-1}^*(\rho_i, u)P_0(\rho_i, u)$, i.e., the

cross-correlation between the first and last subarrays. For $m = 0, 1, \cdots, K-2$, $G_m(\rho_i, u)$ remains the same as (3.11). The discussion in the case of $K < M$ still applies, because

$$P_K(\rho_i, u) = e^{j(N-1)\left(\frac{\rho_i u}{2} - \frac{\pi t}{L} - \frac{\pi K}{K}\right)} \frac{\sin\left(N\left(\frac{\rho_i u}{2} - \frac{\pi t}{L} - \frac{\pi K}{K}\right)\right)}{\sin\left(\frac{\rho_i u}{2} - \frac{\pi t}{L} - \frac{\pi K}{K}\right)}$$
$$= e^{j(N-1)\left(\frac{\rho_i u}{2} - \frac{\pi t}{L}\right)} \frac{\sin\left(N\left(\frac{\rho_i u}{2} - \frac{\pi t}{L}\right)\right)}{\sin\left(\frac{\rho_i u}{2} - \frac{\pi t}{L}\right)} = P_0(\rho_i, u). \tag{3.15}$$

The theorem is proved. □

3.4 Unambiguous Wideband Estimation of Nu

Following Theorem 3.1, we are able to explicitly calibrate the signs of $R_m(f_i)$ in (3.4) at every sub-carrier i, as given by

$$\tilde{R}_m(f_i) = \begin{cases} (-1)^Q R_m(f_i) & \text{if } m \neq m'; \\ (-1)^{Q+1} R_m(f_i) & \text{if } m = m', \end{cases} \tag{3.16}$$

where $m = 0, 1, \cdots, K-1$, $i = 0, 1, \cdots, I-1$, and m' is the index to the cross-correlation with the largest amplitude per sub-carrier, as specified in Lemmas 3.1 and 3.2.

The generalization of (3.16) to the case of $0 \leq m \leq M-1$ is straightforward. Particularly, $P_m(\rho_i, u)$ in (3.10) has the periodicity of K and P for m and t, respectively, since the phase shifts in (3.8) exhibit the same periodicity. Therefore, we can have $R_m(f_i) = R_{\text{mod}\{m,K\}}(f_i)$, $0 \leq m \leq M-1$. Out of the total M cross-correlations, there can be $\lfloor \frac{M}{K} \rfloor$ or $\lceil \frac{M}{K} \rceil$ with the opposite sign to the rest, i.e., one every K. These $\lceil \frac{M}{K} \rceil$ or $\lfloor \frac{M}{K} \rfloor$ cross-correlations are the strongest, the least susceptible to the noise, and therefore most unlikely to be misjudged in terms of sign, as compared to the rest.

Note that Theorem 3.1 is under an implicit assumption that $G_{m'}(\rho_i, u) \neq 0$. In the case of $G_{m'}(\rho_i, u) = 0$, $m' \in [0, K-1]$, we have $\sin^2(N\omega_{m'}^i) = 0$, accord-

ing to (3.12). Combining (3.13) and (3.11), we have $G_m(\rho_i, u) = 0$ and hence $|\tilde{S}(f_i)|^2 G_m(\rho_i, u) e^{j\rho_i Nu} = 0$; see (3.4), for all m's at the current symbol t and sub-carrier i. Therefore, we can still calibrate the signs using (3.16) with no impact on the coherent accumulation, since all useful signal components are zero.

Note that $|\tilde{S}(f_i)|^2 |\tilde{S}(f_{i+\triangle_i})|^2 G_m^*(\rho_i, u) G_m(\rho_{i+\triangle_i}, u)$ in (3.6) becomes consistent in sign, after calibrating the signs of $G_m(\rho_i, u)$ $\forall m, t, i$. Coherent accumulation can be performed across m, f_i and t to improve the estimation SNR and accuracy, as given by

$$\tilde{R} = \sum_{m=0}^{M-2} \sum_{t=0}^{P-1} \sum_{i=0}^{I-\triangle_i-1} \tilde{R}_m(f_i). \tag{3.17}$$

From (3.6), we can estimate Nu by evaluating the phase of $\tilde{R}$, as given by

$$\widehat{Nu} = \frac{\arg\{\tilde{R}\}}{\delta_{\rho_i}}, \tag{3.18}$$

Algorithm 3 summarizes the proposed unambiguous estimation method for Nu.

Note that the estimation error of $\widehat{Nu}$ can be amplified by $\delta_{\rho_i} (= \rho_{i+\triangle_i} - \rho_i = \frac{\triangle_i B}{I f_h}) \ll 1$. To this end, $\delta_{\rho_i} N < 1$ is expected to be as close to 1 as possible. This requires $\triangle_i$ to be large. However, a larger value of $\triangle_i$ can compromise the gain of the coherent accumulation of (3.17), degrading the estimation of $\arg\{\tilde{R}\}$ and $\widehat{Nu}$. $\triangle_i$ needs to be holistically selected, as we do in the following corollary.

Corollary 3.1. *The optimal $\triangle_i$ for $\widehat{Nu}$,* is $\frac{2I}{3}$.

Proof. To prove this corollary, we first derive the mean square error lower bound (MSELB) of $\widehat{Nu}$, based on which the optimal $\triangle_i$ providing to the minimum MSELB of $\widehat{Nu}$ can be identified.

Suppose that the i.i.d zero-mean AWGN in (3.1) has noise power σ_0^2, i.e., $z_m(t_i) \sim \mathcal{CN}(0, \sigma_0^2)$ for $0 \le m \le M-1, 0 \le i \le I-1, 0 \le t \le P-1$. After the I-point DFT, the noise in (3.2) yields $Z_m(f_i) \sim \mathcal{CN}(0, I\sigma_0^2)$. From (3.5), the noise power of $\tilde{Z}_m(f_i)$ is $|\sigma_{\tilde{z}}^i|^2 = |\tilde{S}(f_i)|^2 \left[|P_m(\rho_i, u)|^2 + |P_{m+1}(\rho_i, u)|^2\right] I\sigma_0^2$. After the cross-

Algorithm 3 Optimal Nu Estimation Approach

Initialize: Given N and M, select K ($\in [2, M)$, $\in Z^+$),
$N/K \in Z^+$, $\triangle_i$ and I.

procedure ANALOG BEAMFORMING AND SIGN ALIGNMENT FOR CROSS-CORRELATIONS
  for $t \leftarrow 0 : P-1$ **do**
    $s_m(t_i) \leftarrow$ Analog beamforming using α_m^t in (3.8);
    $\boldsymbol{s}_m^t = [s_m(t_0), s_m(t_1), \cdots, s_m(t_{I-1})]^T$;
    $S_m(f_i) \leftarrow \mathrm{DFT}[\boldsymbol{s}_m^t]$;
    for $i \leftarrow 0, 1, \cdots, I-1$ **do**
      for $m \leftarrow 0, 1, \cdots, M-2$ **do**
        $R_m(f_i) \leftarrow S_m^*(f_i) S_m(f_{i+1})$;
      end for
      Find $R_{m'}(f_i)$ having the largest magnitude
      among $R_m(f_i)$ $(0 \le m \le K-1)$;
      for all $m \in [0, M-2]$ **do**
        $R_m(f_i) \leftarrow (-1)^Q R_m(f_i)$;
      end for
      for $m \leftarrow k' : K : M$ **do**
        $R_m(f_i) \leftarrow -R_m(f_i)$;
      end for
      $\tilde{R}_m(f_i) \leftarrow R_m^*(f_i) R_m(f_{i+\triangle_i})$
    end for
  end for
end procedure

$\tilde{R} \leftarrow \sum_{m=0}^{M-2} \sum_{t=0}^{P-1} \sum_{i=0}^{I-\triangle_i-1} \tilde{R}_m(f_i)$, $\triangle_i = \frac{2I}{3}$;
return $\widehat{Nu} \leftarrow \frac{1}{\delta_{\rho_i}} \arg\{\tilde{R}\}$

correlation between sub-carriers, the noise power of $\breve{Z}_m(f_i)$ in (3.7) can be given by

$$|\sigma_{\breve{z}}|^2 = |\tilde{S}(f_i)|^4 |G_m(\rho_i, u)|^2 |\sigma_{\tilde{z}}^{i+\triangle_i}|^2 + |\tilde{S}(f_{i+\triangle_i})|^4 |G_m(\rho_{i+\triangle_i}, u)|^2 |\sigma_{\tilde{z}}^{i}|^2 \tag{3.19}$$

The power of the signal component of $\tilde{R}_m(f_i)$ in (3.6), denoted by σ_s^2, can be given by

$$\sigma_s^2 = |\tilde{S}(f_i)|^4 |\tilde{S}(f_{i+\triangle_i})|^4 |G_m(\rho_i, u)|^2 |G_m(\rho_{_{i+\triangle_i}}, u)|^2. \tag{3.20}$$

Therefore, the SNR for evaluating the phase of $\tilde{R}_m(f_i)$ in (3.6) at sub-carrier

$$\begin{aligned}\gamma_{\tilde{s},u} = \frac{\sigma_s^2}{|\sigma_{\tilde{z}}|^2} &= \frac{|\tilde{S}(f_i)|^4|\tilde{S}(f_{i+\triangle_i})|^4|G_m(\rho_i,u)|^2|G_m(\rho_{i+\triangle_i},u)|^2}{|\tilde{S}(f_i)|^4|G_m(\rho_i,u)|^2|\sigma_{\tilde{z}}^{i+\triangle_i}|^2 + |\tilde{S}(f_{i+\triangle_i})|^4|G_m(\rho_{i+\triangle_i},u)|^2|\sigma_{\tilde{z}}^{i}|^2} \\ &= \frac{|\tilde{S}(f_i)|^4|\tilde{S}(f_{i+\triangle_i})|^4|G_m(\rho_i,u)|^2|G_m(\rho_{i+\triangle_i},u)|^2}{|\tilde{S}(f_i)|^4|G_m(\rho_i,u)|^2|\tilde{S}(f_{i+\triangle_i})|^2 H_m(\rho_{i+\triangle_i},u) I\sigma_0^2 \\ &\qquad + |\tilde{S}(f_{i+\triangle_i})|^4|G_m(\rho_{i+\triangle_i},u)|^2|\tilde{S}(f_i)|^2 H_m(\rho_i,u) I\sigma_0^2}, \\ &\approx \frac{1}{2I\sigma_0^2}\frac{|\tilde{S}(f_i)|^8|G_m(\rho_j,u)|^4}{|\tilde{S}(f_i)|^6|G_m(\rho_i,u)|^2 H_m(\rho_i,u)} = \frac{|\tilde{S}(f_i)|^2|G_m(\rho_i,u)|^2}{2I\sigma_0^2 H_m(\rho_i,u)} \\ &\leq \frac{|\tilde{S}(f_i)|^2 H_m(\rho_i,u)}{8I\sigma_0^2} \end{aligned} \tag{3.21a}$$

i, subarray m and symbol t, denoted by $\gamma_{\tilde{s},u}$, can be calculated in (3.21) placed on the top of the next page, where $H_m(\rho_i,u) = |P_m(\rho_i,u)|^2 + |P_{m+1}(\rho_i,u)|^2$; the approximation in (3.21a) is taken by assuming that the signal power is uniformly distributed in the passband and $\triangle_i$ is so small that the difference of $|G_m(\rho_i,u)|$ and $H_m(\rho_i,u)$ over different sub-carriers is negligible. The last inequality in (3.21a) is based on the following inequalities:

$$\frac{|G_m(\rho_i,u)|^2}{H_m(\rho_i,u)} = \frac{|P_m^*(\rho_i,u)P_{m+1}(\rho_i,u)|^2}{|P_m(\rho_i,u)|^2 + |P_{m+1}(\rho_i,u)|^2} \leq \frac{|P_m(\rho_i,u)|^2 \cdot |P_{m+1}(\rho_i,u)|^2}{|P_m(\rho_i,u)|^2 + |P_{m+1}(\rho_i,u)|^2} \tag{3.22a}$$

$$\leq \frac{\left(\frac{|P_m(\rho_i,u)|^2 + |P_{m+1}(\rho_i,u)|^2}{2}\right)^2}{|P_m(\rho_i,u)|^2 + |P_{m+1}(\rho_i,u)|^2} \tag{3.22b}$$

$$= \frac{H_m(\rho_i,u)}{4},$$

where (3.22a) is based on the Cauchy-Schwartz inequality and (3.22b) is due to the fact that a geometric mean, i.e., the numerator in (3.22a), is never greater than the corresponding quadratic mean, i.e., the numerator in (3.22b).

By taking the expectation of $\gamma_{\tilde{s},u}$ over $\tilde{S}(f_i)$ and u, we can obtain

$$\gamma^{Nu} = \mathbb{E}\{\gamma_{\tilde{s},u}\} \leq \frac{\mathbb{E}_s\{|\tilde{S}(f_i)|\}}{8I\sigma_0^2}\mathbb{E}_u\left\{H_m(\rho_i,u)\right\} = \frac{I}{4}\gamma_e, \tag{3.23}$$

where $\mathbb{E}_s\{\cdot\}$ and $\mathbb{E}_u\{\cdot\}$ take expectations over $\tilde{s}(t)$ and u, respectively; $\frac{\mathbb{E}_s\{|\tilde{S}(f_i)|\}}{I\sigma_0^2} = \frac{I}{N}\gamma_e$ with γ_e denoting the receive SNR at each antenna in the time domain; and $\mathbb{E}_u\{H_m(\rho_i, u)\} = 2N$ with u uniformly distributed in $[-\pi, \pi]$, as proved in Appendix 7.2.4.

By coherently accumulating over subarrays, symbols and sub-carriers, the unconditional SNR for estimating $\arg\{\tilde{R}\}$ can be finally given by

$$\gamma_\Sigma^{Nu} \leq \frac{(M-1)P(I-\triangle_i)I}{4}\gamma_e. \tag{3.24}$$

In high SNR regions, the MSELB of estimating the phase of a noise-corrupted complex signal, like (3.6) and (3.17), can be approximated by $\frac{1}{2\gamma_x}$ [1, 102], where γ_x denotes the SNR of the signal. By plugging this and (3.24) into (3.18) and exploiting the definitions of δ_{ρ_i} and $\triangle_i$; see (3.6), we have that

$$\text{MSELB}(\widehat{Nu}) \approx \frac{2If_h^2}{B^2(M-1)P\gamma_e} \times \frac{1}{\triangle_i^2(I-\triangle_i)}. \tag{3.25}$$

By setting the first-order derivative of (3.25) with respect to $\triangle_i$ to 0, we can achieve that the minimum of $\text{MSELB}(\widehat{Nu})$ is taken at $\triangle_i = \frac{2I}{3}$. This concludes the proof. □

Note that the MSELB is the modified Cramér-Rao lower bound (CRLB) which, like the standard CRLB, provides the theoretical limit (i.e., a lower bound) to the error variance of any parameter estimator [93, 96]. The modified CRLB is particularly useful when, in addition to the parameter to be estimated, the observed data also depend on other unwanted parameters [96], such as the Rayleigh fading channel gains in the context of this chapter. By comparing to the modified CRLB, the proposed approach is demonstrated to asymptotically approach the theoretical limit of the error variance exhibited by the estimates, as will be shown in Section 3.7.

3.5 Unambiguous Wideband AoA Estimation

Given $\widehat{Nu}$, we are able to estimate u unambiguously at each sub-carrier i, as dictated in the following theorem.

Theorem 3.2. *Provided $e^{j(\widehat{Nu}-Nu)} \to 1$, u can be estimated unambiguously at each individual sub-carrier i ($0 \le i \le I-1$) by using the new phase shifts (3.8) in localized hybrid arrays.*

Proof. Let $n = n' + qK$ $(0 \le n' \le K-1,\ 0 \le q \le Q-1)$. Then (3.10) can be written as

$$\begin{aligned} P_m(\rho_i, u) &= \sum_{n'=0}^{K-1} \sum_{q=0}^{Q-1} e^{j(n'+qK)\rho_i u} e^{-j2\pi \frac{mP+t}{L}(n'+qK)} \\ &= \sum_{n'=0}^{K-1} g_{n'}(\rho_i, u) e^{-j\frac{2\pi n' m}{K}}, m = 0, 1, \cdots, K-1, \end{aligned} \tag{3.26}$$

where

$$g_{n'}(\rho_i, u) = e^{jn'(\rho_i u - \frac{2\pi t}{L})} e^{j(N-K)(\frac{\rho_i u}{2} - \frac{\pi t}{L})} \frac{\sin(\frac{\rho_i N u}{2} - \frac{N\pi t}{L})}{\sin(\frac{\rho_i K u}{2} - \frac{K\pi t}{L})}, n' = 0, 1, \cdots, K-1, \tag{3.27}$$

provides the Fourier coefficients of $P_m(\rho_i, u)$.

In light of this, we can multiply $e^{-jm\rho_i \widehat{Nu}}$ to both sides of (3.2), i.e.,

$$\begin{aligned} \tilde{S}_m(f_i) =& S_m(f_i) e^{-j\rho_i m \widehat{Nu}} \\ =& \tilde{S}(f_i) P_m(\rho_i, u) e^{jm\rho_i(Nu - \widehat{Nu})} + Z_m(f_i) e^{-jm\rho_i \widehat{Nu}}, \end{aligned} \tag{3.28}$$

Assume that $e^{j(\widehat{Nu}-Nu)} \to 1$ we can take the IDFT of $S_m(f_i)e^{-j\rho_i m\widehat{Nu}}$, $m = 0, 1, \cdots, K-1$, and obtain $a_{n'}(\rho_i, u) = \tilde{S}(f_i) g_{n'}(\rho_i, u) + \xi_{n'}(f_i)$, where

$$\xi_{n'}(f_i) = \text{IDFT}\{Z_m(f_i) e^{-jm\rho_i \widehat{Nu}}\}, n' = 0, 1, \cdots, K-1 \tag{3.29}$$

A cross-correlation, denoted by $r_{n'}(\rho_i, u)$, can be taken between any two consec-

utive IDFT points, i.e., $a_{n'}(\rho_i, u)$ and $a_{n'+1}(\rho_i, u)$, as given by

$$\begin{aligned} r_{n'}(\rho_i, u) =& a^*_{n'}(\rho_i, u) a_{n'+1}(\rho_i, u) \\ =& e^{j(\rho_i u - \frac{2\pi t}{L})} |C(\rho_i, u)|^2 |\tilde{S}(f_i)|^2 + \tilde{\xi}_{n'}(f_i), n' = 0, 1, \cdots, K-2, \end{aligned} \tag{3.30}$$

where

$$\begin{aligned} \tilde{\xi}_{n'}(f_i) =& \tilde{S}(f_i) C(\rho_i, u) e^{j(n'+1)(\rho_i u - \frac{2\pi t}{L})} \xi^*_{n'}(f_i) + \tilde{S}^*(f_i) C(\rho_i, u)^* e^{-jn'(\rho_i u - \frac{2\pi t}{L})} \xi_{n'+1}(f_i) \\ &+ \xi^*_{n'}(f_i) \xi_{n'+1}(f_i), \end{aligned} \tag{3.31}$$

and

$$C(\rho_i, u) = e^{j(N-K)(\frac{\rho_i u}{2} - \frac{\pi t}{L})} \frac{\sin(\frac{\rho_i N u}{2} - \frac{N\pi t}{L})}{\sin(\frac{\rho_i K u}{2} - \frac{K\pi t}{L})}. \tag{3.32}$$

Given $\rho_i = \frac{f_l + f_i}{f_h} \leq 1$ when $f_i \leq B$, an unambiguous estimate of $\rho_i u$ and thus u, i.e., $\rho_i \hat{u}_i$ and $\hat{u}_i$, can be obtained by evaluating the phase of $r_{n'}(\rho_i, u)$ at sub-carrier i. This concludes the proof. □

Following Theorem 3.2, u can be estimated independently at each individual sub-carrier. Considering the periodicity of the phase shifts in (3.8) across K consecutive analog subarrays outputs, there are a total of $\lfloor M/K \rfloor$ non-overlapping groups of K consecutive subarrays, including the last group cyclically concatenating the last $(M - \lfloor M/K \rfloor K)$ subarrays and the last $(K - M + \lfloor M/K \rfloor K)$ subarrays of the first group. Coherently accumulating (3.30) over $n' = 0, 1, \cdots, K-2$, $l = 0, 1, \cdots, \lceil M/K \rceil$ and $t = 0, 1, \cdots, P-1$, (i.e., antennas, subarrays and symbols) can enhance the robustness of the AoA estimation to the receive noise, hence improving the estimation accuracy, where l indicates the l-th group of K consecutive subarrays.

In the extremely high frequency, LoS-dominating multibeam satellite applications with $f_h \gg B$, we can further conduct coherent accumulations across subcarriers to increase the estimation robustness and accuracy, as stated in the following

Algorithm 4 The Proposed AoA estimation Approach

1: **Initialize**: Given $\widehat{Nu}$ obtained from Algorithm 3.
2: **for all** $i \in [0, I-1], m = [0, M-1], t \in [0, P-1]$ **do**
3: $\quad \tilde{S}_m(f_i) \leftarrow S_m(f_i) e^{-jm\rho_i \widehat{Nu}}$
4: **end for**
5: **for** $t \leftarrow 0, 1, \cdots, P-1$ **do**
6: $\quad$ **for** $i \leftarrow 0, 1, \cdots, I-1$ **do**
7: $\quad\quad$ **for all** $m \in [0, M-1]$ **do**
8: $\quad\quad\quad \tilde{S}_m(f_i) \leftarrow S_m(f_i) e^{-jm\rho_i \widehat{Nu}}$;
9: $\quad\quad$ **end for**
10: $\quad\quad$ **for** $l \leftarrow 0, 1, \cdots, \lfloor M/K \rfloor - 1$ **do**
11: $\quad\quad\quad m \leftarrow lK$;
12: $\quad\quad\quad \mathbf{S}^l(f_i) \leftarrow \{\tilde{S}_m(f_i), \cdots, \tilde{S}_{m+K-1}(f_i)\}$;
13: $\quad\quad\quad \boldsymbol{a}^l(\rho_i, u) \leftarrow \text{IDFT}\{\mathbf{S}^l(f_i)\}$;
14: $\quad\quad\quad$ **for** $n' = 0, 1, \cdots, K-1$ **do**
15: $\quad\quad\quad\quad \boldsymbol{r}^l_{n'}(\rho_i, u) \leftarrow \boldsymbol{a}^{l*}_{n'}(\rho_i, u) \boldsymbol{a}^l_{n'+1}(\rho_i, u)$
16: $\quad\quad\quad$ **end for**
17: $\quad\quad$ **end for**
18: $\quad\quad$ $\mathbf{S}^{\lfloor M/K \rfloor}(f_i) \leftarrow \{S_{\lfloor M/K \rfloor K}(f_i), \cdots, S_{M-1}(f_i),$
19: $\quad\quad\quad\quad\quad\quad S_{M-\lfloor M/K \rfloor K}(f_i), \cdots S_{K-1}(f_i)\}$;
20: $\quad\quad$ $\boldsymbol{a}^{\lfloor M/K \rfloor}(\rho_i, u) \leftarrow \text{IDFT}\big\{\mathbf{S}^{\lfloor M/K \rfloor}(f_i)\big\}$;
21: $\quad\quad$ **for** $n' = 0, 1, \cdots, K-1$ **do**
22: $\quad\quad\quad \boldsymbol{r}^{\lfloor M/K \rfloor}_{n'}(\rho_i, u) \leftarrow \boldsymbol{a}^{\lfloor M/K \rfloor *}_{n'} \boldsymbol{a}^{\lfloor M/K \rfloor}_{n'+1}$; see (3.30)
23: $\quad\quad$ **end for**
24: $\quad$ **end for**
25: **end for**
26: $\tilde{r}_i \leftarrow \sum_{t=0}^{P-1} e^{j\frac{2\pi t}{L}} \left[\sum_{l=0}^{\lfloor M/K \rfloor} \sum_{n'=0}^{K-2} \boldsymbol{r}^l_{n'}(\rho_i, u) \right]$;
27: $\tilde{r} \leftarrow \sum_{i=i_1}^{i_2} \tilde{r}_i$
28: **return** $\hat{u} \leftarrow \frac{f_h}{f_l + \frac{B(i_2+i_1)}{2I}} \arg\{\tilde{r}\}$ and $\hat{u}_i \leftarrow \frac{1}{\rho_i} \arg\{\tilde{r}_i\}$

proposition.

Proposition 3.1. In a strong LoS environment with $|\tilde{S}(f_i)|^2$ identical for consecutive sub-carriers $i = i_1, i_1+1, \cdots, i_2$, $(i_2 \geq i_1)$, u can be estimated from $\sum_{i=i_1}^{i_2} \tilde{r}_i$, where $\tilde{r}_i$ is given in Algorithm 4 Line 26.

Proof. Substituting (3.30) into $\tilde{r}_i$, we have

$$\begin{aligned}\tilde{r}_i &= \sum_{t=0}^{P-1}\sum_{l=0}^{\lfloor M/K\rfloor}\sum_{n'=0}^{K-2} e^{j\rho_i u}\left|C(\rho_i,u)\right|^2\left|\tilde{S}(f_i)\right|^2+\tilde{\xi}(f_i)\\ &= \lceil M/K\rceil(K-1)|\tilde{S}(f_i)|^2 e^{j\rho_i u}\frac{PN}{K}+\tilde{\xi}(f_i),\end{aligned} \tag{3.33}$$

where $\tilde{\xi}(f_i)=\sum\limits_{t=0}^{P-1} e^{j\frac{2\pi t}{L}}\left[\sum\limits_{l=0}^{\lfloor M/K\rfloor}\sum\limits_{n'=0}^{K-2}\tilde{\xi}_{n'}(f_i)\right]$ and $\tilde{\xi}_{n'}(f_i)$ is given in (3.31). The second equality in (3.33) is based on Lemma 7.3, as provided in Appendix 7.2.5.

On the condition that $\tilde{S}(f_i)$ remains unchanged for sub-carriers $i\in[i_1,i_2]$, we can accumulate $\tilde{r}_i$ over sub-carrier i by directly adding up $\tilde{r}_i$ in (3.33), as given by

$$\tilde{r}=\frac{PN}{K}\lceil M/K\rceil(K-1)|\tilde{S}(f_i)|^2\sum_{i=i_1}^{i_2}e^{j\rho_i u}+\sum_{i=i_1}^{i_2}\tilde{\xi}(f_i); \tag{3.34}$$

$$\begin{aligned}\sum_{i=i_1}^{i_2}e^{j\rho_i u} &= e^{j\frac{f_l}{f_h}u}\sum_{i=i_1}^{i_2}e^{j\frac{Bu}{If_h}i}=e^{j\frac{f_l}{f_h}u}\left(\frac{1-e^{j\frac{Bu}{If_h}(i_2-i_1+1)}}{1-e^{j\frac{Bu}{If_h}}}\right)\\ &= e^{j\frac{f_l+\frac{B(i_2+i_1)}{2I}}{f_h}u}\frac{\sin\left[\frac{Bu}{2If_h}(i_2-i_1+1)\right]}{\sin\left(\frac{Bu}{2If_h}\right)},\end{aligned} \tag{3.35}$$

where $\frac{\sin\left[\frac{Bu}{2If_h}(i_2-i_1+1)\right]}{\sin\left(\frac{Bu}{2If_h}\right)}>0$, since $u\in[-\pi,\pi]$, $0\le i_2-i_1\le I-1$, $\frac{B}{If_h}<\frac{B}{If_h}(i_2-i_1+1)\le\frac{B}{f_h}\ll 1$ and thus the numerator and denominator of (3.35) always take the same sign.

Evaluating the phase of $\tilde{r}$, we can achieve the estimate of u, denoted by $\hat{u}$, as given by

$$\hat{u}=\frac{f_h}{f_l+\frac{B(i_2+i_1)}{2I}}\arg\{\tilde{r}\}, \tag{3.36}$$

which is unambiguous and therefore concludes this proof. □

Remark 3.1. In the numerator of (3.35), $\frac{Bu}{2If_h}(i_2-i_1+1)\le\frac{B\pi}{2f_h}$, since $i_2\le I-1$, $i_1\ge 0$ and $|u|\le\pi$. In a wideband multibeam satellite communication system, B

can typically be $0.1f_h$, in which case, $\frac{Bu}{2If_h}(i_2 - i_1 + 1) \leq 0.05\pi$. Since $\sin x \approx x$ for $0 \leq x \leq 0.05\pi$ with an error less than 6.45×10^{-4}, we have $\frac{\sin\left[\frac{Bu}{2If_h}(i_2-i_1+1)\right]}{\sin\left(\frac{Bu}{2If_h}\right)} \approx (i_2-i_1+1)$. In other words, the coherent accumulation across sub-carriers can bring a gain of about $(i_2 - i_1 + 1)$.

3.6 Analysis and Discussion

3.6.1 Accuracy of the Proposed AoA Estimation

To analyze the accuracy, we first calculate the estimation SNRs for evaluating $\arg\{\tilde{r}_i\}$ and $\arg\{\tilde{r}\}$; refer to Algorithm 4, Lines 26 and 27, and then derive the MSELBs of $\hat{u}_i$ and $\hat{u}$.

From $r_{n'}(\rho_i, u)$ in (3.30), the noise power can be given by

$$|\eta_{\tilde{\xi}}^i|^2 = 2|\tilde{S}(f_i)|^2|C(\rho_i, u)|^2|\eta_\xi|^2,$$

where $|\eta_\xi|^2$ is the noise power of $\xi_{n'}(f_i)$ in (3.29). $|\eta_\xi|^2 = \frac{I\sigma_0^2}{K}$, since $\xi_{n'}(f_i) = \text{IDFT}\{Z_m(f_i)e^{-j\rho_i\widehat{Nu}}\}$ and $Z_m(f_i) \sim \mathcal{CN}(0, I\sigma_0^2)$ $(m = 0, 1, \cdots, K-1)$. The signal power of $r_{n'}(\rho_i, u)$ can be given by $\eta_s^2 = |C(\rho_i, u)|^4|\tilde{S}(f_i)|^4$. Therefore, the estimation SNR of $r_{n'}(\rho_i, u)$ develops as

$$\gamma_{\tilde{s},u} = \frac{\eta_s^2}{|\eta_{\tilde{\xi}}^i|^2} = \frac{K|C(\rho_i, u)|^2|\tilde{S}(f_i)|^2}{2I\sigma_0^2} \tag{3.37}$$

By taking the expectation of $\gamma_{\tilde{s},u}$ over $\tilde{S}(f_i)$ and u, we can obtain the unconditional SNR, as given by

$$\gamma_i = \mathbb{E}\{\gamma_{\tilde{s},u}\} = \frac{K}{2}\mathbb{E}_s\left\{\frac{|\tilde{S}(f_i)|^2}{I\sigma_0^2}\right\}\mathbb{E}_u\left\{|C(\rho_i, u)|^2\right\} = \frac{I}{2}\gamma_e \tag{3.38}$$

where $\mathbb{E}_s\left\{\frac{|\tilde{S}(f_i)|^2}{I\sigma_0^2}\right\} = \frac{I}{N}\gamma_e$ provided the signal powers are constant across sub-carriers; and $\mathbb{E}_u\left\{|C(\rho_i, u)|^2\right\} = \frac{N}{K}$ with u uniformly distributed in $[-\pi, \pi]$; see Appendix 7.2.6. By coherently accumulating over antennas, subarrays and symbols,

the estimation SNR at the i-th sub-carrier can be finally given by

$$\gamma_{\Sigma i}^{u} \leq \frac{P\lceil M/K \rceil (K-1)I}{2}\gamma_e, \tag{3.39}$$

where $i = 0, 1, \cdots, I-1$ and the equality can be taken if $\frac{M}{K} = \lceil \frac{M}{K} \rceil$.

As discussed in Remark 3.1, the gain of constructively combining $\tilde{r}_i$ to achieve $\tilde{r}$ is $(i_2 - i_1 + 1)$. Let γ_Σ denote the estimation SNR of $\hat{u}$. Then we get

$$\gamma_{\Sigma}^{u} < \frac{P\lceil M/K \rceil (K-1)I(i_2 - i_1 + 1)}{2}\gamma_e. \tag{3.40}$$

With reference to the proof of Corollary 3.1, the MSELB of $\hat{u}_i$ and $\hat{u}$ is finally established as

$$\text{MSELB}(\hat{u}_i) = \frac{1}{\rho_i^2} \cdot \frac{1}{2\gamma_{\Sigma i}^{u}}; \tag{3.41a}$$

$$\text{MSELB}(\hat{u}) = \left(\frac{f_h}{f_l + \frac{B(i_2+i_1)}{2I}} \right)^2 \cdot \frac{1}{2\gamma_{\Sigma}^{u}}. \tag{3.41b}$$

3.6.2 Extension to the State-of-the-art [1]

In the state-of-the-art approach [1], u was estimated directly from the second cross-correlation across sub-carriers, i.e., $\tilde{R}_m(f_i)$ in (3.6). We can adopt that approach to estimate the AoA from $\tilde{R}$ in (3.17) after the signs of the cross-correlations between subarrays are calibrated, as given by

$$\hat{u}^1 = \frac{\arg\{\tilde{R}\}}{N\delta_{\rho_i}} = \frac{\widehat{Nu}}{N}, \tag{3.42}$$

where the second equality is obtained based on (3.18).

Given the $\text{MSELB}(\widehat{Nu})$ in (3.25), we have that

$$\text{MSELB}(\hat{u}^1) = \frac{1}{N^2}\text{MSELB}(\widehat{Nu}) = \frac{1}{(N\delta_{\rho_i})^2} \cdot \frac{1}{2\gamma_{\Sigma i}^{Nu}}, \tag{3.43}$$

where $\gamma_{\Sigma i}^{Nu}$ is given in (3.24). It is noteworthy that Corollary 3.1 also holds for $\hat{u}^1$,

i.e., $\triangle_i = \frac{2I}{3}$ is optimal for estimating $\hat{u}^1$.

We note that $\hat{u}^1$ can be seen as an enhanced version of the state-of-the-art [1], since (i) different subarray-specific phase shifts are used to enable the cross-correlations between subarrays to exhibit deterministic property; (ii) the signs of the cross-correlations are fine enough calibrated to enhance robustness to noises, and (iii) the optimal sub-carrier offset $\triangle_i$ is applied to balance error scaling and accumulation gain.

3.6.3 Computational Complexity

The proposed scheme is a cross-correlation based approach like the one developed in [1], and only incurs scalar multiplications and additions to compute cross-correlations and coherent accumulations. There is neither a matrix inversion, nor a singular value or eigenvalue decomposition. As pointed out in [1], the cross-correlation based algorithm developed in [1] has a significantly lower computational complexity than conventional subspace-based methods, such as MUSIC or ESPRIT [88], and maximum likelihood estimators, both of which would require matrix inversions or decompositions. The proposed approach incurs even far lower complexity than the algorithm developed in [1]. As will be shown in Figs. 3.3 and 3.4, the proposed approach can achieve much better estimation performance with only 3 symbols than the algorithm developed in [1] with 10 symbols.

3.7 Numerical and Simulation Results

In this section, simulations are carried out to validate the proposed approach and analysis. We also simulate the state of the art, referred to as double cross-correlation (DCC) [1]. Without loss of generality, uniform linear localized hybrid arrays are considered. The array configurations and other simulation parameters are provided in Table 3.1; unless otherwise specified. The parameters are set up to achieve fair comparisons with DCC. According to [1, Sec. V], DCC is about to converge after 10 iterations, or in other words 10 training symbols, for a localized hybrid arrays with 32 antennas. Therefore, we compare our approach with DCC

after 10 iterations.

Table 3.1 : Configurations of the wideband localized hybrid array and other parameters

Par.	Val.	Par.	Val.	Par.	Val.
M, N	$4, 8$	K	4	N_{trial} †	4×10^4
I	256	P	3	AoA	$U_{[-\pi,\pi]}$ ‡
fc	70 GHz	B	$0.1f_c$	λ	$\frac{c}{f_c+B/2}$
$\triangle_i$	85	f_s	B	SNR††	$-30 : 10 : 40$ dB

† N_{trials} is the number of independent trials for each SNR. The curves in all figures are the average of 4×10^4 independent trials.

‡ $U_{[-\frac{\pi}{2},\frac{\pi}{2}]}$ represents the uniform distribution. Provided $d = \frac{\lambda}{2}$, $[-\frac{\pi}{2}, \frac{\pi}{2}]$ corresponds to $-30° \leq \theta \leq 30°$ of boresight.

†† The SNR is antenna-wise, i.e. γ_e in (3.24), (3.39) and (3.40).

Fig. 3.1 validates the accuracy of the SNR upper bounds (3.24), (3.39) and (3.40) by comparing with the simulation results of unconditional SNRs, where different number of sub-carriers are accumulated in the two sub-figures. We see that the SNR upper bounds for $\gamma_{\Sigma i}^u$ and γ_Σ^u are tight, especially in the case that M is a multiple of K, as plotted in Fig. 3.1. γ_Σ^{Nu} has a relatively loose upper bound due to the inequalities in (3.22). We also see that the SNR of $\hat{u}$ is improved by $3(= 10\log_{10} 2)$ dB and $24(= 10\log_{10} 255)$ dB, compared with $\hat{u}_i$ in both sub-figures. In other words, accumulating across sub-carriers can be approximated to be coherent with negligible errors.

Fig. 3.2a compares the MSELB of $\hat{u}^1$, $\hat{u}_i$ and $\hat{u}$ with the increase of $\triangle_i$ ($i_2 - i_1$ for $\hat{u}$), where the receive SNR per antenna is set to 0 dB. We see that the MSELB of $\hat{u}^1$ first decreases and then increases with $\triangle_i$. This is due to the decreasing unconditional SNR (3.24), and linear scaling coefficient (3.43), as the growth of $\triangle_i$. $\hat{u}_i$ stays unchanged, since it is the estimate per sub-carrier. We also see that $\hat{u}$ decreases monotonically with $(i_2 - i_1)$, as expected. Consider the consistent SNR gap of 11.1 dB for $\hat{u}^1$ in Fig. 3.1. We can compensate for the SNR gap by lifting up the analytical result of SNR by 11.1 dB for the estimation of $\hat{u}^1$. The accuracy of $\hat{u}^1$, despite the gain across sub-carriers, can still be worse than the per-sub-carrier estimate $\hat{u}_i$, due to the adverse effect of the cross-correlation across sub-carriers on

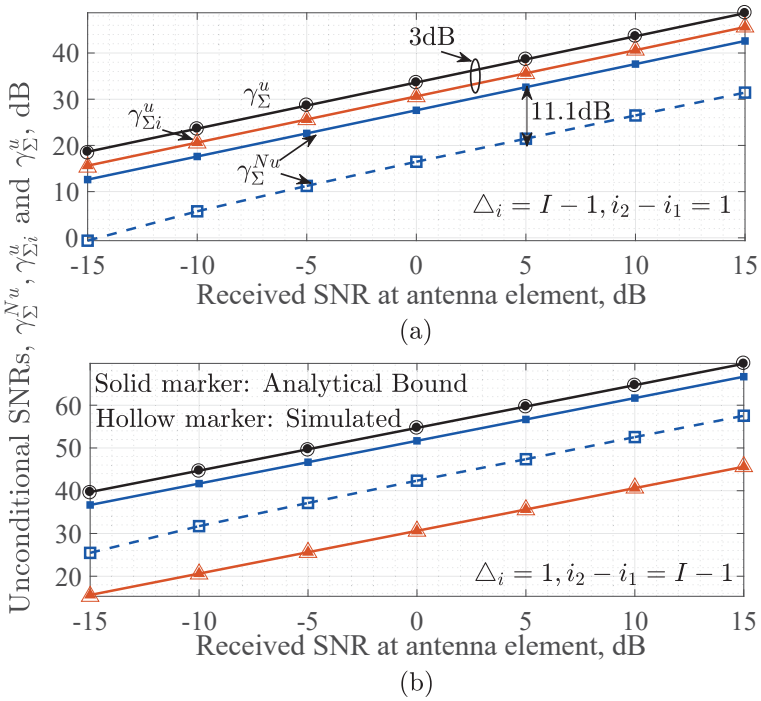

Figure 3.1 : Simulated and analytical estimation SNRs versus the receive SNR at antenna element, where (a) for γ_{Σ}^{Nu}, $\triangle_i = 255$, which means no accumulation across sub-carriers, and for γ_{Σ}^{u}, $i_2 - i_1 = 1$; (b) $\triangle_i$ and $i_2 - i_1$ are changed to 1 and 255, respectively.

the accuracy. Fig. 3.2b plots the MSELB of $\hat{u}^1$ as $\triangle_i$ increases, where different values are tested for I. We show that the optimal estimation accuracy is always achieved at $\frac{2I}{3}$, validating Corollary 3.1.

Fig. 3.3 plots the MSEs of the proposed estimation $\hat{u}_i$, $\hat{u}$ and $\hat{u}^1$, and the state-of-the-art, DCC, as γ_e increases. We show that the proposed approach is able to dramatically and increasingly outperform the state of the art, even with a much smaller number of training symbols. The proposed approach can reduce the MSE by orders of magnitude. Particularly, $\hat{u}_i$ and $\hat{u}$ can have a much better estimation accuracy than DCC in the extremely low SNR region. The MSE of DCC is as high as 12.63 at -30 dB receive SNR; while for $\hat{u}_i$ and $\hat{u}$, the MSEs are 2.094 and 0.0064, respectively. This is due to our discovery of the deterministic signs of $G_m(\rho_i, u)$ and our exploitation of the discovery, which lead to effective coherent accumulation of weak signals across antennas, subarrays, sub-carriers and also symbols.

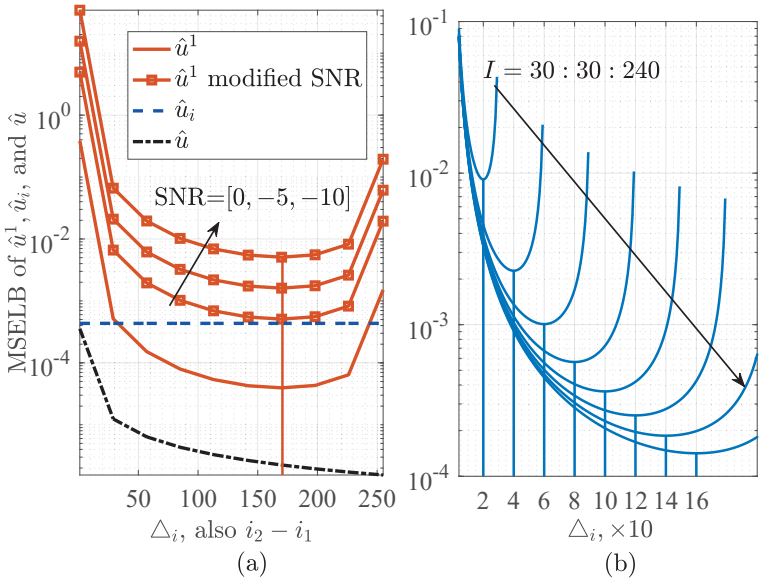

Figure 3.2 : (a) MSELB of $\hat{u}_i$, $\hat{u}$ and $\hat{u}^1$ versus $i_2 - i_1$ or $\triangle_i$, where the receive SNR at antenna is $\gamma_e = 0$ dB unless otherwise specified and the MSELBs of $\hat{u}^1$ with modified SNR are obtained by subtracting the constant gap of 11.1 dB between simulated and analytical estimation SNRs obtained from Fig. 3.1; (b): MSELB of $\hat{u}^1$ versus $\triangle_i$, where the number of sub-carriers, I, ranges from 30 to 240 stepped by 30.

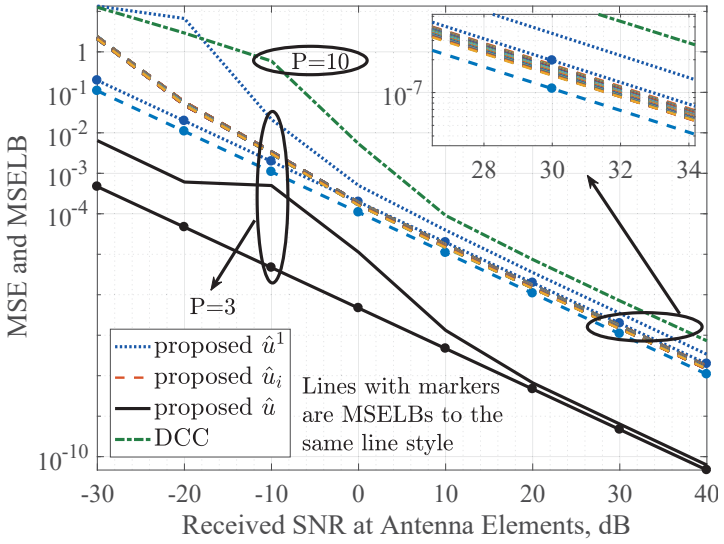

Figure 3.3 : MSE of the state-of-the-art DCC [1], $\hat{u}_i$, $\hat{u}$ and $\hat{u}^1$ versus the receive SNR at antenna, where the MSE of $\hat{u}_i$ of all sub-carriers given.

We also see from Fig. 3.3 that the proposed $\hat{u}^1$ has better but close performance compared with the convergent result of DCC. This is because both $\hat{u}^1$ and DCC are

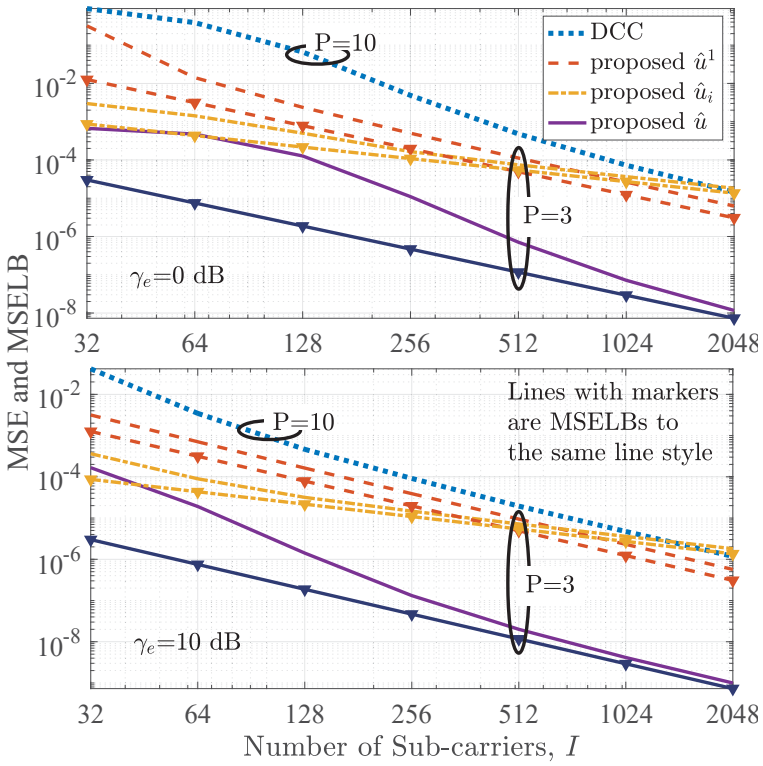

Figure 3.4 : MSE of the state-of-the-art DCC [1], $\hat{u}_i$, $\hat{u}$ and $\hat{u}^1$ versus I.

obtained from the second cross-correlation over sub-carriers, and thus suffer from the low SNR than the proposed $\hat{u}_i$ and $\hat{u}$; c.f. (3.23) and (3.38). Moreover, both $\hat{u}^1$ and DCC suffer from the amplification of estimation error pertaining to the linear scaling coefficient $\frac{1}{N\delta_{\rho_i}}$; see (3.6) and (3.42). Nevertheless, $\hat{u}^1$ provides better accuracy than DCC, due to the optimal frequency interval for cross-correlations across sub-carriers. It is worth noting that the MSE of $\hat{u}^1$ is obtained with only 3 symbols, while for DCC, 10 symbols are required.

Fig. 3.4 compares the MSE of AoA estimation between the proposed approaches and DCC [1], as the number of sub-carriers, I, increases. The superiority of the proposed approach is demonstrated in both MSE and the small number of training symbols. The proposed approach can reduce the MSE by orders of magnitude, especially for $\hat{u}$. In the case of (extremely) low SNRs and small numbers (even single) of sub-carriers, $\hat{u}_i$ and $\hat{u}$ can be far more accurate than $\hat{u}^1$. We also see that $\hat{u}^1$ is persistently better than DCC under different configurations, which again

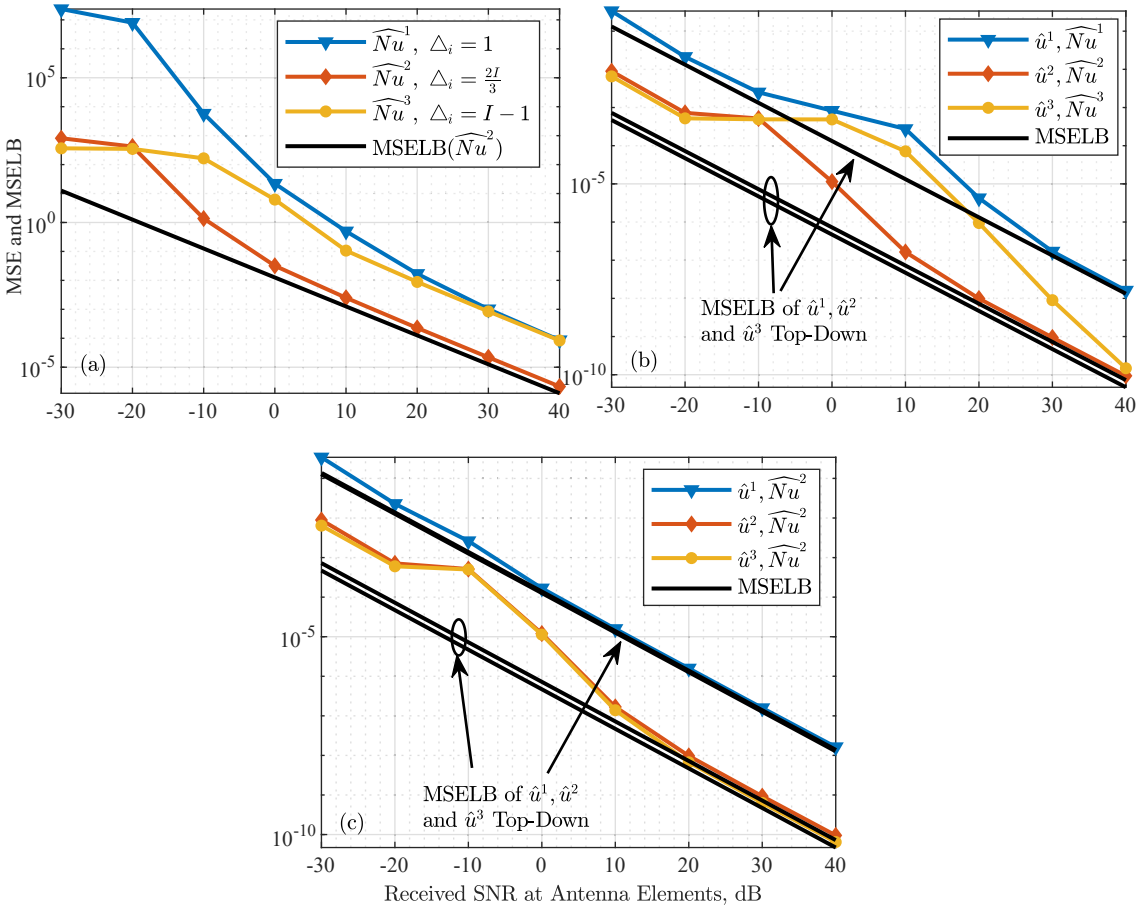

Figure 3.5 : (a) MSE of $\widehat{Nu}$ versus the receive SNR at antenna element, γ_e; (b) MSE of $\hat{u}$ versus γ_e, where $\hat{u}^1, \hat{u}^2$ and $\hat{u}^3$ stands for $\hat{u}$ obtained with $i_2 - i_1 = 1, \frac{2I}{3}$ and $I-1$, respectively, and the three Nu estimations are used in Algorithm 4 Line 8 to obtained the three u estimations; (c) the best Nu estimation $\widehat{Nu}^2$ is used for generating the three u estimations.

highlights the importance of holistic selection of $\triangle_i$.

Fig. 3.5a plots MSE of $\widehat{Nu}$ with the growth of γ_e, where $\triangle_i = 1, \frac{2I}{3}$ and $I-1$ are taken to illustrate the effect of $\triangle_i$ on Nu estimation. We confirm that $\triangle_i = \frac{2I}{3}$ gives the best Nu estimation, as dictated in Corollary 3.1. We also see the selection $\triangle_i = \frac{2I}{3}$ allows for asymptotically approaching the MSELB (3.25), where the constant gap of 11.1 dB between analytical and simulated estimation SNR, observed from Fig. 3.1, has been compensated.

Fig. 3.5b plots the MSE of $\hat{u}$, obtained by using the proposed Algorithm 4 upon the three different approaches for Nu estimation in Fig. 3.5a, and different number

of sub-carriers are combined. We see that the estimation accuracy of Nu can have a strong impact on u estimation. Fig. 3.5c correspondingly plots the MSE of $\hat{u}$ upon the best estimation $\widehat{Nu}^2$ is used, and the three different numbers of sub-carriers are accumulated. As expected, we can see that $\hat{u}$ can be increasingly accurate, as the number of accumulated sub-carriers increases. This indicates the importance of accurate estimation of Nu and holistic selection of $\triangle_i$ on the estimation accuracy of u. Note that the irregular turning point observed in Figs. 3.3 and 3.5 at -10 dB is mainly caused by the different estimation errors of $\widehat{Nu}$ before and after -10 dB.

3.8 Conclusion

In this chapter, we propose new subarray-specific time-varying phase shifts, which enable the cross-correlations of receive signal between subarrays to have deterministic phase changes, and hence be coherently accumulated to deliver unambiguous and noise-tolerant estimation of Nu and u. We also optimize the frequency interval for coherent accumulation across sub-carriers, minimizing the estimation error of Nu and leveraging the accumulation gain. Evident from simulations, our approach is able to dramatically improve the estimation accuracy by orders of magnitudes, with significantly reduced requirements of complexities and training symbols, as well as enhanced robustness against noises.

Chapter 4

Efficient Angle-of-Arrival Estimation of Lens Antenna Arrays

AoA estimation are critical to the efficiency of WIPT. The AoA estimation is challenging for energy-efficient LAAs, due to discrete sets of fixed DFT beams. This chapter presents a novel fast and accurate approach for the AoA estimation of LAAs. The key idea is that we prove the two differential outputs of three adjacent lens beams, referred to as "DFT beam differences (DBDs)", are the strongest at the two sides of an AoA. They are easy to identify and robust to noises, and their powers are proved to provide an accurate estimate of the AoA. Another important aspect is a new beam synthesis technique which produces different beam widths based on DFT beams and practical 1-bit phase shifts in real time. As a result, the angular region containing the AoA can exponentially narrow down, and the two strongest DBDs can be quickly identified. The proposed approach can operate in coupling with successive interference cancellation to estimate the AoAs of multiple paths. Simulations show that the proposed approach is able to outperform the state of the art by orders of magnitude in term of accuracy. The power transfer efficiency can be dramatically improved.

4.1 Introduction

In this chapter, we propose a novel efficient and accurate approach for the AoA estimation of LAAs for WIPT. The key idea is that we propose to use the difference of the outputs between adjacent lens DFT beams, referred to as "DFT beam difference (DBD)". We prove that the two DBDs around an AoA to be estimated have the stronger outputs than other DBDs and any DFT beams. Therefore, they are more robust against noises, and more likely to be correctly identified. We also prove that

the AoA can be accurately estimated based on the signal strengths of the two DBDs, independent of the signal itself. This is particularly valuable to WIPT, suppressing the need for training signals and saving energy consumptions of wirelessly powered devices.

Another important aspect is that we develop a new beam synthesis technique to produce beams with configurable widths based on DFT beams and practical 1-bit phase shifts in the analog domain. As a result, the angular region with the occurrence of a path can be recursively narrowed down to within K contiguous DFT beams by using only $(\log_K^N + 1)$ symbols (as compared to $\frac{N}{K}$ symbols by the state-of-the-art auxiliary beam pair (ABP) [5]). K is the number of RF chains. $(K + 1)$ DBDs can be generated, the two strongest of which can be identified for the accurate estimation of the AoA.

Other contributions of this chapter also include the AoA estimation of multiple paths by conducting the proposed approach in coupling with SIC, where both cases of concurrent WPT and WIT, and SWIPT are considered. Corroborated by extensive simulations and analyses, the proposed approach is able to outperform the state-of-the-art ABP by orders of magnitude in terms of estimation accuracy, especially in low SNR regions. With the exploitation of DBDs, the time-average beamforming gain and hence the power transfer efficiency of a LAA based on the proposed approach are able to dramatically surpass those of a DAA based on ABP.

The rest of this chapter is organized as follows. In Section 4.2, the system model is presented. In Section 4.3, the accurate AoA estimation algorithm is developed based on DBDs, and the estimation accuracy is analyzed. In Section 4.4, the proposed beam synthesis method is proposed, based on which the fast search of the two strongest DBDs is developed. In Section 4.5, the proposed approach is conducted in coupling with SIC to estimate multiple AoAs in both cases of concurrent WPT and WIT, and SWIPT. The superiority of the proposed approach is demonstrated through simulations in Section 4.6, followed by conclusions in Section 4.7.

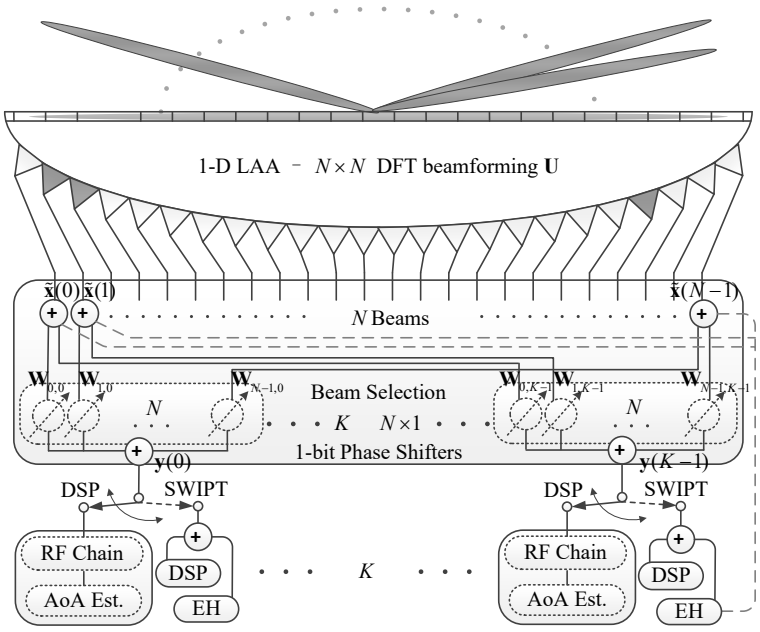

Figure 4.1 : Schematic diagram of the LAA receiver for concurrent WPT and WIT, or SWIPT.

4.2 System Model

Fig. 4.1 illustrates an N-dimensional LAA receiver, where the LAA can generate N DFT beams (this is equivalent to the $N \times N$ DFT beamforming) [40–42]. The N antennas are on the focal surface of the lens. Each antenna corresponds to a DFT beam and outputs the received signals from the beam. We employ the beam selecting network developed in [43] for beam selection, also known as antenna selection in [41, 42]. The output signal of a DFT beam is evenly split between K number of 1-bit phase shifters (c.f., "on/off" switches [39–42]). We further assume that the phase shifters can be turned off, as can be readily implemented in practice [43]. As a result, each phase shifter can take three different states, "$+1(= e^{j\frac{2\pi\times 0}{2}})$", "$-1(= e^{j\frac{2\pi\times 1}{2}})$", or "0 (turned off)". The outputs of N phase shifters connecting the N antennas can be combined and fed to an RF chain for information detection and/or energy harvesting.

As extensively assumed in the literature [31], a narrowband multi-path environment is considered. The N-dimensional received signal at the antennas is given by

[35]

$$\mathbf{x} = \sum_{p=0}^{P-1} \beta_p \mathbf{a}(u_p) s_p + \mathbf{n}^{\mathrm{e}}, \tag{4.1}$$

where u_p is the AoA of the p-th path of the energy signal in the spatial frequency domain, and $u_p = \frac{2\pi d \sin\theta_p}{\lambda} = \pi \sin\theta_p$; θ_p is the AoA in the physical angle domain; λ is the wavelength; and d is the antenna spacing, typically $d = \frac{\lambda}{2}$ [39]. $\mathbf{a}(u_p) \in \mathbb{C}^{N\times 1}$ is the steering vector of the LAA for the p-th path. β_p is the path loss labeled in the descending order of amplitude gain, i.e., $|\beta_0| \geq |\beta_1| \geq \cdots \geq |\beta_{P-1}|$. This labeling is consistent with SIC, where the paths are detected and canceled one after another, from the strong to the weak. s_p is the signal of the p-th path. P is the total number of paths. $\mathbf{n}^{\mathrm{e}} \in \mathbb{C}^{N\times 1}$ is the external additive white Gaussian noise (AWGN). The steering vector $\mathbf{a}(u_p)$ $(p = 0, 1, \cdots, P-1)$ can be given by

$$\mathbf{a}(u_p) = \left[1, e^{-ju_p}, \cdots, e^{-j(N-1)u_p}\right]^T, \ j = \sqrt{-1}. \tag{4.2}$$

- In the case of concurrent WPT and WIT, the transmitter of the energy and the transmitter of information are separate. The strong paths $p = 0, 1, \cdots, P_i - 1 < P-1$ carry energy; and the rest of the paths, $p = P_i, \cdots, P-1$, carry information.
- In the case of SWIPT, a single transmitter sends modulated energy signals, $s_0 = s_1 = \cdots = s_{P-1}$.

We consider single-tone sinewave energy signals which have been widely used for WPT applications [1], [2], [19], [30], [31], due to the fact that the signal has an constant envelope and can therefore maximize the efficiency of power amplifiers at the transmitters [19]. The algorithms developed for singletone sinewave signals in this chapter can be extended to multitone signals (e.g., OFDM signals) by referring to our recent work [25], where the estimation results at each of the tones can be constructively combined to further improve the estimation accuracy.

Let $\mathbf{U} \in \mathbb{C}^{N\times N}$ denote the normalized $N \times N$ DFT matrix, i.e., $\mathbf{U}(a, b) = \frac{1}{N} e^{-j\frac{2\pi ab}{N}}$ $(0 \leq a, b \leq N-1)$. The received signals of the N DFT beams, denoted

by $\tilde{\mathbf{x}} \in \mathbb{C}^{N\times 1}$, can be written as

$$\tilde{\mathbf{x}} = \sum_{p=0}^{P-1} \beta_p \mathbf{U}^{\mathrm{H}} \mathbf{a}(u_p) s_p + \mathbf{U}^{\mathrm{H}} \mathbf{n}^{\mathrm{e}}, \tag{4.3}$$

where $(\cdot)^{\mathrm{H}}$ stands for conjugate transpose.

Let $\mathbf{W} \in \mathbb{R}^{N\times K}$ denote the matrix of the total $N \times K$ states of all the 1-bit phase shifters. The k-th column of $\mathbf{W}$, denoted by $\mathbf{W}_{:,k} \in \mathbb{R}^{N\times 1}$ $(k = 0, 1, \cdots, K-1)$, collects the states of the N phase shifters connected to the k-th RF chain; and $\mathbf{W}_{n,:} \in \mathbb{R}^{1\times K}$ collects the states of the K number of 1-bit phase shifters connecting the n-th antenna and the K RF chains. Note that the received signal of antenna n is split evenly into $\|\mathbf{W}_{n,:}\|_0$ parts if $\|\mathbf{W}_{n,:}\|_0$ out of K phase shifters connected to the n-th antenna are turned on simultaneously. $\|\cdot\|_0$ gives the sparsity of a vector.

The output of the k-th RF chain is given in (4.4),

$$\begin{aligned}\mathbf{y}(k) =& \mathbf{W}_{:,k}^{\mathrm{H}} \Big(\mathbf{b} \odot \tilde{\mathbf{x}} + \mathbf{n}_k^{\mathrm{p}} \Big) + \mathbf{n}^{\mathrm{d}}(k) \\ =& \sum_{p=0}^{P-1} \beta_p \mathbf{W}_{:,k}^{\mathrm{H}} \Big(\mathbf{b} \odot \mathbf{U}^{\mathrm{H}} \mathbf{a}(u_p) \Big) s_p + \underbrace{\mathbf{W}_{:,k}^{\mathrm{H}} \Big(\mathbf{b} \odot \mathbf{U}^{\mathrm{H}} \mathbf{n}^{\mathrm{e}} + \mathbf{n}_k^{\mathrm{p}} \Big)}_{\mathbf{n}^{\mathrm{a}}(k)} + \mathbf{n}^{\mathrm{d}}(k),\end{aligned} \tag{4.4}$$

where $\mathbf{y} \in \mathbb{C}^{K\times 1}$ collects the outputs of the K RF chains, $\mathbf{y}(k)$ is the k-th element of $\mathbf{y}$, $\mathbf{n}_k^{\mathrm{p}} \in \mathbb{C}^{N\times 1}$ collects the noises of the N 1-bit phase shifters connected to the k-th RF chain, and $\mathbf{n}^{\mathrm{d}}(k)$ is the noise at the k-th RF chain. $\odot$ stands for Hadamard product. The n-th element of $\mathbf{b}$ is $\mathbf{b}(n) = \frac{1}{\|\mathbf{W}_{n,:}\|_0} \leq 1$. $\mathbf{b} \in \mathbb{R}^{N\times 1}$ is multiplied due to the splitting of the received signal of the antenna n.

We assume that $\mathbf{n}^{\mathrm{e}}$, $\mathbf{n}_k^{\mathrm{p}}$ $(k = 0, 1, \cdots, K-1)$ and $\mathbf{n}^{\mathrm{d}}$ are independent and identically distributed (i.i.d). σ_{e}^2, σ_{p}^2 and σ_{d}^2 are their respective variances. We have the following findings:

1. *As the system frequency increases,* $\mathbf{n}^{\mathrm{d}}$ *increasingly outgrows* $\mathbf{n}^{\mathrm{e}}$*, i.e.,* $\sigma_{\mathrm{e}}^2 \ll \sigma_{\mathrm{d}}^2$ *[103, 104].* For example, the noise figure of $\mathbf{n}^{\mathrm{d}}$ is 12 dB at 10 GHz; while the noise figure of $\mathbf{n}^{\mathrm{e}}$ is typically -8 dB and -15 dB in urban and suburban areas,

respectively [104, Fig. 3.1];

2. *The power of* $\mathbf{n}_k^{\mathrm{p}}$ $(k = 0, 1, \cdots, K-1)$ *can be far smaller than that of* $\mathbf{n}^{\mathrm{d}}$*, i.e.,* $\sigma_{\mathrm{p}}^2 \ll \sigma_{\mathrm{d}}^2$, since the noises incurred at the 1-bit phase shifters can be much weaker than those from the electronic components in the RF chains [43, 103].

Accordingly, the power of $\mathbf{n}^{\mathrm{a}}(k) + \mathbf{n}^{\mathrm{d}}(k)$ is no greater than $(1 + r\|\mathbf{W}_{:,k}\|_0)\sigma_{\mathrm{d}}^2$ $(k = 0, 1, \cdots, K-1)$, since $\mathbf{b}(n) \leq 1$ $(n = 0, 1, \cdots, N-1)$, where

$$r = \frac{\sigma_{\mathrm{e}}^2 + \sigma_{\mathrm{p}}^2}{\sigma_{\mathrm{d}}^2} \ll 1. \tag{4.5}$$

From (4.4), we notice that there can be mutual interferences between energy and information signals in the case of concurrent WPT and WIT. The energy signal received from the sidelobes, particularly the first several sidelobes, can cause destructive interference to the weak information signal received from the mainlobe. This is due to the large power difference between the energy and information signals, e.g., around typically 90 dB [34]. On the other hand, the mainlobe-to-sidelobe power ratio is -13.26 dB, -17.83 dB, and -20.82 dB for the first three sidelobes [95]. For these reasons, the accurate AoA estimation of the strong WPT paths is critical to the AoA estimations of both the weak WPT paths and weaker WIT paths. In the case of SWIPT, the accurate AoA estimation of the paths is important to accumulate energy of the paths and improve power transfer efficiency.

4.3 Fine-resolution AoA Estimation

In this section, we propose the new AoA estimation method with the focus on the strongest path. For illustration convenience, our elaboration of the proposed method is focused on one path (without loss of generality, the strongest path). The AoA estimations of other paths will be discussed in Section 4.5, separately, for the case of concurrent WPT and WIT, and the case of SWIPT. The estimation of the path is elemental to the estimations of multiple paths.

4.3.1 DBD

Distinctively different from existing studies of high-resolution AoA estimation where the DFT beams are focused on [5, 7, 73], we propose to estimate the AoA of a path by exploiting the difference of the received signals of adjacent DFT beams – DBD. DBDs can be readily obtained based on the DFT beams. Specifically, by setting the N 1-bit phase shifters connected to the k-th RF chain as

$$\mathbf{W}_{:,k} = \left[0, \cdots, 0, \underbrace{1}_{n\text{-th}}, -1, 0, \cdots, 0\right]^{\mathrm{T}}, \tag{4.6}$$

the n-th DBD can be obtained for $n = 0, 1, \cdots, N-2$. For $n = N-1$, we have $\mathbf{W}_{:,k} = \left[\underbrace{-1}_{0\text{th}}, 0, \cdots, 0, 1\right]^{\mathrm{T}}$, based on the cyclicity of the DFT matrix.

Let $g(n, u)$ denote the spatial response of the n-th DFT beam at the AoA of interest u. Based on (4.2), we have

$$\begin{aligned} g(n,u) = \mathbf{U}_{:,n}^{\mathrm{H}}\mathbf{a}(u) &= \frac{1}{N}\sum_{n'=0}^{N-1} e^{-jn'\left(u-\frac{2\pi n}{N}\right)} \\ &= e^{-j\frac{N-1}{2}\left(u-\frac{2\pi n}{N}\right)}\frac{\sin\frac{N}{2}\left(u-\frac{2\pi n}{N}\right)}{N\sin\frac{1}{2}\left(u-\frac{2\pi n}{N}\right)},\ u \in [0, 2\pi), \end{aligned} \tag{4.7}$$

where $\mathbf{U}_{:,n}$ is the n-th column of the normalized DFT matrix. Accordingly, the spatial response of the n-th DBD at u, denoted by $g_{\mathrm{d}}(n, u)$, can be written as

$$\begin{aligned} g_{\mathrm{d}}(n,u) &= g(n,u) - g(n+1,u) \\ &= -\frac{1}{N}e^{-j\frac{N}{2}\left(u-\frac{2\pi n}{N}\right)}\frac{\sin\frac{N}{2}\left(u-\frac{2\pi n}{N}\right)\sin\frac{\pi}{N}}{\sin\frac{1}{2}\left(u-\frac{2\pi n}{N}\right)\sin\frac{1}{2}\left(u-\frac{2\pi(n+1)}{N}\right)}. \end{aligned} \tag{4.8}$$

Fig. 4.2 illustrates the synthesis of a DBD from two consecutive DFT beams, where both the amplitude and phase responses of the two DFT beams are plotted. It is shown that the DFT beams have the same amplitudes at the intersection of their mainlobes, but opposite phases. By synthesizing the DFT beams differentially, the DBD aligns the phases of the DFT beams at the intersection point and enhances

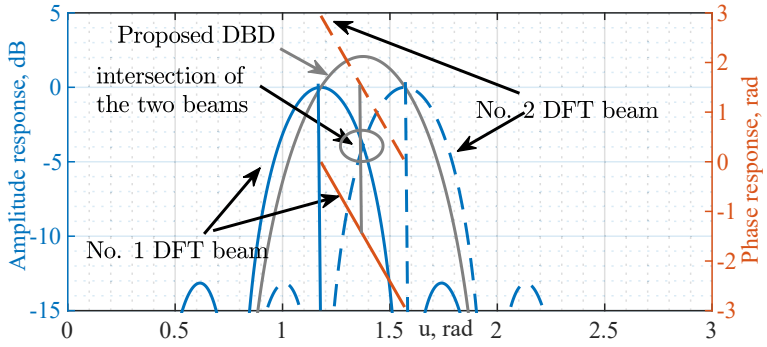

Figure 4.2 : Illustration of two adjacent DFT beams including their amplitude and phase responses.

the amplitude response. Another useful feature of DBDs follows.

Lemma 4.1. *For a path with $u \in \Omega_n$ $(n = 0, 1 \cdots, N-1)$, the n-th and $(n+1)_N$-th DBDs have the larger amplitude gains than the rest of the other DBDs, where $\Omega_n = \left[\frac{2\pi n}{N} + \frac{\pi}{N}, \frac{2\pi(n+1)_N}{N} + \frac{\pi}{N}\right)$ and $(\cdot)_N$ denotes the modulo-N operation.*[*]

Proof. See Appendix 7.3.1. □

From Lemma 4.1, the n-th and $(n+1)$-th DBDs have the larger amplitude gains than the rest of the DBDs for the path with $u \in \Omega_n = [\frac{2\pi n}{N} + \frac{\pi}{N}, \frac{2\pi n}{N} + \frac{3\pi}{N})$. The two DBDs can be identified for AoA estimation by comparing the amplitudes among the received signals of all the N DBDs. Given the highest amplitude gains of the two DBDs, the received signals of the two DBDs are more tolerant to the noises, and less likely to be missed, than those of the other DBDs. Fig. 4.3 shows that the AoA to be estimated must be in the mainlobes of two consecutive DBDs, as asserted in Lemma 4.1. Provided the receiver noises are negligible, the received signals of the two DBDs can be used to deterministically estimate the AoA, as to be established in Theorem 4.1.

In [5], the two DFT beams with the strongest outputs, i.e., the n-th and $(n+1)$-th DFT beams, were used for the AoA estimation of $u \in \Omega_n$. By exploiting DBDs,

[*]The modulo-N is used in (8) to capture the case where $n = N-1$ and the $(N-1)$-th and $(N-1+1)_N$th (i.e., 0th) DBDs form the pair of consecutive DBDs to estimate the AoAs in the region of $\left[\frac{2\pi(N-1)}{N} + \frac{\pi}{N}, 2\pi\right) \cup \left[0, \frac{\pi}{N}\right)$. In the following, for notational simplicity, we suppress the modulo operation.

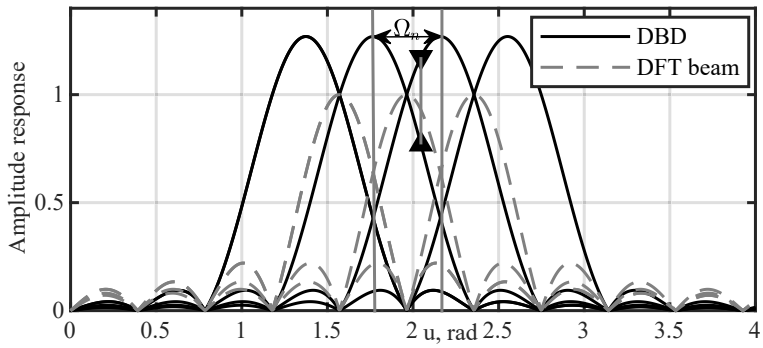

Figure 4.3 : Illustration of amplitude gains of DBDs, where four consecutive DBDs are illustrated with the AoA of $u = 2.05$ rad to be estimated.

the probability of correctly identifying the n-th and $(n+1)$-th DBDs, denoted by $P_{\rm d}^{\rm DBD}$, is greater than the probability of correctly identifying the corresponding DFT beams, denoted by $P_{\rm d}^{\rm DFT}$; i.e., $P_{\rm d}^{\rm DBD} \geq P_{\rm d}^{\rm DFT}$, as proved in Appendix 7.3.5. This improved probability can lead to a better estimation accuracy, as will be shown in Section 4.6.

4.3.2 DBD-based AoA Estimation

We proceed to prove that for any $u \in \Omega_n$, by exploiting the received signals of the n-th and $(n+1)$-th DBDs ($n = 0, 1, \cdots, N-1$), the AoA, u, can be estimated. By first substituting (4.6) into (4.4), and then combining with (4.8), the received signal of the n-th DBD to the k-th RF chain can be written as[†]

$$\mathbf{y}_n(k) = -\frac{\beta s}{N} e^{-j\frac{N}{2}\tilde{u}} \frac{\sin\frac{N}{2}\tilde{u}\sin\frac{\pi}{N}}{\sin\frac{1}{2}\tilde{u}\sin\frac{1}{2}(\tilde{u}-\frac{2\pi}{N})} + \mathbf{n}(k), \tag{4.9}$$

where $\tilde{u} = u - \frac{2\pi n}{N}$ and $\mathbf{n}(k) = \mathbf{n}^{\rm a}(k) + \mathbf{n}^{\rm d}(k)$; see (4.4). The received signal of the $(n+1)$-th DBD to the l-th RF chain is $\mathbf{y}_{n+1}(l)$. $k \neq l$, and $0 \leq k, l \leq N-1$. Given $u \in \Omega_n$, we have $\tilde{u} = u - \frac{2\pi n}{N} \in \left[\frac{\pi}{N}, \frac{3\pi}{N}\right)$. $\tilde{u}$, and in turn $u = \tilde{u} + \frac{2\pi n}{N}$, can be estimated.

In the following, we first establish the proposed AoA estimator in the absence of noises, as given in Theorem 4.1; and then evaluate the impact of non-negligible noises

[†]For notational simplicity, we suppress the subscript for β and s, since the strongest path is estimated at first.

on the estimation accuracy, as provided in Corollary 4.1. Such research approach has been taken by the state-of-the-art ABP [5], as well as the latest AoA estimators for hybrid antenna arrays [7, 73].

Theorem 4.1. *In the absence of noises (i.e.,* $\mathbf{n}(k) = 0$*), for* $\frac{\pi}{N} \leq \tilde{u} = u - \frac{2\pi n}{N} \leq \frac{3\pi}{N}$*,* $\tilde{u}$ *can be evaluated from the ratio* $\rho = \rho(\tilde{u}, n) = \frac{|\mathbf{y}_n(k)|^2 - |\mathbf{y}_{n+1}(l)|^2}{|\mathbf{y}_n(k)|^2 + |\mathbf{y}_{n+1}(l)|^2}$*:*

$$\hat{\tilde{u}} = \frac{2\pi}{N} - \arcsin\left(\frac{\rho \sin\frac{2\pi}{N}\left(\text{sign}\{\rho\}\sqrt{1-\rho^2}\cos\frac{2\pi}{N} - 1\right)}{\sin^2\frac{2\pi}{N} + \rho^2\cos^2\frac{2\pi}{N}}\right), \tag{4.10}$$

where $\text{sign}\{\cdot\}$ *takes sign. The AoA estimate, denoted by* $\hat{u}$*, can be given by* $\hat{u} = \frac{2\pi n}{N} + \hat{\tilde{u}}$*.*

Proof. See Appendix 7.3.2. □

The impact of non-negligible noises on the estimation error is then analyzed in Corollary 4.1:

Corollary 4.1. *The mean square errors (MSEs) of* $\hat{\tilde{u}}$ *and* $\hat{u}$*, denoted by* $\sigma^2_{\hat{\tilde{u}}}$ *and* $\sigma^2_{\hat{u}}$*, respectively, are* $\sigma^2_{\hat{u}} = \sigma^2_{\hat{\tilde{u}}} = \mathbb{E}\{\hat{u}^2\} \approx \frac{\epsilon}{2\gamma_u}$*, where* $\epsilon = \frac{1+\rho^2(\tilde{u},n)}{\eta^2}$*,* $\eta = \frac{\mathrm{d}\rho(\tilde{u},n)}{\mathrm{d}\tilde{u}}\Big|_{\tilde{u}=\frac{2\pi}{N}}$*, and* $\gamma_u = \frac{\sigma^2_{\text{s}}}{\sigma^2_{\text{n}}}$ *is the estimation SNR.*

Proof. See Appendix 7.3.3. □

4.4 Fast Search of Two Strongest DBDs

As described in Section 4.3, the AoA of a path can be accurately estimated from the received signals of two DBDs, if the path is in the overlapped angular region of the mainlobes of the two DBDs. As also proved, the two DBDs can be more reliably identified than the other DBDs, as they have higher amplitude gains in the direction of the path. In this section, we propose a fast search of the two DBDs for AoA estimation (as opposed to enumerating all the N DBDs by using the K RF chains and hence taking at least $\frac{N}{K}$ symbols). This starts with a new beam synthesis technique, as proposed in the following.

Theorem 4.2. *Given the N DFT beams, $g(n,u)$, $n=0,1,\cdots,N-1$, and the $N\times 1$ vector of the states of the N phase shifters connected to the k-th RF chain:*

$$\mathbf{W}_{:,k}=\left[\underbrace{0,\cdots,0}_{m},\underbrace{(-1)^m,\cdots,(-1)^{m+K-1}}_{K},\underbrace{0,\cdots,0}_{N-K-m}\right]^{\mathrm{H}}, \tag{4.11}$$

the beam synthesized by $\sum_{n=0}^{N-1} g(n,u)\mathbf{W}_{n,k}$ approximates the following beam

$$P(u)=\begin{cases} e^{j\frac{N}{2}u}, & \frac{2\pi m}{N}\le u\le\frac{2\pi(m+K-1)}{N};\\ 0, & \text{otherwise,}\end{cases} \tag{4.12}$$

where m is an integer satisfying $1\le m\le N-K$, and $\mathbf{W}_{n,k}$ is the state of the n-th phase shifter connected with the k-th RF chain; see Fig. 4.1.

Proof. See Appendix 7.3.4. □

By exploiting Theorem 4.2, we can recursively steer K simultaneous synthesized beams to probe the angular region of interest (one beam per RF chain), identify the beam with the strongest received signal, and narrow down the angular region by $\frac{1}{K}$ times. By repeating this for up to $(\log_K^N-1)$ rounds (or in other words, symbols), we are able to identify the angular region with the strongest output and the width of K contiguous DFT beams.

In the next two symbols, we propose to use DBDs to scan the angular region with the width of K DFT beams. For illustration convenience, we assume that the angular region of interest spans from the n-th DFT beam through the $(n+K-1)$-th DFT beam. Consider the boundary of the angular region. The $(n-1)$-th and $(n+K)$-th DFT beams are also taken into account for the DBD generation.

- In the first of the two symbols, a set of $(\frac{K}{2}+1)$ DBDs, i.e., the $(n+2k-1)$-th DBD, are generated as the differences between $(n+2k-1)$-th and $(n+2k)$-th DFT beams for $k=0,1,\cdots,\frac{K}{2}$;
- In the second of the two symbols, a set of $\frac{K}{2}$ DBDs, i.e., the $(n+2k)$-th DBD, are

Algorithm 5 Fine-resolution AoA Estimation

Input K and N. **Initialize**: $t = 0$;
Generate K beams to cover $[0, 2\pi)$ using Theorem 4.2;
for $1 \leq t \leq \log_K^N - 2$ **do**
    Update the angular region to be detected to the mainlobe of the beam with the strongest output at $t - 1$;
    Update K beams to cover the reduced angular region;
end for
The angular region is reduced to have the width of K DFT beams. **Identify** the starting index of the DFT beam, denoted by n $(0 \leq n \leq N - 1)$;
At $t = \log_K^N - 1$, the $(n + 2k - 1)$-th DBD, are generated as the differences between $(n + 2k - 1)$-th and $(n + 2k)$-th DFT beams for $k = 0, 1, \cdots, \frac{K}{2}$;
At $t = \log_K^N$, the $(n + 2k)$-th DBD, are generated as the differences between $(n + 2k)$-th and $(n + 2k + 1)$-th DFT beams for $k = 0, 1, \cdots, \frac{K}{2} - 1$;
Let $\mathbf{y}_t \in \mathbb{C}^{(K+1)\times 1}$ collect the $(K+1)$ DBD outputs. **Identify** the two strongest elements in $\mathbf{y}_t$ with the indices denoted by k and $(k + 1)$;
Perform the fine-resolution AoA estimation by exploiting Theorem 4.1 based on $\mathbf{y}_t(k)$ and $\mathbf{y}_t(k + 1)$. **Return** $\hat{u}$.

generated as the differences between $(n + 2k)$-th and $(n + 2k + 1)$-th DFT beams for $k = 0, 1, \cdots, \frac{K}{2} - 1$.

The total $(K + 1)$ received signals of the DBDs, i.e., the $(n - 1)$-th, n-th, $\cdots$, $(n + K - 1)$-th, can be compared.

The two consecutive DBDs with the strongest received signals can be reliably identified and used as the input of (4.10) to the AoA estimation of the path, as discussed in Section 4.3.2. We use the two strongest DBDs, rather than the two strongest DFT beams (as done in [5]), since the probability of correctly identifying the two strongest DBDs is higher, as revealed in Section 4.3.1. Collecting the received signals of the $(K + 1)$ DBDs at two symbols can prevent power splitting between the DBDs; otherwise, the n-th and $(n + 1)$-th DBDs are generated at the same symbol at different RF chains, $\mathbf{b}(n + 1) = \frac{1}{2}$; see (4.4), and the estimation SNR at the symbol would reduce.

Algorithm 5 is formally established to accurately estimate the AoA of a path. It starts with the fast search of the angular region with the width of K DFT beams and the strongest received signal, by using the proposed beam synthesis technique

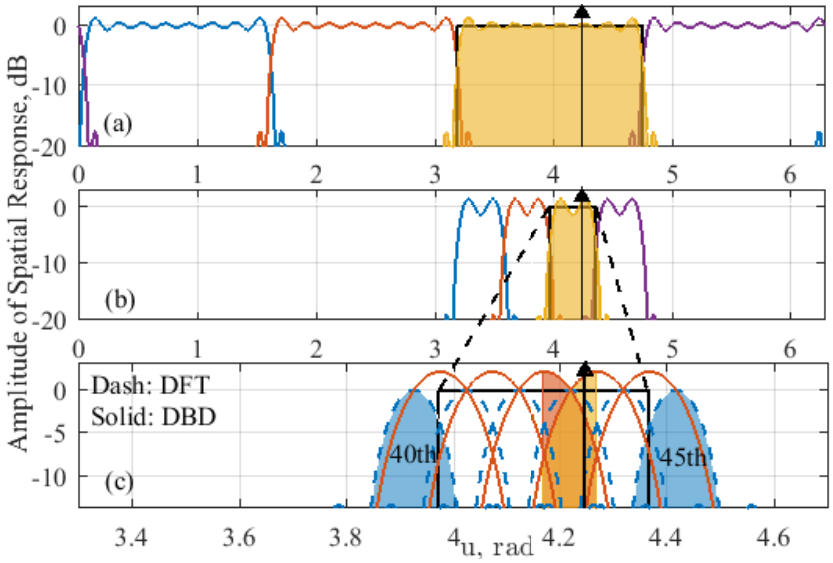

Figure 4.4 : Illustration of the search of the DBDs for AoA estimation using the beams synthesized by Theorem 4.2, where $N = 64$, $K = 4$ and $u = 4.25$ rad.

in Theorem 4.2; see Steps 2 to 6. Then the algorithm generates $(K + 1)$ DBDs to scan the angular region, and identifies the two adjacent DBDs with the strongest received signals based on Lemma 1; see Steps 7 to 9. Finally, the AoA estimation can be accurately estimated by using Theorem 1; see Steps 10 and 11.

Fig. 4.4 illustrates the search of DBDs for AoA estimation, where the amplitudes of the synthesized beams in each iteration are plotted. $N = 64$, $K = 4$ and $u = 4.25$ rad. We see that in Fig. 4.4(a), four synthesized wide beams evenly cover $[0, 2\pi]$. The angular region of interest is recursively reduced by K times per iteration, as shown in Fig. 4.4(b). After the angular region with the width of K DFT beams is identified, DBDs are generated at two symbols and placed together, as shown in Fig. 4.4(c). We see that two adjacent DBDs can output much higher signal strengths in the direction of u, and they are much stronger than the corresponding DFT beams.

The similarities of the proposed approach to ABP are that both approaches identify the two consecutive beams with the strongest received signals, and estimate the AoA based on the received signals. The AoA is estimated based on the ratio between the difference and the sum of the powers of the received signals in both the proposed approach and ABP.

Table 4.1 : Comparison of the proposed approach to existing techniques in terms of computational and signaling overheads.

	ABP [5]	WDFT[2]	Proposed
Complexity†	$\mathcal{O}(N)$	$> \mathcal{O}(N)$	$\mathcal{O}\left(K(\log_K^N + 1)\right)$
Signaling‡	$\frac{N}{K}$	$\log_K^N$	$\log_K^N + 1$

†Computational complexity.
‡Signaling overhead, i.e., no. of training symbols.

A key difference of the proposed approach to ABP is that the proposed approach synthesizes DFT beams differentially to produce beams with configurable beamwidths and improved gains (as compared to the direct use of the DFT beams in ABP). As a result, the proposed approach can recursively narrow down the beams until the beams become the difference of two consecutive DFT beams (i.e., DBDs) and cannot be further narrowed. In contrast, ABP linearly scans the entire angular space by directly using the standard DFT beams with fixed beamwidths.

Another important difference is that the proposed approach estimates the AoA based on the strongest received signals of the two consecutive DBDs, while ABP is based on the strongest received signals of the two consecutive DFT beams. With the improved gain, the two strongest consecutive DBDs are much more likely to be correctly identified in the proposed approach than the two strongest consecutive DFT beams in ABP. This is particularly important in low SNR regions. In the case that both the DBDs and DFTs can be correctly identified in the proposed approach and ABP, e.g., in high SNR regions, the improved gain of the DBDs leads to lower estimation errors in the proposed approach.

Table 4.1 compares the proposed approach and the existing methods in terms of complexity and signaling overhead. The computational complexity of the proposed approach is $\mathcal{O}\left(K(\log_K^N + 1)\right)$, since the approach synthesizes K beams per symbol in the analog domain for $(\log_K^N - 1)$ symbols. With configurable beamwidths, the K beams keep narrowing down by $\frac{1}{K}$ times per symbol around the AoA of a path. The last two symbols are used to synthesize K DBDs out of K consecutive DFT beams in the analog domain. The complexity of ABP [5] is $\mathcal{O}(N)$, since

ABP scans K DFT beams per symbol for a total of $\frac{N}{K}$ symbols to assess all the N possible pairs of consecutive DFT beams. WDFT [2] synthesizes K beams per symbol with configurable beamwidths, and narrows the beams down by $\frac{1}{K}$ times at every symbol for $\log_K^N$ symbols. Different from the proposed approach, the beam synthesis of WDFT is carried out in the digital domain and has a complexity of $\mathcal{O}\left(\sum_{l=0}^{\log_K^N-1}\frac{N}{K^l}\right)=\mathcal{O}\left(\frac{N-1}{1-K^{-1}}\right)>\mathcal{O}(N)$, given the beam synthesis coefficients generated in prior. Here, $\frac{N}{K^l}$ is the number of DFT beams to synthesize each of the K beams at the l-th symbol ($l=0,1,\cdots,\log_K^N-1$). Moreover, WDFT only identifies the DFT beam in which the path is, and does not estimate the AoA of the path.

In the case of multiple paths, say L paths, Algorithm 1 can be extended to estimate the paths one after another by carrying out SIC, as will be described in Section V. The total number of symbols required is no greater than $L(\log_K^N+1)$, which can be still much lower than $\frac{N}{K}$, in the case where the number of significant paths, L, is less than $\frac{N}{K(\log_K^N+1)}$. The accuracy of Algorithm 1 can be much finer than a DFT beamwidth. This is important to implement SIC and recover typically weaker information signals in the case of concurrent WPT and WIT. It is also important to accurately estimate different paths for effective equalization and coherent accumulation of received signals in the case of SWIPT.

4.5 Multipath AoA Estimation

In this section, we elaborate on the application of the proposed Algorithm 5 to the estimation of multi-path AoAs, where both cases of concurrent WPT and WIT, and SWIPT are discussed. For illustration convenience, we consider two paths. In the case of concurrent WPT and WIT, the stronger path delivers energy signals, and the weaker carries information. In the case of SWIPT, both paths carry modulated energy signals. The discussions can be readily extrapolated to the scenarios with more paths.

From (4.4), (4.8) and (4.9), the received signal of n-th DBD to the k-th RF chain

at time t, can be given by

$$\begin{aligned}\mathbf{y}_n^t(k) &= \beta_1 s_t g_{\mathrm{d}}(n,\tilde{u}_1) + \beta_2 s_t g_{\mathrm{d}}(n,\tilde{u}_2) + \mathbf{n}_t(k) \\ &= -\frac{\beta_1 s_t}{N} e^{-j\frac{N}{2}\tilde{u}_1} \frac{\sin\frac{N}{2}\tilde{u}_1 \sin\frac{\pi}{N}}{\sin\frac{1}{2}\tilde{u}_1 \sin\frac{1}{2}(\tilde{u}_1 - \frac{2\pi}{N})} - \frac{\beta_2 s_t}{N} e^{-j\frac{N}{2}\tilde{u}_2} \frac{\sin\frac{N}{2}\tilde{u}_2 \sin\frac{\pi}{N}}{\sin\frac{1}{2}\tilde{u}_2 \sin\frac{1}{2}(\tilde{u}_2 - \frac{2\pi}{N})} + \mathbf{n}_t(k),\end{aligned} \tag{4.13}$$

where the definitions of β_i and $\tilde{u}_i$ $(i = 1, 2)$ can be referred to those of β and $\tilde{u}$ in (4.9), and $|\beta_1| \geq |\beta_2|$. Clearly, the interference between the two paths can affect the accuracy of the AoA estimation. We have the following finding.

Corollary 4.2. *The signal-to-interference ratio (SIR) in the n_1-th DBD for u_1 estimation, denoted by γ, satisfies*

$$\gamma = \frac{\left|\beta_1 s_t g_{\mathrm{d}}\left(n_1, u_1 - \frac{2\pi n_1}{N}\right)\right|^2}{\left|\beta_2 s_t g_{\mathrm{d}}\left(n_1, u_2 - \frac{2\pi n_1}{N}\right)\right|^2} \tag{4.14a}$$

$$\geq \frac{N^2 \sin^2\left(\frac{2\pi(n_2-n_1)+\pi}{2N}\right) \sin^2\left(\frac{2\pi(n_2-n_1)-\pi}{2N}\right)}{\sin^2\frac{\pi}{N}}, \tag{4.14b}$$

where n_1 and n_2 are the indices for the strongest DBDs for the two paths, respectively.

Proof. The lower bound of γ is attained in the case that the numerator on the RHS of (4.14a) takes the minimum while the denominator takes the maximum. According to the proof of Lemma 4.1, we have $\left|g_{\mathrm{d}}\left(n_1, u_1 - \frac{2\pi n_1}{N}\right)\right| \geq 1$, i.e., the minimum of the numerator is $|\beta_1 s_t|^2$.

To derive the maximum value of the denominator, we first prove that

$$\left|g_{\mathrm{d}}\left(n, \frac{2\pi(n'-n)}{N} + \frac{\pi}{N}\right)\right| \text{ for } \forall n' \in [0, N-1],$$

takes the maximums of the sidelobes of the n'-th DBD. The proof can be readily established based on (4.7) and (4.8), and therefore is suppressed for brevity.

As a result, the maximum of the denominator on the RHS of (4.14a) can be taken at the n_1-th DBD if $u_2 = \frac{2\pi n_2}{N} + \frac{\pi}{N}$. By substituting (4.8) and $u_2 = \frac{2\pi n_2}{N} + \frac{\pi}{N}$ into (4.14a), and exploiting $|\beta_1| \geq |\beta_2|$, we can finally obtain (4.14b). □

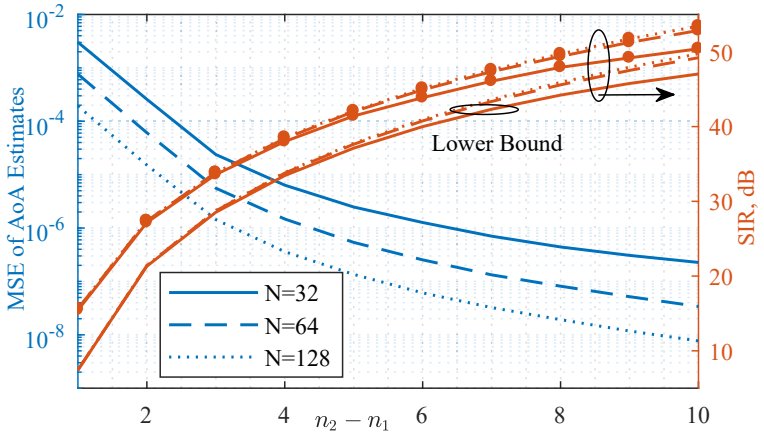

Figure 4.5 : The SIR between two equal-power paths and the MSE of the AoA estimate of one path in the presence of the interference from the other path vs. $(n_2 - n_1)$, where $n_1 = 1$ and the AoAs of the two paths are set as $u_1 \sim \mathcal{U}\left[\frac{\pi}{N}, \frac{3\pi}{N}\right]$, $u_2 \sim \mathcal{U}\left[\frac{2\pi(n_2-1)}{N} + \frac{\pi}{N}, \frac{2\pi(n_2-1)}{N} + \frac{3\pi}{N}\right]$, respectively.

Fig. 4.5 plots the SIR for and the MSE of u_1 estimation as $(n_2 - n_1)$ increases, where $|\beta_1| = |\beta_2|$, $n_1 = 1$, $u_1 \sim \mathcal{U}\left[\frac{\pi}{N}, \frac{3\pi}{N}\right]$, $u_2 \sim \mathcal{U}\left[\frac{2\pi(n_2-1)}{N} + \frac{\pi}{N}, \frac{2\pi(n_2-1)}{N} + \frac{3\pi}{N}\right]$, and $N = 32, 64$ and 128. $\mathcal{U}[a, b]$ denotes the uniform random distribution between a and b. The lower bounds for SIRs are also provided. We see that the SIR increases with the growth of $(n_2 - n_1)$, which is consistent with the lower bound (4.14b). We also see that the MSE of u_1 decreases with the growth of $(n_2 - n_1)$. Specifically, when $n_2 - n_1 > 3$, the MSE is less than 10^{-5} for different values of N. As will be shown in Section 4.6, the interference of the weaker path to the stronger is negligible in the case of $n_2 - n_1 > 3$; and the weaker path can also be estimated with high accuracy by canceling the stronger path in the case of $0 \leq n_2 - n_1 \leq 3$.

4.5.1 Concurrent WPT and WIT

By using the LS approximation [31], we can estimate the path response, as given by $\widehat{\beta_1 s_{t_1}} = \frac{\mathbf{y}_{n_1}^{t_1}(k)}{g_{\mathrm{d}}(n_1, \hat{u}_1 - \frac{2\pi n_1}{N})}$, where $\mathbf{y}_{n_1}^{t_1}(k)$ is strongest DBD output at symbol t_1 for u_1 estimation. By running Algorithm 5, the second strongest DBD output is at either $(t_1 - 1)$ or $(t_1 + 1)$. For illustration convenience, we assume that the second strongest signal is received by the $(n_1 - 1)$-th DBD output to the k-th RF chain at symbol $(t_1 - 1)$, and denoted by $\mathbf{y}_{n_1-1}^{t_1-1}(k)$. Given the single-tone sinewave energy signal,

$s_t = |s|e^{j\omega t+\phi}$, where $|s|$ is the constant envelope of the energy, ω is the baseband frequency, and ϕ is the initial phase (which is assumed to be constant during AoA estimation). ω can be estimated as $\hat{\omega} = \arg\left\{\frac{\mathbf{y}_{n_1}^{t_1}(k)}{\mathbf{y}_{n_1-1}^{t_1-1}(k)}\right\}$, where $\arg\{\cdot\}$ takes angle. The impinging energy signal carried by the first path can be recovered as $\widehat{\beta_1 s}e^{j\hat{\omega}t}$ for $\forall t$, where $s = |s|e^{j\phi}$ and $\widehat{\beta_1 s} = \frac{\widehat{\beta_1 s_{t_1}}}{e^{j\hat{\omega}t_1}}$.

The second path can be estimated at symbol $t_2 = t_1 + \log_K^N + 1$ and $t_2 - 1$ by subtracting the first path. Suppose that $\mathbf{y}_{n_2-1}^{t_2-1}(k)$ and $\mathbf{y}_{n_2}^{t_2}(k)$ are the (n_2-1)-th and n_2-th DBDs (with the first two strongest outputs), respectively. By carrying out SIC, the second path can be estimated based on

$$\tilde{\mathbf{y}}_{n'}^{t'}(k) = \mathbf{y}_{n'}^{t'}(k) - \widehat{\beta_1 s}e^{j\hat{\omega}\cdot t'} g_{\mathrm{d}}\left(n', \hat{u}_1 - \frac{2\pi n'}{N}\right), \tag{4.15}$$

where $t' = t_2 - 1, t_2$ and $n' = n_2 - 1, n_2$. This is because, as a single-tone sinewave, the energy signal is a-priori known to the receiver. The first path can be readily canceled while the information path is being estimated. The receiver does not have to spare an RF chain to monitor the first path.

Algorithm 6 summarizes the proposed multi-AoA estimation in the case of concurrent WPT and WIT. The initial estimation of the strong WPT path in Step 4 is critical for carrying out SIC. As will be evaluated in Figs. 4.11 and 4.12, with substantially improved accuracy of the initial estimation, the proposed approach is able to alleviate error propagations in SIC and provide good estimation accuracy for the WIT paths.

4.5.2 SWIPT

For illustration convenience, we consider the binary phase-shift keying (BPSK) modulated energy signals. During a symbol duration, the modulated energy signal is a single-tone sinewave. The AoA estimation of the first path can be achieved, as described in Section 4.3.2, since the proposed AoA estimation (i.e., Theorem 4.1) only depends on the amplitude of the signals. The amplitude, phase and the modulated signal can be estimated, as described in Section 4.5.1.

Algorithm 6 Multi-AoA Estimation for Concurrent WPT and WIT

1: Run Algorithm 5, producing $\hat{u}_1$, $\mathbf{y}_{n_1-1}^{t_1-1}(k)$ and $\mathbf{y}_{n_1}^{t_1}(k)$;
2: Estimate $\widehat{\beta_1 s_{t_1}} = \frac{\mathbf{y}_{n_1}^{t_1}(k)}{g_{\mathrm{d}}(n_1, \hat{u}_1 - \frac{2\pi n_1}{N})}$;
3: Estimate the sinewave frequency as $\hat{\omega} = \arg\left\{\frac{\mathbf{y}_{n_1}^{t_1}(k)}{\mathbf{y}_{n_1-1}^{t_1-1}(k)}\right\}$;
4: Calculate $\widehat{\beta_1 s} = \frac{\widehat{\beta_1 s_{t_1}}}{e^{j\hat{\omega}t_1}}$ and the received signal from the first path is receiver as $\widehat{\beta_1 s} e^{j\hat{\omega}t}$;
5: Run Algorithm 5 till Step 9;
6: Modify the selected DBD outputs based on (4.15);
7: Perform the last step of Algorithm 5 for u_2 estimation.

Algorithm 7 Multi-AoA Estimation for SWIPT

1: Run Algorithm 6 till Step 5;
2: Test the hypothesis that the modulated signal is "+1";
3: Run Steps 6 and 7 of Algorithm 6 for u_2 estimation, and Steps 2 to 4 for the estimation of the modulated signal;
4: Test the hypothesis that the modulated signal is "−1" by rerunning Step 3;
5: The u_2 estimation under the hypothesis that the modulated signal estimate of the second path is consistent with that of the first path is accepted as the final AoA estimate.

The estimation of the second path of the SWIPT signal is different in the case of SWIPT from it is in the case of concurrent WPT and WIT. This is because the SWIPT signals can be modulated differently at different symbols. While estimating the second path, the signal (on the first path) may have already changed from what it was when the first path was estimated. An uncertainty arises when SIC is used to cancel the first path for estimating the second path. Nevertheless, we can readily eliminate the uncertainty by setting up two different hypotheses of the current SWIPT signal (i.e., "+1" or "−1"), canceling the first path separately under the two hypotheses and estimating the two AoAs of the second path. The modulated signal of the second path can be estimated under the two different hypotheses. Given the fact that the first and second paths carry the same SWIPT signal, we accept the one of the hypotheses, under which the modulated signal of the second path is consistent with the hypothesis of the modulated signal on the first path. Algorithm 7 summarizes the proposed multi-AoA estimation in the case of SWIPT.

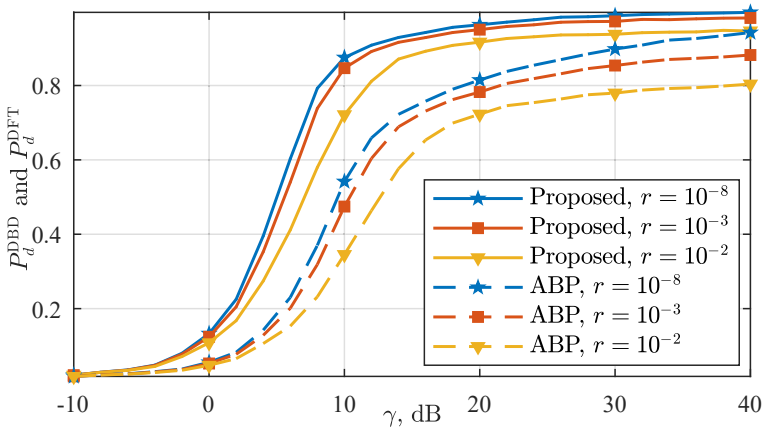

Figure 4.6 : $P_{\mathrm{d}}^{\mathrm{DFT}}$ of ABP [5] and $P_{\mathrm{d}}^{\mathrm{DBD}}$ of the proposed method vs. $\gamma = \frac{1}{\sigma_{\mathrm{d}}^2}$.

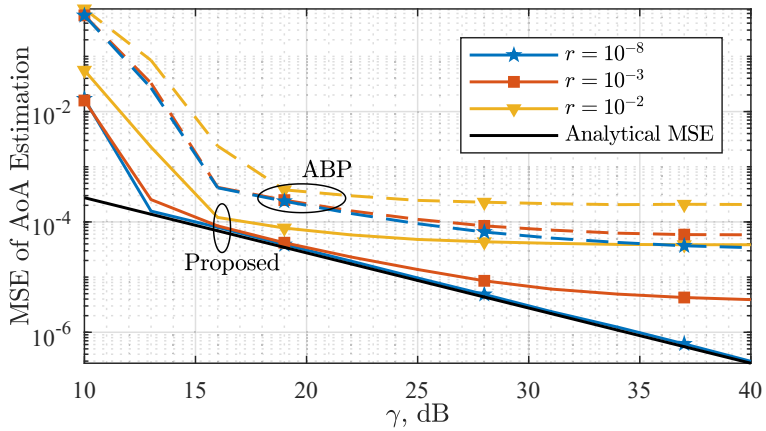

Figure 4.7 : MSE of the single-AoA estimates, using ABP [5] and the proposed algorithm (Algorithm 5), vs. $\gamma = \frac{1}{\sigma_{\mathrm{d}}^2}$.

4.6 Simulation Results

In this section, simulations are carried out to validate the proposed method and analysis, as shown in Fig. 4.1. Without loss of generality, linear LAA receivers are considered. Unless otherwise specified, the array configuration and other simulation parameters are set as follows. $N = 64$, $K = 4$ and $u \sim \mathcal{U}[-\pi, \pi]$. $\tilde{\gamma} = \frac{1}{(1+rL)\sigma_{\mathrm{d}}^2}$ is tested for $r = 10^{-8}$, 10^{-3} and 10^{-2}; refer to (4.5). Here, $\tilde{\gamma}$ has already accounted for the DFT beamforming gain. For fair comparison, the state of the art, namely, ABP [5], DFT-DA [6], and WDFT [2], are also simulated.

Fig. 4.6 compares $P_{\mathrm{d}}^{\mathrm{DFT}}$ and $P_{\mathrm{d}}^{\mathrm{DBD}}$, i.e., the probabilities of correctly identify-

ing the DFT beams and DBDs to carry out ABP [5] and the proposed method, respectively. We see that $P_{\mathrm{d}}^{\mathrm{DBD}}$ is increasingly larger than $P_{\mathrm{d}}^{\mathrm{DFT}}$. Particularly, in the low SNR regions, e.g., $\gamma \leq 10$ dB, $P_{\mathrm{d}}^{\mathrm{DBD}}$ grows much faster than $P_{\mathrm{d}}^{\mathrm{DFT}}$. $P_{\mathrm{d}}^{\mathrm{DBD}}$ is 0.8751, 0.8467 and 0.7214 at $\gamma = 10$ dB for $r = 10^{-8}$, 10^{-3}, and 10^{-2}, respectively. In contrast, $P_{\mathrm{d}}^{\mathrm{DFT}}$ is only 0.5417, 0.4744, or 0.3454. We also see that $P_{\mathrm{d}}^{\mathrm{DBD}}$ converges after $\gamma \geq 20$ dB and reaches 0.995, 0.9816 and 0.9435 at $\gamma = 40$ dB for the three values of r. In contrast, $P_{\mathrm{d}}^{\mathrm{DFT}}$ only reaches 0.9424, 0.882 and 0.8038 at $\gamma = 40$ dB. The reason for the faster convergence of the proposed method to 1 than WDFT is because the proposed approach is able to achieve higher receive SNR by conducting the beam synthesis in the analog domain (before entering RF chains) under low noises. In contrast, WDFT synthesizes the beams in the digital domain, and hence accumulates the noises from all RF chains, as proved in Appendix 7.3.5.

Fig. 4.7 compares the MSE of the AoA estimation between the proposed algorithm, Algorithm 5, and the state-of-the-art ABP [5], as γ increases. We see that Algorithm 5 is able to dramatically and increasingly outperform ABP, and reduces the MSE by orders of magnitude, especially in low SNR regions. In the case of $\gamma = 10$ dB, the MSEs of ABP can be as high as 0.5545 and 0.7252 for $r = 10^{-3}$ and 10^{-2}, respectively. The corresponding MSEs of Algorithm 5 are only 0.01588 and 0.05608. This is because the MSE in the low SNR regions, e.g., $\gamma \leq 16$ dB, is dominated by $P_{\mathrm{d}}^{\mathrm{DBD}}$ (or $P_{\mathrm{d}}^{\mathrm{DFT}}$). According to Fig. 4.6, $P_{\mathrm{d}}^{\mathrm{DBD}}$ is much higher than $P_{\mathrm{d}}^{\mathrm{DFT}}$. The simulations also validate the analysis in Corollary 1. Specifically, the analytical and numerical results of the MSE asymptotically converge and become indistinguishably close for $\gamma \geq 13$ dB, when $P_{\mathrm{d}}^{\mathrm{DBD}} > 0.9$. In other words, the MSE is dominated by the estimation SNR for $P_{\mathrm{d}}^{\mathrm{DBD}} > 0.9$, and the MSE decreases linearly with the growth of γ.

The key reason underlying the superiority of the proposed approach to ABP in terms of AoA estimation accuracy is that the DBDs can produce higher receive SNR than the DFT beams. As proved in Appendix 7.3.5, a DBD has a much higher amplitude response than a DFT beam, while the noise powers of the two schemes are roughly the same. As a result of the improved gain of the DBDs, the two

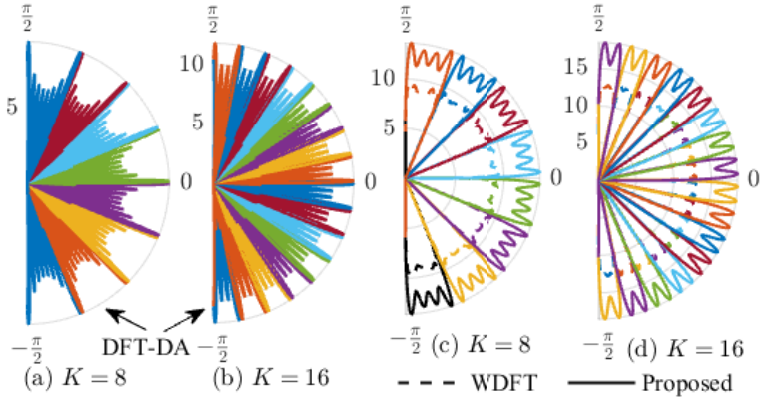

Figure 4.8 : Mainlobe comparison of the synthesized beams using WDFT [2], DFT-DA [6] and the proposed method, where $N = 64$, $K = 8$ and 16.

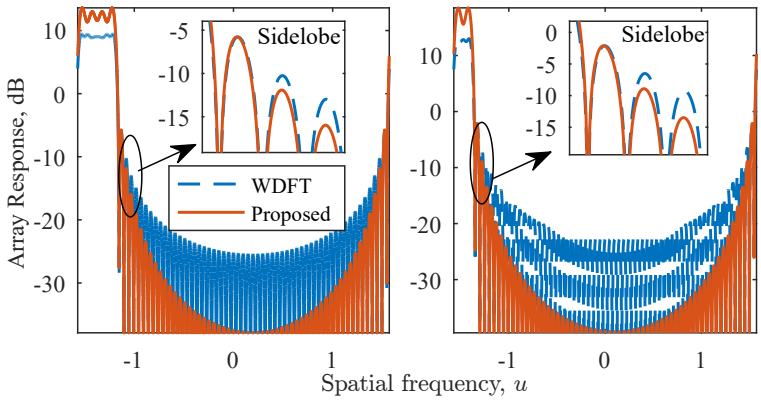

Figure 4.9 : A Detailed comparisons of WDFT [2] and the proposed method by zooming in the beams in Figs. 4.8(c) and 4.8(d), respectively.

strongest consecutive DBDs are much more likely to be correctly identified in the proposed approach than the two strongest consecutive DFT beams in ABP. This is particularly important in low SNR regions. In the case that both the DBDs and DFTs can be correctly identified in the proposed approach and ABP, e.g., in high SNR regions, the improved gain of the DBDs leads to lower estimation errors in the proposed approach.

Fig. 4.8 compares the performance of beam synthesis between DFT-DA [6], WDFT [2], and the proposed method developed in Theorem 2, where, for fair comparison, the parameter configurations in [2] are considered, i.e., $K = 8$ and 16. We see that the synthesized beams of DFT-DA show severe ripples and nulls in the

mainlobes; while the proposed method can eliminate nulls and reduces ripples in the mainlobe, as discussed in Appendix 7.3.6. We also see that, as expected, the ripples in the mainlobes of the proposed method can be slightly more severe than those of WDFT, since WDFT is designed to minimize the variances of the mainlobe directivity. Nevertheless, the mainlobe-to-sidelobe ratio of the proposed method is higher than that of WDFT; see Appendix 7.3.7. By aligning the first sidelobes of the synthesized beams, we see that the beamforming gain of the proposed method is much larger than that of WDFT.

Fig. 4.9 zooms in and shows the detailed differences of the synthesized beams of Figs. 4.8(c) and 4.8(d). We see that the mainlobe beamforming gains of the approach proposed Theorem 4.2 are higher than those of WDFT, and the sidelobe levels of the proposed method are much lower than those of WDFT. For $M = 8$, the maximum mainlobe beamforming gain of WDFT is 2.543 dB lower than the minimum mainlobe beamforming gain of the proposed method, and the levels of the second and third sidelobes of WDFT are 1.653 dB and 3.033 dB higher than those of the method proposed in Theorem 4.2, respectively. Moreover, the sidelobes of the proposed method decrease much faster than WDFT, and achieve gains of up to 16 dB and 22 dB for $M = 8$ and 16, respectively. It is noteworthy that the proposed method linearly combines the DFT beams with the coefficients having only three values, i.e., ± 1 and 0. In contrast, WDFT optimize the phase rotational speed through exhaustive search [2]. To this end, the proposed method incurs far less complexity.

Fig. 4.10 compares the probabilities of correctly identifying the angular regions (or beams) of interest, denoted by $P\{\mathcal{E}\}$, where the beams are synthesized by WDFT [2] and the proposed method. The beams in Figs. 4.4(a) and 4.4(b) are employed for receiving at symbols $t = 0$ and $t = 1$, respectively. We see that $P\{\mathcal{E}\}$ of the proposed method can converge much faster to 1 than those of WDFT, as γ increases. Given $P\{\mathcal{E}\}$, the proposed method can achieve an average SNR improvement of 10 dB over WDFT. For example, in the case of $P\{\mathcal{E}\} = 0.8$, the proposed method has the SNR improvements of 12 dB and 10 dB at $t = 0$ and 1, respectively. This is

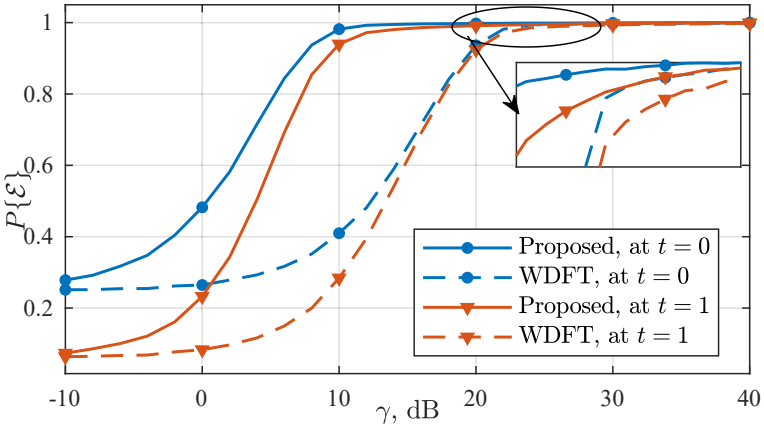

Figure 4.10 : $P\{\mathcal{E}\}$ of WDFT [2] and the proposed beam synthesis vs. $\gamma = \frac{1}{\sigma_{\mathrm{d}}^2}$, where $\mathcal{E}$ refers to the case that the beams of interest are identified correctly.

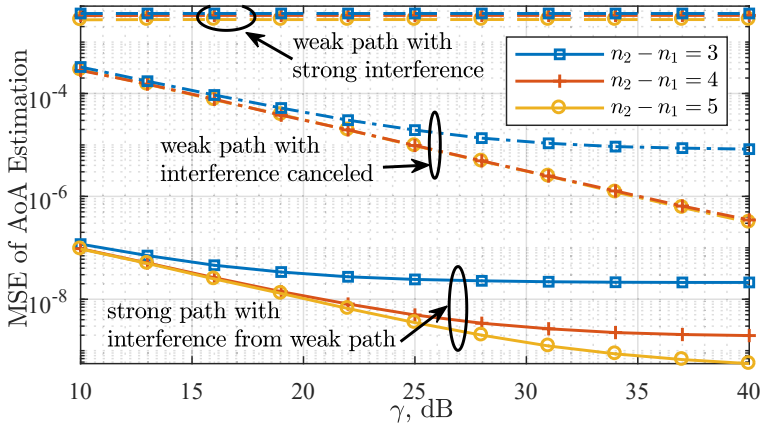

Figure 4.11 : MSEs of the AoA estimates of two paths, where $\gamma = \frac{1}{\sigma_{\mathrm{d}}^2}$, $r = 10^{-3}$, $u_i \sim \mathcal{U}\left[\frac{2\pi n_i}{N}, \frac{2\pi(n_i+1)}{N}\right]$ $(i = 1, 2)$.

because WDFT combines DFT beams in the digital domain, where the noises of the RF chains are also digitized and injected. In contrast, the proposed method combines the DFT beams in the analog domain, and only experiences the RF chain noise, thus providing higher SNRs.

Fig. 4.11 plots the MSEs of AoA estimations for two separate paths, one for WPT and the other for WIT, where the signal power of the WPT and WIT paths are 35 dB and 0 dB, respectively, and $r = 10^{-3}$. The AoA of the WPT path is set as $u_1 \sim \mathcal{U}\left[\frac{2\pi n_1}{N}, \frac{2\pi(n_1+1)}{N}\right]$; and the AoA of the WIT path $u_2 \sim \mathcal{U}\left[\frac{2\pi n_2}{N}, \frac{2\pi(n_2+1)}{N}\right]$. Algorithm 2 is run. Different values of $(n_2 - n_1)$ are considered. We see that the

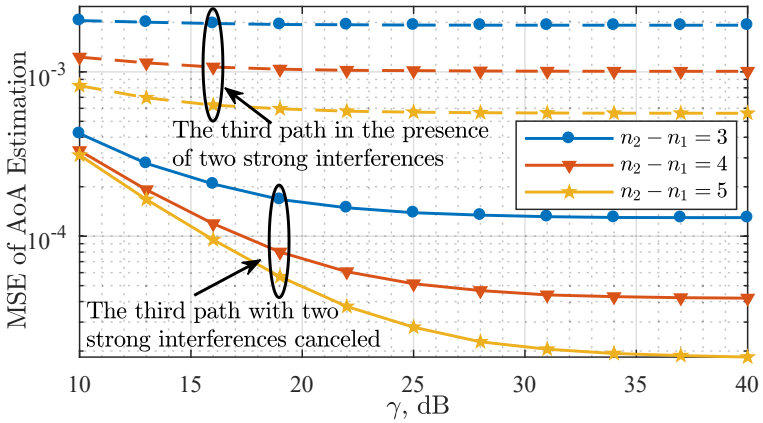

Figure 4.12 : MSE of the AoA estimate (vs. $\gamma = \frac{1}{\sigma_{\rm d}^2}$) of the third path, u_3, in the presence of two strong and close paths, u_1 and u_2, where $r = 10^{-3}$, the powers of the path are 35 dB, 20 dB and 0 dB, respectively; and the AoAs of the three paths satisfy $u_1 \sim \mathcal{U}\left[\frac{2\pi(n_1-1)}{N}, \frac{2\pi n_1}{N}\right]$, $u_2 \sim \mathcal{U}\left[\frac{2\pi n_1}{N}, \frac{2\pi(n_1+1)}{N}\right]$ and $u_3 \sim \mathcal{U}\left[\frac{2\pi n_2}{N}, \frac{2\pi(n_2+1)}{N}\right]$, respectively.

MSEs of the WIT path, for $n_2 - n_1 = 3$, 4 and 5, saturate for any γ, as the result of the interference from the strong path. From Fig. 4.5, we can conclude that the average SIRs between two paths are 33.98dB, 38.57dB and 42.13dB for $n_2 - n_1 = 3$, 4 and 5, respectively. With the setting of 35dB SIR in Fig. 4.11, the strong path can still suffer the interference of 1.02dB, -3.57dB and -7.13dB to the weak path, respectively. With $\gamma = \frac{1}{\sigma_{\rm d}^2} \geq 10$ dB, the noise power is below -10 dB. As a result, the MSE of the weak path is dominated by the interference from the strong path and saturates within the range of γ in the figure. We can still see that the MSE of the weak path can be substantially improved after the strong path is canceled. Moreover, for $n_2 - n_1 = 3$, the accuracy of both the weak and strong paths can be slightly worse than the case of $n_2 - n_1 > 3$. This is consistent with the result of Fig. 4.5.

We proceed to consider a case of three paths, where the two stronger paths are close to each other, i.e., in the mainlobes of two adjacent DBDs. The AoAs of the three paths satisfy $u_1 \sim \mathcal{U}\left[\frac{2\pi(n_1-1)}{N}, \frac{2\pi n_1}{N}\right]$, $u_2 \sim \mathcal{U}\left[\frac{2\pi n_1}{N}, \frac{2\pi(n_1+1)}{N}\right]$ and $u_3 \sim \mathcal{U}\left[\frac{2\pi n_2}{N}, \frac{2\pi(n_2+1)}{N}\right]$. The powers of the three paths are 35dB, 20dB and 0dB, respectively. Fig. 4.12 compares the MSEs of the third path before and after the

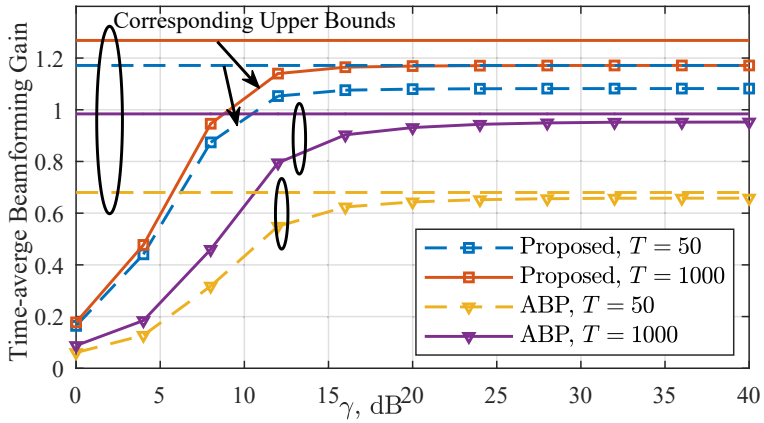

Figure 4.13 : Time-average beamforming gain of an LAA and a DAA receiver based on the AoA estimates obtain by ABP and the proposed algorithm, respectively, where $\gamma = \frac{1}{\sigma_{\mathrm{d}}^2}$, $r = 10^{-2}$, $u \sim \mathcal{U}[0, 2\pi)$.

interferences from the first two paths are canceled. We see that before the interference is canceled, the MSE of the third path quickly saturates with γ. As explained in Fig. 4.11, the saturation is because the interferences have the dominating effect over the noises. We also see that, after SIC, the AoA estimation accuracy of the third path is improved dramatically, especially for $n_2 - n_1 > 3$. This demonstrates the effectiveness of the proposed approach in estimating and canceling the paths.

The accurate AoA estimation of signal paths serves the purpose of forming efficient beams to capture the energy and/or signals carried on the paths. To this end, the time-average beamforming gain of a transmission block, defined as $\frac{(T-\tau)G}{T}$, provides an effective measure of the efficiency of the AoA estimation [78]. T is the duration of a transmission block; τ is the time for AoA estimation, and hence $(T-\tau)$ is the time for power/information transmission; G is the instantaneous beamforming gain based on the estimated AoA. The time-average beamforming gain is used for fair comparisons between the proposed algorithm and ABP [5].

Fig. 4.13 plots the time-average beamforming gains of Algorithm 5 and ABP, where the AoA of a single path is set to $u \sim \mathcal{U}[0, 2\pi)$. We see that the time-average beamforming gains of the proposed algorithm are increasingly and dramatically larger than those of ABP. In the case of $T = 50$, the proposed algorithm, Algorithm 5, converges to 1.082, 64.51 % higher than the maximum gain achieved by ABP. One

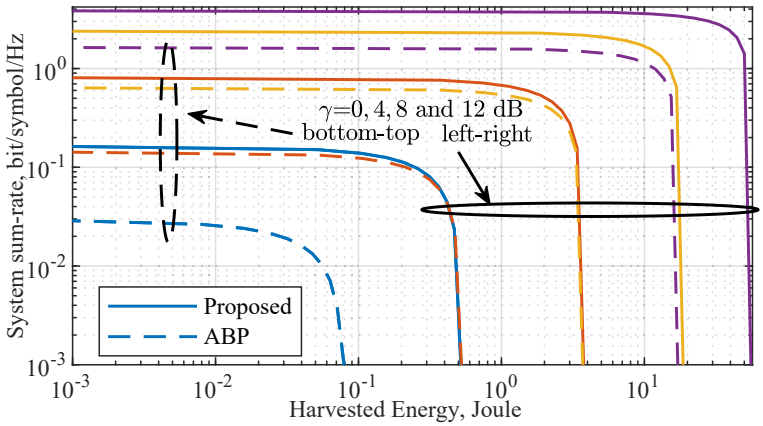

Figure 4.14 : The rate-energy region of the proposed approach and ABP [5] under SWIPT, where $T = 30$.

reason underlying the improvement of Algorithm 5 over ABP is that Algorithm 5 enables fast search of the angular regions of interest based on the new beam synthesis in Theorem 4.2. In this simulation, $\frac{T-\tau}{T} = 0.9$ and 0.68 for Algorithm 5 and ABP, respectively. Another reason for the improvement is that Algorithm 5 can achieve significantly higher accuracy of AoA estimation than ABP; see Fig. 4.7.

In Fig. 4.13, we also plot the upper bounds for the time-average beamforming gains, where the upper bounds are obtained by taking the maximum amplitude gains of a DBD or a DFT beam, i.e., $\frac{4}{\pi}$ and 1, respectively; see Appendix 7.3.6. We can see that the time-average beamforming gain of the proposed Algorithm 5 is close to the corresponding upper bounds. The gap of LAA between the achieved beamforming gain and its upper bound is because the LAA can only produce fixed DFT beams. In contrast, the DAA is able to point at any directions with its continuous adjustable phase shifters. Nevertheless, the proposed algorithm enables the LAA to achieve larger time-average beamforming gains than the upper bounds which ABP allows the DAA to achieve. The significant gain of LAA, resulting from the accurate AoA estimation of the proposed algorithm, is of practical value to WIPT.

Fig. 4.14 plots the rate-energy region of the proposed approach in comparison to that of the state-of-the-art ABP [5] by taking SWIPT as an example. T is partitioned between the AoA estimation (i.e., $(\log_K^N + 1)$ symbols for the proposed approach and

$\frac{N}{K}$ symbols for ABP), and SWIPT by focusing on the AoA (i.e., $(T - \log_K^N - 1)$ symbols for the proposed approach and $(T - \frac{N}{K})$ symbols for ABP). We set $T = 30$ and the RF-to-DC conversion efficiency to 0.1 [31]; and these parameters can be readily changed to other values. The power splitting factor (PSF) of SWIPT is denoted by η; in other words, the ratio of the received power used for information detection and for energy harvesting is $\frac{\eta}{1-\eta}$ [29, eqs. 2&3]. By varying the value of η from 0 to 1 with a step of 0.1, the figure is plotted. We can see that the proposed approach is able to substantially enlarge the rate-energy region and the trade-off between energy harvesting and information detection is improved. This is due to the reduced number of training symbols for the AoA estimation and hence the increased number of symbols for SWIPT. It is also due to the increased accuracy of the AoA estimation and hence the improved beamforming gain for energy transfer and data transmission.

4.7 Conclusion

In this chapter, we propose a fast and accurate approach for the AoA estimation of LAAs. We prove that the AoA of a path can be accurately estimated from the two adjacent DBDs at both sides of the AoA. The two DBDs are proved to have larger amplitude gains than other DBDs and DFT beams, and can be more reliably identified in the presence of non-negligible noises. We design a novel beam synthesis method based on discrete lens beams and 1-bit phase shifts. As a result, the angular region containing the two strongest DBDs can be exponentially narrowed down, and the two DBDs can be quickly identified. Evident from extensive simulations, the proposed approach is able to improve estimation accuracy and the time-average beamforming gain by orders of magnitude, as compared to the state of the art.

Chapter 5

Exploiting Spatial-Wideband Effect for Fast AoA Estimation at Lens Antenna Array

Energy-efficient, highly integrated LAAs have found widespread applications in wideband millimeter wave or terahertz communications, localization and tracking, and wireless power transfer. Accurate estimation of AoA is key to those applications, but has been hindered by a spatial-wideband effect in wideband systems. This chapter proposes to exploit (rather than circumventing) the spatial-wideband effect to develop a fast and accurate AoA estimation approach for LAAs. Specifically, we unveil new spatial-frequency patterns based on the spatial-wideband effect, and establish one-to-one mappings between the patterns and the strongest DFT beam containing the AoA. With the strongest DFT beam identified, we propose to estimate the AoA uniquely and accurately with only a few training symbols, by deriving a new one-to-one mapping between the AoA and the set of DFT beams judiciously selected based on the strongest. In the case that an impinging path is uniformly distributed in $[0, 2\pi]$, simulations show that the proposed algorithm is able to reduce the mean squared error of the AoA estimation by as much as 82.1% while reducing the number of required symbols by 93.2%, as compared to existing techniques. The algorithm can also increase the spectral efficiency by 89% when the average SNR is -20 dB at each antenna of the receiver.

5.1 Introduction

Distinctively different from any existing work, the proposed algorithm utilizes (rather than circumventing) the spatial-wideband effect to achieve a fast and accurate AoA estimation for mmWave wideband LAAs. None of the existing techniques, such as [7], can make use of the spatial-wideband effect for AoA estimation. They

have either focused on narrowband signals [5-7], [18], or overlooked the spatial-wideband effect [19-21]. The proposed algorithm is the first of its kind with promising performances. The main contributions of the chapter are summarized as follows.

1. We unveil new spatial-frequency patterns by exploiting the spatial-wideband effect over a judiciously selected set of frequency sub-carriers, and establish one-to-one mappings between the strongest DFT beam and the patterns. As a result, the strongest DFT beam containing the AoA can be identified within as few as a single symbol at a high probability. To the best of our knowledge, this has never been achieved in the existing literature.

2. With the strongest DFT beam identified and selected, we propose to estimate the AoA uniquely and accurately in another symbol by deriving a one-to-one mapping between the AoA and a set of DFT beams judiciously selected based on the strongest. We prove that, in the presence of a dominant LoS path, the strongest DFT beam contributes predominantly to the AoA estimation accuracy. The other beams contribute asymptotically negligibly, as the number of antennas increases;

3. We analyze the performance of the proposed methods. The average number of symbols for identifying the strongest DFT beam is derived. The number is proved to rapidly decrease, as the system bandwidth or K increases. K denotes the number of RF chains. We also derive the closed-form expression for the MSELB of the proposed AoA estimation.

Corroborated by extensive simulations, the analytical results confirm that, by using an average of only 1.19 symbols, the proposed approach is able to achieve as much as 82.1% accuracy improvement, as compared with the existing work [7] using 32 symbols. Due to the accurate AoA estimation, the proposed approach is able to improve the spectral efficiency by up to 89.04% at the SNR of -20 dB, compared to the state of the art [8].

The rest of the chapter is organized as follows. In Section 5.2, the system model is described. In Section 5.3, the new spatial-frequency pattern is unveiled, and

the fast search of the strongest DFT beam is designed. In Section 5.4, the AoA estimation based on the strongest DFT beam is proposed, followed by the analyses of the proposed algorithms. The superiority of the proposed approaches in terms of accuracy and efficiency is demonstrated via extensive simulations in Section 5.6, followed by conclusions in Section 5.7.

5.2 System Architecture

We consider an uplink mmWave massive SIMO-OFDM system with a single-antenna user*, where the BS is equipped with a large-scale LAA. This section describes the structure of the LAA receiver, followed by the signal model.

5.2.1 LAA Structure

Fig. 5.1 illustrates an N-dimensional LAA receiver, where the LAA can generate N DFT beams (this is equivalent to the $N \times N$ DFT beamforming) [41, 43]. The N antennas are placed on the focal surface of the lens. Each antenna corresponds to a DFT beam and outputs the received signals from the beam. Let K denote the number of RF chains, which is much less than N, i.e., $K \ll N$. A beam selection, also known as antenna selection [41, 74], is required to connect the smaller number of RF chains (than the number of antennas). We employ the beam selection network developed in [8, 43]. The output signal of a DFT beam is evenly split between K number of 1-bit phase shifters (c.f., "on/off" switches [74, 75]). The phase shifters can be turned off, as can be readily implemented in practice [43]. As a result, each phase shifter can take three different states, "$+1(= e^{j\frac{2\pi\times 0}{2}})$", "$-1(= e^{j\frac{2\pi\times 1}{2}})$", or "0 (turned off)". The outputs of N phase shifters connecting the N antennas can be combined and fed to an RF chain.

Let $\mathbf{W}$ denote the beam selection network, where $[\mathbf{W}]_{:,k} \in \mathbb{R}^{N\times 1}$ collects the states of the N phase shifters connected to the k-th RF chain ($k = 0, 1, \cdots, K-1$); see Fig. 5.1. The (n,k)-th element of $\mathbf{W}$ is $[\mathbf{W}]_{nk} = \frac{1}{\sqrt{\|[\mathbf{W}]_{n,:}\|_0}}$, 0 or $-\frac{1}{\sqrt{\|[\mathbf{W}]_{n,:}\|_0}}$.

*The proposed approach can be readily extended to multi-user scenario by applying orthogonal pilot signals in the same way as the existing wideband channel estimation algorithms [8, 66].

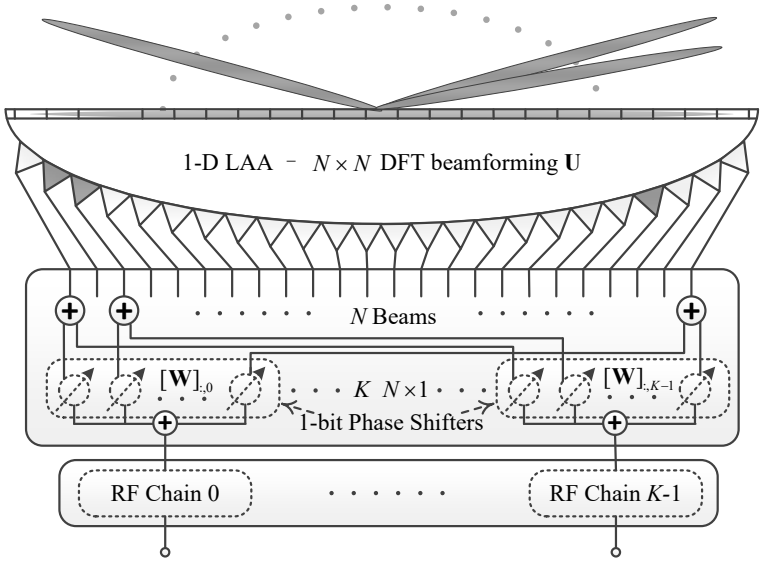

Figure 5.1 : The schematic diagram of an LAA receiver, where "1-bit phase shifter" refers to a phase shifter with the phase shift of either 0 or 180 degree.

The received signal power of antenna n is divided evenly into $\|\mathbf{W}_{n,:}\|_0$ parts, if $\|\mathbf{W}_{n,:}\|_0$ out of K phase shifters connected to the n-th antenna are turned on simultaneously. As widely adopted in the literature [8, 66], we assume that the lens and beam selection network are frequency-flat. In other words, they are invariant in the frequency band of interest.

We note that the beam selection network, comprised of 1-bit phase shifters, was originally proposed in [6, 8]. But the idea of this chapter using the beam selection network to generate spatial-frequency wide beams for AoA estimation (as will be provided in Lemma 1) is novel and distinctively different from [6, 8]. In particular, the beam selection network is now configured deterministically based on the one-to-one mapping we unveil between an inherent spatial-frequency pattern and the strongest DFT beam (as will be provided in Theorem 2). This facilitates estimating the AoA quickly and accurately by utilizing a spatial-wideband effect (or in other words, beam squinting). In contrast, all the 1-bit phase shifters were randomly set to ± 1 in [6, 8], and the channels (not the AoA) were estimated using compressive sensing.

5.2.2 Signal Model

As assumed in [8, 84], the user sends the i-th pilot signal at symbol i, denoted by s_i, and the same signal is conveyed by the M sub-carriers for uplink channel estimation. Let s_{im} denote the signal transmitted by sub-carrier m at symbols i. s_{im} $(\forall m, i)$ has the unit power, i.e., $|s_{im}|^2 = 1$ [8]. At the user, an IDFT is taken over $s_{i0}, s_{i1}, \cdots, s_{i(M-1)}$, the cyclic prefix (CP) is appended to the IDFT results, and an OFDM symbol is transmitted. The m'-th sample of the OFDM symbol is denoted by $x_{m'}$, where $0 \leq m' \leq M' - 1$ and $(M' - M)$ is the number of CP samples.

The m'-th sample is received by the LAA antennas at the BS. The antenna outputs can be collected by the vector $\mathbf{y}_{m'} = \mathbf{U}^{\mathrm{H}}\mathbf{h}_{m'}x_{m'}$, where $\mathbf{U} \in \mathbb{C}^{N\times N}$ is the $N \times N$ DFT matrix and $\mathbf{h}_{m'} \in \mathbb{C}^{N\times 1}$ is the time-domain channel response vector. The (a, b)-th element of $\mathbf{U}$ is given by

$$[\mathbf{U}]_{ab} = \frac{1}{N}e^{-j\frac{2\pi ab}{N}},\ \forall a, b \in [0, N-1]. \tag{5.1}$$

After receiving for the M' samples, the received OFDM symbol can be written as $\hat{\mathbf{Y}} = [\mathbf{y}_0, \mathbf{y}_1, \cdots, \mathbf{y}_{M'-1}] \in \mathbb{C}^{N\times M'}$, where $\mathbf{y}_{m'} \in \mathbb{C}^{N\times 1}$ collects the OFDM samples received by the N-element array at the sample time $m'(= 0, 1, \cdots, M' - 1)$. The first $(M' - M)$ columns of $\hat{\mathbf{Y}}$ are the CP. After the beam selection and RF chains, $\bar{\mathbf{Y}}$ is digitized and $\tilde{\mathbf{Y}} = \mathbf{W}^{\mathrm{T}}\bar{\mathbf{Y}} + \mathbf{N}$, where $\mathbf{W} \in \mathbb{R}^{N\times K}$ is the beam selection matrix and $\mathbf{N} \in \mathbb{C}^{K\times M'}$ is the AWGN from the RF chains.

After the removal of the CP which consists of the first $(M' - M)$ columns of $\tilde{\mathbf{Y}}$, we obtain

$$\mathbf{Y} = \mathbf{W}^{\mathrm{T}}\bar{\mathbf{Y}} + \bar{\mathbf{N}}, \tag{5.2}$$

where $\bar{\mathbf{Y}}$ and $\bar{\mathbf{N}}$ are obtained by removing the first $(M' - M)$ columns of $\hat{\mathbf{Y}}$ and $\hat{\mathbf{N}}$, respectively. Taking the row-wise DFT of $\mathbf{Y}$, i.e., the DFT over each row of M samples, we obtain (5.3), where (5.3a) and (5.3c) are based on the definition of matrix product, and (5.3b) is obtained by applying the linearity of DFT (i.e., $ax_1[n] + bx_2[n] \overset{\mathrm{DFT}}{\longleftrightarrow} aX_1[k] + bX_2[k]$. $x_1[n]$ and $x_2[n]$ are finite sequences. $X_1[k]$ and

$$
\begin{aligned}
\mathbf{X} &= \mathrm{DFT}\{\mathbf{W}^{\mathrm{T}}\bar{\mathbf{Y}} + \mathbf{N}\} = \mathrm{DFT}\{\mathbf{W}^{\mathrm{T}}\bar{\mathbf{Y}}\} + \mathrm{DFT}\{\mathbf{N}\} \\
&= \begin{bmatrix} \mathrm{DFT}\left\{\sum_{n=0}^{N-1}[\mathbf{W}]_{n0}[\bar{\mathbf{Y}}]_{n,:}\right\} \\ \vdots \\ \mathrm{DFT}\left\{\sum_{n=0}^{N-1}[\mathbf{W}]_{n(K-1)}[\bar{\mathbf{Y}}]_{n,:}\right\} \end{bmatrix} + \mathrm{DFT}\{\mathbf{N}\} && (5.3a) \\
&= \begin{bmatrix} \sum_{n=0}^{N-1}[\mathbf{W}]_{n0}\mathrm{DFT}\left\{[\bar{\mathbf{Y}}]_{n,:}\right\} \\ \vdots \\ \sum_{n=0}^{N-1}[\mathbf{W}]_{n(K-1)}\mathrm{DFT}\left\{[\bar{\mathbf{Y}}]_{n,:}\right\} \end{bmatrix} + \mathrm{DFT}\{\mathbf{N}\} && (5.3b) \\
&= \mathbf{W}^{\mathrm{T}}\mathrm{DFT}\{\bar{\mathbf{Y}}\} + \mathrm{DFT}\{\mathbf{N}\} && (5.3c) \\
&= \mathbf{W}^{\mathrm{T}}\mathbf{U}^{\mathrm{H}}[\mathbf{h}_0, \mathbf{h}_1, \cdots, \mathbf{h}_{M-1}]s_i + \tilde{\mathbf{N}},
\end{aligned}
$$

$X_2[k]$ are the DFTs of $x_1[n]$ and $x_2[n]$, respectively.) [24, Eq.(8.76)].

We note that there are two DFTs in the signal model. The first DFT is performed inherently by the lens to produce the DFT beam. The outcome is the product of the conjugate transpose of the DFT matrix $\mathbf{U}$ and the received signal vector; see (1). The second DFT in (3) accounts for an important step of an OFDM receiver to transform the received signal from the time domain to the frequency domain (i.e., sub-carriers). Both are the standard DFT operations, and they are salient parts of lens and OFDM signal processing.

We adopt the dual-wideband channel model proposed in [66]. $\mathbf{h}_m$ is given by

$$
\mathbf{h}_m = \sqrt{\frac{1}{(P+1)}} \sum_{p=0}^{P} \beta_p e^{-j2\pi f_m \tau_p} \mathbf{a}(\mu_p, \rho_m), \tag{5.4}
$$

where there are $(P+1)$ resolvable paths, β_p is the channel response of the p-th path, τ_p is the path delay, and μ_p is the beamspace-domain AoA of the path; $f_m = f_{\mathrm{L}} + \frac{mB}{M-1}$ is the centroid frequency of the m-th sub-carrier ($0 \leq m \leq M-1$), B is the bandwidth, and f_{L} is the lowest system frequency; and $\rho_m = \frac{f_m}{f_{\mathrm{H}}}$ is the normalized frequency, and $f_{\mathrm{H}} = f_{\mathrm{L}} + B$ is the highest system frequency. Clearly, $\frac{f_{\mathrm{L}}}{f_{\mathrm{H}}} \leq \rho_m \leq 1$. $\mathbf{a}(\mu_p, \rho_m)$ is the spatial-frequency-domain array response vector. The n-th element

of $\mathbf{a}(\mu_p, \rho_m)$ is

$$[\mathbf{a}(\mu_p, \rho_m)]_n = e^{-jn\rho_m \frac{2\pi d \sin\theta_p}{c/f_{\mathrm{H}}}} = e^{-jn\rho_m \mu_p}, \tag{5.5}$$

where θ_p is the physical AoA of the p-th incident path, and $\mu_p = \frac{2\pi d \sin\theta_p}{c/f_{\mathrm{H}}}$ is the beamspace-domain AoA defined based on f_{H}. $\mathbf{a}(\mu_p, \rho_m)$ is both spatial- and frequency-dependent, and undergoes the spatial-wideband effect [66, 82].

Note that the beamspace AoA μ_p is distributed in $[-\pi, \pi]$ based on the definition below (5). Our proposed AoA estimation will be achieved by estimating the angle of $e^{j\mu_p}$ which, for $\mu_p \in [-\pi, 0]$, is a one-to-one mapping to $e^{j\mu_p}$, $\mu_p \in [\pi, 2\pi]$, due to the periodicity of $e^{j\mu_p}$. By using variable substitution, we can define

$$u_p = \begin{cases} \mu_p, & \text{if } \mu_p \in [0, \pi]; \\ \mu_p + 2\pi, & \text{if } \mu_p \in [-\pi, 0], \end{cases} \tag{5.6}$$

so that u_p can be consistent with the angular coverage of the complete set of DFT beams in terms of range, i.e., $[0, 2\pi]$ [24]. μ_p can be recovered unambiguously, once u_p is estimated by using the proposed approach.

We note from (5.3) that it is non-trivial to estimate u_p from $\mathbf{X}$, because the key information to estimate the AoA, i.e., the phase difference between adjacent antenna elements, is obscure due to the DFT beamforming $\mathbf{U}$. On the other hand, given the much smaller number of RF chains than the number of antenna ports ($K \ll N$), only partial or combined DFT beam outputs can be obtained through the beam selection $\mathbf{W}$. It is important to establish the relation between the AoA and the (selected) DFT beams to achieve the accurate AoA estimate and expedite the AoA estimation. $\mathbf{W}$ is the only DoF, and needs to be designed judiciously to fast identify the required DFT beams.

5.2.3 Proposed Wideband AoA Estimation Framework

In this chapter, we design the method to unambiguously extract the AoA from the selected DFT beams and $\mathbf{W}$ to fast identify the required DFT beams. Overall,

the proposed AoA estimation has two stages:

1. Identify the strongest DFT beam by exploiting the spatial-wideband effect in Section 5.3. Interestingly, the spatial-wideband effect, generally known to degrade the performance of mmWave massive MIMO systems [8, 15, 66], is shown to be exploitable to the AoA estimation problem;

2. Extract the AoA based on the identified DFT beam using as few as a single symbol in Section 5.4. We prove that, in the LoS-dominant mmWave/THz channel, the strongest DFT beam contributes the most to the AoA estimation.

The strongest DFT beam refers to the DFT beam with the strongest output, which indicates that the AoA is in the angular region of $(\frac{2\pi n^\star}{N} - \frac{\pi}{N}, \frac{2\pi n^\star}{N} + \frac{\pi}{N}]$, where $n^\star$ denotes the index of the strongest DFT beam.

5.3 Fast Search of the Strongest DFT Beam

In this section, the spatial-wideband effect of the spatial response is exploited to achieve the fast search of the strongest DFT beam (as opposed to enumerating all the N DFT beams by taking at least $\frac{N}{K}$ symbols [7]), so that we can suppress the subscript "$_p$" from u_p for illustration convenience. To achieve this, we first unveil a new spatial-frequency pattern underlying the spatial-wideband effect. Then, we reveal that, by constructing the spatial-frequency patterns judiciously based on the selected sub-carriers, the patterns can be exploited to identify the strongest DFT beam uniquely and rapidly.

5.3.1 New Spatial-Frequency Pattern

We reveal two spatial-frequency patterns based on the selection of a single or multiple DFT beams. As will be proved later, the patterns are the one-to-one mappings to the strongest DFT beam at the LAA. The new spatial-frequency patterns can be defined based on the spatial responses of the DFT beams. Based on (5.5),

the spatial response of the n-th DFT beam at sub-carrier m is given by

$$[\mathbf{g}(u,\rho_m)]_n = \mathbf{U}^{\mathrm{H}}\mathbf{a}(u,\rho_m) = \sum_{n'=0}^{N-1}\frac{1}{N}e^{-jn'(\rho_m u-\frac{2\pi n}{N})}$$
$$= e^{-j\frac{N-1}{2}(\rho_m u-\frac{2\pi n}{N})}\frac{\sin\frac{N}{2}(\rho_m u-\frac{2\pi n}{N})}{N\sin\frac{1}{2}(\rho_m u-\frac{2\pi n}{N})}, \tag{5.7}$$

where $\mathbf{g}(u,\rho_m)\in\mathbb{C}^{N\times 1}$ collects the spatial responses of the N DFT beams at sub-carrier m and the AoA u. We see that the spatial response is coupled multiplicatively with the sub-carrier frequency, which is known as the spatial-wideband effect [66, 82]. If $m=M-1$, then $\rho_{M-1}=1$ and $[\mathbf{g}(u,\rho_{M-1})]_n$ is a standard discrete sinc function of u. As m and ρ_m decrease, $[\mathbf{g}(u,\rho_m)]_n$ moves rightwards w.r.t. u, and its pointing direction increases from $\frac{2\pi n}{N}$ to $\frac{2\pi n}{N\rho_m}$.

Single-Beam Selection

In the case that a single DFT beam is selected to be fed to an RF chain, the spatial-frequency pattern can be revealed at the output of the RF chain, as follows.

Definition 5.1. *The $\tilde{m}$-th spatial-frequency pattern ($0\le\tilde{m}\le\tilde{M}-1$) associated with the n-th DFT beam is defined to be the spatial responses of the n-th DFT beam at $u=\frac{2\pi(n+\tilde{m})}{N}$ and the selected set of frequency sub-carriers $\mathcal{M}$. $\tilde{M}=\lfloor\frac{n}{\rho_0}\rfloor-n$ is the maximum number of patterns when DFT beam n is selected. Let $\tilde{\mathbf{g}}_n(\tilde{m},\mathcal{M})\in\mathbb{C}^{\tilde{M}\times 1}$ denote the $\tilde{m}$-th pattern. The $\tilde{m}'$-th element of $\tilde{\mathbf{g}}_n(\tilde{m},\mathcal{M})$ ($0\le\tilde{m}\le\tilde{M}-1,\ 0\le\tilde{m}'\le|\mathcal{M}|-1$) is given by*

$$[\tilde{\mathbf{g}}_n(\tilde{m},\mathcal{M})]_{\tilde{m}'} = \left[\mathbf{g}\left(\frac{2\pi(n+\tilde{m})}{N},\rho_{\mathcal{M}_{\tilde{m}'}}\right)\right]_n \tag{5.8}$$

where the RHS is obtained by substituting $u=\frac{2\pi(n+\tilde{m})}{N}$ and $m=[\mathcal{M}]_{\tilde{m}'}$ into (5.7).

Note in (5.8) that, by selecting different sub-carriers in $\mathcal{M}$, different spatial-frequency patterns can be obtained even the same DFT beam is selected. By properly constructing $\mathcal{M}$, the spatial-frequency pattern can be used for the fast search of the strongest DFT beam, as described in the following theorem.

Theorem 5.1. *Suppose that the n_{L}-th DFT beam is selected, n_{H} is the index of the last DFT beam in the angular region $\mathcal{S}(n_{\mathrm{L}}, n_{\mathrm{H}})$ covered by the n_{L}-th DFT beam*[†]*,*

$$u \in \mathcal{S}(n_{\mathrm{L}}, n_{\mathrm{H}}),\ n_{\mathrm{H}} = \lfloor \frac{n_{\mathrm{L}}}{\rho_0} \rfloor, \tag{5.9a}$$

$$\text{and } [\mathcal{M}]_{\tilde{m}'} = \left\lfloor \frac{(\frac{n_{\mathrm{L}}}{n_{\mathrm{L}}+\tilde{m}'} - \rho_0)(M-1)}{1-\rho_0} \right\rfloor, \tag{5.9b}$$

the index for the strongest DFT beam is a one-to-one mapping of the spatial-frequency pattern $\tilde{\mathbf{g}}_{n_{\mathrm{L}}}(\tilde{m}, \mathcal{M})$, $\tilde{m} \in \{0, 1, \cdots, \tilde{M}-1\}$, because the patterns satisfy

$$\left| \left[\tilde{\mathbf{g}}_{n_{\mathrm{L}}}(0, \mathcal{M}), \tilde{\mathbf{g}}_{n_{\mathrm{L}}}(1, \mathcal{M}), \cdots, \tilde{\mathbf{g}}_{n_{\mathrm{L}}}(\tilde{M}-1, \mathcal{M}) \right] \right| \approx \mathbf{I}_{\tilde{M}}, \tag{5.10}$$

where $\tilde{M} = |\mathcal{M}| = n_{\mathrm{H}} - n_{\mathrm{L}} + 1$ is the number of patterns.

Proof. See Appendix 7.4.1. □

Multi-Beam Selection

The spatial-frequency pattern can also be defined when multiple consecutive DFT beams are activated and synthesized to be a wider beam fed into an RF chain. The beam selection vector has to be carefully designed; otherwise, nulls can be formed in the mainlobe [51]. We propose a new wideband beam synthesis method in LAAs with flat mainlobe response across the frequencies and the angular regions of interest. The flat mainlobe response is the design goal of the newly proposed beam synthesis method in wideband lens antenna arrays in Lemma 1.

Lemma 5.1. *Given the N DFT beams, $\mathbf{g}(u, \rho_m)$, and the $N \times 1$ beam selection vector for the k-th RF chain*

$$[\mathbf{W}]_{:,k} = \left[\underbrace{0, \cdots, 0}_{l}, \underbrace{(-1)^l, \cdots, (-1)^{l+L-1}}_{L}, \underbrace{0, \cdots, 0}_{N-L-l} \right]^{\mathrm{T}}, \tag{5.11}$$

[†]Due to the spatial-wideband effect, the pointing direction of the n_{L}-th DFT beam becomes $\frac{2\pi n_{\mathrm{L}}}{N\rho_0}$; see (5.7). Note that $n_{\mathrm{H}} \leq \frac{n_{\mathrm{L}}}{\rho_0}$. Here, $\mathcal{S}(n_{\mathrm{L}}, n_{\mathrm{H}})$ denotes the angular region spanned between the n_{L}-th and the n_{H}-th DFT beams. $\mathcal{S}(n_{\mathrm{L}}, n_{\mathrm{H}}) := (\frac{(2n_{\mathrm{L}}-1)\pi}{N}, \frac{(2n_{\mathrm{H}}+1)\pi}{N}]$.

the beam synthesized by $\sum_{n=0}^{N-1}[\mathbf{g}(u,\rho_m)]_n[\mathbf{W}]_{n,k}$ *can be approximated by*

$$P(u,\rho_m)=\begin{cases} e^{j\frac{N}{2}\rho_m u}, & \text{if } \frac{2\pi l}{N}\le\rho_m u\le\frac{2\pi(l+L-1)}{N};\\ 0, & \text{otherwise,}\end{cases} \tag{5.12}$$

where l *is a non-negative integer satisfying* $0\le l\le N-L$.

Proof. By treating $\rho_m u$ as a single variable, $\mathbf{g}(u,\rho_m)$ in (5.7) becomes a standard sinc function of $\rho_m u$. The lemma can be proved based on [51, App. D]. □

Based on Lemma 5.1, we define the spatial-frequency pattern for a selected set of multiple consecutive DFT beams for an RF chain.

Definition 5.2. *Given the multiple DFT beams selected by* $[\mathbf{W}]_{:,k}$ *in (5.11), the spatial-frequency pattern is defined as such that the* $\tilde{m}'$*-th element of the* $\tilde{m}$*-th spatial-frequency pattern* $(0\le\tilde{m}\le\tilde{M}-1)$ *associated with the* n*-th DFT beam* $(l\le n\le l+L-1)$*, denoted by* $[\breve{\mathbf{g}}_n(\tilde{m},\mathcal{M})]_{\tilde{m}'}$*, is given by*

$$[\breve{\mathbf{g}}_n(\tilde{m},\mathcal{M})]_{\tilde{m}'}=P\left(\frac{2\pi(n+\tilde{m})}{N},\rho_{\mathcal{M}_{\tilde{m}'}}\right), \tag{5.13}$$

where the RHS is obtained by substituting $u=\frac{2\pi(n+\tilde{m})}{N}$ *and* $m=\mathcal{M}_{\tilde{m}'}$ *into (5.12).* $\tilde{M}=\lfloor\frac{n}{\rho_0}\rfloor-n$ *and* $\mathcal{M}$ *are the same as in Definition 5.1.*

The spatial-frequency pattern of multiple consecutive DFT beams (5.13) can also be used to identify the strongest DFT beam, as stated in the following theorem.

Theorem 5.2. *Consider that* $[\mathbf{W}]_{:,k}$ *in (5.11) is set for beam selection. By setting* $n_{\mathrm{L}}=l+L-1$ *(*$L\ge 2$ *and* $0\le l\le N-L$*), the index for the strongest DFT beam for* $u\in\mathcal{S}(n_{\mathrm{L}},n_{\mathrm{H}})$ *is a one-to-one mapping of the spatial-frequency pattern* $\breve{\mathbf{g}}_{n_{\mathrm{L}}}(\tilde{m},\mathcal{M})$*,* $\tilde{m}\in\{1,\cdots,\tilde{M}-1\}$*, if (5.9) holds.*[‡]

Proof. See Appendix 7.4.2. □

[‡] $\breve{\mathbf{g}}_{n_{\mathrm{L}}}(0,\mathcal{M})$ is precluded in the spatial-frequency patterns under multi-beam selection because of its non-uniqueness. In contrast, $\tilde{\mathbf{g}}_{n_{\mathrm{L}}}(0,\mathcal{M})$ is unique in the spatial-frequency patterns under single-beam selection, as stated in Theorem 5.1.

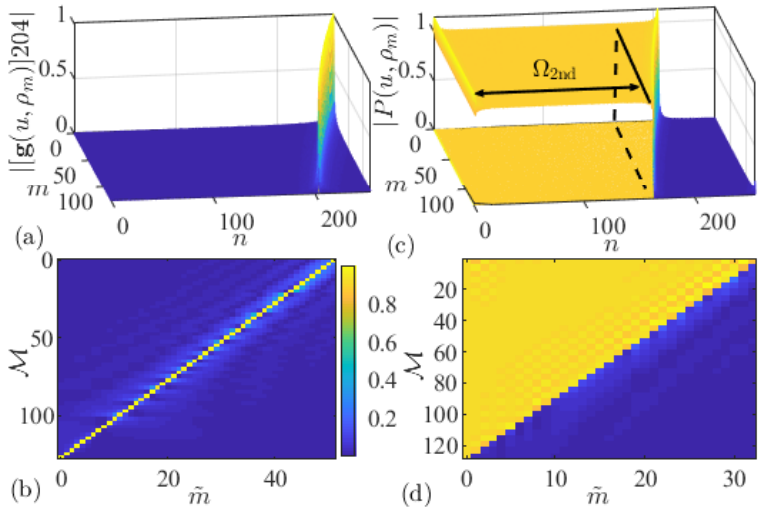

Figure 5.2 : Illustration of the spatial-frequency patterns, where $N = 256$, $M = 128$, and $\rho_m \in [0.8, 1]$. Fig. 5.2(a) plots $[\mathbf{g}(u, \rho_m)]_{204}$ in (5.7) across $u \in \mathcal{S}(0, 255)$ and sub-carrier $m \in [0, M-1]$; Fig. 5.2(c) plots $P(u, \rho_m)$ in (5.12) by taking $l = 0$ and $L = 164$ in Lemma 5.1; and Figs. 5.2(b) and 5.2(d) plot $\tilde{\mathbf{g}}_n(\tilde{m}, \mathcal{M})$ and $\breve{\mathbf{g}}_n(\tilde{m}, \mathcal{M})$ constructed based on the beams in Figs. 5.2(a) and 5.2(b) by applying Theorems 5.1 and 5.2, respectively.

Fig. 5.2 illustrates the spatial-frequency patterns, where Figs. 5.2(a) and 5.2(b) show the beams for constructing the patterns, Theorems 5.1 and 5.2 are applied to construct $\mathcal{M}$ based on the beams, and Figs. 5.2(c) and 5.2(d) show the patterns based on $\mathcal{M}$. Each column in Fig. 5.2(b) or 5.2(d) represents a pattern specified by (5.8) or (5.13). We see from Figs. 5.2(c) and 5.2(d) that, in either case, the patterns are unique, which validates the one-to-one mappings in Theorems 5.1 and 5.2.

5.3.2 Fast Search of the Strongest DFT Beam

By exploiting the new spatial-frequency patterns, the strongest DFT beam can be rapidly identified via a pattern matching. Take the single-beam selection described in Section 5.3.1 for instance. The n_{L}-DFT beam is selected to feed an RF chain by setting $[\mathbf{W}]_{:,k}$ in (5.3) as

$$[\mathbf{W}]_{:,k} = \Big[0, \cdots, 0, \underset{n_{\text{L}}}{1}, 0, \cdots, 0\Big]^{\text{T}}, \tag{5.14}$$

and hence the k-th RF chain output can be written as

$$[\mathbf{X}]_{:,k} = \beta[\mathbf{U}_{:,n_{\mathrm{L}}}]^{\mathrm{H}}[\mathbf{h}_0, \mathbf{h}_1, \cdots, \mathbf{h}_{M-1}] + [\tilde{\mathbf{N}}]_{:,k}, \tag{5.15a}$$

$$= \beta\big[[\mathbf{g}(u, \rho_0)]_{n_{\mathrm{L}}}, \cdots, [\mathbf{g}(u, \rho_{M-1})]_{n_{\mathrm{L}}}\big]^{\mathrm{T}} \odot \mathbf{b} + [\tilde{\mathbf{N}}]_{:,k}, \tag{5.15b}$$

where (5.15b) is obtained by substituting (5.7) into (5.15a), s_i is suppressed since the pilot sequence is known to BS, $\odot$ stands for Hadamard product, and $\mathbf{b} \in \mathbb{C}^{M\times 1}$ collects the phase difference caused by the path delay τ, as given by $[\mathbf{b}]_m = e^{-j2\pi f_m \tau}$; see (5.4).

In (5.15), only the LoS is considered for illustration convenience[§], and hence the subscript "$_0$" of β, u and τ is suppressed. By substituting $n_{\mathrm{L}} = 128$ into (5.15), the frequency response of the received signal is shown in Fig. 5.2(c), where the AoA $u = 5.65$ and $|\beta| = 1$.

Provided $|\beta|$ is known and the condition (5.9) is satisfied, the index for the strongest DFT beam is $n^\star = n_{\mathrm{L}} + \tilde{m}^\star$, where $\tilde{m}^\star$ can be obtained through the following pattern matching:

$$\min_{\tilde{m}} \ \Big\| \big|[\mathbf{X}]_{\tilde{\mathcal{M}},k}\big| - \big|\beta \tilde{\mathbf{g}}_{n_{\mathrm{L}}}(\tilde{m}, \tilde{\mathcal{M}})\big| \Big\|_2^2. \tag{5.16}$$

Here, $\tilde{\mathcal{M}}$ is obtained from (5.9b), and $[\mathbf{X}]_{\tilde{\mathcal{M}},k}$ rearranges the elements of $[\mathbf{X}]_{:,k}$ according to $\tilde{\mathcal{M}}$; see the pentagram markers in Fig. 5.3(a). $\tilde{\mathbf{g}}_{n_{\mathrm{L}}}(\tilde{m}, \mathcal{M})$, is constructed by substituting $n = n_{\mathrm{L}}$ and $\mathcal{M} = \tilde{\mathcal{M}}$ into (5.8). Due to $|\mathbf{b}| = \mathbf{1}_M$, $\mathbf{b}$ has no impact on the pattern matching (5.16). Problem (5.16) can be depicted by matching the RF chain output in Fig. 5.3(a) with the patterns in Fig. 5.2(b) that scaled by $|\beta|$. By replacing (5.14) with (5.11) and $\tilde{\mathbf{g}}_{n_{\mathrm{L}}}(\tilde{m}, \tilde{\mathcal{M}})$ with (5.13), Problem (5.16) can be used to describe the pattern matching under multi-beam selection described in Section 5.3.1.

Note that Problem (5.16) is intractable due to the lack of the knowledge on n_{L}

[§]The mmWave/THz wireless channels are typically dominated by the LoS paths over much weaker NLoS paths [44, 66]. The applicability of the proposed approach in the multi-path case is validated in Section 5.6.

and β. With no *a-priori* information on the AoA, n_{L} can take any of the values $0, 1, \cdots, N-1$. Moreover, it is non-trivial to estimate the channel response β, since, in the RF chain output, β is coupled with the spatial response which relies on the unknown AoA; see (5.7) and (5.15). Next, we propose a viable solution to (5.16) by first identifying n_{L}, then estimating $|\beta|$ based on the output of the selected beam, and finally substituting the estimates of n_{L} and $|\beta|$ into (5.16) to perform pattern matching. To identify n_{L}, we propose to iteratively and rapidly search for the angular region of interest. Based on Theorems 5.1 and 5.2 and Lemma 5.1, we design an efficient way to partition the angular region of interest, such that n_{L} and, in turn, the strongest DFT beam can be identified using as few as a single symbol.

Let $[\mathbf{n}_{\mathrm{L}}]_k (k \in \{0, 1, \cdots, K-2\})$ denote the index for the DFT beam selected by the k-th RF chain. $[\mathbf{n}_{\mathrm{H}}]_k (k \in \{0, 1, \cdots, K-2\})$ denotes the index for the last DFT beam in the angular region covered by the $[\mathbf{n}_{\mathrm{L}}]_k$-th DFT beam (see Theorem 5.1). $\Omega_{\mathrm{T}} \in \mathcal{S}(b_{\mathrm{L}}, b_{\mathrm{H}})$ denotes the angular region of interest. b_{L} and b_{H} are the first and last beam of Ω_{T}, respectively.

Proposition 5.1. *Given K RF chains and Ω_{T}, the following beam selection scheme,*

$$[\mathbf{n}_{\mathrm{H}}]_0 = b_{\mathrm{H}},\ [\mathbf{n}_{\mathrm{L}}]_0 = \lceil [\mathbf{n}_{\mathrm{H}}]_0 \times \rho_0 \rceil; \tag{5.17a}$$

$$[\mathbf{n}_{\mathrm{H}}]_k = [\mathbf{n}_{\mathrm{L}}]_{k-1} - 1, [\mathbf{n}_{\mathrm{L}}]_k = \lceil [\mathbf{n}_{\mathrm{H}}]_k \rho_0 \rceil, k \in [1, K-2]; \tag{5.17b}$$

$$l = b_{\mathrm{L}},\ L = \lceil ([\mathbf{n}_{\mathrm{L}}]_{K-2} - 1)\rho_0 \rceil - b_{\mathrm{L}} + 1, \tag{5.17c}$$

results in the highest probability of identifying the strongest DFT beam at a symbol. A wide beam is synthesized at RF chain $(K-1)$ by substituting (5.17c) into Lemma 5.1.

Proof. See Appendix 7.4.3. □

Based on Theorems 5.1 and 5.2, and Proposition 5.1, a fast search for the strongest DFT beam is proposed in Algorithm 8. Step 1 lists the input parameters, where Ω_{T} is the initial angular region of interest, e.g., $(0, 2\pi]$. Steps 2 to 7

Algorithm 8 Fast Search of the Strongest DFT Beam

1: **Input** $\Omega_{\mathrm{T}} = \mathcal{S}(b_{\mathrm{L}}, b_{\mathrm{H}})$, K, M and N.
2: **Initialize** $[\mathbf{n}_{\mathrm{H}}]_0 = b_{\mathrm{H}}$ and $[\mathbf{n}_{\mathrm{L}}]_0 = \lceil b_{\mathrm{H}}\rho_0 \rceil$;
3: **for** $k = 1 : K-2$ **do**
4: **Update** $[\mathbf{n}_{\mathrm{H}}]_k = [\mathbf{n}_{\mathrm{L}}]_{k-1} - 1$ and $[\mathbf{n}_{\mathrm{L}}]_k = \lceil [\mathbf{n}_{\mathrm{H}}]_k \rho_0 \rceil$;
5: **Set** $\tilde{\mathbf{W}}_{:,k}$ by substituting $n_{\mathrm{L}} = [\mathbf{n}_{\mathrm{L}}]_k$ into (5.14);
6: **end for**
7: **Substitute** $L = \tilde{L} = \lceil ([\mathbf{n}_{\mathrm{L}}]_{K-2} - 1)\rho_0 \rceil - b_{\mathrm{L}} + 1$ and $l = \tilde{l} = b_{\mathrm{L}}$ into (5.11) to obtain $\tilde{\mathbf{W}}_{:,K-1}$;
8: Substituting $\tilde{\mathbf{W}}$ into (5.3) yields $\tilde{\mathbf{X}}$;
9: **Estimate** $|\beta|$ via $\{|\hat{\beta}|, k^\star\} = \max_k \left\{ \left\| [\tilde{\mathbf{X}}]_{:,k} \right\|_\infty \right\}$;
10: **if** $k^\star \neq (K-1)$ **then**
11: **Calculate** $\tilde{\mathcal{M}}$ by setting $n_{\mathrm{L}} = [\mathbf{n}_{\mathrm{L}}]_{k^\star}$ into (5.9b);
12: **Construct** $\tilde{\mathbf{g}}_{[\mathbf{n}_{\mathrm{L}}]_{k^\star}}(\tilde{m}, \tilde{\mathcal{M}}),\ \forall \tilde{m} \in [0, |\tilde{\mathcal{M}}| - 1]$ by
13: substituting $n_{\mathrm{L}} = [\mathbf{n}_{\mathrm{L}}]_{k^\star}$ and $\mathcal{M} = \tilde{\mathcal{M}}$ into (5.8);
14: **Solve** (5.16) to obtain $\tilde{m}^\star$;
15: **Return** $n^\star = n_{\mathrm{L}k^\star} + \tilde{m}^\star$.
16: **else**
17: **Calculate** $\tilde{\mathcal{M}}$ by setting $n_{\mathrm{L}} = \tilde{l} + \tilde{L} - 1$ into (5.9b);
18: **Construct** $\breve{\mathbf{g}}_{\tilde{l}+\tilde{L}-1}(\tilde{m}, \tilde{\mathcal{M}}),\ \forall \tilde{m} \in [0, |\tilde{\mathcal{M}}| - 1]$ by
19: substituting $n_{\mathrm{L}} = \tilde{l} + \tilde{L} - 1$ and $\mathcal{M} = \tilde{\mathcal{M}}$ into (5.13);
20: **Solve** (5.16) to obtain $\tilde{m}^\star$;
21: **if** $m^\star \neq 0$ **then**
22: **Return** $n^\star = \tilde{l} + \tilde{L} - 1 + \tilde{m}^\star$.
23: **else**
24: **Update** $b_{\mathrm{H}} = \tilde{l} + \tilde{L} - 1$. **Go** to Step 2.
25: **end if**
26: **end if**

partition Ω_{T} into K sub-regions, denoted by sub-regions $0, \cdots, K-1$, by applying Proposition 5.1. In Step 9, $|\beta|$ is estimated based on the RF chain outputs obtained from Step 8. Combining (5.7) and (5.15), we see that $\|[\mathbf{X}]_{:,k}\|_\infty \leq |\beta|$ in the absence of AWGNs, since we have $|[\mathbf{g}(u, \rho_0)]_{n_{\mathrm{L}}}| \leq 1$. To this end, we can use $\max_{k \in [0, K-1]} \left\{ \left\| [\mathbf{X}]_{:,k} \right\|_\infty \right\}$ as the $|\beta|$ estimate. The index for the strongest RF chain output is denoted by $k^\star$.

In the case of $k^\star \neq K-1$, the strongest DFT beam can be uniquely identified by matching the spatial-frequency pattern; refer to Theorem 5.1. First, the sub-carrier set corresponding to $[\mathbf{n}_{\mathrm{L}}]_{k^\star}$ can be calculated, leading to $\tilde{\mathcal{M}}$; see Step 11. Next, the

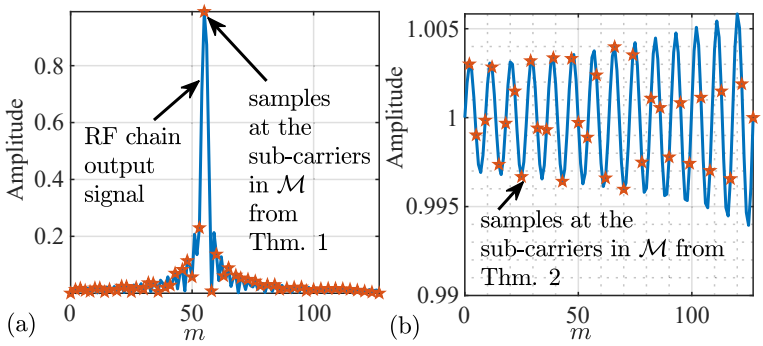

Figure 5.3 : Illustration of RF chain outputs based on (a) the single-beam selection, where $n_{\mathrm{L}} = 204$ and $u = 5.65$; and (b) the multi-beam selection, where $l = 0$, $L = 164$ and $u = 2.9$. $N = 256$, $M = 128$, and $\rho_m \in [0.8, 1]$.

spatial-frequency patterns corresponding to DFT beam $[\mathbf{n}_{\mathrm{L}}]_{k^\star}$ can be constructed by substituting $n_{\mathrm{L}} = [\mathbf{n}_{\mathrm{L}}]_{k^\star}$ and $\mathcal{M} = \tilde{\mathcal{M}}$ into (5.8); see Step 12. Then, by substituting $\mathcal{M} = \tilde{\mathcal{M}}$, $|\beta| = |\hat{\beta}|$, $n_{\mathrm{L}} = [\mathbf{n}_{\mathrm{L}}]_{k^\star}$, $\tilde{\mathbf{g}}_{n_{\mathrm{L}}}(\tilde{m}, \mathcal{M}) = \tilde{\mathbf{g}}_{[\mathbf{n}_{\mathrm{L}}]_{k^\star}}(\tilde{m}, \tilde{\mathcal{M}})$, and $\mathbf{X} = \tilde{\mathbf{X}}$ into (5.16), the pattern matching can be performed, leading to $\tilde{m}^\star$. Finally, the index for the strongest DFT beam, denoted by $n^\star$, is obtained in Step 15.

In the case of $k^\star = K - 1$, the strongest DFT beam can be uniquely identified if $u \in \mathcal{S}(\tilde{l} + \tilde{L}, [\mathbf{n}_{\mathrm{L}}]_{K-2} - 1)$; refer to Theorem 5.2. Therefore, we proceed to check whether the strongest DFT beam falls into the angular region. Step 17 evaluates the sub-carrier set corresponding to DFT beam $\tilde{l} + \tilde{L} - 1$. Step 18 constructs the spatial-frequency patterns. Step 20 performs pattern matching, leading to $\tilde{m}^\star$. We further check whether $\tilde{m}^\star$ is acceptable, since the first spatial-frequency pattern needs to be excluded in the case of multi-beam selection, as proved in Theorem 5.2. If $m^\star \neq 0$, the strongest DFT beam is uniquely identified; see Step 22; otherwise, we can assert $u \in \mathcal{S}(b_{\mathrm{L}}, \tilde{l} + \tilde{L} - 1)$. Accordingly, we update $b_{\mathrm{H}} = \tilde{l} + \tilde{L} - 1$ and return to Step 2; see Step 24.

An example of Algorithm 1 is provided in Figs. 5.2 and 5.3. The beams selected and synthesized at the two RF chains by following Steps 2 to 7 are shown in Figs. 5.2(a) and 5.2(c). *In the case of* $u = 5.65$ rad, the first RF chain produces the stronger output with the frequency response plotted in Fig. 5.3(a). This leads to $k^\star = 0 (\neq K - 1)$ in Step 9. By implementing Steps 11 and 12, the spatial-frequency

patterns are obtained in Fig. 5.2(b). By matching Fig. 5.3(a) with the patterns in Fig. 5.2(b), $\tilde{m}^\star$ and $n^\star$ can be identified in Steps 14 and 15, respectively. *In the case of* $u = 2.9$ rad, the second RF chain produces the stronger output, leading to $k^\star = K - 1 = 1$ in Step 9. The frequency response of the second RF chain output is shown in Fig. 5.3(b). Thus, the pattern matching based on the multi-beam selection needs to be performed. The spatial-frequency patterns, associated with the wide beam in Fig. 5.2(b), are constructed in Steps 17 and 18, and plotted in Fig. 5.2(d). By matching Fig. 5.3(b) with the patterns in Fig. 5.2(d), $\tilde{m}^\star = 0$ can be obtained in Step 20, followed by Step 24 to restart the search of a narrower angular region, denoted by Ω_{2nd}, in Fig. 5.2(c).

5.4 Estimating AoA from DFT Beam Outputs

With the strongest DFT beam identified from Algorithm 8, we design an efficient algorithm to estimate the AoA with a fine accuracy. The research approach taken is to first develop a one-to-one mapping between the AoA and a general set of DFT beams which may not include the strongest DFT beam; and then prove that, if the strongest DFT beam is activated, the AoA estimation accuracy can asymptotically approach what is achieved with all the DFT beams. The one-to-one mapping between the AoA and the selected DFT beams is established under a noise-free condition. With the noise-free mapping, we reveal that in the presence of non-negligible noises, the estimation problem can be converted to a maximum-likelihood estimator with additive noises. The maximum-likelihood estimator is known to be optimal and achieve the Cramér-Rao lower bound in the presence of additive noises. Moreover, we derive the analytical lower bound of the mean squared error (MSE) of the AoA estimate to evaluate the noise-resistance performance of the proposed method. This research approach has been taken by state-of-the-art studies on AoA/channel estimation, e.g., [13, 14] and [21].

For illustration convenience, we consider that K is a factor of N, i.e., $Q = \frac{N}{K}$ is an integer. At the i-th symbol, the $\big((i)_Q + kQ\big)$-th DFT beam ($i \geq 0$ and $0 \leq k \leq K-1$) is selected at the k-th RF chain. $(\cdot)_Q$ denotes modulo-Q. The beam selection can

be achieved by configuring $\mathbf{W}$. Take $N = 4$ and $K = 2$ for instance. We can set $\mathbf{W} = \left[\begin{smallmatrix} 1, & 0, & 0, & 0 \\ 0, & 0, & 1, & 0 \end{smallmatrix}\right]^{\mathrm{T}}$ at symbol $i = 0$ to select the 0th and 3rd DFT beams at the two RF chains; and set $\mathbf{W} = \left[\begin{smallmatrix} 0, & 1, & 0, & 0 \\ 0, & 0, & 0, & 1 \end{smallmatrix}\right]^{\mathrm{T}}$ at symbol $i = 1$ to select the other two DFT beams.

Let $\tilde{\mathbf{h}}_m = \mathbf{U}^{\mathrm{H}}\mathbf{h}_m \in \mathbb{C}^{N\times 1}$; see (5.3), and refer to $\tilde{\mathbf{h}}_m$ as the beamspace channel response vector at sub-carrier m. By substituting (5.4) and (5.5) into (5.3), and assuming $0 \le i \le Q-1$ (such that the modulo operator $(\cdot)_Q$ can be suppressed), the $\big((i)_Q + kQ\big)$-th element of $\tilde{\mathbf{h}}_m$ is given by

$$[\tilde{\mathbf{h}}_m]_{i+kQ} = \beta e^{-j2\pi f_m \tau}[\mathbf{U}^{\mathrm{H}}\mathbf{a}(u,\rho_m)]_{i+kQ} \tag{5.18a}$$

$$= \frac{1}{N}\beta e^{-j2\pi f_m \tau}\sum_{n'=0}^{N-1} e^{-jn'(\rho_m u - \frac{2\pi(i+kQ)}{N})} \tag{5.18b}$$

$$= \frac{1}{N}\beta e^{-j2\pi f_m \tau}\sum_{k'=0}^{K-1}\sum_{q=0}^{Q-1} e^{-j(k'+qK)(\rho_m u - \frac{2\pi(i+kQ)}{N})} \tag{5.18c}$$

$$= \sum_{k'=0}^{K-1} \underbrace{\frac{1}{N}\beta e^{-j2\pi f_m \tau}\mathcal{G}(u,\rho_m,i)e^{-jk'(\rho_m u - \frac{2\pi i}{N})}}_{[\mathbf{f}(u,\rho_m,i)]_{k'}} e^{j\frac{2\pi k' k}{K}}, \tag{5.18d}$$

where (5.18c) is obtained by substituting $n' = k' + qK$ in (5.18b) for $0 \le k' \le K-1$ and $0 \le q \le Q-1$; and

$$\mathcal{G}(u,\rho_m,i) = \sum_{q=0}^{Q-1} e^{-jqK(\rho_m u - \frac{2\pi(i+kQ)}{N})} = e^{j\frac{Q-1}{2}K(\rho_m u - \frac{2\pi i}{N})}\frac{\sin\frac{QK}{2}(\rho_m u - \frac{2\pi i}{N})}{\sin\frac{K}{2}(\rho_m u - \frac{2\pi i}{N})}. \tag{5.19}$$

Clearly, $\mathcal{G}(u,\rho_m,i)$ is independent of k and k'.

From (5.18d), we notice that the beamspace channel vector $[\tilde{\mathbf{h}}_m]_{i+kQ}$ ($k = 0,1,\cdots,K-1$) is the IDFT of $[\mathbf{f}(u,\rho_m,i)]_{k'}$ ($k' = 0,1,\cdots,K-1$). $[\mathbf{f}(u,\rho_m,i)]_{k'}$ are the Fourier coefficients. The cross-correlation of any two consecutive Fourier coefficients has a constant angular component $\rho_m u$, and can be used for unambiguous u estimation. In practice, only the scaled version of $[\tilde{\mathbf{h}}_m]_{i+kQ}$ is available, i.e., $\mathbf{X}$ in (5.15). Therefore, we can take the DFT of $\mathbf{X}$ w.r.t. k, i.e., column-wise DFT, to estimate the AoA u. At sub-carrier m, by taking the DFT of the m-th column of

Algorithm 9 Wideband AoA Estimation

1: **Input** $n^\star$ (obtained by running Algorithm 8), K and N.
2: **Calculate** $i^\star = (n^\star)_Q$, and select the DFT beams for AoA estimation: $\mathcal{I} = \{i^\star, i^\star + Q, \cdots, i^\star + (K-1)Q\}$;
3: Set the beam selection matrix $\tilde{\mathbf{W}}$ based on $\mathcal{I}$; see (5.23);
4: Substituting $\mathbf{W} = \tilde{\mathbf{W}}$ into (5.3) yields $\tilde{\mathbf{X}}$;
5: Take DFT of each column of $\tilde{\mathbf{X}}$; see (5.18);
6: Take cross-correlation of the DFT coefficients; see (5.20);
7: Accumulating $[\mathbf{r}(u, \rho_m, i^\star)]_{k'}$ across RF chain and sub-carrier yields $\tilde{r}(u, i^\star)$; see (5.21);
8: Taking the angle of $\tilde{r}(u, i^\star)$ for u estimate, i.e., (5.22);

$\mathbf{X}$, the DFT coefficient becomes $K[\mathbf{f}(u, \rho_m, i)]_{k'} s_i + [\tilde{\mathbf{n}}_m]_{k'}$, where $\tilde{\mathbf{n}}_m$ is the DFT of the m-th column of $\tilde{\mathbf{N}}$, and K is multiplied to $[\mathbf{f}(u, \rho_m, i)]_{k'}$ based on the definition of IDFT [24].

Taking the cross-correlation between the consecutive coefficients, we obtain

$$\begin{aligned}[\mathbf{r}(u, \rho_m, i)]_{k'} &= [\mathbf{f}(u, \rho_m, i)]_{k'} [\mathbf{f}(u, \rho_m, i)]^*_{k'+1} |s_i|^2 + [\breve{\mathbf{n}}_m]_{k'} \\ &= \frac{K^2 |\beta s_i|^2}{N^2} |\mathcal{G}(u, \rho_m, i)|^2 e^{j(\rho_m u - \frac{2\pi i}{N})} + [\breve{\mathbf{n}}_m]_{k'}, \end{aligned} \tag{5.20}$$

where $\breve{\mathbf{n}}_m$ is the noise component after the cross-correlation. As the signal component of $[\mathbf{r}(u, \rho_m, i)]_{k'}$ is independent of k', we can coherently accumulate $[\mathbf{r}(u, \rho_m, i)]_{k'}$ across k' to improve the estimation SNR for u, i.e., $\sum_{k'=0}^{K-2} [\mathbf{r}(u, \rho_m, i)]_{k'} = (K-1)[\mathbf{r}(u, \rho_m, i)]_{k'}$. Since $\rho_m \leq 1$, we can take the angle of $(K-1)[\mathbf{r}(u, \rho_m, i)]_{k'}$ at any sub-carrier m and symbol i for the unambiguous u estimation.

By taking $m_1 \leq m \leq m_2$, $|\mathcal{G}(u, \rho_m, i)|^2$ is approximately uniform across m, and $(K-1)[\mathbf{r}(u, \rho_m, i)]_{k'}$ can be coherently accumulated across m [7, Proposition 1], leading to

$$\begin{aligned}\tilde{r}(u, i) &= \underbrace{\frac{(K-1)K^2 |\beta s_i|^2}{N} |\mathcal{G}(u, \rho_m, i)|^2 e^{-j\frac{2\pi i}{N}}}_{\mathcal{C}(i)} \sum_{m=m_1}^{m_2} e^{j\rho_m u} \\ &= \mathcal{C}(i) \times e^{j\frac{f_{\mathrm{L}} + \frac{B(m_2+m_1)}{2(M-1)}}{f_{\mathrm{H}}} u} \frac{\sin \frac{Bu(m_2-m_1+1)}{2(M-1)f_{\mathrm{H}}}}{\sin \frac{Bu}{2(M-1)f_{\mathrm{H}}}} \end{aligned} \tag{5.21}$$

where $0 \leq m_2 - m_1 \leq M-1$, $\frac{B}{(M-1)f_{\mathrm{H}}} < \frac{B}{(M-1)f_{\mathrm{H}}}(m_2 - m_1 + 1) \leq \frac{B}{f_{\mathrm{H}}} \ll 1$, and hence the numerator and denominator of (5.21) always take the same sign. The accumulation gain of (5.21) is approximately $(m_2 - m_1 + 1)$ [7, Remark 1]. For illustration convenience, the noise term is suppressed in (5.21), and its impact on the AoA estimation will be analyzed later.

By taking the angle of $\tilde{r}(u,i)$, the one-to-one mapping of u to the selected DFT beams is finally established, as given by

$$\hat{u}_i = \frac{f_{\mathrm{H}}}{f_{\mathrm{L}} + \frac{B(m_2+m_1)}{2(M-1)}} \arg\{\tilde{r}(u,i)\}. \tag{5.22}$$

Note that the estimation accuracy of $\hat{u}_i$ is affected, not only by the noise component in (5.20), but $\mathcal{C}(i)$ in (5.21) as well. From (5.19) and (5.21), we can see that the amplitude of $\mathcal{C}(i)$ depends on the selected DFT beams. By selecting the DFT beams with the strongest received signal power, the largest $|\mathcal{C}(i)|$ and, in turn, the largest AoA estimation SNR can be attained.

Corollary 5.1. *In the case of $N \to \infty$, the AoA estimate (5.22) is dominated by the DFT beam $i^\star + k^\star Q$ which is the strongest DFT beam; and the contribution of other DFT beams is asymptotically negligible.*

Proof. See Appendix 7.4.4. □

Algorithm 9 summarizes the proposed AoA estimation technique by applying Corollary 5.1 in the one-to-one mapping established in (5.22). Step 1 lists the input parameters, where $n^\star$ is obtained from Algorithm 8. In Step 2, $i^\star$ is calculated via $i^\star = (n^\star)_Q$ and the set of DFT beams for the AoA estimation is identified, denoted by $\mathcal{I}$. The index for the strongest DFT beam is in $\mathcal{I}$. In Step 3, the beam selection matrix is configured based on $\mathcal{I}$. The k-th column of $\tilde{\mathbf{W}}$ can be given by

$$[\tilde{\mathbf{W}}]_{:,k} = \Big[0, \cdots, 0, \underset{[\mathcal{I}]_k-\mathrm{th}}{1}, 0, \cdots, 0\Big]^{\mathrm{T}}. \tag{5.23}$$

In Step 4, $[\tilde{\mathbf{W}}]_{:,k}$ is set at RF chain k to perform beam selection. In Step 5, the

column-wise DFT is taken on $\tilde{\mathbf{X}}$ based on (5.18). The DFT coefficients are cross-correlated and accumulated in Steps 6 and 7, respectively. Finally, the AoA is estimated unambiguously in Step 8.

Algorithm 9 can be carried out by applying the same $\mathcal{I}$ for receiving over the consecutive $N^u_{\rm sym}(\geq 1)$ symbols, in the case that there are $N^u_{\rm sym}$ symbols for u estimation. In this case, Steps 4 to 7 are carried out individually at each symbol. Let $\tilde{r}_{i'}(u, i^\star)$ denote the output of Step 7 at any symbol $i' \in [1, N^u_{\rm sym}]$. In Step 8, before taking the angle for AoA estimation, $\tilde{r}_{i'}(u, i^\star)$ can be accumulated across i', i.e., $\sum_{i'=1}^{N^u_{\rm sym}} \tilde{r}_{i'}(u, i^\star)$. This is a coherent accumulation, because $\tilde{r}_{i'}(u, i^\star)$ $\forall i'$ is identical by applying the same $\mathcal{I}$; see (5.21).

Note that (5.22) is derived with reference to [7]; nevertheless, the asymptotic analysis in Corollary 5.1 is new: in the LoS-dominant mmWave channel, the strongest DFT beam having the strongest output, i.e., satisfying $\rho_m u - \frac{2\pi i^\star}{N} + \frac{2\pi k^\star}{K} \in \left(-\frac{\pi}{N}, \frac{\pi}{N}\right)$, contributes the most to the AoA estimation. In other words, given $(i^\star, k^\star)$, we can use *a single symbol* to estimate the AoA of interest accurately. Without this, the complete set of DFT beams have to be enumerated in [7], for which $\frac{N}{K}$ symbols would be required.

Also note that all the three values of the beam selection, i.e., 0 and ± 1, are used in the proposed AoA estimation method across this chapter. Specifically, during the wide beam synthesis, all the three beam selection values are all used; see (5.11). The synthesized beams are used in Algorithm 1 for quick search of strongest DFT beams, and hence contribute to the estimation of the AoA in Algorithm 2.

5.5 Performance Analysis

This section analyzes the average number of symbols used by Algorithm 8, and the MSELB of the proposed AoA estimation, $\hat{u}_{i^\star}$, from Algorithm 9.

5.5.1 Average Number of Symbols Used by Algorithm 8

We start by deriving the probability at which Algorithm 8 requires x symbols, where x can be as small as one. This is because, by exploiting the new spatial-

frequency pattern, Algorithm 8 can identify the strongest DFT beam at any symbol, if the AoA u sits in the unambiguous angular region probed at the symbol[¶].

Lemma 5.2. *Assume that beamspace-domain AoA u is uniformly distributed in $[0, 2\pi]$. The probability at which Algorithm 8 uses x symbols asymptotically converges to*

$$\lim_{N\to\infty} f(x) = (1-\rho_0^K)\rho_0^{(x-1)K},\ x \geq 1. \tag{5.24}$$

Proof. See Appendix 7.4.5. □

Based on Lemma 5.2, the average number of symbols used by Algorithm 8 for the fast search of the strongest DFT beam can be evaluated, as follows.

Theorem 5.3. *The average number of symbols used by Algorithm 8, denoted by $\bar{I}$, can be asymptotically given by*

$$\lim_{N\to\infty} \bar{I} = \frac{1}{1-\rho_0^K}. \tag{5.25}$$

Proof. Based on the results from Lemma 5.2, we have

$$\begin{aligned}\lim_{N\to\infty} \bar{I} &= \lim_{N\to\infty} \sum_{x=1}^{\infty} x f(x) = \lim_{N\to\infty} \sum_{x=1}^{\infty} (1-\rho_0^K) x \rho_0^{(x-1)K} \\ &\overset{\tilde{\rho}_0:=\rho_0^K}{=\!=} \lim_{N\to\infty} (1-\tilde{\rho}_0) \underbrace{\sum_{x=1}^{\infty} x \tilde{\rho}_0^{x-1}}_{g(\tilde{\rho}_0)} \qquad (5.26a) \\ &= (1-\tilde{\rho}_0) \lim_{x\to\infty} \frac{\mathrm{d}\left(\tilde{\rho}_0 + \tilde{\rho}_0^2 + \cdots + \tilde{\rho}_0^x\right)}{\mathrm{d}\tilde{\rho}_0}. \qquad (5.26b)\end{aligned}$$

To calculate $g(\rho_0)$ in (5.26a), the sum of the geometric series in (5.26b) is first calculated, leading to $\tilde{\rho}_0 + \tilde{\rho}_0^2 + \cdots + \tilde{\rho}_0^x = \frac{\tilde{\rho}_0(1-\tilde{\rho}_0^x)}{1-\tilde{\rho}_0}$; and then taking the derivative of the series sum w.r.t. $\tilde{\rho}_0$ results in $g(\tilde{\rho}_0) = \frac{1-\tilde{\rho}_0^{x+1}-(x+1)\tilde{\rho}_0^x(1-\tilde{\rho}_0)}{(1-\tilde{\rho}_0)^2}$ which asymptotically

[¶]The unambiguous angular region consists of the sub-regions probed by the first $(K-1)$ RF chains; see Steps 11 to 15, and a part of the sub-region probed by RF chain $(K-1)$, given by $\mathcal{S}(\tilde{L}+\tilde{l}, [\mathbf{n}_{\mathrm{L}}]_{K-2})$; see Step 17.

$$\begin{aligned}\sum_{k'=0}^{K-2}[\breve{\mathbf{n}}_m]_{k'} &= \sum_{k'=0}^{K-2}[\tilde{\mathbf{n}}_m]_{k'}[\tilde{\mathbf{n}}_m]^*_{k'+1} + \sum_{k'=0}^{K-2} KM[\mathbf{f}(u,\rho_m,i^\star)]^*_{k'} s^*_{i^\star}[\tilde{\mathbf{n}}_m]_{k'+1} \\ &\quad + \sum_{k'=0}^{K-2} KM[\mathbf{f}(u,\rho_m,i^\star)]_{k'+1} s_{i^\star}[\tilde{\mathbf{n}}_m]^*_{k'} \\ &\geq \sum_{k'=0}^{K-2}[\tilde{\mathbf{n}}_m]_{k'}[\tilde{\mathbf{n}}_m]^*_{k'+1} + KM[\mathbf{f}(u,\rho_m,i^\star)]^*_{K-2} s^*_{i^\star}[\tilde{\mathbf{n}}_m]_{K-1} \\ &\quad + KM[\mathbf{f}(u,\rho_m,i^\star)]_1 s_{i^\star}[\tilde{\mathbf{n}}_m]^*_0 + 2\mathcal{R}\{\xi\} \end{aligned} \tag{5.27}$$

$$\begin{aligned}\bar{\gamma}_{\text{up}} = \frac{\sigma_{\text{s}}^2}{\sigma_{\text{n}}^2} &\overset{(a)}{=} \frac{\frac{K^4}{N^4}|M\beta s_{i^\star}|^4 |\mathcal{G}(u,\rho_m,i^\star)|^4 (K-1)^2}{\frac{2K^2M^2}{N^2}|\beta s_{i^\star}|^2 |\mathcal{G}(u,\rho_m,i^\star)|^2 KM\sigma_0^2} \\ &\overset{(b)}{=} \frac{KM|\mathcal{G}(u,\rho_m,i^\star)|^2(K-1)^2}{2N^2} \underbrace{\frac{|\beta s_{i^\star}|^2}{\sigma_0^2}}_{\gamma_0} \overset{(c)}{=} \frac{M(K-1)^2\gamma_0}{2K} \end{aligned} \tag{5.28}$$

converges to $\frac{1}{(1-\tilde{\rho}_0)^2}$. This is because $\tilde{\rho}_0 < 1$ and $\lim_{x\to\infty} \tilde{\rho}_0^x = 0$. By substituting $g(\tilde{\rho}_0) = \frac{1}{(1-\tilde{\rho}_0)^2}$ into (5.26b), (5.25) is obtained, which concludes the proof. □

As revealed in Lemma 5.2, there is a high probability given by $(1 - \rho_0^K)$, at which Algorithm 8 can identify the strongest DFT beam with a single symbol. By substituting the typical wideband configuration $\rho_0 \approx 0.2$ and $K = 8$ [8] into (5.24), $f(1) \approx 0.832$. Moreover, the probability increases with the bandwidth and the number of RF chains. To the best of our knowledge, no existing methods can identify the strongest DFT beam with a single symbol.

Theorem 5.3 indicates that the average number of symbols required by Algorithm 8 depends asymptotically on the system bandwidth and the number of RF chains, K. We note that $\bar{I}$ decreases, as ρ_0 or K increases. In the above example, $\bar{I} \approx 1.2$. Although Lemma 5.2 and Corollary 5.1 are derived based on the asymptotic condition that $N \to \infty$, we will show in Section 5.6 that the results are accurate for a typical massive MIMO antenna with $N = 256$.

5.5.2 MSELB of Algorithm 9

With reference to [7], the MSELB can be approximated by $\frac{1}{2\bar{\gamma}_{\rm up}}$ in high SNR regions, where $\bar{\gamma}_{\rm up}$ is the upper bound of the estimation SNR for u. Therefore, we first derive $\gamma_{\rm up}$. From (5.22) and (5.21), the signal power is $\sigma_{\rm s}^2 = |\mathcal{C}(i^\star)|^2(m_2 - m_1 + 1)^2 = \frac{K^4}{N^4}|M\beta s_{i^\star}|^4|\mathcal{G}(u,\rho_m,i^\star)|^4(K-1)^2$, where $|Ms_{i^\star}|$ is the amplitude of the signal component in (5.3), $\mathcal{G}(u,\rho_m,i^\star)$ is given in (5.19), and $(m_2 - m_1 + 1)$ is the gain of the accumulation across sub-carriers; see (5.21).

Based on (5.20), the noise term in the cross-correlation of consecutive DFT coefficients can be given by $[\breve{\mathbf{n}}_m]_{k'} = [\tilde{\mathbf{n}}_m]_{k'}[\tilde{\mathbf{n}}_m]^*_{k'+1} + KM[\mathbf{f}(u,\rho_m,i^\star)]^*_{k'}s^*_{i^\star}[\tilde{\mathbf{n}}_m]_{k'+1} + KM[\mathbf{f}(u,\rho_m,i^\star)]_{k'+1}s_{i^\star}[\tilde{\mathbf{n}}_m]^*_{k'}$, where $\tilde{\mathbf{n}}_m$ is the DFT of the m-th column of $\tilde{\mathbf{N}}$ in (5.15) and $\tilde{\mathbf{N}}$ is the row-wise DFT of $\mathbf{N}$ in (5.2). The accumulation in (5.21) also accumulates the noise $[\breve{\mathbf{n}}_m]_{k'}$ across k', leading to (5.27), where

$$\xi = \sum_{k'=0}^{K-3} KM[\mathbf{f}(u,\rho_m,i^\star)]^*_{k'}s^*_{i^\star}[\tilde{\mathbf{n}}_m]_{k'+1}.$$

The inequality in (5.27) is obtained due to the removal of the imaginary parts of ξ, provided $\rho_m u - \frac{2\pi i^\star}{N} = 0$ and $[\mathbf{f}(u,\rho_m,i^\star)]_{k'}$ is invariant with k'; see (5.18d).

By suppressing the real component in (5.27), the power of the accumulated noise, denoted by $\sigma_{\rm n}^2$, is lower bounded by

$$\begin{aligned}\sigma_{\rm n}^2 \geq& K^2M^2|[\mathbf{f}(u,\rho_m,i^\star)]_{K-2}s_{i^\star}|^2\sigma^2_{\tilde{\mathbf{n}}} + K^2M^2|[\mathbf{f}(u,\rho_m,i^\star)]_1 s_{i^\star}|\sigma^2_{\tilde{\mathbf{n}}}\\ =& \frac{2K^2M^2}{N^2}|\beta s_{i^\star}|^2|\mathcal{G}(u,\rho_m,i^\star)|^2KM\sigma_0^2,\end{aligned}$$

where $\sigma^2_{\tilde{\mathbf{n}}}$ is the variance of every element of the AWGN $\tilde{\mathbf{N}}$ and σ_0^2 is the variance of every element of $\mathbf{N}$.

By assuming the AWGN $\mathbf{N}$ satisfies $[\mathbf{N}]_{km} \sim \mathcal{CN}(0,\sigma_0^2)\ \forall k,m$, the power of the AWGN $\tilde{\mathbf{N}}$ is $[\tilde{\mathbf{N}}]_{km} \sim \mathcal{CN}(0,KM\sigma_0^2)\ \forall k,m$; see (5.2) and (5.15). Given $\sigma_{\rm s}^2$ and $\sigma_{\rm n}^2$, the SNR of the AoA estimation is upper bounded by (5.28), where γ_0 is the receive SNR at each antenna element, (a) is obtained by replacing $\sigma_{\rm s}^2$ and $\sigma_{\rm n}^2$ with their expressions derived earlier in this section, and (b) is obtained by suppressing the

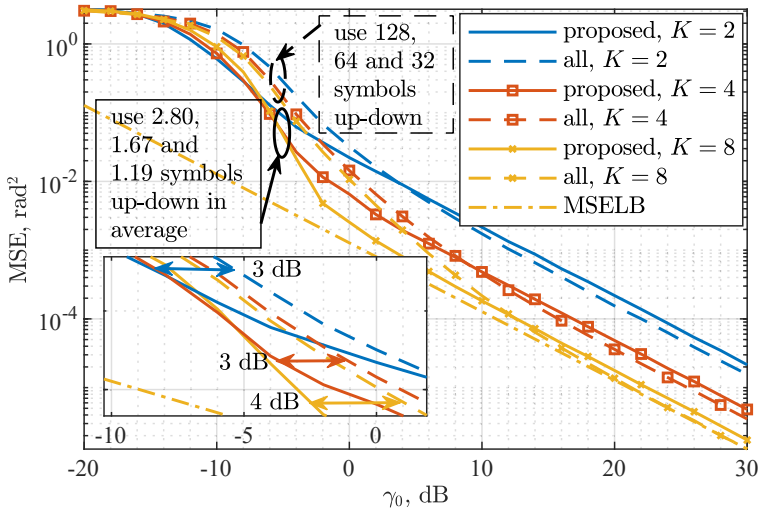

Figure 5.4 : MSE of the AoA estimate, where the state of the art [7], referred to as "all" (since the method uses all the DFT beams), is simulated for comparison.

common factors in the numerator and denominator. In (5.28), (c) is obtained by plugging (7.31) into (b). Specifically, as $N \to \infty$, we have

$$\begin{aligned}\lim_{N\to\infty} \bar{\gamma}_{\text{up}} &= \lim_{N\to\infty} \frac{KM|\mathcal{G}(u,\rho_m,i^\star)|^2(K-1)^2}{2N^2}\gamma_0 \\ &= \frac{KM\left[\lim_{N\to\infty}|\mathcal{G}(u,\rho_m,i^\star)|^2\right](K-1)^2}{2N^2}\gamma_0 = \frac{KMQ^2(K-1)^2}{2N^2}\gamma_0 \\ &= \frac{MK^2Q^2(K-1)^2}{2KN^2}\gamma_0 = \frac{M(K-1)^2\gamma_0}{2K},\end{aligned} \tag{5.29}$$

where the third equality is obtained based on (7.31), the fourth is obtained by multiplying K to both the enumerator and denominator, and the last equality is obtained due to the setting $N = QK$. In the case that $N^u_{\text{sym}}(\geq 1)$ symbols are used in Algorithm 2, the estimation SNR for u is $N^u_{\text{sym}} \times \bar{\gamma}_{\text{up}}$. Finally, the MSELB of the proposed AoA estimation is given by

$$\text{MSELB}(\hat{u}_{i^\star}) = \frac{1}{2N^u_{\text{sym}}\bar{\gamma}_{\text{up}}} = \frac{K}{N^u_{\text{sym}}M(K-1)^2\gamma_0}. \tag{5.30}$$

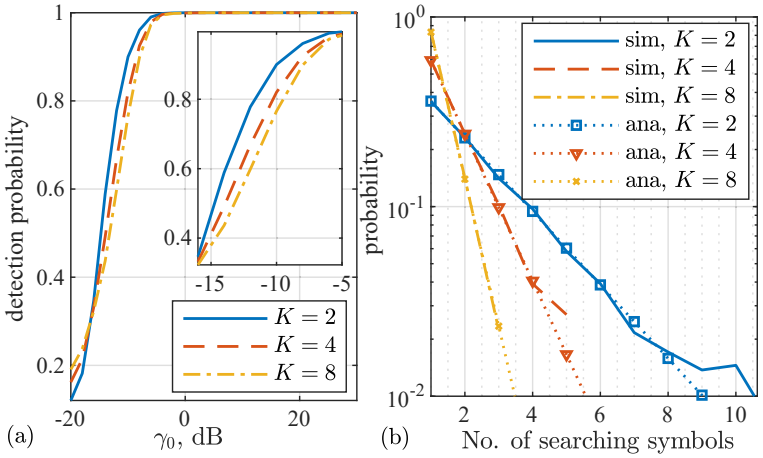

Figure 5.5 : (a) the detection probability of Algorithm 8; (b) the simulated and analytical probability that Algorithm 8 uses x symbols ($x = 1, 2, \cdots$), referred to as "sim" and "ana", respectively.

5.6 Simulation Results

In this section, simulations are carried out to validate the proposed method and analysis. Without loss of generality, linear LAAs are considered. Unless otherwise specified, the array configuration and other simulation parameters are set as: $N = 256$, $M = 128$, $f_{\mathrm{H}} = 1$, $f_{\mathrm{L}} = 0.8$, $B = 0.2$, $K = 4$ and $N^u_{\mathrm{sym}} = 1$. We first consider LoS-dominant channels with negligible NLoS paths, where the beamspace-domain AoA of the LoS satisfies $u_0 \in \mathcal{U}[0, 2\pi]$ and $|\beta_0| = 1$. Then the multi-path scenario in [8] is simulated to validate the applicability of the proposed algorithms, where $|\beta_0| = 1$ and $\beta_p \sim \mathcal{CN}(0, -10 \text{ dB})$, $p = 1, 2$; and $u_p \sim \mathcal{U}[0, 2\pi]$, $p = 0, 1, 2$. $N_{\mathrm{trial}} = 2 \times 10^4$ independent trials are carried out to get each of the results in this section.

Fig. 5.4 compares the MSE of the AoA estimate between the proposed method and the existing work [7], as γ_0 increases, where the number of symbols used by Algorithm 8 is listed in Table 5.1, and $N^u_{\mathrm{sym}} = 1$ symbol is adopted to perform Algorithm 9. We see that, using a much smaller number of symbols on average, the proposed method outperforms the state of the art in the low SNR regions. To achieve the same MSE of 4.975×10^{-3} rad^2, the proposed algorithm can operate at a 4 dB lower estimation SNR than the existing method [21], when $K = 8$. . We also

Table 5.1 : Simulated occurrence of using x symbols by Algorithm 8

x	1	2	3	4	5	6	SA	AA
$K=2$	7169	4631	2821	1944	1162	785	2.80	2.78
$K=4$	11799	4840	2039	783	539	0	1.67	1.69
$K=6$	16723	2830	447	0	0	0	1.19	1.20

[1] The content of the six columns corresponding to $x = 1, 2, \cdots, 6$ is the occurrence times based on 2×10^4 independent trials;
[2] SA and AA refer to the "Simulated Average number of symbols" and the "Analytical Average number of symbols" used by Algorithm 8, respectively;
[3] x can take up to 11 in the case of $K = 2$, the occurrence times of which are less than 450 and suppressed due to limited space.

see that, despite using a single symbol, Algorithm 9 attains the considerably close performance to the existing work [7], using 32, 64 and 128 symbols in the case of $K = 2,\ 4$ and 8, respectively.

The key reason for the superiority of the proposed method is that the strongest DFT beam can be fast identified by exploiting the new spatial-frequency patterns in Algorithm 8. As a result, only a single symbol is required to carry out Algorithm 9. In contrast, without the *a-priori* information on the AoA, the existing work [7] had to enumerate all the N DFT beams. To this end, the AWGNs are accumulated by the proposed method not as much as the existing work [7] in the noise-limited low SNR regions. In the high SNR regions, the strongest DFT beams dominates the AoA estimation performance, as proved in Corollary 5.1.

Fig. 5.5(a) plots the probability of correctly detecting the strongest DFT beam using Algorithm 8, referred to as "detection probability". We see that the detection probability asymptotically approaches 1 as γ_0 increases, and achieves 1 when $\gamma_0 \geq 0$ dB. By jointly inspecting Figs. 5.4(a) and 5.5, we see that the AoA estimation performance is affected by the detection probability when $\gamma_0 \leq -2$ dB. The high detection probability contributes to the high estimation accuracy of Algorithm 9 in high SNR regions, i.e., $\gamma_0 \geq 0$ dB. From Fig. 5.5(a), the detection probability decreases as K increases, for $-15 \leq \gamma_0 \leq -5$ dB. This leads to the better estimation accuracy at smaller K in the same SNR region; see Fig. 5.4.

Fig. 5.5(b) plots the probability that Algorithm 8 uses $x(= 1, 2, \cdots)$ symbols. As

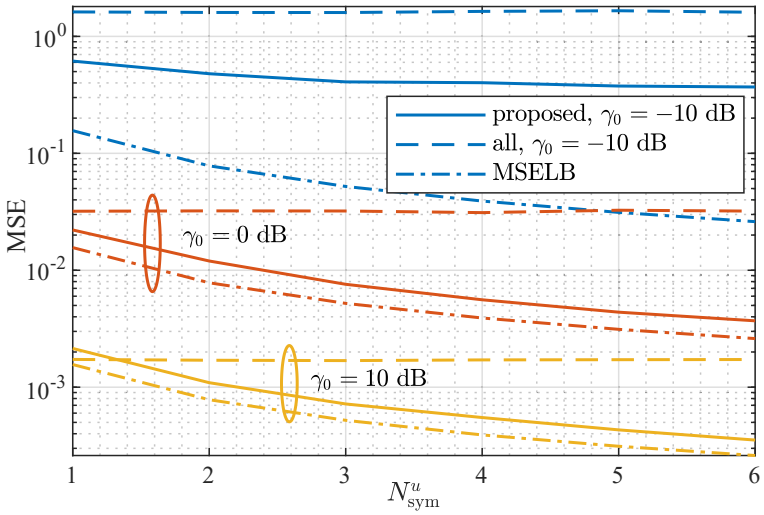

Figure 5.6 : MSE of the AoA estimate against the number of symbols used by Algorithm 9, denoted by N^u_{sym}, where $K = 4$.

shown in the figure, the probability of a single symbol being sufficient for Algorithm 1 to correctly identify the strongest DFT beam grows with K. The probability can be as high as 83.2%. Based on the simulated probability, the average numbers of symbols used by Algorithm 8 are 2.80, 1.67, and 1.19 in the case of $K = 2$, 4, and 8, respectively. Correspondingly, the analytical average numbers of symbols are 2.78, 1.69 and 1.20 by applying Theorem 5.3. The conclusion drawn is that Fig. 5.5(b) validates the applicability of Lemma 5.2 and Theorem 5.3 in typical massive MIMO settings with a large finite number of antennas, e.g., 256. Hence they can serve as a practical guideline for the design of LAA-based wideband massive MIMO transceivers.

Fig. 5.6 compares the MSE of the AoA estimations between the proposed algorithm, i.e., Algorithm 9, and the existing work [7], as N^u_{sym} increases. $K = 4$. We see that the estimation accuracy of the proposed AoA estimation technique increases with N^u_{sym}, especially in high SNR regions, i.e., $\gamma_0 \geq 0$ dB. The MSE can be reduced by over 80% from 2.14×10^{-3} rad^2 to 3.53×10^{-4} rad^2, as N^u_{sym} grows from 1 to 6. Algorithm 9 can also achieve as large as 88% improvement of the MSE in the case of $N^u_{\text{sym}} = 6$. Even in the case of $N^u_{\text{sym}} = 6$, the total number of symbols used by the proposed method is still much smaller than that of the existing work [7]. In Fig. 5.6,

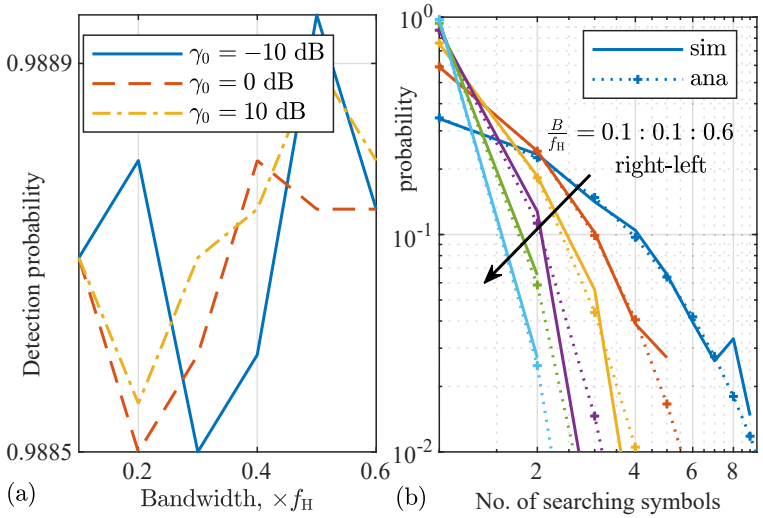

Figure 5.7 : (a) detection probability of the strongest DFT beam against bandwidth B; (b) probability of Algorithm 8 using x symbols given various B.

we also see that the MSE improvement in the case of $\gamma_0 = -10$ dB is not as obvious as that in the other two cases. This indicates that the AoA estimation accuracy is limited by the detection probability of the strongest DFT beam; see Fig. 5.5(a). To improve AoA estimation accuracy in low SNR regions, we need to enhance the detection probability.

Fig. 5.7 presents the performance of Algorithm 8 with regards to the system bandwidth, where $K = 4$. We see that the detection probability remains around 1 as the bandwidth B changes. We see that the probability of a single symbol being sufficient for Algorithm 1 to correctly identify the strongest DFT beams grows with B. The simulated results are consistent with the analytical results. This again validates the applicability of Lemma 5.2 and Theorem 5.3. Moreover, the probability of using a single symbol can be as large as 0.9744 under $B = 0.6 f_{\mathrm{H}}$. This indicates that the proposed algorithm has the potential to be applied to ultra-wideband systems.

Fig. 5.8 compares the MSE of the LoS AoA estimation between the proposed method and the existing work [7] in the presence of two NLoS paths, where $B = 0.4$ and $M = 256$. We see that the proposed method achieves the substantially improved robustness against the NLoS paths, as compared to the method [7]. Specifically, to

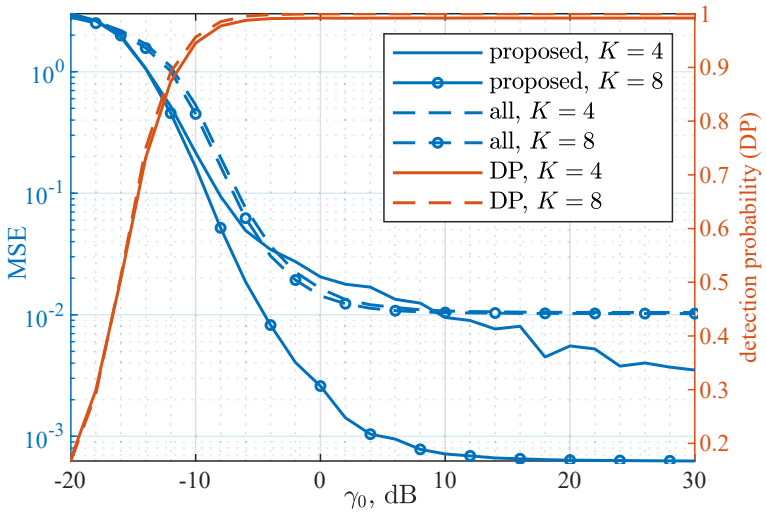

Figure 5.8 : left: MSE of LoS AoA estimate against γ_0 in the presence of two NLoS paths; right: the DP of the strongest DFT beam for the LoS path. The existing work [7], referred to as "all", is provided as the benchmark.

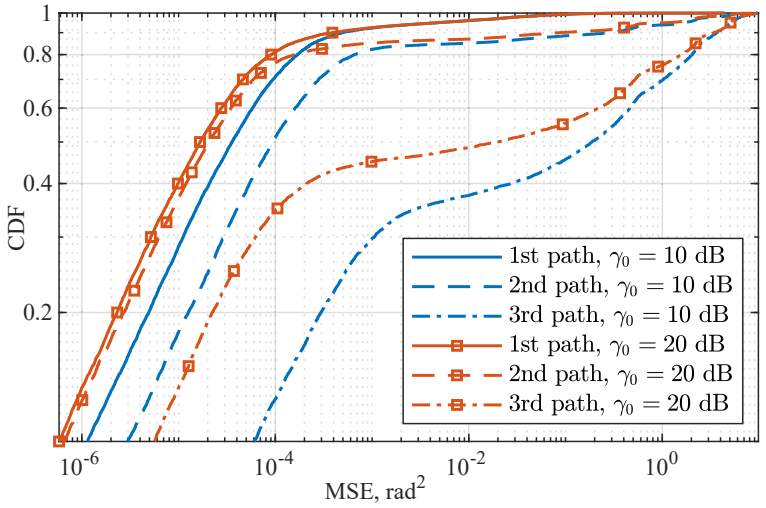

Figure 5.9 : CDF of the squared error of the AoA estimates obtained by running the proposed algorithm in conjunction with SIC.

achieve the MSE of 0.01 rad^2, the proposed method can operate at a 14 dB lower estimation SNR than the existing technique [21]. Moreover, the proposed method can reduce the MSE of the AoA estimation for the LoS path by orders of magnitude in high SNR regions (e.g., $\gamma_0 \geq 10$ dB). This is because the proposed Algorithm 9 only selects the strongest DFT beams, and can suppress the NLoS paths with the

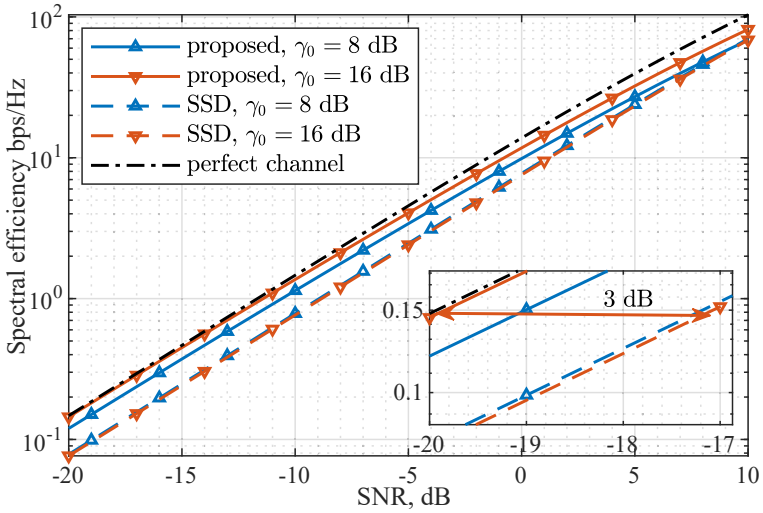

Figure 5.10 : Spectral efficiency comparison between the proposed method and the state of the art [8], referred to as "SSD", where $K = 8$.

asymptotic probability $(1 - \frac{K}{N})^2$ as $N \to \infty$.[‖] On the other hand, the superiority of the proposed method to the existing work [7] is also due to the resilience of Algorithm 8 against the NLoS paths. This can be confirmed by jointly inspecting the detection probabilities in Figs. 5.5(b) and 5.8.

Fig. 10 plots the performance of the proposed algorithm in the presence of multipath, where three paths, an LoS path and two NLoS paths, are considered and the cumulative density functions (CDFs) of the squared estimation errors of their AoAs are plotted. The estimation SNR is set to be $\gamma_0 = 10$ or 20 dB for the LoS path. With reference to [8], the two NLoS paths are are set to be 5 and 10 dB weaker than the LoS, respectively. The AoAs of the three paths are independently and randomly uniformly distributed within $[0, \pi]$.

The proposed AoA estimation method can operate in coupling with serial iterative cancellation (SIC), especially in a typical multipath mmWave scenario where

[‖]The probability that the NLoS paths are not in the K DFT beams selected in Step 2, Algorithm 9 asymptotically approaches the probability that the NLoS paths are incident to any of the other $(N - K)$ DFT beams, as $N \to \infty$. The asymptotic condition $N \to \infty$ guarantees no leak of the signal received by any DFT beam. Given $u_p \sim \mathcal{U}[0, 2\pi]$, $\forall p = 1, 2$, the probability that the p-th NLoS path is incident to one of the $(N-K)$ DFT beams is $\frac{N-K}{N}$; moreover, the probability becomes $(\frac{N-K}{N})^2$ in the presence of two independent NLoS paths. Based on the simulation configurations, the probability is 0.969 and 0.9385 for $K = 4$ and 8, respectively.

an NLoS path is much weaker than the LoS path, e.g., lower by 5 dB or more [1, 2]. One after another from the strongest path to the weakest, the multiple AoAs can be estimated by using the proposed algorithm and the associated paths can be accordingly estimated and canceled from the received signals. Specifically, after obtaining the AoA, we can pick up the received signal of the corresponding DFT beams (with the beam selection matrix specified in Lemma 1). Given the estimated AoA of a path, the complex gain of the path, β_p in (4), can be estimated by reusing the same received symbol with which the AoA has been estimated. As the spatial responses of the DFT beams in the direction of the AoA can be calculated, β_p can be estimated by using an LS estimator [13, 14].

As we can see in Fig. 10, the proposed algorithm in conjunction with SIC is effective in the presence of multipath components. The LoS path can achieve much better estimation accuracy than the NLoS paths. This is due to the weaker signal strengths of the NLoS paths and the error propagation which is typical to SIC.

Last but not least, we provide the spectral efficiency achieved by the proposed method and the state of the art [8], referred to as "SSD". As in [8], the transmission scheme in [82] is adopted to evaluate the spectral efficiency based on the AoA estimated in Fig. 5.8. The beam selection matrix can be configured to activate multiple DFT beams capturing the estimated AoAs in their mainlobes to receive the signals, as designed in Lemma 1.

Fig. 11 plots the special efficiency, under the condition that the SNR of the uplink AoA estimation $\gamma_0 = 8$ or 16 dB (and the AoAs can be reasonably accurately estimated), while the SNR of the downlink transmission varies from -20 dB to 10 dB. The condition is consistent with [6, Figs. 8&9] and [8, Fig. 4], and adopted for the purpose of fair comparisons with [8]. Each curve in Fig. 11 has a consistent uplink SNR for AoA (and channel) estimation, i.e., $\gamma_0 = 8$ or 16 dB, and hence a consistent AoA estimation accuracy. Given the AoA estimation, the curve displays the growth in spectral efficiency with the increasing downlink SNR. The spectral efficiency based on the perfect channel is also provided as an upper bound. We see that the proposed method outperforms SSD [8] markedly, especially in low SNR

regions. The spectral efficiency can be improved by up to 89.04% at -20 dB SNR. This is due to the accurate AoA estimation achieved by the proposed method. By jointly assessing Figs. 5.8 and 5.10, we can see that the better AoA estimation can result in the larger spectral efficiency, confirming the value of the accurate AoA estimation to the wideband LAA-based massive MIMO systems.

As shown in Fig. 11, a curve under $\gamma_0 = 16$ dB is above its counterpart under $\gamma_0 = 8$ dB, indicating that channel estimation is worse in a low region of the uplink SNR used for AoA estimation, than in a high region of the SNR. The spectral efficiency of the proposed approach (under $\gamma_0 = 16$ dB) is indistinguishably close to that under the perfect channel condition, when the downlink SNR is low. This is because the downlink SNR is too low and contributes negligibly (by Shannon's Law).

We notice in Fig. 11 that spectral efficiency is close to zero, with and without the perfect channel conditions. The spectral efficiency of the proposed approach deviates from that under the perfect channel, when the downlink SNR is high. The reason is that the received SNR in the downlink has a dominating effect on the spectral efficiency, when it is high (by Shannon's Law as well). The channel estimation error of the proposed approach under $\gamma_0 = 16$ dB can lead to a consistent gap (in dB) between its spectral efficiency and the spectral efficiency under the perfect channel condition, as shown at the right ends of the curves.

5.7 Conclusion

In this chapter, we propose a fast and accurate approach for the AoA estimation in wideband LAAs. This is achieved by unprecedentedly exploiting (rather than circumventing) the spatial-wideband effect. New spatial-frequency pattern are uncovered to establish one-to-one mappings between the strongest DFT beam containing the AoA of interest and the patterns constructed judiciously based on selected sub-carriers. As a result, a new algorithm is designed to identify the strongest DFT beam within a single symbol, set up the relation between the AoA and the strongest DFT beam, and accordingly estimate the AoA uniquely and accurately. Closed-

form expressions are derived to analyze the accuracy of the algorithm. Evident from simulations, the proposed approach is able to improve the estimation accuracy and spectral efficiency substantially using a much smaller number of symbols, as compared to the state of the art.

Chapter 6

Conclusions and Future Works

In this thesis, the AoA estimation in popular mmWave hybrid arrays is studied. Four approaches are proposed for different types of hybrid arrays, which are accomplished by new discoveries, derivations and analyses. To be concluded, the following has been accomplished.

1. For narrowband hybrid array of phased subarrays, we propose high-accuracy AoA estimation with new subarray-specific time-varying phase shifts in general localized hybrid arrays. We discover that the signs of the cross-correlations between consecutive subarrays are deterministic, and only the strongest cross-correlation takes a different sign from the rest. We propose to align the cross-correlations, achieving constructive combinations and improving tolerance to noises. Evident from extensive simulations, the estimation accuracy can be substantially improved by orders of magnitude through our design of phase shifts and our discovery, and asymptotically approach the MSELB.

2. For wideband hybrid array of phased subarrays, we propose new subarray-specific time-varying phase shifts, which enable the cross-correlations of receive signal between subarrays to have deterministic phase changes, and hence be coherently accumulated to deliver unambiguous and noise-tolerant estimation of Nu and u. We also optimize the frequency interval for coherent accumulation across sub-carriers, minimizing the estimation error of Nu and leveraging the accumulation gain. Evident from simulations, our approach is able to dramatically improve the estimation accuracy by orders of magnitudes, with significantly reduced requirements of complexities and training symbols, as well as enhanced robustness against noises.

3. For narrowband LAAs, we propose a fast and accurate approach for the AoA estimation. We prove that the AoA of a path can be accurately estimated from the two adjacent DBDs at both sides of the AoA. The two DBDs are proved to have larger amplitude gains than other DBDs and DFT beams, and can be more reliably identified in the presence of non-negligible noises. We design a novel beam synthesis method based on discrete lens beams and 1-bit phase shifts. As a result, the angular region containing the two strongest DBDs can be exponentially narrowed down, and the two DBDs can be quickly identified. Evident from extensive simulations, the proposed approach is able to improve estimation accuracy and the time-average beamforming gain by orders of magnitude, as compared to the state of the art.

4. For wideband LAAs, we propose a fast and accurate approach for the AoA estimation. This is achieved by unprecedentedly exploiting (rather than circumventing) the spatial-wideband effect. New spatial-frequency pattern are uncovered to establish one-to-one mappings between the strongest DFT beam containing the AoA of interest and the patterns constructed judiciously based on selected sub-carriers. As a result, a new algorithm is designed to identify the strongest DFT beam within a single symbol, set up the relation between the AoA and the strongest DFT beam, and accordingly estimate the AoA uniquely and accurately. Closed-form expressions are derived to analyze the accuracy of the algorithm. Evident from simulations, the proposed approach is able to improve the estimation accuracy and spectral efficiency substantially using a much smaller number of symbols, as compared to the state of the art.

The following future works are expected to further validate the proposed in real-life scenarios, and to improve the proposed approaches in more complicated scenarios.

1. The proposed approach can have the potential to operate in the presence of NLoS, despite LoS overwhelmingly dominates many mmWave bands [1, 3, 4, 15–17]. For instance, the difference of SNR between LoS and NLoS paths is

more than 5 dB on average in the 28, 38, 60, and 73 GHz mmWave bands [91], and 20 dB in the sub-mmWave bands, e.g., Tera Hertz [92]. In the case that NLoS is much weaker than LoS, e.g., by 5 dB or more, as typically is in practice [91], SIC can be conducted in AoA estimation [21, 22]. Sequentially, we can run the proposed algorithms (i.e., Algorithms 1 and 2) to estimate the AoA of one of the paths, $u_i = 2\pi d \sin(\theta_i)/\lambda_c$ (u_i, or more precisely, θ_i, is the AoA of the i-th strongest path); see (2.7), steer the beams of all subarrays towards the path to measure its strength and phase, regenerate and subtract the signal component of the path from the received signals at all subarrays. One path after another, all paths can be estimated and subtracted. We note that, as part of the inputs to SIC, the AoAs of the estimated paths take sinusoidal formats, i.e., $u_i = 2\pi d \sin(\theta_i)/\lambda_c$. The sinusoidal formats result from the array responses (or in other words, the geometry of the arrays). The pilot signal can take a different waveform. By conducting SIC, we can successively reconstruct and cancel the received pilot signals along the estimated paths based on the waveform and the estimates of the AoAs, and then estimate the AoAs of the other paths. In the case that the multiple paths are of similar strengths, which is unlikely though in practice for mmWave, the proposed approach could be carried out in coupling with parallel interference cancellation (PIC). The only difference from SIC is that the multiple AoAs would be estimated and canceled in parallel. Initial results of combining the proposed approaches with SIC have been achieved, e.g., in Figs. 2.7, 4.12, 4.11 and 5.9. More analysis is required on the SIC-based AoA estimation approaches, in terms of the convergence rate and performance.

2. Butler matrix, as a promising massive MIMO transceiver front-end, has increasingly attracted interest, due to its high gain, low complexity and low cost, as compared to conventional DAAs and LAAs [2, 105, 106]. Similar to a LAA, a Butler matrix is an integrated, passive beamforming network with $N_{\rm a}$ evenly-spaced antennas and $N_{\rm b}$ feeders (also referred to as "beam ports"), where typically $N_{\rm a} = N_{\rm b} = 2^Z$ (Z is a nonnegative integer) [106]. A Butler ma-

trix is more energy-efficient than a conventional DAA. An N_{a}-dimensional Butler matrix can readily produce N_{a} Butler beams [2], while an N_{a}-dimensional linear DAA requires N_{a}^2 number of $\log_2^{N_{\mathrm{a}}}$-bit phase shifters to generate N_{a} DFT beams [106]. The Butler matrix can also provide effective spatial interference suppression, and separate signals with different angle-of-arrivals (AoAs) by exploiting different mainlobes of the Butler beams [105, 107].* An array of Butler matrices can readily form a massive MIMO transceiver. This can help by either reducing the numbers of antennas and beam ports of each Butler matrix, or improving the array gain, as compared to a single large-dimensional Butler matrix [2, 105]. The reduced number of antennas and beam ports can also reduce cross-overs and mutual couplings [106, 107]. To fulfill the expected benefits of Butler matrix arrays (BMAs), accurate estimations of AoAs of incident paths are the key. BMAs still face the key challenge which DAAs face in AoA estimation. Specifically, the received signals of the antennas of an analog subarray (or a Butler matrix) are added up (via RF combining) before being fed into an RF chain. The key information to estimate the AoA, which is the phase offset between adjacent antenna elements, becomes obscure. There is no existing research on the AoA estimation for BMAs. For future works, we would explore the feasibility of applying the proposed approaches onto Butler matrices, or develop new methods based on the proposed ones in this thesis.

*It is noteworthy that Butler matrices incur lower implementation complexity and cost, and can cover wider angular region and generate more orthogonal beams, as compared to LAAs [106].

Chapter 7

Appendices

7.1 Proof, Analysis and Derivation in Chapter 2

7.1.1 Proof of Lemma 2.1

Proof. We start by proving the existing of m' satisfying $\sin(\omega_{m'}^t)\sin(\omega_{m'+1}^t) < 0$. To prove this, we hypothesize that $\sin(\omega_{\tilde{m}}^t) = 0$ for a given $\tilde{m} \in [0, M-1]$, so that m' does not exist. Then $\sin(N\omega_{\tilde{m}}^t) = 0$, since $\omega_{\tilde{m}}^t \in \{2a\pi, 2a\pi \pm \pi, a \in \mathbb{N}\}$. For $m = 0, 1, \cdots, K-1$, we have

$$\begin{aligned}\sin(N\omega_m^t) &= \sin(N\omega_{\tilde{m}+m-\tilde{m}}^t) \\ &= \sin(N\omega_{\tilde{m}}^t - N\pi(m-\tilde{m})/K) \\ &= (-1)^{(m-\tilde{m})Q}\sin(\omega_{\tilde{m}}^t) = 0.\end{aligned}$$

From (2.13), $G_m^t(u) = 0$, $\forall m \in [0, K-1]$. In other words, if $\sin(\omega_m^t)$ intersects with the x-axis at any integer point, all the cross-correlations of the received signals between consecutive subarrays would be zero. This contradicts with the condition of Theorem 2.1 that $G_{m'}^t(u) \neq 0$, and therefore, $\tilde{m}$ does not exist. The existence of m' satisfying $\sin(\omega_{m'}^t)\sin(\omega_{m'+1}^t) < 0$.

The uniqueness of m' is confirmed since $\sin(\omega_m^t)$ $(0 \leq k \leq K-1)$ only covers half a cycle of a sine function. □

7.1.2 Proof of Lemma 2.2

Proof. ω_m^t is a function of u and can be written as $\omega_m^t(u)$. Given m_1 and $m_2(= m_1 + 1)$, satisfying $\sin(\omega_{m_1+1}^t(u)) = \sin(\omega_{m_2}^t(u)) = 0$, we have $|\sin(\omega_{m_1}^t(u))| = |\sin(\omega_{m_2+1}^t(u))|$ according to Lemma 7.1 in Appendix 7.1.3. Without loss of generality, we assume that $\sin(\omega_{m_1}^t) > 0$ and $\sin(\omega_{m_2+1}^t) < 0$. An adequate $\delta_u > 0$ can

be selected to make $\sin(\omega_{m_2}^t(u+\delta_u)) > 0$ and $\sin(\omega_{m_2+1}^t(u+\delta_u)) < 0$. As a result, $|\sin(\omega_{m_1}^t(u+\delta_u))| > |\sin(\omega_{m_1}^t(u))|$ and $|\sin(\omega_{m_2+1}^t(u))| > |\sin(\omega_{m_2+1}^t(u+\delta_u))|$. Subsequently, we have

$$\begin{aligned} |\sin(\omega_{m_1}^t(u+\delta_u))| &> |\sin(\omega_{m_1}^t(u))| \\ &= |\sin(\omega_{m_2+1}^t(u))| \\ &> |\sin(\omega_{m_2+1}^t(u+\delta_u))|. \end{aligned} \tag{7.1}$$

Since $m_2 = m_1 + 1$, we have $|\sin(\omega_{m_1+1}^t(u+\delta_u))| = |\sin(\omega_{m_2}^t(u+\delta_u))|$, which, combined with (7.1), leads to

$$\begin{aligned} &|\sin(\omega_{m_1}^t(u+\delta_u))\sin(\omega_{m_1+1}^t(u+\delta_u))| \\ &\quad > |\sin(\omega_{m_2+1}^t(u+\delta_u))\sin(\omega_{m_2}^t(u+\delta_u))|. \end{aligned} \tag{7.2}$$

Let $m_2 = m'$, we obtain that $|\sin(\omega_{m'}^t)\sin(\omega_{m'-1}^t)| > |\sin(\omega_{m'+1}^t)\sin(\omega_{m'}^t)|$.

Likewise, we can prove that $|\sin(\omega_{m'+1}^t)\sin(\omega_{m'+2}^t)| > |\sin(\omega_{m'+1}^t)\sin(\omega_{m'}^t)|$, by selecting an adequate $\delta_u < 0$.

From Lemma 2.1, m' is unique. Since $\sin(\omega_m^t)$, $0 \le m \le K-1$, only covers half a cycle of a sine function, we confirm that $|\sin(\omega_{m'}^t)\sin(\omega_{m'+1}^t)|$ has the smallest amplitude. This concludes this proof. □

7.1.3 Lemma 7.1 and its proof

Lemma 7.1. *If there exist* m_1 *and* $m_2 \in [0, K-1]$, *satisfying* $\sin(\omega_{m_1+1}^t) = \sin(\omega_{m_2}^t) = 0$ *and* $m_2 = m_1 + 1$, *then* $|\sin(\omega_{m_1}^t)| = |\sin(\omega_{m_2+1}^t)|$.

Proof. If $\sin(\omega_{m_1+1}^t) = 0$ and $\sin(\omega_{m_2}^t) = 0$, then $|\cos(\omega_{m_1+1}^t)| = |\cos(\omega_{m_2}^t)| = 1$. Since $2\omega_{m+1}^t = 2\omega_m^t - 2\pi/K$, it is easy to obtain that

$$\begin{aligned} |\sin(\omega_{m_1}^t)| &= |\sin(\omega_{m_1+1}^t + \pi/K)| \\ &= |\sin(\omega_{m_1+1}^t)\cos(\pi/K) + \cos(\omega_{m_1+1}^t)\sin(\pi/K)| \\ &= |\sin(\pi/K)|. \end{aligned} \tag{7.3}$$

Likewise, $|\sin(\omega^t_{m_2+1})| = |\sin(\omega^t_{m_2} - \pi/K)| = |\sin(\pi/K)|$. Therefore, Lemma 7.1 is proved. □

7.1.4 Derivation of $\mathbb{E}_u\{|P^t_m(u)|^2\}$

Given that u is uniformly distributed in $[-\pi, \pi]$ and $P^t_m(u)$ in (2.11), we have $\mathbb{E}_u\{|P^t_m(u)|^2\} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin(N\omega^t_m)}{\sin(\omega^t_m)}\right|^2 \mathrm{d}u$. The integral on the RHS can be calculated by using the property of Fourier transform. Specifically, we can define a rectangular window function as $r_N(n) = 1,\ \forall n \in [0, N-1]$. The discrete-time Fourier transform (DTFT) of $r_N(mK)$, termed $f(u)$, can be given by $f(u) = \sum_{n=0}^{N-1} r_N(n)e^{-jun} = e^{j\frac{N-1}{2}u}\frac{\sin(Nu/2)}{\sin(u/2)}$. Exploiting the frequency-shift property of DTFT [95], the inverse DTFT of $P^t_m(u)$ is the phase shifted version of $r_N(n)$, i.e., $r_N(n)e^{j(\frac{2\pi tn}{L}+\frac{2m\pi n}{K})}$. This is because $P^t_m(u) = f(u - \frac{2\pi t}{L} - \frac{2m\pi}{K})$ and is a shifted version of $f(u)$ in the (angular) frequency domain. As a result, we have $\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|f\left(u - \frac{2\pi t}{L} - \frac{2\pi m}{K}\right)\right|^2 \mathrm{d}u = \sum_{n=-\infty}^{\infty}|r_N(n)|^2 = N$, based on Parseval's theorem [95]. Therefore, we prove that $\mathbb{E}_u\{|P^t_m(u)|^2\} = N$ for $m = 0, 1, \cdots, M-1$ and $t = 0, 1, \cdots, P-1$.

7.1.5 Saturation Analysis of Fig. 2.7

The proposed operations of AoA estimation (or more specifically, Algorithm 4), such as the IDFT of all subarrays, the cross-correlations of consecutive IDFT points, and the coherent accumulation of the cross-correlations, can have different impacts on the suppression of the noise and the NLoS interference. In the presence of a LoS and a NLoS paths, denoted respectively by $\tilde{s}(t)$ and $\tilde{i}(t)$, the cross-correlation between every K consecutive IDFT points in Algorithm 2 can be given in (7.4), where u_s and u_i capture the AoAs of the LoS and NLoS paths, respectively, $a_{n'}(t) = \tilde{s}(t)g^t_{n'}(u_s) + \tilde{i}(t)g^t_{n'}(u_i) + z_{n'}(t)$, $n' = 0, 1, \cdots, K-1$, stands for the Fourier coefficients of the outputs of K consecutive subarrays, $g^t_{n'}(u)$ and $C(u)$ are given in (2.22) and (2.24), respectively.

The first and second terms on the RHS of (7.4), denoted by $f(t)$, provide the effective signal with a deterministic phase which is evaluated to estimate the AoA in Algorithm 2, for any given u_s, u_i, and the input per-antenna SIR, $\mathbb{E}_{\tilde{s},\tilde{i}}\left\{\frac{|\tilde{s}(t)|^2}{|\tilde{i}(t)|^2}\right\}$, of 5dB in Fig. 2.7. $\mathbb{E}_{\tilde{s},\tilde{i}}\{\cdot\}$ takes expectations over $\tilde{s}(t)$ and $\tilde{i}(t)$. The estimated AoA

$$r_{n'}(t) = a^*_{n'}(t)a_{n'+1}(t) = \underbrace{e^{j(u_s-\frac{2\pi t}{L})}\left|C(u_s)\right|^2\left|\tilde{s}(t)\right|^2 + e^{j(u_i-\frac{2\pi t}{L})}\left|C(u_i)\right|^2\left|\tilde{i}(t)\right|^2}_{f(t)}$$
$$+ \underbrace{\tilde{s}^*(t)g^{t*}_{n'}(u_s)\tilde{i}(t)g^t_{n'+1}(u_i) + \tilde{s}(t)g^t_{n'+1}(u_s)\tilde{i}^*(t)g^{t*}_{n'}(u_i)}_{I(t)} + \tilde{z}_{n'}(t). \tag{7.4}$$

$$\begin{aligned}\tilde{z}_{n'}(t) =& z^*_{n'}(t)z_{n'+1}(t)\\ &+ \underbrace{\tilde{s}(t)C(u_s)e^{j(n'+1)(u_s-\frac{2\pi t}{L})}z^*_{n'}(t) + \tilde{s}(t)^*C(u_s)^*e^{-jn'(u_s-\frac{2\pi t}{L})}z_{n'+1}(t)}_{\text{denoted by } \tilde{z}_{\tilde{s}}}\\ &+ \underbrace{\tilde{i}(t)C(u_i)e^{j(n'+1)(u_i-\frac{2\pi t}{L})}z^*_{n'}(t) + \tilde{i}(t)^*C(u_i)^*e^{-jn'(u_i-\frac{2\pi t}{L})}z_{n'+1}(t)}_{\text{denoted by } \tilde{z}_{\tilde{i}}}.\end{aligned} \tag{7.5}$$

$$\begin{aligned}\sigma^2_{I(t)} =& \left\lceil\frac{M}{K}\right\rceil^2\\ &\times \mathbb{E}_{\tilde{s},\tilde{i},u_s,u_i}\left\{\left(\sum_{t=0}^{P-1}\sum_{n'=0}^{K-2}\left[\tilde{s}^*(t)g^{t*}_{n'}(u_s)\tilde{i}(t)g^t_{n'+1}(u_i) + \tilde{s}(t)g^t_{n'+1}(u_s)\tilde{i}^*(t)g^{t*}_{n'}(u_i)\right]e^{\frac{j2\pi t}{L}}\right)^2\right\}\\ &\cong 44.787 \text{ Watts.}\end{aligned} \tag{7.6}$$

can incur a noise-independent bias resulting from $\tilde{i}(t)$. The third and fourth terms on the RHS of (7.4), denoted by $I(t)$, capture the cross-correlations between the LoS and NLoS paths, and bear randomness resulting from the random instantaneous phases and amplitudes of $\tilde{s}(t)$ and $\tilde{i}(t)$. In Fig. 2.7, uniformly distributed phases and Rayleigh distributed amplitudes are taken into account. $I(t)$ can be interpreted as an additional noise source. The last noise term on the RHS of (7.4) can be written as (7.5).

From (7.4), we can see three separate causes of error to the estimation of u_s, namely, the bias from the estimation of the phase of $f(t)$, the randomness-bearing interference $I(t)$, and the noise $\tilde{z}_{n'}(t)$. $f(t)$ is independent of $\tilde{z}_{n'}(t)$ and has a different impact on the estimation of u from $I(t)$ and $\tilde{z}_{n'}(t)$, while both $I(t)$ and $\tilde{z}_{n'}(t)$ act as additive noises. We can study the saturation behavior of the AoA estimation with respect to γ_0 by comparing the coherently accumulated powers of $I(t)$ and $\tilde{z}_{n'}(t)$,

denoted by $\sigma^2_{I(t)}$ and $\sigma^2_{\tilde{z}}$, respectively. The threshold of γ_0, which leads to $\sigma^2_{I(t)} \geq \sigma^2_{\tilde{z}}$ is evaluated. The estimation accuracy of u_s starts to saturate beyond the threshold, since $\tilde{z}_{n'}(t)$ starts to be, and increasingly is, dominated by $I(t)$.

Consider the settings of Fig. 2.7, i.e., u_s and u_i are uniformly distributed on $\left[\frac{\pi}{4}, \frac{\pi}{3}\right]$ and $\left[-\frac{\pi}{3}, -\frac{\pi}{4}\right]$, respectively; $N = 8$; and $M = P = K = 4$. $\sigma^2_{I(t)}$ can be given in (7.6), where $\mathbb{E}_{\tilde{s},\tilde{i},u_s,u_i}\{\cdot\}$ takes expectations over all randomness, i.e., $\tilde{s}(t), \tilde{i}(t), u_s$ and u_i. The gain of coherent accumulation is accounted for by the coefficient $\lceil \frac{M}{K} \rceil$ and the summations over t and n'. The result is achieved numerically.

On the other hand, $\sigma^2_{\tilde{z}}$ can be attained by separately evaluating the accumulated powers of $\tilde{z}_{\tilde{s}}$ and $\tilde{z}_{\tilde{i}}$ in (7.5), denoted by $\sigma^2_{\tilde{z}_{\tilde{s}}}$ and $\sigma^2_{\tilde{z}_{\tilde{i}}}$, respectively. By referring to (2.40), the ratio of the coherently accumulated signal power of $\lceil \frac{M}{K} \rceil (K-1) \sum_{t=0}^{P-1} |C(u_s)|^2 |\tilde{s}(t)|^2 e^{ju_s}$ to $\sigma^2_{\tilde{z}_{\tilde{s}}}$ can be approximated to $\frac{P(K-1)\lceil M/K \rceil}{2}\gamma_0 = 6\gamma_0$, while the coherently accumulated signal power can be given by

$$\mathbb{E}_{\tilde{s},u_s}\left\{\left(\left\lceil\frac{M}{K}\right\rceil (K-1)|\tilde{s}(t)|^2 \sum_{t=0}^{P-1} |C(u_s)|^2\right)^2\right\}$$
$$= \left\lceil\frac{M}{K}\right\rceil^2 (K-1)^2 \times \underbrace{\mathbb{E}_{\tilde{s}}\{|\tilde{s}(t)|^4\}}_{8\sigma_s^4} \times \underbrace{\mathbb{E}_{u_s}\left\{\left(\sum_{t=0}^{P-1} |C(u_s)|^2\right)^2\right\}}_{\frac{12}{\pi}\int_{\frac{\pi}{4}}^{\frac{\pi}{3}} \left(\sum_{t=0}^{P-1}|C(u_s)|^2\right)^2 \mathrm{d}u_s = 64} = 4608\sigma_s^4. \quad (7.7)$$

Here, σ_s is the scale parameter of the Rayleigh distributed LoS path. $\mathbb{E}_{\tilde{s}}\{|\tilde{s}(t)|^4\} = 8\sigma_s^4$ is the fourth raw moment of the Rayleigh distributed channel gain, i.e., $|\tilde{s}(t)|$, of the LoS path. Therefore, $\sigma^2_{\tilde{z}_{\tilde{s}}} = \frac{4608\sigma_s^4}{6\gamma_0}$.

Likewise, we have $\sigma^2_{\tilde{z}_{\tilde{i}}} = \frac{4608\sigma_i^4}{6\gamma_0 \frac{\sigma_i^2}{\sigma_s^2}}$, where $\gamma_0 \frac{\sigma_i^2}{\sigma_s^2}$ gives the per-antenna interference-to-noise ratio at each individual antenna element and σ_i is the scale parameter of the Rayleigh distributed NLoS path.

As discussed, the threshold of γ_0, beyond which the estimation accuracy of u_s starts to saturate, can be achieved by solving $\sigma^2_{\tilde{z}_{\tilde{s}}} + \sigma^2_{\tilde{z}_{\tilde{i}}} \leq \sigma^2_{I(t)}$, which, with $\sigma_s^2 = 10^{\frac{1}{2}}$

Watts and $\sigma_i^2 = 1$ Watts in Fig. 2.7, leads to

$$\frac{4608\sigma_s^4}{6\gamma_0} + \frac{4608\sigma_i^4}{6\gamma_0 \frac{\sigma_i^2}{\sigma_s^2}} \leq 44.787 \Rightarrow \gamma_0 \geq 23.5354 \text{ dB}. \tag{7.8}$$

This is consistent with Fig. 2.7, where the estimation accuracy of "LoS with NLoS" starts to saturate over 23 dB.

7.2 Proof and Derivation in Chapter 3

7.2.1 Proof of Lemma 3.1

Proof. We start by proving the existing of m' satisfying $\sin(\omega_{m'}^t)\sin(\omega_{m'+1}^t) < 0$. To prove this, we assume that $\sin(\omega_{\tilde{m}}^t) = 0$ for a given $\tilde{m} \in [0, M-1]$, so that m' does not exist. Then $\sin(N\omega_{\tilde{m}}^t) = 0$, since $\omega_{\tilde{m}}^t \in \{2a\pi, 2a\pi \pm \pi, a \in \mathbb{N}\}$. For $m = 0, 1, \cdots, K-1$, we have

$$\begin{aligned}\sin(N\omega_m^i) &= \sin(N\omega_{\tilde{m}+m-\tilde{m}}^t) = \sin(N\omega_{\tilde{m}}^t - N\pi(m-\tilde{m})/K) \\ &= (-1)^{(m-\tilde{m})Q}\sin(\omega_{\tilde{m}}^t) = 0.\end{aligned}$$

From (3.11), $G_m(\rho_i, u) = 0$, $\forall m \in [0, K-1]$. In other words, if $\sin(\omega_m^i)$ intersects with the x-axis at any integer point, all the cross-correlations of the receive signals between consecutive subarrays would be zero. This contradicts with the condition of Theorem 3.1 that $G_{m'}^t(u) \neq 0$, and therefore, $\tilde{m}$ does not exist. The existence of m' satisfies $\sin(\omega_{m'}^t)\sin(\omega_{m'+1}^t) < 0$.

The uniqueness of m' can be readily confirmed since $\sin(\omega_m^i)$ $(0 \leq k \leq K-1)$ only covers half a cycle of a sine function. □

7.2.2 Proof of Lemma 3.2

Proof. ω_m^t is a function of u and can be written as $\omega_m^t(u)$. Given m_1 and $m_2(= m_1 + 1)$, satisfying $\sin(\omega_{m_1+1}^t(u)) = \sin(\omega_{m_2}^t(u)) = 0$, we have $|\sin(\omega_{m_1}^t(u))| = |\sin(\omega_{m_2+1}^t(u))|$ according to Lemma 7.2 in Appendix 7.2.3. Without loss of generality, we assume that $\sin(\omega_{m_1}^t) > 0$ and $\sin(\omega_{m_2+1}^t) < 0$. An adequate $\delta_u > 0$ can

be selected to ensure $\sin(\omega_{m_2}^t(u+\delta_u)) > 0$ and $\sin(\omega_{m_2+1}^t(u+\delta_u)) < 0$. As a result, $|\sin(\omega_{m_1}^t(u+\delta_u))| > |\sin(\omega_{m_1}^t(u))|$ and $|\sin(\omega_{m_2+1}^t(u))| > |\sin(\omega_{m_2+1}^t(u+\delta_u))|$. Subsequently, we have $|\sin(\omega_{m_1}^t(u+\delta_u))| > |\sin(\omega_{m_1}^t(u))| = |\sin(\omega_{m_2+1}^t(u))| > |\sin(\omega_{m_2+1}^t(u+\delta_u))|$. Since $m_2 = m_1+1$, we have $|\sin(\omega_{m_1+1}^t(u+\delta_u))| = |\sin(\omega_{m_2}^t(u+\delta_u))|$, which leads to

$$|\sin(\omega_{m_1}^t(u+\delta_u))\sin(\omega_{m_1+1}^t(u+\delta_u))| > |\sin(\omega_{m_2+1}^t(u+\delta_u))\sin(\omega_{m_2}^t(u+\delta_u))|. \tag{7.9}$$

Let $m_2 = m'$, we obtain that $|\sin(\omega_{m'}^i)\sin(\omega_{m'-1}^t)| > |\sin(\omega_{m'+1}^i)\sin(\omega_{m'}^i)|$.

Likewise, we can prove that $|\sin(\omega_{m'+1}^i)\sin(\omega_{m'+2}^t)| > |\sin(\omega_{m'+1}^i)\sin(\omega_{m'}^i)|$, by selecting an adequate $\delta_u < 0$.

From Lemma 3.1, m' is unique. Since $\sin(\omega_m^i)$, $0 \le m \le K-1$, only covers half a cycle of a sine function, we confirm that $|\sin(\omega_{m'}^i)\sin(\omega_{m'+1}^i)|$ has the smallest amplitude. This concludes this proof. □

7.2.3 Lemma 7.2 and its proof

Lemma 7.2. *If there exist* m_1 *and* $m_2 \in [0, K-1]$, *satisfying* $\sin(\omega_{m_1+1}^t) = \sin(\omega_{m_2}^t) = 0$ *and* $m_2 = m_1 + 1$, *then* $|\sin(\omega_{m_1}^t)| = |\sin(\omega_{m_2+1}^t)|$.

Proof. If $\sin(\omega_{m_1+1}^t) = 0$ and $\sin(\omega_{m_2}^t) = 0$, then $|\cos(\omega_{m_1+1}^t)| = |\cos(\omega_{m_2}^t)| = 1$. Since $2\omega_{m+1}^t = 2\omega_m^i - 2\pi/K$, it is easy to obtain that

$$\begin{aligned} |\sin(\omega_{m_1}^t)| &= |\sin(\omega_{m_1+1}^t + \pi/K)| \\ &= |\sin(\omega_{m_1+1}^t)\cos(\pi/K) + \cos(\omega_{m_1+1}^t)\sin(\pi/K)| \\ &= |\sin(\pi/K)|. \end{aligned} \tag{7.10}$$

Likewise, $|\sin(\omega_{m_2+1}^t)| = |\sin(\omega_{m_2}^t - \pi/K)| = |\sin(\pi/K)|$. Therefore, Lemma 7.2 is proved. □

7.2.4 Calculation of $\mathbb{E}_u\{H_m(\rho_i, u)\}$ $(= 2N)$

By substituting the definition of $H_m(\rho_i, u)$, we have

$$\mathbb{E}_u\{H_m(\rho_i, u)\} = \mathbb{E}_u\{|P_m(\rho_i, u)|^2\} + \mathbb{E}_u\{|P_{m+1}(\rho_i, u)|^2\}.$$

Given that u is uniformly distributed in $[-\pi, \pi]$ and $P_m(\rho_i, u)$ in (3.10), we have

$$\mathbb{E}_u\left\{|P_m(\rho_i, u)|^2\right\} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin(N\omega_m^i)}{\sin(\omega_m^i)}\right|^2 \mathrm{d}u.$$

The integral on the RHS can be calculated by using the property of Fourier transform. Specifically, we can define a rectangular window function as $r_N(n) = 1,\ \forall n \in [0, N-1]$. The discrete-time Fourier transform (DTFT) of $r_N(mK)$, termed $f(u)$, can be given by $f(u) = \sum_{n=0}^{N-1} r_N(n)e^{-jun} = e^{j\frac{N-1}{2}u}\frac{\sin(Nu/2)}{\sin(u/2)}$.

Exploiting the frequency-shift property of DTFT [95], the inverse DTFT of $P_m(\rho_i, u)$ is the phase shifted version of $r_N(n)$, i.e., $r_N(n)e^{j(\frac{2\pi tn}{L}+\frac{2m\pi n}{K})}$. This is because $P_m(\rho_i, u) = f(u - \frac{2\pi t}{L} - \frac{2m\pi}{K})$ at $\rho_i = 1$ and is a shifted version of $f(u)$ in the (angular) frequency domain. As a result, we have $\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|f\left(u - \frac{2\pi t}{L} - \frac{2\pi m}{K}\right)\right|^2 \mathrm{d}u = \sum_{n=-\infty}^{\infty}|r_N(n)|^2 = N$, based on Parseval's theorem [95] is applied. Therefore, we prove that $\mathbb{E}_u\left\{|P_m(\rho_i, u)|^2\right\} = N$ at $\rho_i = 1$.

On the other hand, ρ_i in (3.3) is typically close to 1. In the case of $B = 0.05f_c$, we have $\rho_i \in [0.9512, 1]$, according to (3.3). Finally, we have $\mathbb{E}_u\left\{|P_m(\rho_i, u)|^2\right\} = \mathbb{E}_u\left\{|P_{m+1}(\rho_i, u)|^2\right\} = N$ and $\mathbb{E}_u\{H_m(\rho_i, u)\} = 2N$.

7.2.5 Proof of Lemma 7.3

Lemma 7.3. For any sub-carrier i and random u, we have $\sum_{t=0}^{P-1}|C(\rho_i, u)|^2 = \frac{PN}{K}$, where $C(\rho_i, u)$ is given in (3.32).

Proof. Given (3.26) and (3.32), we can express $C(\rho_i, u)$ as matrix multiplication, i.e., $C(\rho_i, u) = \mathbf{Ax}$, where $\mathbf{A}$ is a $P \times Q$ matrix with the element at the t-th row and q-th column given by $\mathbf{A}(t, q) = e^{-j\frac{2\pi tq}{P}}$ $(0 \le t \le P-1,\ 0 \le q \le Q-1)$ and $\mathbf{x}$ is a $Q \times 1$ vector with the q-th element $e^{j\rho_i Kuq}$. Clearly, $\mathbf{A}$ is the first Q columns of

a $P \times P$ DFT matrix [95]; in other words, $\mathbf{A}$ has full column rank. Therefore, we have $|C(\rho_i, u)|^2 = \mathbf{x}^H \mathbf{A}^H \mathbf{A}\mathbf{x} = PQ = \frac{PN}{K}$. This concludes the proof. □

7.2.6 Calculation of $\mathbb{E}_u\{|C(\rho_i, u)|^2\}$

Given (3.32) and a uniform distribution of u over $[-\pi, \pi)$, we have

$$\mathbb{E}_u\left\{|C(\rho_i,u)|^2\right\} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|C(\rho_i,u)|^2\right|^2 \mathrm{d}u = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin(\frac{N\rho_i u}{2}-\frac{N\pi t}{L})}{\sin(\frac{K\rho_i u}{2}-\frac{K\pi t}{L})}\right|^2 \mathrm{d}u.$$

Like in Appendix 7.2.4, we can calculate the integral through the property of Fourier transform. Note that $\frac{\sin(\frac{N\rho_i u}{2}-\frac{N\pi t}{L})}{\sin(\frac{K\rho_i u}{2}-\frac{K\pi t}{L})}$ can be seen as the DTFT of a K-decimation of a rectangular window function, denoted by $r_N(mK)$. $r_N(mK) = 1$, if $m = 0, 1, \cdots, N/K-1$; or $r_N(mK) = 0$, otherwise [95]. By applying Parseval's theorem [95], we have $\mathbb{E}_u\{|C(\rho_i,u)|^2\} = \sum_{m=0}^{N/K-1}|r_N(mK)|^2 = \frac{N}{K}$.

7.3 Proof and Analysis in Chapter 4

7.3.1 Proof of Lemma 4.1

Proof. It is clear that $\Omega_n \cap \Omega_{n'} = \emptyset$ for $n' \neq n$, and $\Omega_0 \cup \Omega_1 \cup \cdots \Omega_{N-1} = [0, 2\pi)$. We first prove the lemma in the case of $u \in \Omega_n^{\mathrm{L}}$ by showing that $|g_{\mathrm{d}}(n,u)|$ monotonically decreases and $|g_{\mathrm{d}}(n+1,u)|$ monotonically increases with the growth of u, respectively; and the maximum and the minimum of $|g_{\mathrm{d}}(n,u)|$ and $|g_{\mathrm{d}}(n+1,u)|$ are taken at $u = \frac{2\pi(n+1)}{N}$. $\Omega_n^{\mathrm{L}} = \left[\frac{2\pi n}{N} + \frac{\pi}{N}, \frac{2\pi(n+1)}{N}\right)$ is the left half of Ω_n.

From (4.8), $|g_{\mathrm{d}}(n,u)| = -\frac{1}{N}\frac{\sin\frac{N}{2}\left(u-\frac{2\pi n}{N}\right)\sin\frac{\pi}{N}}{\sin\frac{1}{2}\left(u-\frac{2\pi n}{N}\right)\sin\frac{1}{2}\left(u-\frac{2\pi(n+1)}{N}\right)}$ for $u \in \Omega_n^{\mathrm{L}}$, where the negative sign is due to the fact that $\sin\frac{1}{2}\left(u-\frac{2\pi(n+1)}{N}\right) < 0$. The first-order derivative of $|g_{\mathrm{d}}(n,u)|$ with respect to (w.r.t) u is given by (7.11), from which we can obtain $\frac{\mathrm{d}|g_{\mathrm{d}}(n,u)|}{\mathrm{d}u} \leq 0$. The equality only takes at $u = \frac{2\pi n}{N} + \frac{\pi}{N}$. In other words, $|g_{\mathrm{d}}(n,u)|$ decreases monotonically with the growth of u. Likewise, we can prove that $\frac{\mathrm{d}|g_{\mathrm{d}}(n+1,u)|}{\mathrm{d}u} > 0$, i.e., $|g_{\mathrm{d}}(n+1,u)|$ monotonically increases with u for $u \in \Omega_n^{\mathrm{L}}$. In addition, we have $\left|g_{\mathrm{d}}\left(n, \frac{2\pi(n+1)}{N}\right)\right| = \left|g_{\mathrm{d}}\left(n+1, \frac{2\pi(n+1)}{N}\right)\right| = 1$ by substituting $u = \frac{2\pi(n+1)}{N}$ into (4.8). As a result, $|g_{\mathrm{d}}(n,u)| > 1 > |g_{\mathrm{d}}(n+1,u)|$ for $u \in \Omega_n^{\mathrm{L}}$.

$$\frac{\mathrm{d}|g_{\mathrm{d}}(n,u)|}{\mathrm{d}u} = -\frac{\frac{N}{2}\cos\frac{N}{2}\left(u-\frac{2\pi n}{N}\right)\sin\frac{1}{2}\left(u-\frac{2\pi n}{N}\right)\sin\frac{1}{2}\left(u-\frac{2\pi(n+1)}{N}\right) - \frac{1}{2}\sin\frac{N}{2}\left(u-\frac{2\pi n}{N}\right)\sin\left(u-\frac{2\pi(n+1)}{N}\right)}{\frac{1}{\sin\frac{\pi}{N}}\sin^2\frac{1}{2}\left(u-\frac{2\pi n}{N}\right)\sin^2\frac{1}{2}\left(u-\frac{2\pi(n+1)}{N}\right)} \tag{7.11}$$

On the other hand, by exploiting (4.8), we can write

$$\frac{|g_{\mathrm{d}}(n+1,u)|}{|g_{\mathrm{d}}(n',u)|} = \frac{\left|\sin\frac{1}{2}\left(u-\frac{2\pi n'}{N}\right)\sin\frac{1}{2}\left(u-\frac{2\pi(n'+1)}{N}\right)\right|}{\left|\sin\frac{1}{2}\left(u-\frac{2\pi(n+1)}{N}\right)\sin\frac{1}{2}\left(u-\frac{2\pi(n+2)}{N}\right)\right|} > 1,$$

since the numerator of the RHS is greater than the denominator for $u \in \Omega_n^{\mathrm{L}}$ and $n' \neq n, n+1$. Finally, we obtain that $|g_{\mathrm{d}}(n,u)| > |g_{\mathrm{d}}(n+1,u)| > |g_{\mathrm{d}}(n',u)|$ for $u \in \Omega_n^{\mathrm{L}}$.

The case of $u \in \Omega_n^{\mathrm{R}}$ can be proved in the same way and therefore is suppressed for brevity. $\Omega_n^{\mathrm{R}} = \left[\frac{2\pi(n+1)}{N}, \frac{2\pi(n+1)}{N} + \frac{\pi}{N}\right)$ is the right half of Ω_n. Therefore, we attain

$$\begin{cases} |g_{\mathrm{d}}(n,u)| > |g_{\mathrm{d}}(n+1,u)| > |g_{\mathrm{d}}(n',u)|, & u \in \Omega_n^{\mathrm{L}}; \\ |g_{\mathrm{d}}(n',u)| < |g_{\mathrm{d}}(n,u)| \leq |g_{\mathrm{d}}(n+1,u)|, & u \in \Omega_n^{\mathrm{R}}, \end{cases} \tag{7.12}$$

where $|g_{\mathrm{d}}(n,u)| = |g_{\mathrm{d}}(n+1,u)|$ holds at $u = \frac{2\pi(n+1)}{N}$.

It can be concluded from (7.12) that, given $u \in \Omega_n$, the spatial response of either the n-th or the $(n+1)$-th DBD has the maximum amplitude gain among the total N DBDs, and the other one has the second maximum amplitude gain. Lemma 4.1 is proved. □

7.3.2 Proof of Theorem 4.1

Proof. After suppressing the noise in (4.9), we substitute (4.9) into $\rho(\tilde{u}, n)$ and have

$$\rho(\tilde{u}, n) = \frac{\sin\frac{2\pi}{N}\sin(\frac{2\pi}{N} - \tilde{u})}{1 - \cos\frac{2\pi}{N}\cos(\frac{2\pi}{N} - \tilde{u})}. \tag{7.13}$$

$$\sin\left(\frac{2\pi}{N}-\tilde{u}\right)=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$$
$$=\frac{-\rho(\tilde{u},n)\sin(\frac{2\pi}{N})\pm\text{sign}\{\rho(\tilde{u},n)\}\rho(\tilde{u},n)\sin\frac{2\pi}{N}\cos\frac{2\pi}{N}\sqrt{1-\rho^2(\tilde{u},n)}}{\sin^2\frac{2\pi}{N}+\rho^2(\tilde{u},n)\cos^2\frac{2\pi}{N}}. \quad (7.16)$$

Given $\rho(\tilde{u},n)$, $\tilde{u}$ can be uniquely solved from (7.13) if $\rho(\tilde{u},n)$ is strictly monotonic w.r.t $\tilde{u}$, i.e., it is a one-to-one mapping between $\rho(\tilde{u},n)$ and $\tilde{u}$. The derivative of $\rho(\tilde{u},n)$ w.r.t $\tilde{u}$ can be given by

$$\frac{\mathrm{d}\rho(\tilde{u},n)}{\mathrm{d}\tilde{u}}=\frac{\frac{1}{2}\sin\frac{4\pi}{N}-\sin\frac{2\pi}{N}\cos(\tilde{u}-\frac{2\pi}{N})}{\left(\cos\frac{2\pi}{N}\cos(\tilde{u}-\frac{2\pi}{N})-1\right)^2}. \quad (7.14)$$

For $\frac{\pi}{N}\leq\tilde{u}\leq\frac{3\pi}{N}$, we have $|\tilde{u}-\frac{2\pi}{N}|\leq\frac{\pi}{N}$, leading to $\cos(\tilde{u}-\frac{2\pi}{N})>\cos\frac{2\pi}{N}$ and in turn $\sin\frac{2\pi}{N}\cos(\tilde{u}-\frac{2\pi}{N})>\frac{1}{2}\sin\frac{4\pi}{N}$. Therefore, $\frac{\mathrm{d}\rho(\tilde{u},n)}{\mathrm{d}\tilde{u}}<0$ for $\frac{\pi}{N}\leq\tilde{u}\leq\frac{3\pi}{N}$, which indicates the strict monotonicity of $\rho(\tilde{u},n)$.

To derive the inverse function of (7.13), we replace $\cos(\frac{2\pi}{N}-\tilde{u})$ in (7.13) with $\sqrt{1-\sin^2(\frac{2\pi}{N}-\tilde{u})}$ and collapse the resulting terms. (7.13) can be rewritten as

$$a\sin^2\left(\frac{2\pi}{N}-\tilde{u}\right)+b\sin\left(\frac{2\pi}{N}-\tilde{u}\right)+c=0, \quad (7.15)$$

where $a=\sin^2(\frac{2\pi}{N})+\rho^2(\tilde{u},n)\cos^2(\frac{2\pi}{N})$, $b=2\rho(\tilde{u},n)\sin(\frac{2\pi}{N})$ and $c=\rho^2(\tilde{u},n)\sin^2(\frac{2\pi}{N})$. The solution for (7.15) is given in (7.16). Since $|\sin\left(\frac{2\pi}{N}-\tilde{u}\right)|\leq\sin\frac{2\pi}{N}$, we finally obtain (4.10). This concludes the proof. □

7.3.3 Proof of Corollary 4.1

Proof. Given $u=\tilde{u}+\frac{2\pi n}{N}$ ($\hat{u}$ and $\hat{\tilde{u}}$ are the estimates of u and $\tilde{u}$, respectively), the MSE of $\hat{u}$ is equal to that of $\hat{\tilde{u}}$. To derive the MSE of $\hat{\tilde{u}}$, we first calculate the signal power in $|\mathbf{y}_n(k)|^2-|\mathbf{y}_{n+1}(l)|^2$, denoted by σ_{s}^2, and the noise power in $|\mathbf{y}_n(k)|^2+|\mathbf{y}_{n+1}(l)|^2$, denoted by σ_{n}^2. By exploiting (4.9), σ_{s}^2 can be calculated in (7.17). As for σ_{n}^2, we notice that the cross terms between the signal and noise components in $|\mathbf{y}(k)|^2$ are AWGNs with zero means. Therefore, σ_{n}^2 is twice the

$$\sigma_{\mathrm{s}}^2 = \left||\mathbf{y}(k)-\mathbf{n}(k)|^2 - |\mathbf{y}_{n+1}(l)-\mathbf{n}(k+1)|^2\right| = \frac{|\beta s|^2 \sin^2\frac{\pi}{N}\sin^2\frac{N}{2}\tilde{u}}{N^2} \times \left|\frac{\sin^2\frac{1}{2}(\tilde{u}-\frac{4\pi}{N}) - \sin^2\frac{1}{2}\tilde{u}}{\sin^2\frac{1}{2}\tilde{u}\sin^2\frac{1}{2}(\tilde{u}-\frac{2\pi}{N})\sin^2\frac{1}{2}(\tilde{u}-\frac{4\pi}{N})}\right|. \tag{7.17}$$

power of $\mathbf{n}(k)$ in (4.9), i.e., $\sigma_{\mathrm{n}}^2 = 2(1+2r)^2\sigma_{\mathrm{d}}^2$. r is given in (4.5).

Combining (4.10) and (7.13), we have $\mathbb{E}\{\hat{\tilde{u}}\} = 0$. The MSE of $\hat{\tilde{u}}$ can be approximated by [5]

$$\sigma_{\hat{\tilde{u}}}^2 = \mathbb{E}\{\hat{\tilde{u}}^2\} - \left[\mathbb{E}\{\hat{\tilde{u}}\}\right]^2 \approx \frac{1}{2\frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2}}\frac{[1+\rho^2(\tilde{u},n)]}{\eta^2}, \tag{7.18}$$

where $\eta = \frac{\mathrm{d}\rho(\tilde{u},n)}{\mathrm{d}\tilde{u}}\Big|_{\tilde{u}=\frac{2\pi}{N}} = \frac{\frac{1}{2}\sin\frac{4\pi}{N}-\sin\frac{2\pi}{N}}{\left(\cos\frac{2\pi}{N}-1\right)^2}$, as given in (7.14), □

7.3.4 Proof of Theorem 4.2

Proof. From (4.7), $g(n,u)$ is the beamforming gain of the n-th DFT beam in the AoA of u. Apparently, $g(n,u)$ is also the discrete-time Fourier transform (DTFT) of $\frac{1}{N}e^{j\frac{2\pi nn'}{N}}$, $n' = 0,1,\cdots,N-1$. In other words, $g(n,u)$, $n = 0,1,\cdots,N-1$, is a complete set of orthogonal bases spanning the space of continuous band-limited signals [95]. Here, the "band" is the "spatial frequency band". Therefore, $P(u)$ in (4.12) can be linearly combined through $P(u) = \sum_{n=0}^{N-1}\mathbf{W}_{n,k}g(n,u)$, and $\mathbf{W}_{n,k}$ is the projection coefficient of $P(u)$ on the n-th orthogonal basis. $\mathbf{W}_{n,k} = \frac{\langle P(u),g(n,u)\rangle}{\langle g(n,u),g(n,u)\rangle}$. $\langle\cdot,\cdot\rangle$ stands for inner product, i.e.,

$$\langle a(u), b(u)\rangle = \int_0^{2\pi} a(u)b^{\dagger}(u)\,\mathrm{d}u, \tag{7.19}$$

where $(\cdot)^{\dagger}$ denotes conjugate, $a(u)$ and $b(u)$ are band-limited functions of u in the spatial frequency domain.

Accordingly, $\langle P(u), g(n,u)\rangle$ can be calculated as

$$\langle P(u), g(n,u)\rangle = \int_0^{2\pi} P(u)g^{\dagger}(n,u)\ \mathrm{d}u \tag{7.20a}$$

$$= \int_0^{2\pi} P(u)e^{j\frac{N-1}{2}\left(u-\frac{2\pi n}{N}\right)}\frac{\sin\left[\frac{N}{2}\left(u-\frac{2\pi n}{N}\right)\right]}{\sqrt{N}\sin\left[\frac{1}{2}\left(u-\frac{2\pi n}{N}\right)\right]}\ \mathrm{d}u \tag{7.20b}$$

$$= \int_0^{2\pi} P(u)e^{-j\frac{N-1}{2}(\tau-u)}\frac{\sin\left[\frac{N}{2}\left(\tau-u\right)\right]}{\sqrt{N}\sin\left[\frac{1}{2}\left(\tau-u\right)\right]}\ \mathrm{d}u\ \Bigg|_{\tau=\frac{2\pi n}{N}} \tag{7.20c}$$

$$= \int_0^{2\pi} P(u)g(0,\tau-u)\ \mathrm{d}u\ \Bigg|_{\tau=\frac{2\pi n}{N}}. \tag{7.20d}$$

(7.20b) is obtained by plugging (4.7) into (7.20a). By rewriting $\left(u-\frac{2\pi n}{N}\right)$ as $-\left(\frac{2\pi n}{N}-u\right)$ and letting $\tau=\frac{2\pi n}{N}$, (7.20b) becomes (7.20c) which is the linear convolution of $P(u)$ and $g(0,u)=e^{-j\frac{N-1}{2}u}\frac{\sin\frac{N}{2}u}{\sqrt{N}\sin\frac{1}{2}u}$. Finally, $\langle P(u), g(n,u)\rangle$ can be written as (7.20d), the discrete sampling of the linear convolution $\int_0^{2\pi} P(u)g(0,\tau-u)$ at $\tau=\frac{2\pi n}{N}$, $n=0,1,\cdots,N-1$.

Likewise, by substituting (4.7) into (7.19), we have $\langle g(n,u), g(n,u)\rangle=\frac{1}{N}$. Applying the convolution property of DTFT [95], we have

$$\begin{aligned}\mathbf{W}_{n,k} &= N\cdot\int_0^{2\pi} P(u)g(0,\tau-u)\ \mathrm{d}u\ \Bigg|_{\tau=\frac{2\pi n}{N}} = N\cdot\mathrm{DTFT}\{p(n)\odot q(n)\}|_{\tau=\frac{2\pi n}{N}}\,,\\ &= P(\tau)|_{\tau=\frac{2\pi n}{N}} = \begin{cases}(-1)^n & \text{if } m\le n\le m+K-1;\\ 0 & \text{for other } n,\end{cases}\end{aligned} \tag{7.21}$$

where $p(n)$ and $q(n)$ are the inverse DTFT of $P(u)$ and $g(0,u)$, respectively. Given (4.7), $q(n)=\frac{1}{N}$. The convolution of $P(u)$ and $g(0,u)$ is $P(\tau)$ in the spatial frequency domain, which is achieved by replacing u in $P(u)$ with τ.

By using (4.11) and (4.7), the synthesized beam pattern can be written as

$$\widehat{P}(u)=\sum_{n=0}^{N-1}\mathbf{W}_{n,k}g(n,u)=\sum_{n=m}^{m+K-1}(-1)^n g(n,u). \tag{7.22}$$

It is easy to find out the analogy of (7.22) to the synthesis of a time-/frequency-domain rectangular window based on a limited number of sinc kernels [95]. There-

fore, the synthesized beam pattern in (7.22) can suffer from the well-known Gibbs phenomenon [95]. As a result, the synthesized beam pattern takes the exactly same value as (4.12) for $u = \frac{2\pi n}{N},\ m \leq n \leq m + K - 1$, and approximates (4.12) for the rest of the values of u. This concludes the proof. □

7.3.5 Proof of $P_{\mathrm{d}}^{\mathrm{DBD}} \geq P_{\mathrm{d}}^{\mathrm{DFT}}$

Proof. We first consider $\tilde{u} \in \left[\frac{\pi}{N}, \frac{2\pi}{N}\right]$, i.e., $u \in \Omega_n^{\mathrm{L}}$. From Theorem 4.1 and (7.12), we have that the n-th and $(n+1)$-th DBDs were used for AoA estimation, and $|g_{\mathrm{d}}(n,u)| \geq |g_{\mathrm{d}}(n+1,u)|$. From [5], $|g(n+1,u)|$ and $|g(n,u)|(\leq |g(n+1,u)|)$ are used for AoA estimation. To prove $P_{\mathrm{d}}^{\mathrm{DBD}} \geq P_{\mathrm{d}}^{\mathrm{DFT}}$ is equivalent to proving that the SNRs of the n-th and $(n+1)$-th DBD outputs are greater than those of the $(n+1)$-th and n-th DFT beam outputs, respectively.

We start by showing that the noise power in the n-th ($n = 0, 1, \cdots, N-1$) DFT beam output approximates to the noise power in the n-th DBD output, if $r \ll 1$, with the error of $r\sigma_{\mathrm{d}}^2$. From (4.4), the n-th DBD output can be obtained at the k-th RF chain by configuring (4.6) in the LAA receiver. The n-th DFT beam output can be obtained at the k-th RF chain by setting $\mathbf{W}_{:,k}^{\mathrm{DFT}} = [0, \cdots, 0, 1, 0, \cdots, 0]^{\mathrm{T}}$. As discussed in Section 4.2, the powers of the noises at the n-th DBD and the n-th DFT beam outputs are $(1+2r)\sigma_{\mathrm{d}}^2$ and $(1+r)\sigma_{\mathrm{d}}^2$, respectively, since $\|\mathbf{W}_{:,k}\|_0 = 2$ and $\|\mathbf{W}_{:,k}^{\mathrm{DFT}}\|_0 = 1$. Given $r \ll 1$, we have $(1+2r)\sigma_{\mathrm{d}}^2 \cong (1+r)\sigma_{\mathrm{d}}^2$.

We next prove that $|g_{\mathrm{d}}(n,u)| > |g(n+1,u)|$ and $|g_{\mathrm{d}}(n+1,u)| \geq |g(n,u)|$. As shown in the proof of Lemma 4.1, $|g_{\mathrm{d}}(n,u)| > 1$ for $u \in \Omega_n^{\mathrm{L}}$. Given the sinc function $|g(n,u)|$, we have $|g(n,u)| < 1$. From (4.7) and (4.8), it is easy to establish that $|g_{\mathrm{d}}(n+1,u)| \geq |g(n,u)|$ (the details are suppressed for brevity). Therefore, we conclude that the first and second strongest DBD outputs have greater amplitudes than the first and second DFT beam outputs. Given the equal noise-level, the proof concludes. □

7.3.6 Analysis on Fig. 4.8: Null Elimination in Mainlobes

Theorem 4.2 eliminates nulls in the mainlobe. From (4.7), nulls can only occur at $u = \frac{\pi(2n+1)}{N}$ when adjacent beams have identical beamforming amplitudes. The angles of $g(n,u)$ and $g(n+1,u)$ at $u = \frac{\pi(2n+1)}{N}$ satisfy

$$\begin{aligned}&\arg\{g(n,u)\} + \arg\{g(n+1,u)\}\\ &= \frac{N-1}{2}\left(u - \frac{2\pi n}{N}\right) + \frac{N-1}{2}\left(u - \frac{2\pi(n+1)}{N}\right) = 0.\end{aligned}$$

For this reason, directly adding adjacent DFT beams for beam synthesis, as done in [6], can result in nulls in the mainlobe.

In contrast, the proposed beam synthesis in Theorem 4.2 eliminates the nulls. As proved for Lemma 4.1, $|g_{\rm d}(n,u)|$ monotonically decreases with the growth of u in the case of $u \in \Omega_n^{\rm L}$. Therefore, the maximum of $|g_{\rm d}(n,u)|$ is taken at $\frac{\pi(2n+1)}{N}$ (i.e., the left boundary of $\Omega_n^{\rm L}$). As a matter of fact, the maximum is the global maximum of $|g_{\rm d}(n,u)|$, This can be readily established based on (7.11) by verifying that the second derivative of $|g_{\rm d}(n,u)|$ w.r.t u is negative and $\frac{{\rm d}|g_{\rm d}(n,u)|}{{\rm d}u}\Big|_{u=\frac{\pi(2n+1)}{N}} = 0$. By substituting $u = \frac{\pi(2n+1)}{N}$ into (4.8), the maximum amplitude gain of a DBD can be calculated as $\frac{1}{N}\frac{\sin\frac{\pi}{N}}{\sin^2\frac{\pi}{2N}} \approx \frac{1}{N}\frac{\frac{\pi}{N}}{\frac{\pi^2}{4N^2}} = \frac{4}{\pi}$, where the approximation is due to $\sin\left(\frac{\pi}{2N}\right) \approx \frac{\pi}{2N}$ for $N \gg 1$. The approximation error is negligible. Take $N = 32$ for instance, the approximation error is less than 1.971×10^{-5}. Therefore, by controlling the phases of adjacent beams, the proposed beam synthesis turns the nulls in [6] into the maximums.

7.3.7 Analysis on Fig. 4.9: Larger Mainlobe-to-sidelobe Ratio than WDFT [2]

The proposed beam synthesis method in Theorem 4.2 has a larger mainlobe-to-sidelobe ratio than WDFT [2]. Let the element-wise constant-modulus $\tilde{\mathbf{w}} \in \mathbb{C}^{N\times 1}$ and $\mathbf{w} \in \mathbb{R}^{N\times 1}$ denote the coefficient vectors for WDFT and the proposed method to combine the DFT beams in the beam synthesis. According to [2] and (4.11), to approach $P(u)$ in (4.12), the indices for the nonzero elements of $\tilde{\mathbf{w}}$ are identical to those

of $\mathbf{w}$. Given (4.7) and (7.22), the gains of the synthesized beams are $\tilde{\mathbf{w}}^{\mathrm{H}}\mathbf{U}^{\mathrm{H}}\mathbf{a}(u)$ and $\mathbf{w}^{\mathrm{H}}\mathbf{U}^{\mathrm{H}}\mathbf{a}(u)$, i.e., the DTFT of the DFT of $\tilde{\mathbf{w}}$ and $\mathbf{w}$, respectively. From Parseval's Theorem of DTFT and DFT [95], $\tilde{\mathbf{w}}^{\mathrm{H}}\mathbf{U}^{\mathrm{H}}\mathbf{a}(u)$ and $\mathbf{w}^{\mathrm{H}}\mathbf{U}^{\mathrm{H}}\mathbf{a}(u)$ have the same overall power in the spatial frequency domain. In [2], $\tilde{\mathbf{w}}$ is obtained by minimizing the directivity of the synthesized mainlobe, i.e., $\min_{\tilde{\mathbf{w}}} \left| \int_{\frac{2\pi m}{N}}^{\frac{2\pi(m+K-1)}{N}} \tilde{\mathbf{w}}^{\mathrm{H}}\mathbf{U}^{\mathrm{H}}\mathbf{a}(u) \right|^2$ [2, Eq. (36)].* WDFT can achieve smaller ripples, i.e., power, in the mainlobe than the proposed method. Therefore, given the same overall powers of the beams, the sidelobe level of the proposed beam synthesis can be lower than that of WDFT [2]. According to the Gibbs phenomenon [95], the reduction of the sidelobes is greater than the growth of ripples, leading to the larger mainlobe-to-sidelobe ratio of the proposed method than WDFT in Fig. 4.9.

7.4 Proof in Chapter 5

7.4.1 Proof of Theorem 5.1

By combining (5.7) and (5.8), the $\tilde{m}'$-th element ($0 \leq \tilde{m}' \leq \tilde{M}-1$) of the $\tilde{m}$-th spatial-frequency pattern can be given by

$$\begin{aligned} [\tilde{\mathbf{g}}_{n_{\mathrm{L}}}(\tilde{m},\mathcal{M})]_{\tilde{m}'} &= e^{-j\frac{(N-1)\rho_{[\mathcal{M}]_{\tilde{m}'}}\Delta}{2}} \frac{\sin\frac{N\rho_{[\mathcal{M}]_{\tilde{m}'}}\Delta}{2}}{\sin\frac{\rho_{[\mathcal{M}]_{\tilde{m}'}}\Delta}{2}}; \\ \Delta &= \left(\frac{2\pi(n_{\mathrm{L}}+\tilde{m})}{N} - \frac{2\pi n_{\mathrm{L}}}{N\rho_{[\mathcal{M}]_{\tilde{m}'}}}\right). \end{aligned} \tag{7.23}$$

Based on (5.9b) and the definition of $\rho_m = \frac{f_m}{f_{\mathrm{H}}}$, we have

$$\rho_{[\mathcal{M}]_{\tilde{m}'}} = \frac{f_{\mathrm{L}} + \frac{[\mathcal{M}]_{\tilde{m}'}B}{M-1}}{f_{\mathrm{H}}} \approx \rho_0 + \frac{B}{(M-1)f_{\mathrm{H}}} \frac{(\frac{n_{\mathrm{L}}}{n_{\mathrm{L}}+\tilde{m}'} - \rho_0)(M-1)}{1-\rho_0} = \frac{n_{\mathrm{L}}}{n_{\mathrm{L}}+\tilde{m}'}, \tag{7.24}$$

*We reformat the expression to keep consistency with our notations. This does not change the original design object in [2].

where the approximation is due to the suppression of the rounding operator. By substituting (7.24) into (7.23), we obtain

$$\Delta \approx \begin{cases} 0 & \text{if } \tilde{m}' = \tilde{m} \\ \frac{2\pi(\tilde{m}-\tilde{m}')}{N} & \text{if } \tilde{m}' \neq \tilde{m} \end{cases}, \tag{7.25}$$

and, in turn,

$$|[\tilde{\mathbf{g}}_{n_\mathrm{L}}(\tilde{m},\mathcal{M})]_{\tilde{m}'}| = \left|\frac{\sin\frac{N\rho_{[\mathcal{M}]_{\tilde{m}'}}\Delta}{2}}{\sin\frac{\rho_{[\mathcal{M}]_{\tilde{m}'}}\Delta}{2}}\right| \approx \begin{cases} 1 & \text{if } \tilde{m}' = \tilde{m} \\ \left|\frac{\sin\rho_{[\mathcal{M}]_{\tilde{m}'}}\pi(\tilde{m}-\tilde{m}')}{\sin\frac{\rho_{[\mathcal{M}]_{\tilde{m}'}}\pi(\tilde{m}-\tilde{m}')}{N}}\right| \approx 0 & \text{if } \tilde{m}' \neq \tilde{m} \end{cases}. \tag{7.26}$$

The last approximation in the case of $\tilde{m}' \neq \tilde{m}$ is because $\pi(\tilde{m}-\tilde{m}')$ is the null of the sinc function $f(x) = \frac{\sin Nx}{\sin x}$ and $\rho_{[\mathcal{M}]_{\tilde{m}'}}\pi(\tilde{m}-\tilde{m}')$ is close to null, as $\rho_{[\mathcal{M}]_{\tilde{m}'}}$ is small; see (7.24). (7.26) leads to (5.10), which indicates that every pattern is unique. Since $\tilde{m}$ is the index of the pattern and $n_\mathrm{L}+\tilde{m}$ is the index of the DFT beams in the angular region $\mathcal{S}(n_\mathrm{L}, n_\mathrm{H})$, the pattern is a one-to-one mapping of the DFT beam. This concludes the proof.

7.4.2 Proof of Theorem 5.2

By substituting $u = \frac{2\pi(n_\mathrm{L}+\tilde{m})}{N}$, $\rho_m = \rho_{\mathcal{M}_{\tilde{m}'}}$ and (5.9b) into (5.12), we have

$$\begin{aligned} \rho_{\mathcal{M}_{\tilde{m}'}}\frac{2\pi(n_\mathrm{L}+\tilde{m})}{N} &= \frac{2\pi n_\mathrm{L}}{N} \times \frac{n_\mathrm{L}+\tilde{m}}{n_\mathrm{L}+\tilde{m}'} \\ &\in \begin{cases} \left[\frac{2\pi l}{N}, \frac{2\pi n_\mathrm{L}}{N}\right], & \text{if } \tilde{m} \leq \tilde{m}' \leq \frac{n_\mathrm{L}(n_\mathrm{L}+\tilde{m})}{l} - n_\mathrm{L} \\ (0, \frac{2\pi l}{N}] \cup (\frac{2\pi n_\mathrm{L}}{N}, 2\pi], & \text{otherwise} \end{cases}. \end{aligned} \tag{7.27}$$

where $\rho_{\mathcal{M}_{\tilde{m}'}} \approx \frac{n_\mathrm{L}}{n_\mathrm{L}+\tilde{m}'}$, obtained based on (5.9b) in Appendix 7.4.1, is used to attain the first equality. Combining (5.12), (5.13) and (7.27) yields

$$[\breve{\mathbf{g}}_{n_\mathrm{L}}(\tilde{m},\mathcal{M})]_{\tilde{m}'} = P\left(\frac{2\pi(n_\mathrm{L}+\tilde{m})}{N}, \rho_{\mathcal{M}_{\tilde{m}'}}\right) \approx \begin{cases} 1, & \text{if } \tilde{m} \leq \tilde{m}' \leq \frac{n_\mathrm{L}(n_\mathrm{L}+\tilde{m})}{l} - n_\mathrm{L} \\ 0, & \text{otherwise} \end{cases}. \tag{7.28}$$

As $\tilde{m}$ varies from 0 to $\tilde{M}-1$, each pattern $\breve{\mathbf{g}}_{n_{\mathrm{L}}}(\tilde{m},\mathcal{M})$ is unique, which proves the one-to-one mapping.

By taking $\tilde{m}=0$ in (7.28), $\breve{\mathbf{g}}_{n_{\mathrm{L}}}(0,\mathcal{M})=\mathbf{1}_{\tilde{M}\times 1}$ can happen provided a small l. However, $\breve{\mathbf{g}}_{n_{\mathrm{L}}}(0,\mathcal{M})$ may be non-unique in the case of multi-beam selection. By assuming $l=0$, $u\in\mathcal{S}(\lceil\frac{l}{\rho_0}\rceil,n_{\mathrm{L}})$ always leads to $|\mathbf{p}(\mathcal{M})|=\mathbf{1}_{\tilde{M}\times 1}$, where

$$\mathbf{p}(\mathcal{M})=\left[P\left(u,\rho_{\mathcal{M}_0}\right),P\left(u,\rho_{\mathcal{M}_1}\right),\cdots,P\left(u,\rho_{\mathcal{M}_{\tilde{M}-1}}\right)\right]^{\mathrm{T}} \tag{7.29}$$

has the same structure as the spatial-frequency pattern; see (7.28). Therefore, we need to exclude $\tilde{\mathbf{g}}_{n_{\mathrm{L}}}(0,\mathcal{M})$

7.4.3 Proof of Proposition 5.1

To maximize the probability of correctly identifying the strongest DFT beam at a symbol, we should maximize the angular regions probed by the single-beam selections at the first $(K-1)$ RF chains and synthesize a wide beam to probe the remaining angular region. This is the because the single DFT beam can be used for unambiguous estimation of the strongest DFT beam if the AoA of the strongest path is in the angular region covered by the beam; see Theorem 5.1. In contrast, a wide beam may fail to detect the strongest DFT beam, since the AoA cannot be guaranteed to fall into the detectable angular region of Theorem 5.2. The reason we synthesize the wide beam is to make sure the strongest path is not overlooked.

Based on the spatial response of a DFT beam in (5.7), the angular region probed by the k-th RF chain ($k\in[0,K-2]$) is given by $\Omega_k=\mathcal{S}([\mathbf{n}_{\mathrm{L}}]_k,[\mathbf{n}_{\mathrm{H}}]_k)$. The larger $[\mathbf{n}_{\mathrm{L}}]_k$ is, the wider angular region can be probed by selecting the $[\mathbf{n}_{\mathrm{L}}]_k$-th DFT beam. This is because the width of Ω_k is $|\Omega_k|=\frac{2\pi([\mathbf{n}_{\mathrm{H}}]_k-[\mathbf{n}_{\mathrm{L}}]_k+1)}{N}$. Based on (5.9a), we have $|\Omega_k|\approx\frac{2\pi\left([\mathbf{n}_{\mathrm{L}}]_k(\frac{1}{\rho_0}-1)\right)+1}{N}$ which is a monotonically increasing function of $[\mathbf{n}_{\mathrm{L}}]_k$. The angular partition scheme in (5.17) guarantees that the maximum $[\mathbf{n}_{\mathrm{L}}]_k(k\in[0,K-2])$ is taken. This concludes the proof.

7.4.4 Proof of Corollary 5.1

Given the amplitude relation among DFT beams, the fact that the $(i^\star + k^\star Q)$-th DFT beam has the strongest output indicates that $\rho_m u - \frac{2\pi i^\star}{N} + \frac{2\pi k^\star}{K} \in \left(-\frac{\pi}{N}, \frac{\pi}{N}\right)$. Asymptotically, we have

$$\lim_{N\to\infty} \left(\rho_m u - \frac{2\pi i^\star}{N} - \frac{2\pi k^\star}{K}\right) = 0, \tag{7.30}$$

which indicates that $\rho_m u$ can be represented by any Butler beam pointing direction in the case of $N \to \infty$. Substituting (7.30) into (5.19), we obtain

$$\lim_{N\to\infty} |\mathcal{G}(u, \rho_m, i^\star)| = \lim_{N\to\infty} \left| \frac{\sin \frac{QK}{2}\left(\rho_m u - \frac{2\pi i^\star}{N}\right)}{\sin \frac{K}{2}\left(\rho_m u - \frac{2\pi i^\star}{N}\right)} \right| = \left| \frac{\sin k^\star Q\pi}{\sin k^\star \pi} \right| = Q, \tag{7.31}$$

where the last equality is based on *L'Hospital's rule* [94]. Likewise, the amplitude of $\mathcal{G}(u, \rho_m, i)$ at $i \neq i^\star$ can be calculated by

$$\lim_{N\to\infty} |\mathcal{G}(u, \rho_m, i)| = \left| \frac{\sin \frac{QK}{2}\left(\frac{2\pi i^\star}{N} + \frac{2\pi k^\star}{K} - \frac{2\pi i}{N}\right)}{\sin \frac{K}{2}\left(\frac{2\pi i^\star}{N} + \frac{2\pi k^\star}{K} - \frac{2\pi i}{N}\right)} \right| = 0, \tag{7.32}$$

where the last equality is because the enumerator in (7.32) is zero while the denominator is non-zero.

By substituting (7.31) and (7.32) into (5.21), we see that, in the case of $N \to \infty$, the AoA estimate (5.22) can only be obtained at $i \neq i^\star$ with the largest estimation SNR achieved; and the other DFT beams make an asymptotically negligible contribution. This concludes the proof.

7.4.5 Proof of Lemma 5.2

From Algorithm 8, the number of DFT beams probed at RF chain 0 and symbol 0 is $N - \lceil N\rho_0 \rceil + 1 \approx N(1-\rho_0)$, where the approximation is based on $N \to \infty$. Likewise, at symbol 0, the 2nd RF chain asymptotically probes $N\rho_0(1-\rho_0)$ DFT beams. More generally, the k-th RF chain asymptotically probes $N\rho_0^k(1-\rho_0)$ DFT beams. As such, at symbol 1, the first RF chain asymptotically probes $N\rho_0^K(1-\rho_0)$ DFT beams. More generally, the k-th RF chain asymptotically probes

$N\rho_0^{K+k}(1-\rho_0)$ DFT beams. To this end, we conclude that the k-th RF chain asymptotically probes $N\rho_0^{(x-1)K+k}(1-\rho_0)$ DFT beams at symbol $(x-1)$, as $N \to \infty$; and, accordingly, the number of DFT beams probed at symbol $(x-1)$ can be given by $I(x) = \sum_{k=0}^{K-1} N\rho_0^{(x-1)K+k}(1-\rho_0) = N\rho_0^{(x-1)K}(1-\rho_0^K)$.

The fact that Algorithm 8 uses x symbols indicates that the strongest DFT beam is one of the DFT beams probed at symbol $(x-1)$. Therefore, the PDF of Algorithm using x symbols is equivalent to that the strongest DFT beam is one of the $I(x)$ beams probed at symbol $(x-1)$. Finally, as u is uniformly distributed in $[0, 2\pi]$, the probability of any beam being the strongest DFT beam is $\frac{1}{N}$, which leads to (5.24).

Bibliography

[1] X. Huang and Y. J. Guo, "Frequency-domain AoA estimation and beamforming with wideband hybrid arrays," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2543–2553, August 2011.

[2] S. Noh, M. D. Zoltowski, and D. J. Love, "Multi-resolution codebook and adaptive beamforming sequence design for millimeter wave beam alignment," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5689–5701, Sept 2017.

[3] X. Huang, Y. J. Guo, and J. D. Bunton, "A hybrid adaptive antenna array," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1770–1779, 2010.

[4] J. A. Zhang, W. Ni, P. Cheng, and Y. Lu, "Angle-of-arrival estimation using different phase shifts across subarrays in localized hybrid arrays," *IEEE Commun. Lett.*, vol. 20, no. 11, pp. 2205–2208, Nov 2016.

[5] D. Zhu, J. Choi, and R. W. Heath, "Auxiliary beam pair enabled AoD and AoA estimation in closed-loop large-scale millimeter-wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4770–4785, July 2017.

[6] J. Seo, Y. Sung, G. Lee, and D. Kim, "Training beam sequence design for millimeter-wave MIMO systems: A POMDP framework," *IEEE Trans. Signal Process.*, vol. 64, no. 5, pp. 1228–1242, March 2016.

[7] K. Wu, W. Ni, T. Su, R. P. Liu, and Y. J. Guo, "Fast and accurate estimation of angle-of-arrival for satellite-borne wideband communication system," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 2, pp. 314–326, Feb 2018.

[8] X. Gao, L. Dai, S. Zhou, A. M. Sayeed, and L. Hanzo, "Beamspace channel estimation for wideband millimeter-wave MIMO with lens antenna array," in *IEEE ICC*, May 2018, pp. 1–6.

[9] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, February 2014.

[10] W. Roh *et al.*, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, February 2014.

[11] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmwave) for 5G: opportunities and challenges," *Wireless Netw.*, vol. 21, no. 8, pp. 2657–2676, 2015.

[12] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[13] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of RF chains and phase shifters, and scheduling design with hybrid analog–digital beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3311–3326, May 2016.

[14] D. Dawn, P. Sen, S. Sarkar, B. Perumana, S. Pinel, and J. Laskar, "60-GHz integrated transmitter development in 90-nm CMOS," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 10, pp. 2354–2367, Oct 2009.

[15] J. A. Zhang, X. Huang, V. Dyadyuk, and Y. J. Guo, "Massive hybrid antenna array for millimeter-wave cellular communications," *IEEE Trans. Wireless Commun.*, vol. 22, no. 1, pp. 79–87, 2015.

[16] Y. J. Guo, X. Huang, and V. Dyadyuk, "A hybrid adaptive antenna array for long-range mm-wave communications [antenna applications corner]," *IEEE Antennas Propag. Mag.*, vol. 54, no. 2, pp. 271–282, 2012.

[17] J. A. Zhang, X. Huang, and Y. J. Guo, "Adaptive searching and tracking algorithm for AoA estimation in localized hybrid array," in *Int. Conf. Commun. Workshop (ICCW)*. IEEE, June 2015, pp. 1095–1100.

[18] S. Kutty and D. Sen, "Beamforming for millimeter wave communications: An inclusive survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 949–973, 2016.

[19] S.-F. Chuang, W.-R. Wu, and Y.-T. Liu, "High-resolution AoA estimation for hybrid antenna arrays," *IEEE Trans. Antennas Propag.*, vol. 63, no. 7, pp. 2955–2968, 2015.

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

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

[22] M. Kokshoorn, H. Chen, P. Wang, Y. Li, and B. Vucetic, "Millimeter wave MIMO channel estimation using overlapped beam patterns and rate adaptation," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 601–616, Feb 2016.

[23] R. C. de Lamare, "Massive MIMO systems: Signal processing challenges and future trends," *URSI Radio Sci. Bulletin*, vol. 86, no. 4, pp. 8–20, Dec 2013.

[24] P. D. Arapoglou, K. Liolis, M. Bertinelli, A. Panagopoulos, P. Cottis, and R. D. Gaudenzi, "MIMO over satellite: A review," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 1, pp. 27–51, First 2011.

[25] B. G. Evans, "The role of satellites in 5G," in *2014 7th Advanced Sat. Multimedia Syst. Conf. & 13th Signal Process. for Space Commun. Workshop (ASMS/SPSC)*, Sept 2014, pp. 197–202.

[26] W. Ni and I. B. Collings, "A new adaptive small-cell architecture," *IEEE J. Select. Areas Commun.*, vol. 31, no. 5, pp. 829–839, May 2013.

[27] E. Cianca, T. Rossi, A. Yahalom, Y. Pinhasi, J. Farserotu, and C. Sacchi, "EHF for satellite communications: The new broadband frontier," *Proc. IEEE*, vol. 99, no. 11, pp. 1858–1881, Nov 2011.

[28] Z. Fang, W. Ni, F. Liang, P. Shao, and Y. Wu, "Massive MIMO for full-duplex cellular two-way relay network: A spectral efficiency study," *IEEE Access*, vol. 5, pp. 23 288–23 298, 2017.

[29] C. Qin *et al.*, "Joint beamforming and user selection in multiuser collaborative MIMO SWIPT systems with non-negligible circuit energy consumption," *IEEE Trans. on Veh. Tech.*, vol. PP, no. 99, pp. 1–1, June 2017.

[30] Y. Huang and B. Clerckx, "Waveform design for wireless power transfer with limited feedback," *IEEE Trans. on Wireless Commun.*, vol. 17, no. 1, pp. 415–429, Jan 2018.

[31] Y. Zeng, B. Clerckx, and R. Zhang, "Communications and signals design for wireless power transmission," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2264–2290, May 2017.

[32] R. Zhang and C. K. Ho, "Mimo broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[33] T. D. P. Perera, D. N. K. Jayakody, S. K. Sharma, S. Chatzinotas, and J. Li, "Simultaneous wireless information and power transfer (swipt): Recent advances and future challenges," *IEEE Commun. Surveys Tut.*, vol. PP, no. 99, pp. 1–1, 2017.

[34] K. Huang and X. Zhou, "Cutting the last wires for mobile communications by microwave power transfer," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 86–93, June 2015.

[35] L. Yang, Y. Zeng, and R. Zhang, "In-band wireless information and power transfer with lens antenna array," *IEEE Communications Letters*, vol. 21, no. 1, pp. 100–103, Jan 2017.

[36] G. Zhu and K. Huang, "Analog spatial cancellation for tackling the near-far problem in wirelessly powered communications," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3566–3576, Dec 2016.

[37] A. Li and C. Masouros, "Hybrid analog-digital millimeter-wave mu-mimo transmission with virtual path selection," *IEEE Communications Letters*, vol. 21, no. 2, pp. 438–441, Feb 2017.

[38] L. Zhao, D. W. K. Ng, and J. Yuan, "Multi-user precoding and channel estimation for hybrid millimeter wave systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1576–1590, July 2017.

[39] A. Sayeed and J. Brady, "Beamspace MIMO for high-dimensional multiuser communication at millimeter-wave frequencies," in *2013 IEEE GLOBECOM*, Dec 2013, pp. 3679–3684.

[40] J. Brady, N. Behdad, and A. M. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, July 2013.

[41] Y. Zeng, R. Zhang, and Z. N. Chen, "Electromagnetic lens-focusing antenna enabled massive MIMO: Performance improvement and cost reduction," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1194–1206, June 2014.

[42] Y. Zeng and R. Zhang, "Millimeter wave MIMO with lens antenna array: A new path division multiplexing paradigm," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1557–1571, April 2016.

[43] X. Gao, L. Dai, S. Han, C. L. I, and X. Wang, "Reliable beamspace channel estimation for millimeter-wave massive MIMO systems with lens antenna array," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6010–6021, Sept 2017.

[44] K. Guan, G. Li, T. Kürner, A. F. Molisch, B. Peng, R. He, B. Hui, J. Kim, and Z. Zhong, "On millimeter wave and THz mobile radio channel for smart rail mobility," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 5658–5674, July 2017.

[45] B. Ai *et al.*, "Challenges toward wireless communications for high-speed railway," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2143–2158, Oct 2014.

[46] K. Guan *et al.*, "Towards realistic high-speed train channels at 5G millimeter-wave band—part II: Case study for paradigm implementation," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9129–9144, Oct 2018.

[47] F. Hasegawa *et al.*, "High-speed train communications standardization in 3GPP 5G NR," *IEEE Commun. Standards Mag.*, vol. 2, no. 1, pp. 44–52, MARCH 2018.

[48] Z. Miao *et al.*, "A 400-GHz high-gain quartz-based single layered folded reflectarray antenna for Terahertz applications," *IEEE Trans. Terahertz Sci. Technol.*, vol. 9, no. 1, pp. 78–88, Jan 2019.

[49] X. Gao *et al.*, "Low RF-complexity technologies to enable millimeter-wave MIMO with large antenna array for 5G wireless communications," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 211–217, APRIL 2018.

[50] C. Lin and G. Y. L. Li, "Terahertz communications: An array-of-subarrays solution," *IEEE Comm. Mag.*, vol. 54, no. 12, pp. 124–131, December 2016.

[51] K. Wu *et al.*, "Efficient Angle-of-Arrival estimation of lens antenna arrays for wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 1, pp. 116–130, 2019.

[52] V. Joroughi, M. Á. Vázquez, A. I. Pérez-Neira, and B. Devillers, "Onboard beam generation for multibeam satellite systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3714–3726, June 2017.

[53] V. Joroughi, M. Á. Vázquez, and A. I. Pérez-Neira, "Generalized multicast multibeam precoding for satellite communications," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 952–966, Feb 2017.

[54] M. Hasan and C. Bianchi, "Ka band enabling technologies for high throughput satellite (HTS) communications," *Int. J. Sat. Commun. Netw.*, vol. 34, no. 4, pp. 483–501, Nov 2016, sAT-15-0016.R2. [Online]. Available: http://dx.doi.org/10.1002/sat.1161

[55] Y. Vasavada, R. Gopal, C. Ravishankar, G. Zakaria, and N. BenAmmar, "Architectures for next generation high throughput satellite systems," *Int. J. Sat. Commun. Netw.*, vol. 34, no. 4, pp. 523–546, Jan 2016, sat.1175. [Online]. Available: http://dx.doi.org/10.1002/sat.1175

[56] O. Montenbruck and E. Gill, *Satellite Tracking and Observation Models*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 193–232. [Online]. Available: https://doi.org/10.1007/978-3-642-58351-3{_}6

[57] S. D. Ilcev, *Global mobile satellite communications for maritime, land, and aeronautical applications*. Dordrecht : Springer, 2005. [Online]. Available: http://rave.ohiolink.edu/ebooks/ebc/1402027842

[58] M. A. Cervera, A. Ginesi, and K. Eckstein, "Satellite-based vessel automatic identification system: A feasibility and performance analysis," *Int. J. Sat. Commun. Netw.*, vol. 29, no. 2, pp. 117–142, 2011. [Online]. Available: http://dx.doi.org/10.1002/sat.957

[59] G. Su and M. Morf, "The signal subspace approach for multiple wide-band emitter location," *IEEE Trans. Acoust., Speech, and Signal Process.*, vol. 31, no. 6, pp. 1502–1522, Dec 1983.

[60] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources," *IEEE Trans. Acoust., Speech, and Signal Process.*, vol. 33, no. 4, pp. 823–831, Aug 1985.

[61] H. Hung and M. Kaveh, "Focussing matrices for coherent signal-subspace processing," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 36, no. 8, pp. 1272–1281, Aug 1988.

[62] M. A. Doron and A. J. Weiss, "On focusing matrices for wide-band array processing," *IEEE Trans. Signal Process.*, vol. 40, no. 6, pp. 1295–1302, Jun 1992.

[63] F. Sellone, "Robust auto-focusing wideband DOA estimation," *Signal Process.*, vol. 86, no. 1, pp. 17–37, 2006.

[64] A. Shahmansoori *et al.*, "Tracking position and orientation through millimeter wave lens MIMO in 5g systems," *CoRR*, vol. abs/1809.06343, 2018. [Online]. Available: http://arxiv.org/abs/1809.06343

[65] S. A. Shaikh and A. M. Tonello, "Localization based on angle of arrival in EM lens-focusing massive mimo," in *2016 IEEE 6th ICCE-Berlin*, Sep. 2016, pp. 124–128.

[66] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial- and frequency-wideband effects in millimeter-wave massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 66, no. 13, pp. 3393–3406, July 2018.

[67] Z. Guo, X. Wang, and W. Heng, "Millimeter-wave channel estimation based on 2-D] beamspace MUSIC method," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5384–5394, Aug 2017.

[68] M. R. Akdeniz *et al.*, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, June 2014.

[69] G. Yang, C. K. Ho, and Y. L. Guan, "Dynamic resource allocation for multiple-antenna wireless power transfer," *IEEE Trans. Signal Process.*, vol. 62, no. 14, pp. 3565–3577, July 2014.

[70] X. Chen, C. Yuen, and Z. Zhang, "Wireless energy and information transfer tradeoff for limited-feedback multiantenna systems with energy beamforming," *IEEE Trans. Veh. Technol.*, vol. 63, no. 1, pp. 407–412, Jan 2014.

[71] K. W. Choi, D. I. Kim, and M. Y. Chung, "Received power-based channel estimation for energy beamforming in multiple-antenna rf energy transfer system," *IEEE Transactions on Signal Processing*, vol. 65, no. 6, pp. 1461–1476, March 2017.

[72] S. Abeywickrama, T. Samarasinghe, C. K. Ho, and C. Yuen, "Wireless energy beamforming using received signal strength indicator feedback," *IEEE Transactions on Signal Processing*, vol. 66, no. 1, pp. 224–235, Jan 2018.

[73] K. Wu, W. Ni, T. Su, R. P. Liu, and Y. J. Guo, "Robust unambiguous estimation of angle-of-arrival in hybrid array with localized analog subarrays," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2987–3002, May 2018, accepted.

[74] L. Yang, Y. Zeng, and R. Zhang, "Efficient channel estimation for millimeter wave mimo with limited RF chains," in *2016 IEEE ICC*, May 2016, pp. 1–6.

[75] J. Hogan and A. Sayeed, "Beam selection for performance-complexity optimization in high-dimensional MIMO systems," in *2016 Annual CISS*, March 2016, pp. 337–342.

[76] H. Son and B. Clerckx, "Joint beamforming design for multi-user wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6397–6409, Nov 2014.

[77] J. Park and B. Clerckx, "Joint wireless information and energy transfer with reduced feedback in mimo interference channels," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1563–1577, Aug 2015.

[78] J. Xu and R. Zhang, "Energy beamforming with one-bit feedback," *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5370–5381, Oct 2014.

[79] ——, "A general design framework for mimo wireless energy transfer with limited feedback," *IEEE Transactions on Signal Processing*, vol. 64, no. 10, pp. 2475–2488, May 2016.

[80] B. Wang, F. Gao, S. Jin, H. Lin, G. Y. Li, S. Sun, and T. S. Rappaport, "Spatial-wideband effect in massive MIMO with application in mmwave systems," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 134–141, December 2018.

[81] M. Cai, K. Gao, D. Nie, B. Hochwald, J. N. Laneman, H. Huang, and K. Liu, "Effect of wideband beam squint on codebook design in phased-array wireless systems," in *2016 IEEE GLOBECOM*, Dec 2016, pp. 1–6.

[82] J. H. Brady and A. M. Sayeed, "Wideband communication with high-dimensional arrays: New results and transceiver architectures," in *2015 IEEE ICCW*, June 2015, pp. 1042–1047.

[83] Z. Gao *et al.*, "Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1259–1262, June 2016.

[84] K. Venugopal *et al.*, "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, Sep. 2017.

[85] C. Zhang *et al.*, "Doppler shift estimation for millimeter-wave communication systems on high-speed railways," *IEEE Access*, pp. 1–1, 2018.

[86] I. A. Hemadeh *et al.*, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 870–913, Secondquarter 2018.

[87] H. Song *et al.*, "Millimeter-wave network architectures for future high-speed railway communications: Challenges and solutions," *IEEE Wireless Commun.*, vol. 23, no. 6, pp. 114–122, December 2016.

[88] L. C. Godara, "Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations," *Proc. IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.

[89] J. P. González-Coma *et al.*, "Channel estimation and hybrid precoding for frequency selective multiuser mmwave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 2, pp. 353–367, May 2018.

[90] J. Rodríguez-Fernández, N. González-Prelcic, K. Venugopal, and R. W. Heath, "Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2946–2960, May 2018.

[91] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029–3056, Sept 2015.

[92] C. Han, A. O. Bicen, and I. F. Akyildiz, "Multi-ray channel modeling and wideband characterization for wireless communications in the Terahertz band," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2402–2412, May 2015.

[93] X. Huang and Y. J. Guo, "Closed-form MSE performance for phase estimation from Gaussian reference signals," in *Commununications and Information Technologies (ISCIT), 2011 11th International Symposium on.* IEEE, 2011, pp. 154–158.

[94] S. G. Krantz, *A handbook of real variables: with applications to differential equations and Fourier analysis.* Springer Science Business Media, 2011.

[95] A. V. Oppenheim, *Discrete-time signal processing.* Pearson Education India, 1999.

[96] R. Reggiannini, "A fundamental lower bound to the performance of phase estimators over Rician-fading channels," *IEEE Trans. Commun.*, vol. 45, no. 7, pp. 775–778, July 1997.

[97] Y. Yang, X. Zhu, and M. Heimlich, "A broadside-coupled meander-line resonator (BCMLR) for ultra-compact millimeter-wave MMIC designs in 0.13 μm SiGe technology," in *Electromagnetics: Applications and Student Innovation Competition (iWEM), 2016 IEEE International Workshop on.* IEEE, 2016, pp. 1–3.

[98] Y.-S. Yeh and B. A. Floyd, "A 55-GHz power-efficient frequency quadrupler with high harmonic rejection in 0.1-μm sige BiCMOS technology," in *Radio Frequency Integrated Circuits Symposium (RFIC), 2015 IEEE.* IEEE, 2015, pp. 267–270.

[99] H.-C. Yeh, C.-C. Chiong, M.-T. Chen, and H. Wang, "Review of millimeter-wave MMIC mixers." *IEEE Des. Test*, vol. 31, no. 6, pp. 38–45, 2014.

[100] J. A. Shaw, "Radiometry and the Friis transmission equation," *Amer. J. Physics*, vol. 81, no. 1, pp. 33–37, 2013.

[101] A. Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.

[102] A. N. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer-Rao bound and its application to synchronization problems," *IEEE Trans. Commun.*, vol. 42, no. 234, pp. 1391–1399, Feb 1994.

[103] F. Z. T. K. Y. Gao, M. Khaliel, "Rotman lens based hybrid analog-digital beamforming in massive MIMO systems: Array architectures, beam selection algorithms and experiments," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9134–9148, 2017.

[104] A. F. Molisch, *Noise and Interference Limited Systems.* Wiley-IEEE Press, 2011, pp. 37–44.

[105] Y. Han, S. Jin, J. Zhang, J. Zhang, and K. K. Wong, "DFT-based hybrid beamforming multiuser systems: Rate analysis and beam selection," *IEEE J. Sel. Topics Signal Process.*, pp. 1–1, 2018.

[106] W. Hong, Z. H. Jiang, C. Yu, J. Zhou, P. Chen, Z. Yu, H. Zhang, B. Yang, X. Pang, M. Jiang, Y. Cheng, M. K. T. Al-Nuaimi, Y. Zhang, J. Chen, and S. He, "Multibeam antenna technologies for 5g wireless communications," *IEEE Trans. Antennas and Propag.*, vol. 65, no. 12, pp. 6231–6249, Dec 2017.

[107] Y. J. Guo and B. Jones, *Chapter 40. Base Station Antennas.* New York, NY, USA: McGraw-Hill, 2018.