

Modern Antenna

Guolong He · Xin Gao · Rentian Zhang ·
Liangliang Sun · Hui Zhou

Multibeam Phased Array Antennas as Satellite Constellation Ground Station

 Springer

Modern Antenna

Editors-in-Chief

Junping Geng, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China

Jiadong Xu, School of Electronics and Information, Northwestern Polytechnical University, Xi'an, Shaanxi, China

Series Editors

Yijun Feng, School of Electronic Science and Engineering, Nanjing University, Nanjing, Jiangsu, China

Xiaoxing Yin, School of Information Science and Engineering, Southeast University, Nanjing, Jiangsu, China

Gaobiao Xiao, Electronic Engineering Department, Shanghai Jiao Tong University, Shanghai, China

Anxue Zhang, Institute of Electromagnetic and Information Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, China

Zengrui Li, Communication University of China, Beijing, China

Kaixue Ma , School of Microelectronics, Tianjin University, Tianjin, China

Xiuping Li, School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing, China

Yanhui Liu, School of Electronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, China

Shiwei Dong, National Key Laboratory of Science and Technology on Space Microwave, China Academy of Space Technology (Xi'an), Xi'an, Shaanxi, China

Mingchun Tang, College of Microelectronics and Communication Engineering, Chongqing University, Chongqing, China

Qi Wu, School of Electronics and Information Engineering, Beihang University, Beijing, China

The modern antenna book series mainly covers the related antenna theories and technologies proposed and studied in recent years to solve the bottleneck problems faced by antennas, including binary coded antenna optimization method, artificial surface plasmon antenna, complex mirror current equivalent principle and low profile antenna, generalized pattern product principle and generalized antenna array, cross dielectric transmission antenna, metamaterial antenna, as well as new antenna technology and development. This series not only presents the important progress of modern antenna technology from different aspects, but also describes new theoretical methods, which can be used in modern and future wireless communication, radar detection, internet of things, wireless sensor networks and other systems.

The purpose of the modern antenna book series is to introduce new antenna concepts, new antenna theories, new antenna technologies and methods in recent years to antenna researchers and engineers for their study and reference. Each book in this series is thematic. It gives a comprehensive overview of the research methods and applications of a certain type of antenna, and specifically expounds the latest research progress and design methods. As a collection, the series provides valuable resources to a wide audience in academia, the engineering research community, industry and anyone else who are looking to expand their knowledge of antenna methods. In addition, modern antenna series is also open. More antenna researchers are welcome to publish their new research results in this series.

Guolong He · Xin Gao · Rentian Zhang ·
Liangliang Sun · Hui Zhou

Multibeam Phased Array Antennas as Satellite Constellation Ground Station

Guolong He
Beijing Institute of Tracking
and Telecommunication Technology
Beijing, China

Xin Gao
Beijing Institute of Tracking
and Telecommunication Technology
Beijing, China

Rentian Zhang
Beijing Institute of Tracking
and Telecommunication Technology
Beijing, China

Liangliang Sun
Beijing Institute of Tracking
and Telecommunication Technology
Beijing, China

Hui Zhou
Beijing Institute of Tracking
and Telecommunication Technology
Beijing, China

ISSN 2731-7986

ISSN 2731-7994 (electronic)

Modern Antenna

ISBN 978-981-99-7909-7

ISBN 978-981-99-7910-3 (eBook)

<https://doi.org/10.1007/978-981-99-7910-3>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.

The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Paper in this product is recyclable.

Preface

Whether you believe it or not, although this is my first time trying to write an academic book, and the topic of this book is about phased array antennas, however, my professional career began with parabolic reflector antennas. When I earned my bachelor's degree in electronic engineering in 2008, I was fortunately awarded a job that involved the design, construction, and maintenance of the Chinese deep space stations [1]. In the following ten years, our team has developed three globally distributed ground stations with large reflector antennas (66 m in the Jiamusi station and 35 m in the Kashi and Argentina stations), providing continuous telemetry, tracking and command (TT&C), reliable communication, and precise navigation services for all Chinese deep space missions, such as the ambitious Lunar exploration missions Chang'E and the first Chinese Mars exploration mission Tianwen-1. These deep space stations are giant and powerful by integrating with state-of-the-art technologies, including low-loss beam waveguide, centralized high power transmitter, and cryogenic low noise amplifier. I even became the Project Manager and was responsible for all technical matters of the 4×35 m Kashi deep space antenna array [2].

Along with these projects for deep space missions, I also participated in some other projects for Earth orbit missions. There are many aspects of mission demand differences between them. The deep space missions involve receiving extremely weak signals from vast interplanetary distances, hence its focus requires the ground station with better accuracy and higher sensitivity. However, with the increasing number of satellites in the low earth orbit (LEO), one of the most urgent requirements is multiple mission support capability for the LEO satellite ground station. Recently, the concept of LEO satellite constellation accelerates this problem.

The traditional reflector antennas are powerful but can only produce one beam once at a time. On the contrary, phased array antennas have the natural capability to produce multiple beams simultaneously, thus meeting the requirements for ordinary management and control of future satellite constellation. Compared to the reflector antennas, the architecture of phased array antennas are quite different with thousands of antenna elements, integrated transmit/receive (T/R) modules, decentralized power amplifiers, and agile beamforming networks. Furthermore, phased array antennas have many other advantages over reflector antennas, such as better performance,

higher reliability, and lower cost. All of these make it a very attractive candidate as next generation LEO satellite constellation ground station.

There are already a large number of excellent books about phased array antennas available in the open literature, the readers are encouraged to understand the basics of them first (see e.g. [3–5], I also learn and benefit a lot from them!). However, phased array antennas are still new as satellite ground stations. This book is mainly focused on some engineering practices and fills the gap between the well-known phased array theory and satellite ground station application, such as array excitation error impact analysis and calibration methods, array geometry design for hemispherical coverage, multiple beam resource management, etc. Hope you can find something useful. Finally, if you have any comments or questions, please do not hesitate to contact me.

Beijing, China
June 2023

Guolong He
heguolong@alumni.sjtu.edu.cn

References

1. G. He, M. Liu, X. Gao, et al., Chinese deep space stations: a brief review. *IEEE Antennas Propag. Mag.* **64**(1), 102–111 (2022)
2. G. He, M. Liu, X. Du, et al., The Kashi 4×35 -m Antenna Array for Deep Space Telecommunication and Radio Science observation. *IEEE Aerosp. Electron. Syst. Mag.* (2023)
3. A.K. Bhattacharyya, *Phased Array Antennas: Floquet Analysis, Synthesis, BFNs and Active Array Systems*, (John Wiley and Sons, Inc., Hoboken, 2006)
4. R.J. Mailloux, *Phased Array Antenna Handbook*, 3rd edn. (Artech House, Boston, 2017)
5. R.L. Haupt, *Antenna Arrays: A Computational Approach* (New York, Wiley, 2010)

Contents

1	Introduction	1
1.1	The New Era of Small Satellite Constellation	1
1.2	Simultaneous Multiple Satellite Support Requirement on Ground Station	4
1.3	Advantages of Phased Array Antennas as Satellite Ground Station	5
1.3.1	Better Performance	5
1.3.2	Higher Reliability	6
1.3.3	Lower Cost	6
1.4	Outline of This Book	7
	References	8
2	Phased Array Antenna Basics	9
2.1	Evolution of Phased Array Antenna Architecture	9
2.2	Major Components of a Phased Array Antenna	10
2.2.1	Antenna Element	10
2.2.2	T/R Module	12
2.2.3	Beamforming Network	17
2.2.4	Digital Signal Processor	18
2.2.5	Cooling Equipment	19
2.3	Array Pattern Synthesis	19
2.3.1	Array Factor Basics	20
2.3.2	Analytical Algorithms for Array Amplitude Tapers	21
2.3.3	Genetic Algorithm	22
2.3.4	Particle Swarm Optimization	27
2.3.5	Convex Optimization	32
2.4	Summary	34
	References	36

3	Array Excitation Error Analysis	39
3.1	Probabilistic Methods	39
3.1.1	The Mathematical Model	39
3.1.2	The Expectation of the Array Pattern	41
3.1.3	The PDF of the Array Pattern	44
3.2	Interval Arithmetic Methods	46
3.2.1	Interval Arithmetic Basics	46
3.2.2	The Cartesian IA (rIA) Method	47
3.2.3	The Circular IA (cIA) Method	48
3.2.4	The Polygonal IA (pIA) Method	50
3.2.5	The Matrix IA (mIA) Method	51
3.2.6	Numerical Examples for Performance Comparison	53
3.3	Summary	55
	References	58
4	Array Excitation Error Calibration	59
4.1	Calibration Methods	59
4.1.1	The Near-Field Scanning Probe Method	59
4.1.2	The Peripheral Fixed Probe Method	61
4.1.3	The Mutual Coupling Method	63
4.1.4	The Built-In Network Method	64
4.2	Calibration Signal Processing Techniques	65
4.2.1	Serial and Coherent Techniques	66
4.2.2	Serial and Noncoherent Techniques	67
4.2.3	Parallel and Coherent Techniques	70
4.2.4	Parallel and Noncoherent Techniques	70
4.2.5	Machine Learning Based Techniques	71
4.3	Array Element Failure Diagnosis	72
4.3.1	The Backward Transform Method	72
4.3.2	The Pattern Mapping Method	72
4.3.3	The Matrix Inversion Method	73
4.3.4	The Compressed Sensing/Sparse Recovering Method	74
4.4	Summary	77
	References	78
5	Some Key Design Issues for Satellite Ground Station Application	81
5.1	Array Geometry Design for Hemispherical Coverage	81
5.1.1	Pyramidal Frustum	81
5.1.2	Sphere	84
5.1.3	Geodesic Dome	86
5.1.4	Geodesic Dome with Cylinder Base	87
5.2	Multiple Beam Resource Management	89
5.2.1	The Traditional Work Flowchart of Reflector Antenna	89
5.2.2	The Novel Work Flowchart of Phased Array Antenna	90

5.3	Space-Ground Link Analysis	94
5.3.1	The Link Budget Equation	94
5.3.2	Case Study for Satellite-to-Ground Downlink	95
5.3.3	Case Study for Ground-to-Satellite Uplink	97
5.3.4	Design Tradeoffs for the Phased Array Antenna	98
5.4	Summary	100
	References	101

About the Authors

Guolong He received the B.S. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2008, and the Ph.D. degree in aerospace engineering from Tsinghua University, Beijing, China, in 2016. Since 2008, he has been with the Beijing Institute of Tracking and Telecommunication Technology, Beijing, China. He is a vital team member that is in charge of design, construction, and maintenance of the Chinese deep space stations. He first reveals the correlation between the transmit and reflect waves of the plasma sheath, and proposes a novel adaptive transmission method to alleviate the radio blackout problem for spacecraft reentry. Currently, he is the Project Manager of several satellite ground stations with both parabolic reflector antennas and phased array antennas, including the 4×35 m Kashi deep space antenna array.

Xin Gao received the B.S. degree in electronic engineering from Northwest University, Xi'an, Shaanxi, China, in 1995, the M.S. degree in optics engineering from Northwestern Polytechnical University, Xi'an, Shaanxi, China, in 1998, and the Ph.D. degree in optics engineering from Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, Shaanxi, China, in 2007. Currently, he is a Full Professor in Beijing Institute of Tracking and Telecommunication Technology, Beijing, China.

Rentian Zhang received the B.S. degree in electronic engineering from Huazhong University of Science and Technology, Wuhan, China, in 2009, and the M.S. degree in electronic engineering from Beihang University, Beijing, China, in 2012. Currently, he is an Assistant Professor in Beijing Institute of Tracking and Telecommunication Technology, Beijing, China.

Liangliang Sun received the B.S. degree in electronic engineering from Jilin University, China, in 2003. Currently, he is an Associate Professor in Beijing Institute of Tracking and Telecommunication Technology, Beijing, China.

Hui Zhou received both the B.S. and Ph.D. degrees in electronic engineering from National University of Defense Technology, Changsha, Hunan, China, in 2004 and 2010, respectively. Currently, he is an Associate Professor in Beijing Institute of Tracking and Telecommunication Technology, Beijing, China.

Chapter 1

Introduction



1.1 The New Era of Small Satellite Constellation

Since the beginning of the new millennium, the whole aerospace community has evolved rapidly into a new era by revolutionized small satellite technologies [1–3]. Traditionally, satellites were rather complex, heavy, bulk, and also expensive. Benefiting from the increasing availability of advanced hardware and technologies, satellites now can be made with much smaller size in the order of centimeters and much less weight of only a few kilograms. These small satellites are often classified according to their weights, i.e., mini (100–500 kg), micro (10–100 kg), nano (1–10 kg), pico (0.1–1 kg), and femto (<0.1 kg) [3]. A well-known type of small satellite is called CubeSat, which is made of several units with the fundamental size of $10 \times 10 \times 10 \text{ cm}^3$ cube (1U). For their low size and weight, small satellites have attracted not only traditional institutional players such as governments, universities, and space agencies but also intensive investment from private companies.

The main reason that making small satellites so popular is the dramatic overall cost reduction of a satellite project than ever, including shorter development time, advanced manufacturing process, and more affordable launch. First, widely commercial off-the-shelf (COTS) components for small satellites significantly reduce the development time and the total cost of a project. Even a complete satellite platform can be purchased directly from the commercial market now. Therefore, scientists who are not familiar with satellite engineering can keep their focus on mission goals, concept design, payload development, and data analysis. Compared to may over a decade from proposal to launch for a traditional satellite project, the average time for a small satellite project can be shorten to one year or even less. Second, advanced streamlined manufacturing processes accelerate the mass production of small satellites. Standard and highly automated production lines are set up to manufacture small satellites up to several satellites per day. Integration, assembly, and testing tasks are conducted with very high efficiency. Moreover, development of fully reusable rockets is the main reason to dramatically reduce the launch cost. Besides this, minimization

and standardization of small satellites also in a parallel effort to ease the deployment process. Utilizing spare launch capacity as second payloads, much more frequently shared launch chances can be gained access to the space. For a dedicated small satellite launch, more than one hundred small satellites can be deployed into space in a standardized container once.

Unlike the classical single big satellite in geostationary orbit (GEO) and medium earth orbit (MEO), small satellites in low Earth orbit (LEO) are typically organized as a satellite constellation. The satellite number of a LEO satellite constellation would be several hundred or even tens of thousands to realize a fully global coverage. Groups of satellites are deployed at the same altitude and inclination in several orbital planes, ensuring a continuous coverage of any location on Earth. There are a large number of advantages for LEO satellite constellation, such as small propagation delay, low propagation loss, and high fault tolerance as well as network robustness. With the low orbit up to two thousand kilometers, the one way propagation delay of LEO satellites is only up to 15 ms, which is commensurate with that of terrestrial links. Meanwhile, much less power is required to overcome propagation loss and other atmospheric attenuation. Antennas of user terminals can be made as portable or hand-set to meet the link budget easily. Moreover, such modularization and network structure of LEO satellite constellation provides resilience as a space network, which can never be possible with single large satellite.

The concept of LEO satellite constellation has the potential to open exciting innovation opportunities for space exploration and lead to unpredictable scientific breakthroughs. Table 1.1 lists some of recent LEO satellite constellation projects cov-

Table 1.1 Some recent small satellite constellation projects

Name	Number of satellites	Status	Goals
Lemur-2	100	Operational	Weather monitoring
OLFAR	50	Demonstration	Astronomy cosmic radiation observation
QB50	50	Completed	Earth's upper atmosphere observation
Flock	475	Completed	Earth observation
ExactEarth	9	Completed	Ship tracking and maritime situational awareness
Doves	200	Partially operational	Earth observation and remote sensing
StarLink	4,2000	Partially operational	Global broadband Internet access
OneWeb	6372	In deployment	Global broadband Internet access
Kuiper	3236	In deployment	Global broadband Internet access
TeleSat	298	In deployment	Global broadband Internet access

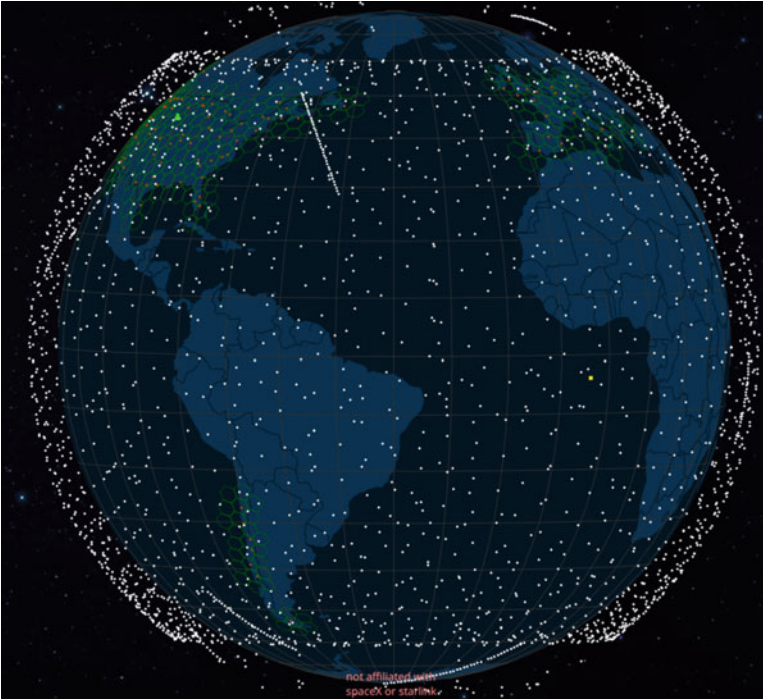


Fig. 1.1 Live Starlink satellite constellation coverage map, it contains 3447 active satellites on service by the end of April 2023. Figure reproduced with permission from <https://satellitemap.space>

ering a wide range of possible applications, including space exploration [4, 5], Earth remote sensing [6, 7], weather monitoring [8], navigation, worldwide telecommunications [9–11], and connectivity to Internet of Things (IoTs) [12, 13]. For example, the orbiting low-frequency antennas for radio astronomy (OLFAR) project is composed of a swarm of 50 or more nano-satellites to develop the large aperture required for very low frequency (below 30 MHz) observation of the cosmic noise [14]. The QB50 constellation is a recent model of international space cooperation contributed by many countries. The project includes 50 CubeSats, 40 of them for scientific exploration and the other 10 for in-orbit technology demonstration, which aims to perform simultaneous and distributed measurements of the Earth’s upper atmosphere [15]. Moreover, there is a booming trend of newly emerging LEO satellite constellation projects for providing global broadband Internet access from space, especially in remote areas that are hard to access via a terrestrial infrastructure. Interestingly, most of them are invested by giant Internet companies [16]. For example, SpaceX conducts the manufacturing of the satellites and launch them using their own Falcon vehicles. Figure 1.1 illustrates a live Starlink satellite constellation coverage map. As of April 2023, 4370 satellites of the Starlink constellation have been deployed, and 3447 of them are active and on service.

1.2 Simultaneous Multiple Satellite Support Requirement on Ground Station

For ordinary management and control operation of a LEO satellite constellation, the ground station is required to communicate with the satellites and provide telemetry, tracking and command (TT&C) services. It steers the antenna beam to follow the satellite during the brief interval from its rise above the horizon until it sets below the horizon. The period of a pass is relatively short for a LEO satellite, with a typical duration of a few minutes, depending on the relative position between the satellite and the ground station. Figure 1.2 shows a typical parabolic reflector antenna that is used as satellite ground station with high performance and medium cost. However, it can only support one satellite in its very narrow beam. This problem can be alleviated by increasing the number of reflector antennas at the mission control site. However, the total cost of land, infrastructure, and antennas forbid such a mode of ‘antenna farms’. Hence, reflector antennas is no longer proper for ordinary management and control operation of future LEO satellite constellation.

To solve the aforementioned problem, phased array antennas are considered as an attractive candidate for future LEO satellite constellation ground station. The primary reason is that phased array antennas can produce numbers of concurrent but independent beams by the beamforming network thus support multiple satellites simultaneously. Previously the cost of phased array antennas are very expensive, thus early phased array applications were confined to military and defense areas only. However, this has changed in recent years. Thanks to advancements in highly reliable solid state devices and Microwave Monolithic Integrated Circuits (MMIC)



Fig. 1.2 A 18 meter in diameter reflector antenna built by the authors' team at Qingdao, China

technologies, phased array antennas now become much affordable for commercial and industrial applications by increasingly matured mass production and dramatically reduced cost.

1.3 Advantages of Phased Array Antennas as Satellite Ground Station

There are many advantages of phased array antennas compared to parabolic reflector antennas as satellite ground station. As shown in Fig. 1.3, they can be generally divided into three categories, i.e., better performance, higher reliability, and lower cost.

1.3.1 Better Performance

Multiple beams. Reflector antennas usually have only one beam at once. Theoretically, phased array antennas can generate an arbitrary number of beams as desired by the beamforming network (this is restricted by the actual capacity of hardware). In the past, multiple analog microwave components are used to realize analog beamforming for phased array antennas. The connections and interfaces are very complicated, strict microwave parameters and specifications are required. With the advance of digital beamforming (DBF) technology, the beamforming process is much more straightforward to be implemented in the digital domain, where digitized data from antenna elements can be duplicated and combined easily.

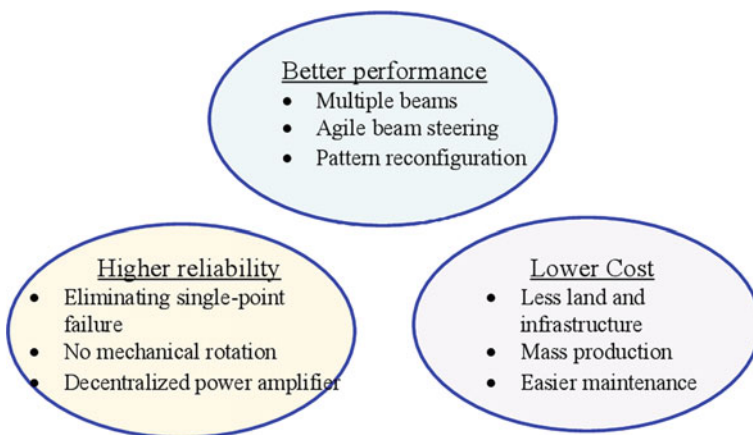


Fig. 1.3 Advantages of phased array antennas as satellite ground station

Agile beam steering. The reflector antennas rely on the mechanical driving system (including motors, gearbox, and servo) to steer the parabolic dish and the corresponding beam direction. Satellites in LEO move rapidly with respect to the Earth's surface, at up to 7.6 km/s for the altitude of 500 km, imposing a large beam shift speed requirement on the antenna mechanical driving system. Sometimes it would cause loss of tracking if the satellite moves too fast. A long preparation time is required if the reflector antenna changes to track another satellite with a large direction difference, possibly two or three minutes. For phased array antennas, the beam direction is electronically steered by setting the value of phase shifts of the antenna elements, with a much faster speed in the order of milliseconds.

Pattern reconfiguration. Due to the parabolic geometry of the reflector surface, the reflector antennas usually produce a pencil beam and have little way to reduce the sidelobe. The first sidelobe level (SLL) may be as high as -13 dB. Such high SLL sometimes even causes mis-tracking by the sidelobe rather than the main lobe. By amplitude tapering and phase control of the antenna elements, phased array antennas can control the SLL and present the nulls on the interfere directions conveniently.

1.3.2 Higher Reliability

Eliminating single-point failure. Some possible single-point failures exist in the reflector antennas. When one part of them fails, it will cause the whole ground station unable to work. However, when some antenna elements of the phased array antenna fail, it will only suffer part of performance degradation without loss of functionality, making them much more reliable than reflector antennas.

No mechanical rotation. As the structure of phased array antennas are stationary and no mechanical rotation is required, it is much more stable than that of parabolic reflector antennas, whose motors and servo will suffer friction and loss during daily use.

Decentralized power amplifier. For reflector antennas, centralized high power amplifiers are used to boost the output signal. Usually, the capability of several hundreds of Watts is required for a typical satellite ground station. While in phased array antennas, each antenna element has its own power amplifier of only several Watts. Such a decentralized mode is much more reliable than the centralized one.

1.3.3 Lower Cost

Less land and infrastructure. To support multiple missions with reflector antennas, a large area of the site and many antenna bases need to be constructed, along with concrete roads, water pipes, electrical supply, and optical fibers, etc. Since phased array antennas can produce multiple beams simultaneously, it reduces the need for land and infrastructure construction to meet the same requirements.

Mass production. Over the years, the high cost and complexity of phased array antennas were the major obstacles to use in commercial applications. Recently, the cost of phased array antennas has been reduced drastically by leveraging the advanced manufacturing technology, volume production, and adaptation of COTS components. Significant cost reduction of the transmit/receive (T/R) module is contributed by the advanced integration technologies. The area of the multilayer RF boards and the number of connectors and cables are dramatically reduced to save the cost.

Easier maintenance. The maintenance and impairment processes for reflector antennas are complicated. Sometimes it requires a down time, dedicated tools (like derrick, crane, and scaffold), and also crews of highly skilled workers. However, if some antenna elements of the phased array antenna fail, they can be easily non-stop replaced without down time. The maintenance process is rather simple, one worker with basic training is usually enough to deal with the ordinary maintenance work. A much lower life cycle cost and budget saving can be expected.

1.4 Outline of This Book

Reflector antennas have been used as satellite ground station for a long history [17, 18]. The phased array antennas, although extensively used for radar applications, are still in the infancy for this purpose. Before its truly operational in practice, there are some technical issues and challenges that must be carefully considered. To the authors' personal opinion, it has come to a point that phased array antennas can replace reflector antennas as future satellite ground station. However, the aerospace industry seem to still hesitate to do so. Even some of the most recent LEO satellite constellation projects still use traditional reflector antennas as their satellite ground stations. This book aims to provide a tutorial-like review by providing researchers and engineers with summarization of some key points in a convenient way. This book is mainly based on the authors' experience gained in the processes of design and construction of several phased array antennas. Because there are so many good introductory books existed in the open literature [19–23], we try to present some considerations for engineering practice only.

The rest of this book is organized as follows. First, some basic principles and theories about phased array antennas are shortly presented in Chap. 2, including the architecture evolution, major components, and array pattern synthesis. Then array excitation errors are treated in detail in Chaps. 3 and 4, with probabilistic and interval arithmetical analysis methods, state-of-the-art calibration schemes, and some innovative signal process techniques in a comparative manner. In Chap. 5, some designing issues and key technologies required as satellite ground station application are studied, such as array geometry design for hemispherical coverage, multiple beam resource management, and space-ground link analysis.

References

1. R.M. Millan, R. Steiger, M. Ariel, et al., Small satellites for space science A COSPAR scientific roadmap. *Adv. Space Res.* **64**, 1466–1517 (2019)
2. F. Davoli, C. Kourgiorgas, M. Marchese, A. Panagopoulos, F. Patrone, Small satellites and CubeSats: survey of structures, architectures, and protocols. *Int. J. Satell. Commun. Netw.* 1–17 (2018)
3. M.N. Sweeting, Modern small satellites-changing the economics of space. *Proc. IEEE* **106**(3), 343–361 (2018)
4. S. Bandyopadhyay, R. Foust, G.P. Subramanian, S.J. Chung, F.Y. Hadaegh, Review of formation flying and constellation missions using nanosatellites. *J. Spacecr. Rock.* **53**(3), 567–578 (2016)
5. S. Marcuccio, S. Ullo, M. Carminati, O. Kanoun, Smaller satellites, larger constellations: trends and design issues for earth observation systems. *IEEE Aerosp. Electron. Syst. Mag.* **34**(10), 50–59 (2019)
6. G. Curzi, D. Modenini, P. Tortora, Large constellations of small satellites: a survey of near future challenges and missions. *Aerospace* **7**(133) (2020)
7. E. Peral, E. Im, L. Wye, et al., Radar technologies for earth remote sensing from CubeSat platforms. *Proc. IEEE* **106**(3), 404–418 (2018)
8. O. Kodheli, E. Lagunas, N. Maturo, et al., Satellite communications in the new space era: a survey and future challenges. *IEEE Commun. Surv. Tutor.* **23**(1), 70–109 (2021)
9. B. Di, L. Song, Y. Li, H.V. Poor, Ultra-dense LEO: integration of satellite access networks into 5G and beyond. *IEEE Wirel. Commun.* **26**(2), 62–69 (2019)
10. I. Leyva-Mayorga, B. Soret, M. Roper, et al., LEO small-satellite constellations for 5G and beyond-5G communications. *IEEE Access* **8**, 184955–184964 (2020)
11. N. Saeed, A. Elzanaty, H. Almorad, H. Dahrouj, T.Y. Al-Naffouri, M.-S. Alouini, CubeSat communications: recent advances and future challenges. *IEEE Commun. Surv. Tutor.* **22**(3), 1839–1862(2020)
12. M. De Sanctis, E. Cianca, G. Araniti, I. Bisio, R. Prasad, Satellite communications supporting internet of remote things. *IEEE Internet Things J.* **3**(1), 113–123 (2016)
13. Z. Qu, G. Zhang, H. Cao, J. Xie, LEO satellite constellation for internet of things. *IEEE Access* **5**, 18391–18401 (2017)
14. M.J. Bentum, M.K. Verma, R.T. Rajan, et al., A roadmap towards a space-based radio telescope for ultra-low frequency radio astronomy. *Adv. Space Res.* **65**, 856–867 (2020)
15. E. Gill, P. Sundaramoorthy, J. Bouwmeester, B. Zandbergen, R. Reinhard, Formation flying within a constellation of nano-satellites: the QB50 mission. *Acta Astro.* **82**, 110–117 (2013)
16. I. Portillo, B. Cameron, E. Crawley, A technical comparison of three low earth orbit satellite constellation systems to provide global broadband. *Acta Astron.* **159**, 123–135 (2019)
17. S. Rao, L. Shafai, S.K. Sharma, *Handbook of Reflector Antennas and Feed Systems* (Artech House, Boston, 2013)
18. Y. Rahmat-Samii, R. Haupt, Reflector antenna developments: a perspective on the past, present and future. *IEEE Antennas Propag. Mag.* **57**(2), 85–95 (2015)
19. R.J. Mailloux, *Phased Array Antenna Handbook*, 3rd edn. (Artech House, Boston, 2017)
20. R.L. Haupt, *Antenna Arrays: A Computational Approach* (Wiley, New York, 2010)
21. R.L. Haupt, Y. Rahmat-Samii, Antenna array developments: a perspective on the past, present and future. *IEEE Antennas Propag. Mag.* **57**(1), 86–96 (2015)
22. A.J. Fenn, D.H. Temme, W.P. Delaney, W.E. Courtney, The development of phased array radar technology. *Lincoln Labor. J.* **12**(2), 321–340 (2000)
23. B. Sadhu, X. Gu, A. Valdes-Garcia, The more (antennas), the merrier: a survey of silicon-based mm-wave phased arrays using multi-IC scaling. *IEEE Microw. Mag.* **20**(12), 32–50 (2019)

Chapter 2

Phased Array Antenna Basics

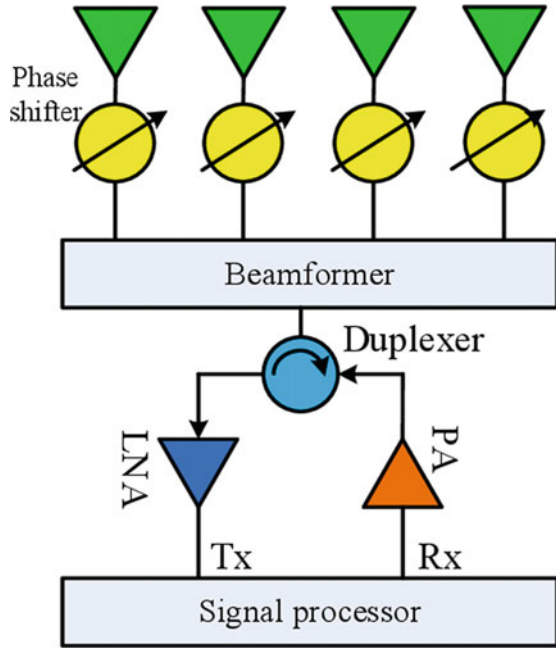


2.1 Evolution of Phased Array Antenna Architecture

The architecture of the phased array antennas has undergone a steady evolution over the years [1–3]. In the first generation of phased array antennas, as shown in Fig. 2.1, the antenna aperture was divided into a large number of antenna elements to realize beam agility, each with an electronically controlled phase shifter. However, the centralized high-power transmit and receive amplifiers were still used. In the 1980s, thanks to advancements in highly reliable solid state devices and MMIC technologies, solid-state transmit amplifiers were distributed and moved much closer to the antenna elements. As shown in Fig. 2.2, all the transmit amplifiers, low-noise receive amplifiers, phase shifters, and attenuators are integrated together and known as transmit/receive (T/R) module. Analog microwave components are used to realize analog beamforming. It has the advantages of wide bandwidth and low power consumption. However, connections and interfaces for this type of beamforming are very complicated, strict microwave parameters and specifications are required. To form M simultaneous beams, the phase shifters and analog beamformers must be implemented in M times. Thus phased array antennas with analog beamforming were inherently constrained by the front-end beamforming electronics. As the number of beams increases, so does the analog components and the cost of a phased array antenna. As a result, it's very difficult to generate multiple independent beams by analog beamforming in actual application.

By the end of the 1990s, subarray level digital (or hybrid analog-digital) beamforming was extensively investigated, which performs partial beamforming digitally and maintains the low overall cost. As shown in Fig. 2.3, the antenna elements on a subarray are first combined by a dedicated analog beamforming network. Using a second-stage digital beamforming network, the subarray outputs are weighted and summed digitally to produce a set of digital beams. After entering the 21st century, there is a trend of fully digital beamforming (DBF) as shown in Fig. 2.4, signals are

Fig. 2.1 The passive architecture of a phased array antenna



digitalized at the element level. The T/R module still performs the usual amplification and filtering operations. However, phase shifting, signal distribution, and core beamforming are shifted to the digital domain. Now a phased array antenna is a comprehensive system combining RF engineering in the front and massive data flow and computational power on the back end. DBF is implemented by the combination of antenna array technology and modern digital signal processing technology. The most significant advantage of DBF is its ability to generate a large number of independent beams in the digital domain. Increasing the number of beams is as simple as adding some more beamforming processing processes in the program, thus simplify the physical design of a phased array antenna while maintaining performance.

2.2 Major Components of a Phased Array Antenna

2.2.1 Antenna Element

The antenna elements are used to transmit/receive electromagnetic waves from the T/R module to the free space. They are usually arranged in lattice or conformal with the surface to radiate the energy with high efficiency. The satellite ground station is required to have dual polarization capability that can support left-hand circular polarization (LHCP) and right-hand circular polarization (RHCP) simultaneously.

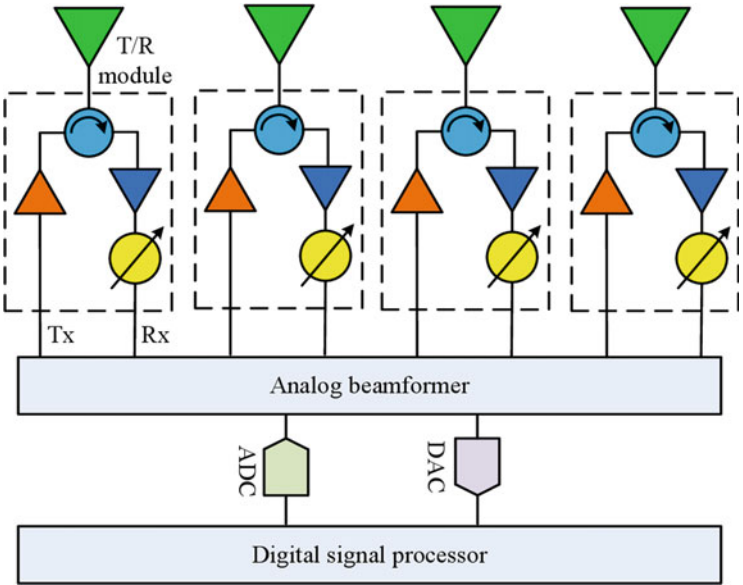


Fig. 2.2 The active analog architecture of a phased array antenna

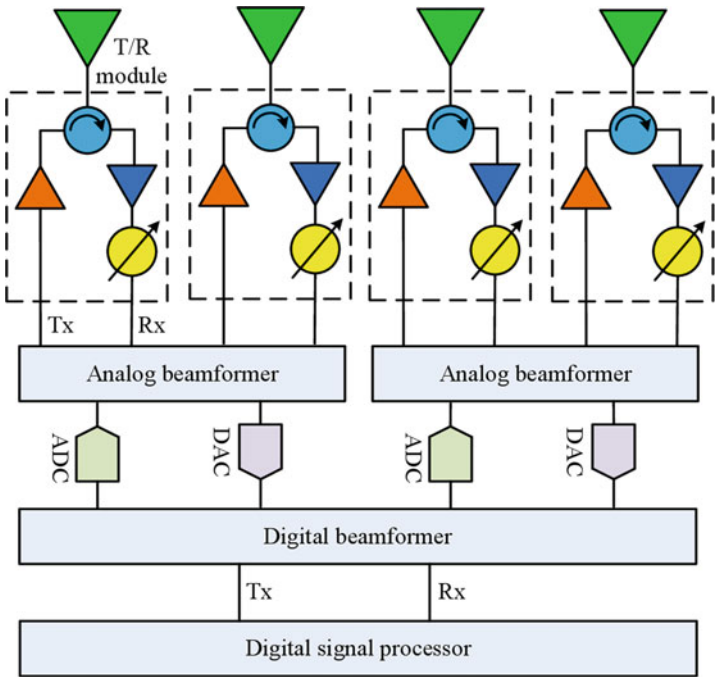


Fig. 2.3 The subarray level digital beamforming architecture of a phased array antenna

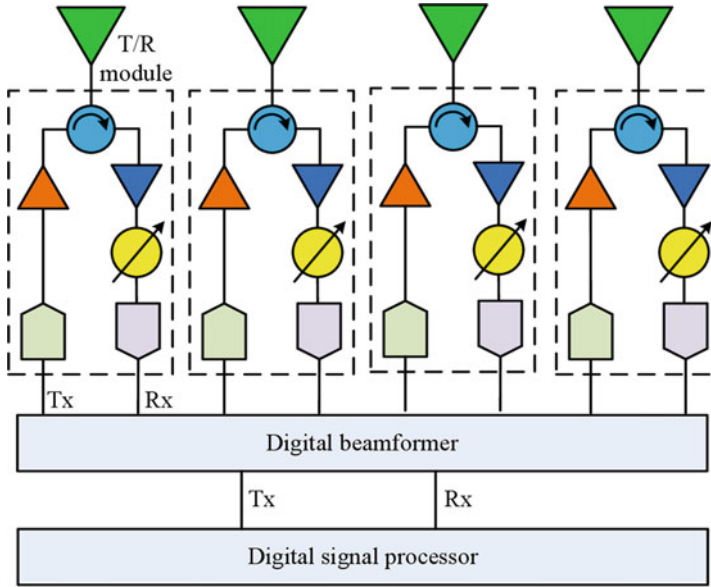


Fig. 2.4 The element level digital beamforming architecture of a phased array antenna

However, high isolation between cross-polarization are not easily achieved in dual-polarized antenna element. In [4], cross-coupling mechanisms between orthogonal polarizations are discussed according to cross coupling between radiation elements. The most classical dual polarization antennas, such as crossed dipole and crossed slot, is to duplicate a radiation element and position it at a right angle to the original one [5]. Recently, there are also very popular with patch antennas excited with two orthogonal modes, each corresponding to an individual polarization [6].

2.2.2 T/R Module

The T/R module is usually viewed as the most important part and accounts for approximately half of the overall cost of a phased array antenna. Over the past few decades, the T/R modules have undergone significant changes in terms of material, design, and layout technologies [7, 8]. Figure 2.5 illustrates the typical layout of a modern T/R module. It contains several basic components such as duplexer, filter, phase shifter, power amplifier, low noise amplifier (LNA), mixer, local oscillator (LO), analog-to-digital converter (ADC), and digital-to-analog converter (DAC).

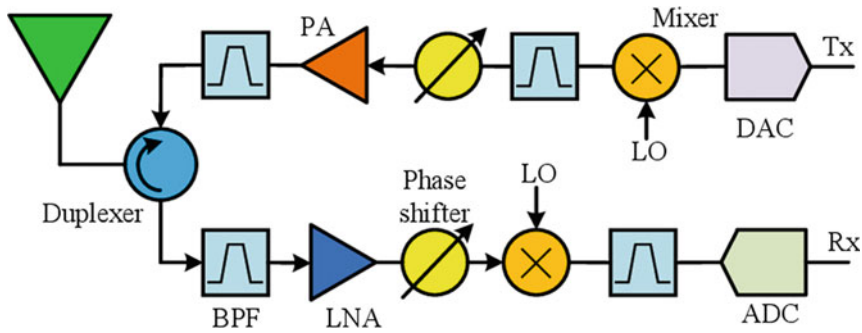


Fig. 2.5 The typical layout of a modern T/R module

2.2.2.1 Duplexer

As a satellite ground station, full duplex communication is required for simultaneous uplink and downlink with the satellite. This is achieved by employing a component called the duplexer (or circulator). It isolates the Rx path from the Tx path with interference suppression while permitting them to share a common antenna element.

2.2.2.2 Filter

A filter is a frequency selective component that allows to pass signals of certain frequency band while suppressing signals outside the band of interest. For example, the filter follows the duplexer in the Rx path is a bandpass filter (BPF), to pass the received signal and attenuate the transmitted signal leaked and coupled from the Tx path. This is very important especially for satellite communication application, which works as frequency multiplex mode rather than time multiplex mode in the radar application.

Figure 2.6 illustrates the transfer function of a typical BPF. Ideally, it's desired to pass a range of frequencies and completely eliminate those signals outside the band. The insertion loss of the filter means the level of attenuation in the passband. A gradual transition region exists between the passband and the suppressed region. The goal in filter design is to provide sharp transitions while maintaining an acceptable insertion loss.

2.2.2.3 Power Amplifier

The power amplifier is to boost the input signal before it is transmitted to the free space via the antenna element. Due to the nonlinearity of the power amplifier, it generates various distortions of the output signal, containing high-order harmonic

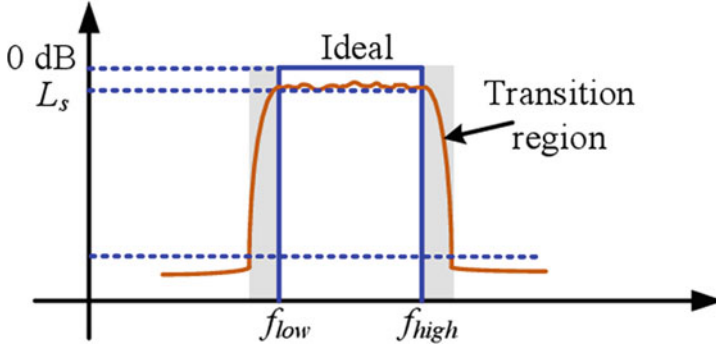


Fig. 2.6 The transfer function of a typical BPF

distortions and intermodulation products. Suppose the transfer function of the power amplifier has the following expression,

$$v_{out}(t) = k_0 + k_1 v_{in}(t) + k_2 v_{in}^2(t) + k_3 v_{in}^3(t) + \dots \quad (2.1)$$

where k_0, k_1, k_2, k_3 are constants.

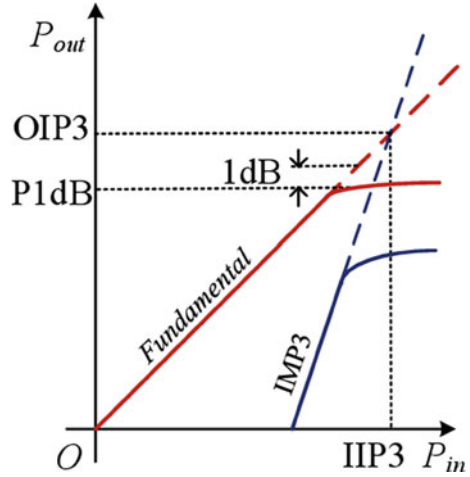
Consider the input signal is comprised of two tones in different carrier frequencies,

$$v_{in}(t) = a_1 \cos(w_1 t + \phi_1) + a_2 \cos(w_2 t + \phi_2) \quad (2.2)$$

Then the output signal is derived as,

$$\begin{aligned} v_{out}(t) = & k_0 + 0.5k_2a_1^2 + 0.5k_2a_2^2 \\ & + (k_1a_1 + 0.75k_3a_1^3 + 1.5k_3a_1a_2^2) \cos(w_1t + \phi_1) \\ & + (k_1a_2 + 0.75k_3a_2^3 + 1.5k_3a_1^2a_2) \cos(w_2t + \phi_2) \\ & + 0.5k_2a_1^2 \cos(2w_1t + 2\phi_1) \\ & + 0.5k_2a_2^2 \cos(2w_2t + 2\phi_2) \\ & + k_2a_1a_2 \cos((w_1 + w_2)t + (\phi_1 + \phi_2)) \\ & + k_2a_1a_2 \cos((w_1 - w_2)t + (\phi_1 - \phi_2)) \\ & + 0.25k_3a_1^3 \cos(3w_1t + 3\phi_1) \\ & + 0.25k_3a_2^3 \cos(3w_2t + 3\phi_2) \\ & + 0.75k_3a_1^2a_2 \cos((2w_1 + w_2)t + (2\phi_1 + \phi_2)) \\ & + 0.75k_3a_1a_2^2 \cos((2w_1 - w_2)t + (2\phi_1 - \phi_2)) \\ & + 0.75k_3a_1a_2^2 \cos((w_1 + 2w_2)t + (\phi_1 + 2\phi_2)) \\ & + 0.75k_3a_1a_2^2 \cos((w_1 + 2w_2)t - (\phi_1 + 2\phi_2)) + \dots \end{aligned} \quad (2.3)$$

Fig. 2.7 The typical input and output relationship of a power amplifier



From Eq. (2.3), it shows that, due to the quadratic and cubic terms, the output signal gives rise to second- and third-order harmonic distortions (HD2 and HD3), and intermodulation products (IMP2 and IMP3). Among these, the third-order intermodulation product (IMP3) is the most important, because two interfering signals spaced in the frequency domain such that their IMP3 will fall in the frequency band of interest.

Figure 2.7 illustrates the typical input and output power relationship of a power amplifier. the fundamental line have a slope of 1, which means a 1 dB increase in input power causes a 1 dB increase in the output power. However, the third-order products have a slope of 3. The third-order intercept point (OIP3 referred to the output and IIP3 referred to the input) is used to quantify nonlinearity of the amplifier. It occurs when the ideal linear IMP3 equals the ideal linear, uncompressed fundamental output. The 1 dB compression point (P1dB) is another descriptive measure of the amplifier linearity. It occurs when the actual output power is 1 dB less than it would be in the ideal linear value. It is a figure of merit for the maximum power for linear operation.

2.2.2.4 LNA

The main function of the LNA is to amplify the signal received from the antenna element and then sent it to subsequent components, e.g., mixers, filters. Each component in the Rx path contributes additional noise over the signal. The signal to noise ratio (SNR) at the output SNR_{out} of a given device will be less than the SNR at the input SNR_{in} . This noise factor, F , is characterized as the loss of SNR caused by the device,

$$F = \frac{SNR_{in}}{SNR_{out}} \quad (2.4)$$

The overall noise factor of a cascaded set of components is given by

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \dots \quad (2.5)$$

where F_i is the noise factor and G_i is the gain of the i th component.

What this equation indicates is that the first element in the RF cascade dominates the resulting noise figure for the system. This indicates that all passive components (cables and filters) prior to the first amplifier will have a negative impact on the noise figure. Likewise, components that follow a high gain amplifier in the cascade will have a minimal effect on the overall noise figure. For a high performance LNA, it has a relative low noise factor and high gain, thus the total noise factor is mainly determined by the first component LNA. To achieve good link sensitivity, the LNA has to be placed as close as possible to the antenna element.

2.2.2.5 Phase Shifter

Phase shifter is used to change the phase of the signal. An ideal phase shifter should have low and near-equal insertion loss for all phase states. Phase shift is a constant with respect to frequency, hence phase shifter is usually used in the narrow-band beamforming implementation [9]. For wideband beamforming, true time delay (TTD) units are required, where time delay is a linear function of frequency [10].

2.2.2.6 Mixer and Local Oscillator

A local oscillator (LO) generates the desired frequency using a phase lock loop (PLL) combined with the crystal oscillator. In the Rx path, the mixer is used to downconvert the input signal from the RF band to the intermediate frequency (IF) band. While in the Tx path, the mixer is used to upconvert the output signal from IF band to RF band. The output of a mixer is the sum and difference frequencies of two input signals. The second BPF following the mixer is used to select the desired difference frequency component and minimize any intermodulation products from the mixing. This heterodyne style of carrier frequency conversion is well proven in engineering practice and has been used for many years [11]. However, many filters are required to keep the challenging frequency bandwidth. Nowadays, the mixers and the frequency conversion process are not mandatory. The direct sampling approach allows direct RF sampling and achieves large input bandwidth [12].

2.2.2.7 ADC and DAC

The ADC is used to convert the signal from the analog domain to digital samples, while the DAC implements the reverse process [13, 14]. The ADC dynamic range is determined by the effective number of bits (ENOB),

$$SNR(dB) = 6.02 * ENOB + 1.76 \quad (2.6)$$

which gives the maximum dynamic range that can be achieved if the signal level is at the full-scale value of the ADC. Notice that the ENOB is used in the calculation because it accounts for the nonideal factors in the ADC, which limits the number of bits achieved.

Sample clock synchronization for ADC and DAC must be carefully designed to minimize the time difference. For analog beamforming and subarray level digital beamforming, signals from several antenna elements are combined first and then digitalized to save the number of ADCs and DACs. With element level digital beamforming, every antenna element has its own ADC and DAC. Such high level of digitization allows increasing functionality and better performance compared to their predecessors.

2.2.3 Beamforming Network

The beamforming network is to combine signals from multiple antenna elements. It forms the main beam in the desired direction of the signal, and places nulls in the direction of interferences or clutters.

For satellite tracking, it needs to determine the angular position of the target. They are based on comparison of the so-called sum and difference signals. Towards this aim, the phased array antenna is required to provide both sum and difference patterns for estimating the azimuth and elevation coordinates to identify the target direction. The former has one main lobe along the target direction and the latter exhibits a null in the same direction. The basic principle of monopulse tracking is shown in Fig. 2.8. The sum and two difference patterns are obtained by subdividing the array into four symmetric quadrants,

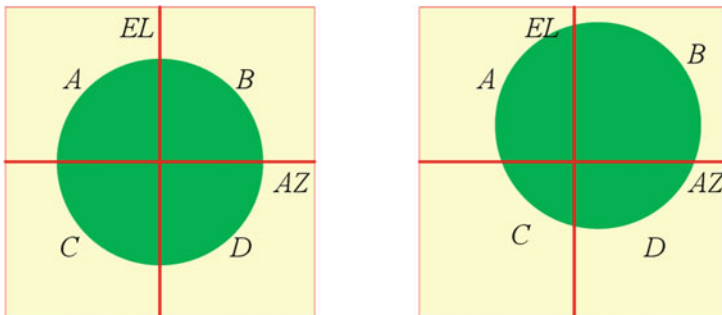


Fig. 2.8 Schematic view of monopulse tracking. **a** the target is located at the center of the beam, and **b** the target is off the center of the beam

$$\begin{aligned}
F_{sum} &= A + B + C + D \\
F_{diff}^{AZ} &= (A + C) - (B + D) \\
F_{diff}^{EL} &= (A + B) - (C + D)
\end{aligned} \tag{2.7}$$

where the sum mode is generated when all quadrants are excited with the same phase, while the two difference patterns are obtained by adding a phase shift to the excitations of a couple of quadrants.

As shown in Fig. 2.8a, when the target is located at the center of the beam, the quadrants will receive the same target echo because they are equidistant from the target. Azimuth and elevation differences will ideally give zero values. If the target is shifted from the center as shown in Fig. 2.8b, then quadrants A and C receive different target echos compared to those received by quadrants B and D. The azimuth difference now gives a nonzero value. The situation is the same for the elevation difference pattern.

If the sum and difference beam patterns are synthesized independently, it results in different excitation coefficients. This requires the adoption of two separate beam-forming networks, thus unavoidably increasing the complexity and cost. These difficulties can be alleviated if both sum and difference beams share full or partial common antenna elements, by aggregating the elements into clusters or sub-arrays where each one is controlled by single amplitude coefficients. A set of numerical examples is reported and discussed to show the versatility and effectiveness of the proposed approach [15]. This can be achieved through various methods, for example as discussed in the Sect. 2.3.

2.2.4 Digital Signal Processor

Digital signal processors usually adopt the general purpose hardware platform by using field-programmable gate array (FPGA) and digital signal processing (DSP) technologies, and the advanced software defined radio (SDR) architecture. Sophisticated functions are implemented by software to realize satellite communication and TT&C service. For satellite communication, digital signal processors handle various modulation and coding schemes to interchange the information with the satellite. By receiving and decoding the telemetry (TM) data from the satellite, one can know the status, configuration, and health state of the satellite. Telecommands (TC) are sent from the ground station to the satellite for altitude adjustment, mission operation control, and onboard software upgrade.

For satellite tracking, the speed, range, and angle of the satellite are the most common measurement parameters. The speed of the satellite can be easily estimated from the Doppler frequency of the carrier signal. By sending a series of tones [16] or a pseudo-noise (PN) code [17] modulated on the carrier signal, the distance between the satellite and the ground station is measured equal to the speed of light times the travel round time. The angular position of the satellite is obtained by the monopulse

tracking technique based on the comparison of the so-called sum and difference signals.

2.2.5 Cooling Equipment

With the progress of solid-state microwave technologies, the radiated power of amplifiers increase steadily. Meanwhile, miniaturization of T/R modules results in very compact size. Both of these two trends lead to increased heat generation and a rise in the working temperature of the T/R module. These would negatively affect the performance and reliability of the module. Experimental studies have shown a correlated relationship between the failure rate of electronic components with the working temperature of the T/R module [18]. A model coupling the temperature variation and performance of the phased array is also proposed and verified in [19].

In order to keep the reliable operation of the T/R module, it is necessary to dissipate the generated heat and below the temperature at a reasonable value. Air and liquid cooling are two commonly used techniques to cool the T/R module. Although with simpler equipment and relatively low cost, air cooling has limited capacity for heat dissipation. For liquid cooling, the T/R module is cooled by pumping cold liquid through cooling channels of the cold plate, which is made of high performance heat conducting material and installed closely with heat generating components [20]. The cooling channel guides the cold fluid through first the region with active components (power amplifiers) then the low-heat region, thus transforms the heat generated by the T/R module. The cost of a liquid cooling system is relatively expensive because of more complicated system with pumps, pipe networks, non-spill connectors, and heat exchanging units.

2.3 Array Pattern Synthesis

The phased array antennas generate desired array patterns such as nulls toward the directions of jamming signals or clutter returns, amplitude and or phase tapering for sidelobe reduction, subarray for beamforming simplification, thinned array, etc. Antenna array synthesis is a process to trade off between different aspects of performances in order to meet specified design goals, we can not improve one aspect significantly without scarifying another. This section briefly introduces the basics of array factor analysis, and presents several popular methods for synthesizing the array pattern.

2.3.1 Array Factor Basics

Considering a phased array antenna with N isotropic elements, each element has the position $\vec{r}_n$ and the complex excitation $w_n = a_n \exp(jp_n)$, where a_n and p_n are the amplitude and phase of w_n , respectively, $n = 1, \dots, N$. The array radiation pattern or the array factor $f(\theta, \phi)$ can be mathematically expressed as follows,

$$\begin{aligned} f(\theta, \phi) &= \sum_{n=1}^N f_n(\theta, \phi) \\ &= \sum_{n=1}^N a_n \exp(jk\vec{r}_0(\theta, \phi) \cdot \vec{r}_n + jp_n) \\ &= \sum_{n=1}^N a_n \exp(jb_n) \end{aligned} \quad (2.8)$$

where $k = 2\pi/\lambda$ is the wavenumber, $\vec{r}_0(\theta, \phi) = [\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta]^T$ is the unit vector of the observation direction (θ, ϕ) , $f_n(\theta, \phi)$ is the array factor contributed by the n th antenna element and b_n is its phase.

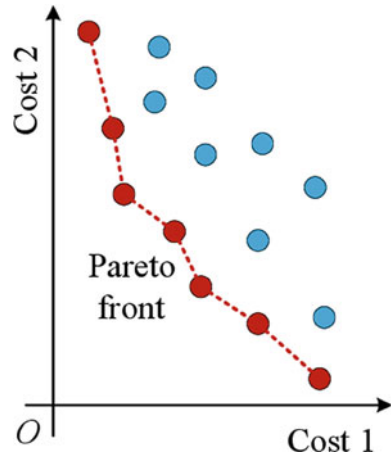
Linear and planar arrays are the most common array configurations. For linear case, $\vec{r}_n = [0, 0, z_n]^T$, then the array factor is represented as,

$$f(\theta, \phi) = \sum_{n=1}^N a_n \exp(jkz_n \cos \theta + p_n) \quad (2.9)$$

For planar case, $\vec{r}_n = [x_n, y_n, 0]^T$, then the array factor is represented as,

$$f(\theta, \phi) = \sum_{n=1}^N a_n \exp(jk(x_n \sin \theta \cos \phi + y_n \sin \theta \sin \phi) + p_n) \quad (2.10)$$

From Eq. (2.8), the array factor is a function of several parameters, including the element amplitudes, phases, positions, and the working frequency. Apparently, array pattern synthesis is an optimization process, which can be single-objective or multi-objective. If there are multiple goals considered in the synthesis process, then these goals can be summed with weight to formulate as a single cost function. However, it is usually very difficult to know a priori what values to choose for the weights, and this approach does not provide any insights into potential tradeoffs. In contrast, multi-objective optimization can show right balance between multiple conflicting objectives, and allow the engineer to choose the best design for a given application. Its ultimate goal is to efficiently find a set of points called the Pareto front with a set of solutions [21]. Every solution in the Pareto front is nondominated, which means that improving one objective leads to degradation of at least one other objective (Fig. 2.9).

Fig. 2.9 The Pareto front

2.3.2 Analytical Algorithms for Array Amplitude Tapers

If all antenna elements have uniform weights, the highest gain and the narrowest half-power beamwidth (HPBW) can be achieved, but unequal and relatively high SLL will be suffered. Often it is desirable to lower the highest sidelobe, at the expense of raising the lower sidelobes. There are many approaches to synthesize amplitudes of the array elements that reduce the SLL, analytical algorithms are one of the simplest kinds of methods among them, which can be directly calculated from mathematical expression easily. The binomial array that follows the binomial coefficients as the array amplitude taper can produce the array pattern with no sidelobe. However, errors in real practice will always produce some sidelobe. Moreover, the taper efficiency for the binomial array is very low. Hence, it's more desirable to balance the sidelobe level and the taper efficiency.

The optimal sidelobe level (for a given beamwidth) will occur when the sidelobes are all equal in magnitude. For uniformly spaced linear arrays, a popular weighting method named the Dolph-Chebyshev taper, the SLL of the array can be set to a specified level by mapping the array factor to a Chebyshev polynomial, and the minimum possible null-null beamwidth can be obtained. Figure 2.10 illustrates a 10-element, half-wavelength equally spaced, Dolph-Chebyshev tapered linear antenna array with $SLL^{\text{ref}} = -20$ dB, and the corresponding excitation amplitudes are $a_1 = a_{10} = 0.642$, $a_2 = a_9 = 0.594$, $a_3 = a_8 = 0.778$, $a_4 = a_7 = 0.921$, and $a_5 = a_6 = 1.000$. Note that the increasing amplitude weight at the array edge ($a_1 > a_2$), hence this taper method is usually used for small arrays only.

Although the Chebyshev window has the narrowest possible mainlobe for a specified sidelobe level, the Taylor window allows one to make tradeoffs between the mainlobe width and the sidelobe level. The Taylor taper reduces the first $n-1$ sidelobes at a specified level by moving the nulls on either side of the main beam, while other remaining sidelobes decrease at the same rate as the corresponding sidelobes

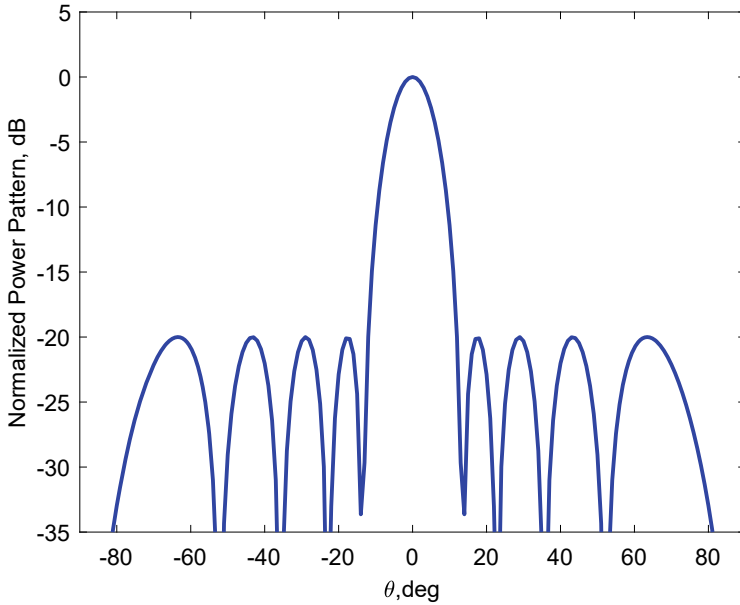


Fig. 2.10 A 10-element, half-wavelength equally spaced, Dolph-Chebyshev tapered linear antenna array with $SLL^{\text{ref}} = -20$ dB

in a uniform array. Figure 2.11 illustrates the same linear antenna array with Taylor tapered $SLL^{\text{ref}} = -20$ dB and $\bar{n} = 2$, and the corresponding excitation amplitudes are $a_1 = a_{10} = 0.542$, $a_2 = a_9 = 0.629$, $a_3 = a_8 = 0.771$, $a_4 = a_7 = 0.913$, and $a_5 = a_6 = 1.000$. The Taylor taper avoids edge discontinuities, and the sidelobes decrease monotonically.

2.3.3 Genetic Algorithm

The aforementioned analytical algorithms can synthesize amplitude weights for linear or grid planar array only. For other geometries or more complicated array synthesis problems, it must be synthesized by other numerical methods. Two very popular evolutionary algorithms, genetic algorithm (GA) is shortly introduced here, and particle swarm optimization (PSO) is introduced in next subsection.

Genetic algorithm (GA) is an optimization technique based on Charles Darwin's theory of natural selection. It has been one of the most widely used optimization algorithms in modern nonlinear optimization [22]. Let us take an thinned array pattern synthesis problem as an example. Since each element of the phased array has a state 'on' or 'off', array pattern synthesis can be realized via state adjustments of the array

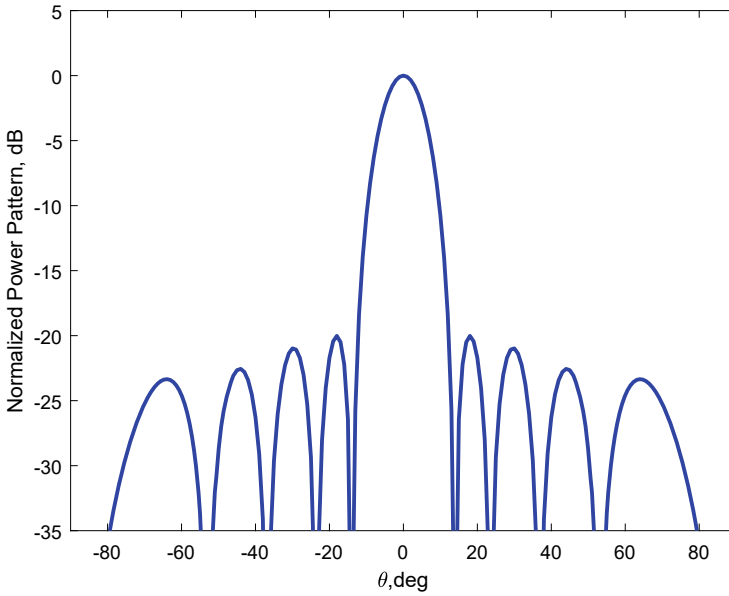


Fig. 2.11 A 10-element, half-wavelength equally spaced, Taylor tapered linear antenna array with $SLL^{ref} = -20$ dB and $\bar{n} = 2$

elements, which allows simplified uniform amplitude excitation feed networks while avoiding additional hardware and cost.

Figure 2.12 illustrates a general flowchart of GA algorithm. It uses binary sequences to encode parameters as genes, and a set of genes is a chromosome. For example, a chromosome with N_{gene} bits is represented as,

$$chromosome = [b_1 b_2 \dots b_{N_{gene}}]. \quad (2.11)$$

where b_n is the n th bit of the chromosome, and ‘1’ and ‘0’ represent state ‘on’ and ‘off’ of the n th array element, respectively.

The GA algorithm begins with defining a fixed number of chromosomes with random genes N_{chrom} , which is called population sometimes. The fitness value $fitval$ of each chromosome, here the fitness value is the sidelobe level in sidelobe region, is evaluated,

$$fitval(chrom) = SLL \text{ in sidelobe region} \quad (2.12)$$

Then these chromosomes are ranked by the fitness value. Under evolutionary biology of natural selection, chromosomes with low fitness are discarded, and superior chromosomes are survived,

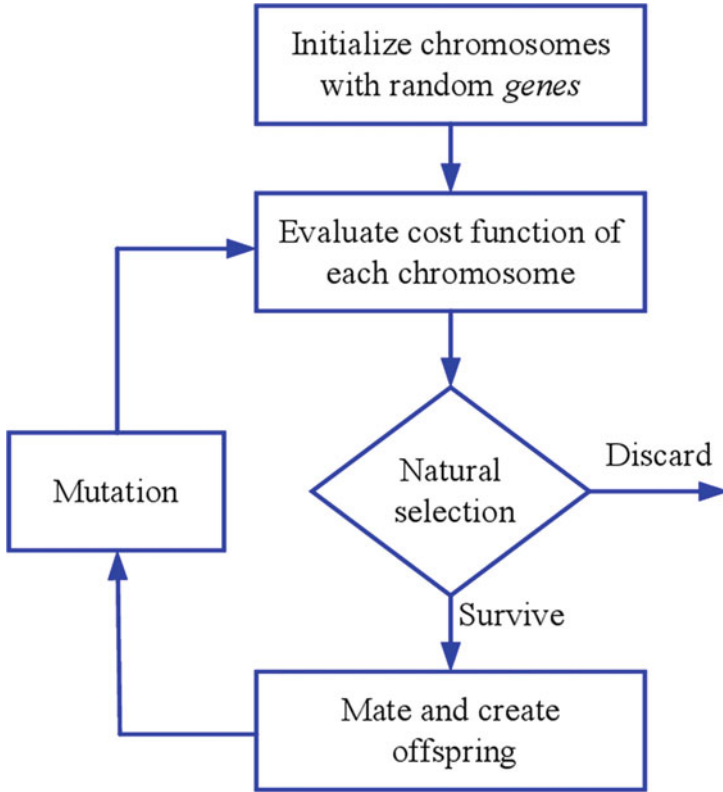


Fig. 2.12 The general flowchart of the GA

$$chrom_i = \begin{cases} \text{survived} & \text{if } fitval(chrom_i) < \text{mean}[fitval(\text{all chromosomes})] \\ \text{discarded} & \text{if } fitval(chrom_i) \geq \text{mean}[fitval(\text{all chromosomes})] \end{cases} \quad (2.13)$$

Only the first N_{surv} chromosomes are survived in the ascending way.

$$fitval(chrom_i) < fitval(chrom_j) \quad \text{for } i < j \quad (2.14)$$

The survived chromosomes become as parents. Usually, the chromosomes with better fitness value have higher probability to be chosen. A simple strategy can be adopted,

$$p(chrom_i) = \frac{N_{sur} + 1 - i}{\frac{(N_{sur} + 1)N_{sur}}{2}} \quad (2.15)$$

where

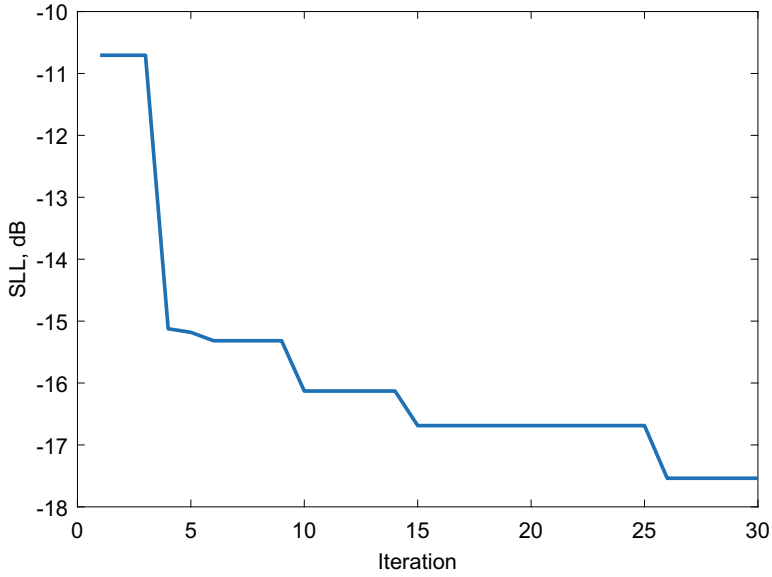


Fig. 2.13 The SLL convergence process of a 40×40 thinned array by the GA algorithm

$$\sum_{i=1}^{N_{sur}} p(chrom_i) = 1 \quad (2.16)$$

Two chromosomes are randomly chosen as *father* and *mother*, then they are mating and creating children,

$$\begin{aligned} &mask = \text{random a new chromosome} \\ &offspring1 = mask.*mother + not(mask).*father \\ &offspring2 = not(mask).*mother + mask.*father \end{aligned} \quad (2.17)$$

These children are created to offset the discarded chromosomes and keep the total number of chromosomes in the population remains constant.

Finally, mutations cause small random variations in the survived chromosome. By defining the mutation rate p_{mut} and generating a random value p between 0 and 1, the bits of a chromosome are updated as,

$$b_j = \begin{cases} b_j & \text{if } p > p_{mut} \\ not(b_j) & \text{if } p \leq p_{mut} \end{cases} \quad (2.18)$$

The fitness values are evaluated for all the offsprings and the mutated chromosomes, and the process is repeated. The algorithm is terminated if an acceptable solution is found, or the maximum times of iteration is reached. Figure 2.13 illustrates the convergence process of a 40×40 thinned array. The best SLL decreases

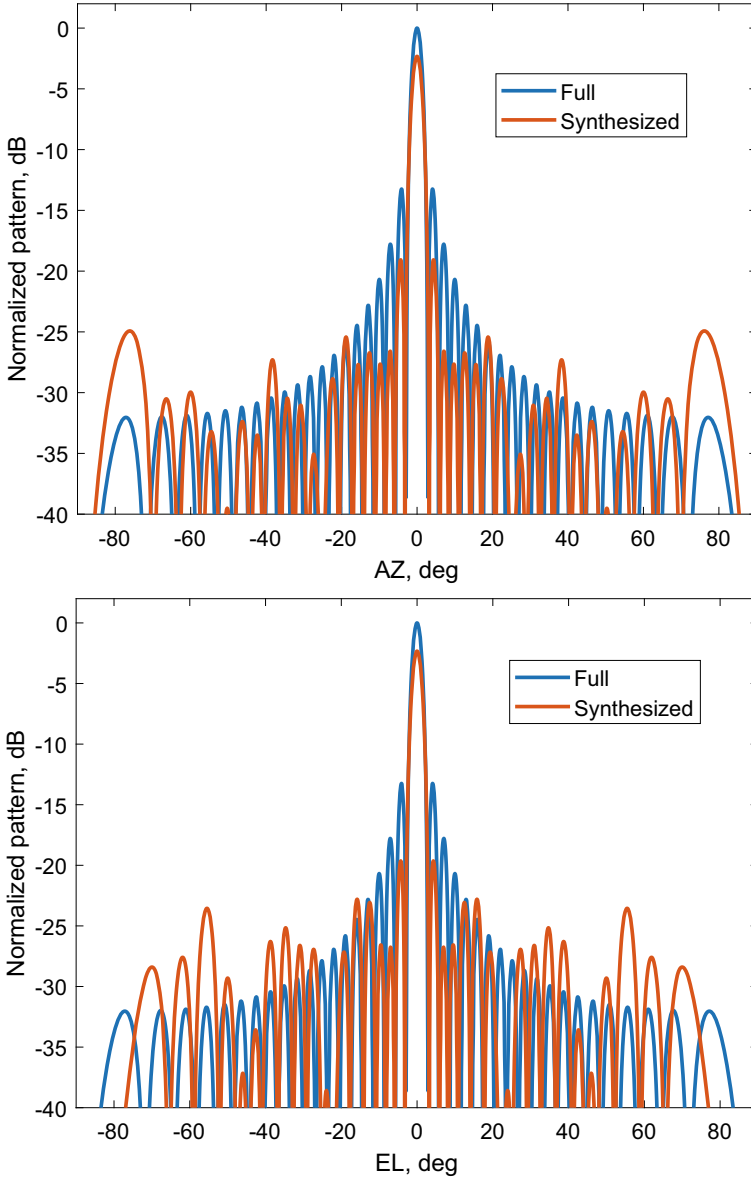


Fig. 2.14 The array pattern of a 40×40 thinned array obtained by the GA algorithm, **a** AZ cut, and **b** EL cut

gradually during the iteration, means that better solutions are found. Figures 2.14 and 2.15 present the synthesized array pattern and the array layout after 30 iterations, respectively. It has the SLL of -17.5 dB and a filtrate of 76.5%.

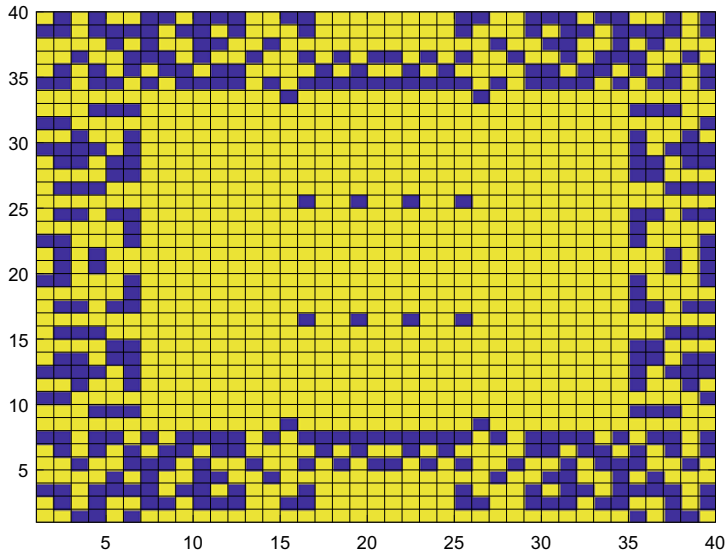


Fig. 2.15 The layout of a 40×40 thinned array obtained by the GA algorithm

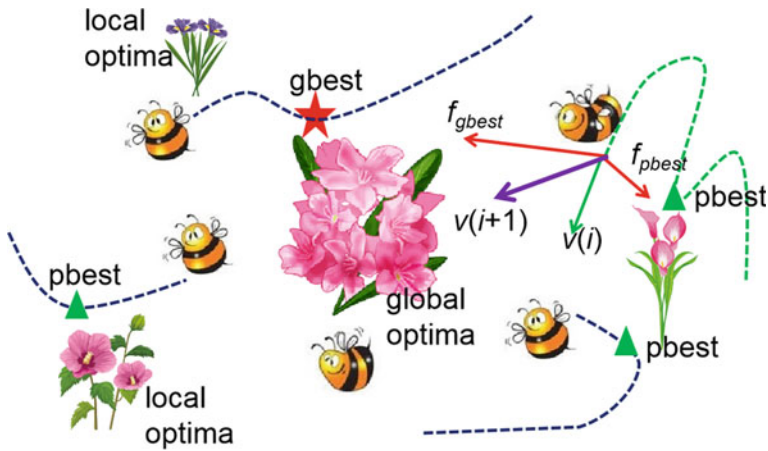


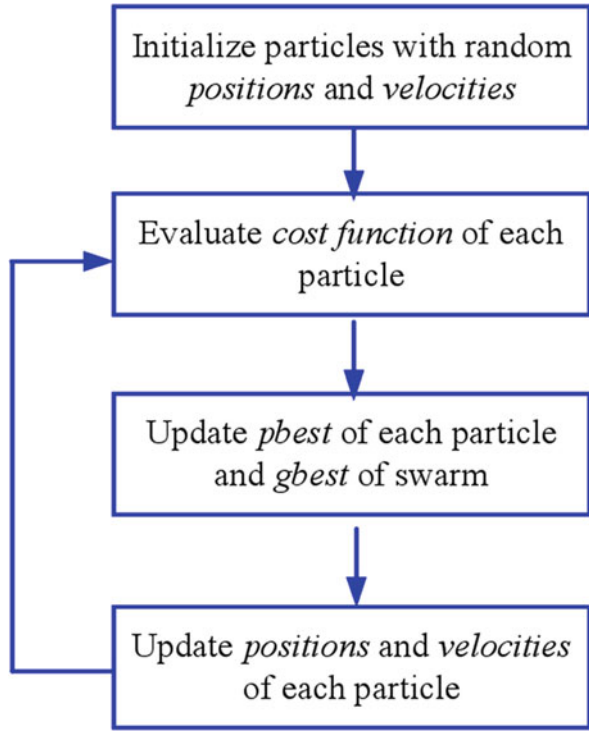
Fig. 2.16 A Cartoon view of the PSO algorithm

2.3.4 Particle Swarm Optimization

2.3.4.1 The Classical PSO Algorithm

As shown in Fig. 2.16, Particle swarm optimization (PSO), a swarm based stochastic evolutionary algorithm, was first developed by Kennedy and Eberhart in 1995, and introduced to the antenna community by Robinson and Rahmat-Samii in 2004 [23,

Fig. 2.17 The general flowchart of the PSO algorithm



24]. Believed to be very robust and effective in multidimensional and nonlinear problems, it has been applied to many different applications. Compared to other evolutionary methods, PSO is much easier to understand and implement and requires the least computational costs. In the PSO scheme, particles fly throughout the problem hyperspace to search for the optimal solution. Suppose there are N undetermined parameters in the optimization problem, then the position of each particle is denoted by an N -dimensional vector $\mathbf{x} = \{x_1, \dots, x_n, \dots, x_N\}$ to represent a possible solution with a velocity vector $\mathbf{v} = \{v_1, \dots, v_n, \dots, v_N\}$. As shown in Fig. 2.17, first without any prior knowledge of the optimal solution, particles are initialized with random positions and velocities. Then for the k th iteration, these two vectors are updated according to the following equations,

$$\begin{aligned} v_n^{k+1} &= wv_n^k + c_1\phi_1(pbest_n^k - x_n^k) + c_2\phi_2(gbest_n^k - x_n^k) \\ x_n^{k+1} &= x_n^k + v_n^k \end{aligned} \quad (2.19)$$

where w is the inertial weight, c_1 and c_2 are acceleration constants that specify how much each particle is influenced by the personally encountered best position, $pbest$, and the global best position ever found by the entire swarm, $gbest$, which control the relative proportion of cognition and social interaction in the swarm, respectively, ϕ_1

and ϕ_2 are two random variables uniformly distributed on $[0, 1]$. Once an updating process is implemented, the new solution found by each particle is evaluated by a fitness function, then both the $pbest$ of each particle and the $gbest$ of the entire swarm are updated. If a criterion is met (usually a sufficiently good fitness value or a maximum number of iterations), the iteration process is terminated and the results are obtained.

It's generally believed that PSO is superior to deal with optimization problems having only real variables, while genetic algorithm (GA) is more effective than PSO to handle optimization problems with integer variables, because PSO was intended to deal with real variables only while GA can translate a binary string to integer value directly. Therefore, how to modify the classical real PSO (RPSO) algorithm suitable for integer optimization problems while maintaining its simplicity and outstanding performance has become great of interest in recent years.

2.3.4.2 The Unified PSO Algorithm

Usually, in many mixed integer optimization problems, real and integer variables are treated separately and different updating mechanisms are applied. Here we propose a method that works with real and integer variables in the same updating scheme. First a unified vector $\mathbf{u} = [\mathbf{u}_r, \mathbf{u}_n]$ having continuous values between 0 and 1 is defined, where $\mathbf{u}_r$ and $\mathbf{u}_n$ are the real and integer part that will be mapped to the corresponding real and integer parameters, $\mathbf{r}$ and $\mathbf{n}$, respectively. During the optimization process, the unified vector is updated by Eq. (2.19), and then the mapping and rounding operations are implemented during fitness evaluation process,

$$\mathbf{r} = r_{min} + \mathbf{u}_r \times (r_{max} - r_{min}) \quad (2.20a)$$

$$\mathbf{n} = n_{min} + \text{round} \{ \mathbf{u}_n \times (n_{max} - n_{min} + 1) - 0.5 \} + \mathbf{m} \quad (2.20b)$$

where $\mathbf{m} = \text{random} \{1, 0, -1\}$

where min and max are the lower and upper bounds for the variables, $r_{min} \leq r \leq r_{max}$, $n_{min} \leq n \leq n_{max}$. The round function is rounding the real value to the nearest integer. In former mapping methods, integers in the range have different possible probabilities to be selected, thus the performance of the algorithms will be influenced. Here we add the term -0.5 in the bracket of Eq. (2.20b) to make each integer in the range can be selected with equal probability.

For integer variable mapping in Eq. (2.20b), a new vector $\mathbf{m}$ consisting of random ternary variables is introduced. Generally speaking, there are two aspects for one to evaluate the performance of an optimization algorithm. The first aspect is how much time the algorithm costs to find a suitable solution, whereas the second aspect is concerned with the best possible solution ever found by the algorithm. The random ternary vector will influence the algorithm in both aspects. At the beginning stage of the iterations, it may (only may) have some negative effects on the convergence speed

of the algorithm. Because updating of the integer part $\mathbf{u}_r$ is the main contribution in Eq. (2.20b) during this period, so the effect of the ternary vector is considered to be very tiny. But at the latter stage of the iterations, the ternary vector can be regarded as additional searching operations nearby the ever found solutions, which leads to better solutions and improved topology exploration capability.

In order to give better understandings of these explanations and demonstrate the effectiveness of the proposed method, we revisit three previous problems in the following sections. We use UPSO-m and UPSO to represent the unified PSO scheme with and without random ternary variables for simplicity, respectively. During the optimization process, the parameters of PSO algorithm are set as suggested by previous literatures, w is decreased linearly from 0.9 to 0.4 during iterations, c_1 and c_2 are chosen to be 2 for better convergence performance, and number of particles are a little larger than the dimension of the optimization problem.

2.3.4.3 Synthesis of the Linear Array with Subarray Configuration

It is a well-known technique to partition a large antenna array into contiguous subarrays in order to reduce cost via common use of components and simplified feed networks. Unfortunately, placing amplitude weights at the subarray ports will create grating lobes (GLs) due to the periodicity and larger spacing between subarrays. A $2N$ -element linear array is centro-symmetrically placed along the x -axis, and each side is divided into Q subarrays, with different amplitude weights at the subarray output ports. The elements are equally spaced and the distance d between adjacent elements is 0.5λ . The corresponding array factor $f(\theta)$ is given by,

$$f(\theta) = \sum_{q=1}^Q w_q \sum_{i=1}^{n_q} \cos(kx_n \sin\theta) \quad (2.21)$$

where x_n is the position of the n th element in wavelength, θ is the angle relative to boresight, n_q and w_q are number of elements and the amplitude weight of the q th subarray, respectively.

First let us consider each subarray having the same number of elements. In RPSO, a Q -dimensional vector is defined to represent the corresponding amplitude weights. A 10-agent swarm is used in optimization for 1000 iterations. Figure 2.18 illustrates the optimized pattern of a 128-element linear array divided into 16 equal sized subarrays, along with the pattern of the uniform array. The obtained peak sidelobe level (SLL) is -31.4 dB. Table 2.1 summaries the best solutions obtained by different optimization methods. The same array with uniform tapering has a peak SLL of -13.3 dB. For a 31 dB and $n = 5$ Taylor tapering, the peak SLL is -30.5 dB. The result is -30.9 dB obtained by GA with Nelder Mead downhill simplex method. Although PSO has found the best solution than any other algorithms, but it has been clearly showed that the GLs prevent the optimization process to find better solutions.

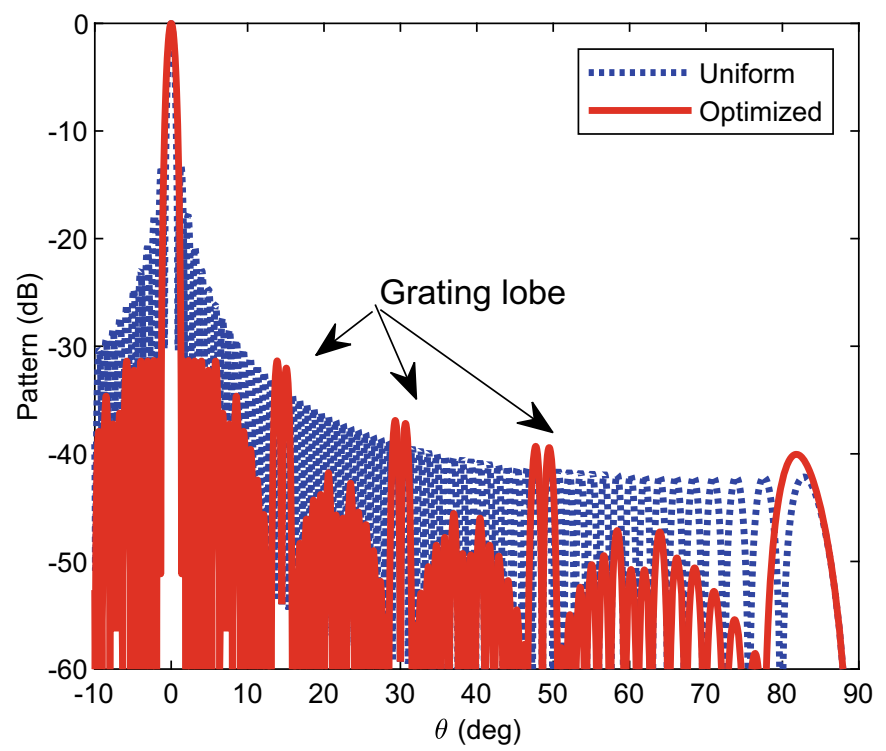


Fig. 2.18 The optimized radiation pattern for the 128-element linear array with equal subarray sizes

Table 2.1 The best results obtained by different optimization algorithms

Method	Subarray sizes	SLL (dB)
Uniform	Equal	−13.3
Taylor tapering	Equal	−30.5
GA	Equal	−30.9
	Unequal	−35.1
UPSO-m	Equal	−31.4
	Unequal	−36.1

Many methods have been proposed to break the periodicity and redistribute the energy of GLs through the array pattern. One of these methods is to allow each subarray having a different number of elements. Since there are two different types of parameters in this optimization problem, i.e., integers for the subarray sizes and real numbers for the amplitude weights, UPSO-m is applied for further simulation rather than RPSO. A $2Q$ -dimensional unified vector $\mathbf{u} = [\mathbf{u}_w, \mathbf{u}_n] = \{u_1, u_2, \dots, u_{2Q}\}$ is designated to each candidate design. The first part $\mathbf{u}_w = \{u_1, u_2, \dots, u_Q\}$ repre-

Table 2.2 Comparison of *gbest* convergence performance of different PSO algorithms over 10 independent trials

Method	Average (dB)	Best (dB)
IPSO	-33.67	-35.04
UPSO	-34.24	-35.30
UPSO-m	-35.19	-36.10

sent the normalized amplitude weights of the subarrays, and the latter part $\mathbf{u}_n = \{u_Q, u_{Q+1}, \dots, u_{2Q}\}$ are mapped and rounded to the subarray sizes. In order to demonstrate the effectiveness of the method proposed in this paper, UPSO-m is validated and compared with other existing PSO schemes. For each method, a 20-agent swarm is used in optimization for 5000 iterations. Figure 2.19 illustrates convergence of *gbest* values averaged over 10 independent trials, and the performances are summarized in Table 2.2. At the end of iteration, an averaged *gbest* value of -35.19 dB is obtained by UPSO-m, compared with -33.67 dB and -34.24 dB by IPSO and UPSO, respectively. As can be seen in Fig. 2.19, at the beginning stage of the iterations, convergence performance of UPSO-m is a little worse than UPSO. But in the latter stage of the iterations, while the other two PSO schemes have little improvements, UPSO-m continues to find better solutions.

During the 10 independent trials, the best solution ever found by UPSO-m is -36.1 dB, which is also better than IPSO (-35.0 dB), UPSO (-35.3 dB) and GA (-35.1 dB). These results have been clearly showed that, UPSO-m has the best convergence performance and topology exploration capability, and most of the improvements should be attributed to the random ternary variables. Figure 2.20 illustrates the optimized subarray sizes and amplitude weights, which $\mathbf{n} = [10, 6, 8, 6, 8, 6, 10, 10]$ and $\mathbf{w} = [1.0000, 0.9264, 0.8463, 0.7018, 0.5748, 0.4308, 0.2964, 0.1529]$. Since the array is symmetric, only the right hand side is shown. The corresponding optimized array pattern is illustrated in Fig. 2.21, which has nearly equal sidelobe levels without GLs.

2.3.5 Convex Optimization

A convex optimization problem is one of the form [25],

$$\begin{aligned} \min & f_0(x) \\ \text{s.t.} & f_i(x) \leq b_i, i = 1, \dots, m \end{aligned} \quad (2.22)$$

where the objective and constraint functions $f_0, \dots, f_m : R_n \rightarrow R$ are convex, which means they satisfy the inequality,

$$f_i(\alpha x_1 + \beta x_2) \leq \alpha f_i(x_1) + \beta f_i(x_2) \quad (2.23)$$

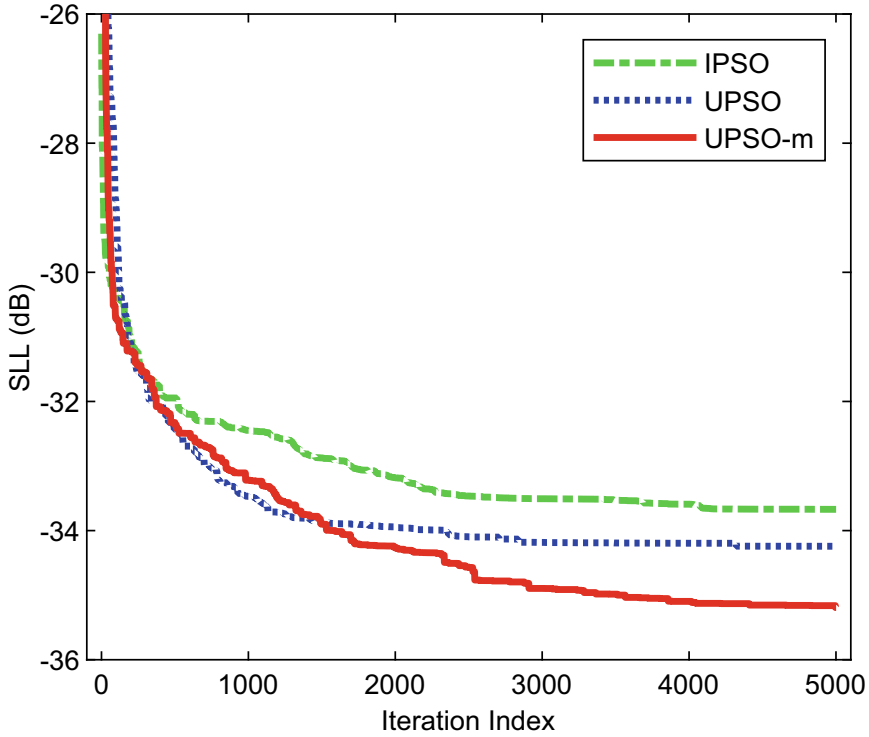


Fig. 2.19 Comparison of convergence performance of different PSO algorithms over 10 independent trials

for all $x \in R_n$ and all $\alpha, \beta \in R$ with $\alpha + \beta = 1, \alpha \geq 0, \beta \geq 0$. There are very effective algorithms that can reliably and efficiently solve even large convex problems.

As an example, consider the minimization of the beampattern level over a given zone with the possibility of level constraints in other areas. The way to express it as a standard convex optimization problem is rather natural and very similar to a framework [26],

$$\begin{aligned} \min & -|f(\theta_0, \phi_0)| \\ \text{s.t. } & |f(\theta_s, \phi_s)| \leq UB_s \end{aligned} \quad (2.24)$$

where (θ_0, ϕ_0) indicates the main beam direction, (θ_s, ϕ_s) and UB_s represent the sidelobe region and upper bound of the desired SLL, respectively.

As an important special case, the convex optimization is called a linear programming if the objective and constraint functions $f_0, \dots, f_m$ are linear, i.e., and satisfy

$$\begin{aligned} \min & c^T x \\ \text{s.t. } & a_i^T x \leq b_i, i = 1, \dots, m. \end{aligned} \quad (2.25)$$

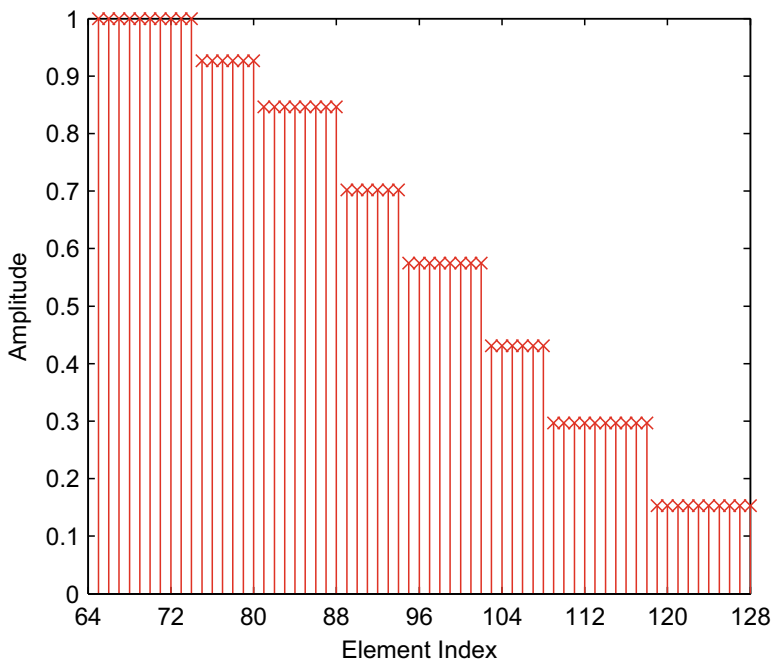


Fig. 2.20 The optimized subarray sizes and amplitude weights for the 128-element linear array

By adopting the conjugate symmetric weight condition of the phased array, the expression of the array factor becomes from the abs form to the linear form, thus the synthesis process is reduced to an even more powerful linear programming.

Figure 2.22 illustrates an example of the 40-element linear array synthesized by convex optimization. The sidelobe level is limited below -25 dB for all regions. Suppose there is an interfere from the direction of 30° , then a null is placed at 30° , and the SLL is suppressed less than -40 dB for the region from 20° to 40° .

2.4 Summary

This chapter reviews some basics of phased array antennas. First, the general architecture and major components of a phased array antenna are introduced. The antenna elements are used to transmit/receive electromagnetic waves from the T/R module to the free space. The T/R module integrates several basic components such as duplexer, filter, phase shifter, power amplifier, LNA, mixer, LO, ADC, and DAC, and is viewed as the most important part of a phased array antenna. The beamforming network is to combine signals from antenna elements. It forms the main beam in the desired direction of the signal. Sophisticated functions are implemented in digital signal

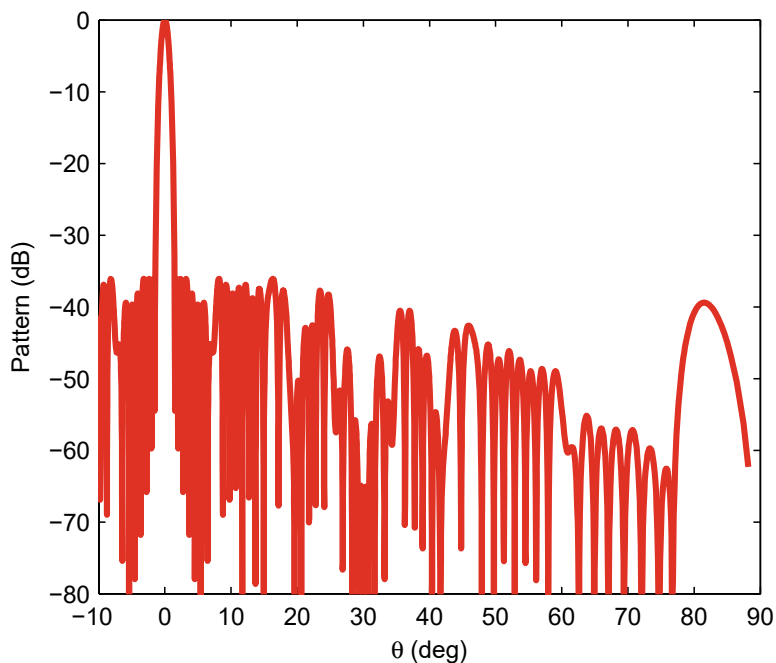


Fig. 2.21 The optimized radiation pattern of the 128-element linear array with unequal subarray sizes

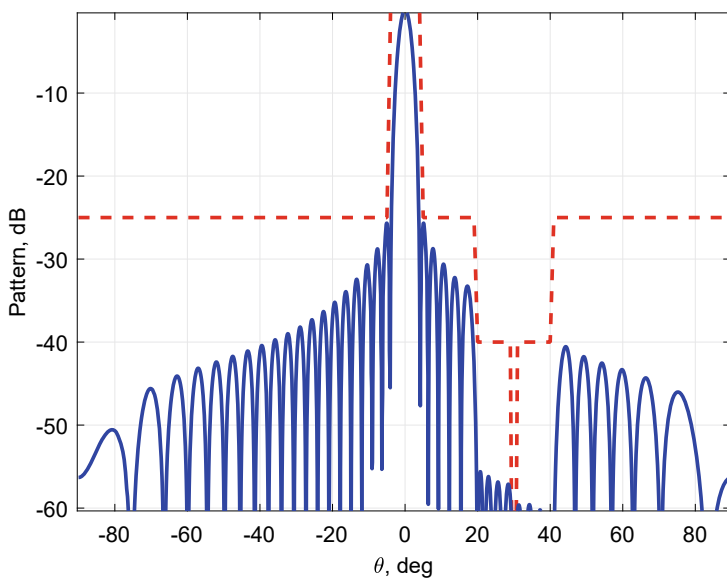


Fig. 2.22 The radiation pattern of the 40-element linear array synthesized by convex optimization with sidelobe level control and null at $\theta = 30^\circ$

processors by software to realize satellite communication and TT&C service. In order to keep the reliable operation of the phased array antenna, air or liquid cooling equipments is necessary to dissipate the generated heat of the the T/R module and below the temperature at a reasonable value. The array pattern synthesis is to produce desired beam that places the nulls in the direction of interferences or clutters. Several popular methods to synthesize the array pattern are introduced. Analytical algorithms are one of the simplest kinds of methods among them, which can be directly calculated from mathematical expression easily. For other geometries or more complicated array synthesis problems, it must be synthesized by other numerical methods or evolutionary algorithms such as GA and PSO. Convex optimization is also found be very effective for array pattern synthesis.

References

1. J.S. Herd, M.D. Conway, The evolution to modern phased array architectures. *Proc. IEEE* **104**(3), 519–529 (2016)
2. C. Fulton, M. Yearly, D. Thompson, J. Lake, A. Mitchell, Digital phased arrays: challenges and opportunities. *Proc. IEEE* **104**(3), 487–503 (2016)
3. S.H. Talisa, K.W. O'Haver, T.M. Comberiate, M.D. Sharp, O.F. Somerlock, Benefits of digital phased array radars. *Proc. IEEE* **104**(3), 530–543 (2016)
4. M. Mirmozafari, G. Zhang, C. Fulton, R.J. Doviak, Dual-polarization antennas with high isolation and polarization purity: a review and comparison of cross-coupling mechanisms. *IEEE Antennas Propag. Mag.* **61**(1), 50–63 (2019)
5. S.X. Ta, I. Park, R.W. Ziolkowski, Crossed dipole antennas: a review. *IEEE Antennas Propag. Mag.* **57**(5), 107–122 (2015)
6. C.-H. Lee, S.-Y. Chen, P. Hsu, Isosceles triangular slot antenna for broadband dual polarization applications. *IEEE Trans. Antennas Propag.* **57**(10), 3347–3351 (2009)
7. B.A. Kopp, M. Borkowski, G. Jerinic, Transmit/receive modules. *IEEE Trans. Microw. Theory Techn.* **50**(3), 827–834 (2002)
8. S. Rathod, K. Sreenivasulu, K.S. Beenamole, K.P. Ray, Evolutionary trends in transmit/receive module for active phased array radars. *Def. Sci. J.* **68**(6), 553–559 (2018)
9. A. Chakraborty, B. Gupta, Paradigm phase shift: RF MEMS phase shifters: an overview. *IEEE Microw. Mag.* **18**(1), 22–41 (2017)
10. R. Rotman, M. Tur, L. Yaron, True time delay in phased arrays. *Proc. IEEE* **104**(3), 504–518 (2016)
11. P.-I. Mak, S.-P. U, R.P. Martins, Transceiver architecture selection: review, state-of-the-art survey and case study. *IEEE Circuits Syst. Mag.* **7**(2), 6–25 (2007)
12. D. Sifarikas, J.L. Volakis, Toward direct RF sampling: implications for digital communications. *IEEE Microw. Mag.* **21**(9), 43–52 (2020)
13. B. Murmann, The race for the extra decibel: a brief review of current ADC performance trajectories. *IEEE Solid-State Circuits Mag.* **7**(3), 58–66 (2015)
14. B. Murmann, ADC performance survey 1997–2021. <http://web.stanford.edu/~murmann/adcsurvey.html>
15. P. Rocca, A. Morabito, Optimal synthesis of reconfigurable planar arrays with simplified architectures for monopulse radar applications. *IEEE Trans. Antennas Propag.* **63**(3), 1048–1058 (2015)
16. J. Kreng, M. Sue, S. Do, Y. Krikorian, S. Raghavan, Telemetry, tracking, and command link performance using the USB/STDN waveform. *IEEE Aerospace Conf.* **2007**, 1–15 (2007)

17. J.B. Berner, S.H. Bryant, P.W. Kinman, Range measurement as practiced in the deep space network. *Proc. IEEE* **95**(11), 2202–2214 (2007)
18. S. Hosseinizadeh, F. Tan, S. Moosania, Experimental and numerical studies on performance of PCM-based heat sink with different configurations of internal fins. *Appl. Therm. Eng.* **31**, 3827–3838 (2011)
19. Y. Wang, C. Wang, P. Lian, et al., Effect of temperature on electromagnetic performance of active phased array antenna. *Electronics* **9**, article no. 1211 (2020)
20. Y.E. Nikolaenko, A.V. Baranyuk, S.A. Reva, E.N. Pismennyi, F.F. Dubrovka, Numerical simulation of the thermal and hydraulic characteristics of the liquid heat exchanger of the APAA transmitter-Receiver module. *Therm. Sci. Eng. Progress* **17**, Article no. 100499, 1–11 (2020)
21. J. Nagar, D.H. Werner, Multiobjective optimization for electromagnetics and optics: an introduction and tutorial based on real-world applications. *IEEE Antennas Propag. Mag.* **60**(6), 58–71 (2018)
22. R.L. Haupt, An introduction to genetic algorithms for electromagnetics. *IEEE Antennas Propag. Mag.* **37**(2), 7–15 (1995)
23. J. Robinson, Y. Rahmat-Samii, Particle swarm optimization in electromagnetics. *IEEE Trans. Antennas Propag.* **52**(2), 397–407 (2004)
24. J. Nanbo, Y. Rahmat-Samii, Advances in particle swarm optimization for antenna designs: real-number, binary, single-objective and multi-objective implementations. *IEEE Trans. Antennas Propag.* **55**(3), 556–567 (2007)
25. S. Boyd, L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, 2009)
26. H. Lebrecht, S. Boyd, Antenna array pattern synthesis via convex optimization. *IEEE Trans. Signal Process.* **45**(3), 526–532 (1997)

Chapter 3

Array Excitation Error Analysis



3.1 Probabilistic Methods

The array pattern synthesis requires precise excitation of each antenna element by various synthesis techniques to obtain desired array pattern features, e.g., high directivity, low SLL, and null control. However, the realistic antenna element excitations inevitably differ from their expected values in practice. These errors would cause array pattern distortion, gain degradation, worse SLL, and null drifts [1–3]. Therefore, knowledge about the impacts of the excitation errors and the corresponding calibration methods are of high importance for phased array antennas. Many methods have been proposed to analyze the impacts of excitation errors on array performance. These analysis methods fall into mainly two different types, probabilistic methods and interval arithmetic (IA) based methods, which are presented in the following two subsections.

3.1.1 The Mathematical Model

Recalling Eq. (2.8) that a phased array antenna with N isotropic elements, each element has the excitation amplitude $\{a_n\}$ and phase $\{p_n\}$, $n = 1, \dots, N$. The array radiation pattern or the array factor $f(\theta, \phi)$ can be mathematically expressed as the sum of radiation pattern of individual antenna elements $f_n(\theta, \phi)$,

$$\begin{aligned} f(\theta, \phi) &= \sum_{n=1}^N f_n(\theta, \phi) \\ &= \sum_{n=1}^N a_n \exp[j(k\vec{r}_0(\theta, \phi) \cdot \vec{r}_n + p_n)] \\ &= \sum_{n=1}^N a_n \exp(jb_n) \end{aligned} \quad (3.1)$$

where $\vec{r}_0(\theta, \phi)$ and $\vec{r}_n$ represent the unit direction and the position of the n -th antenna element, respectively.

Note that the phase term $\{b_n\}$ is also a function of the observation angle (θ, ϕ) , however, to simplify the notation this angle dependence is not shown in the following discussion. Without considering position errors of the antenna elements, the term $\vec{r}_0(\theta, \phi) \cdot \vec{r}_n$ is constant at any given direction, then the two phases $\{b_n\}$ and $\{p_n\}$ have the same statistics and can be treated interchangeably.

With excitation errors, the real antenna element excitation can be described according to its referenced value,

$$\begin{aligned} a_n^{\text{real}} &= a_n(1 + \gamma_n) \\ p_n^{\text{real}} &= p_n + \delta_n \end{aligned} \quad (3.2)$$

and without position errors,

$$b_n^{\text{real}} = b_n + \delta_n \quad (3.3)$$

where a_n^{real} and p_n^{real} are the real excitation amplitude and phase of the n -th antenna element, the two random variables γ_n and δ_n represent its relative amplitude error and absolute phase error in percentage and degree, respectively.

There are generally two types of excitation errors, namely correlated errors and random errors. Usually, it is the intent of the designer to ensure that the correlated errors are appropriately removed once identified, so that all remained residuals are uncorrelated and random, which are difficult to compensate due to their randomness. Hence, only random excitation errors are considered in the following discussion. Without loss of generality, the variances in the amplitude and phase error are assumed to be identical among antenna elements. According to the central limit theorem, both of these random variables are Gaussian distributed and have zero mean,

$$\begin{aligned} \gamma_n &\approx N(0, \sigma_a^2) \\ \delta_n &\approx N(0, \sigma_p^2) \end{aligned} \quad (3.4)$$

where σ_a and σ_p are the standard deviations of the excitation amplitude and phase error, respectively.

Figure 3.1 illustrates an example of a half-wavelength equally spaced ($d = \lambda/2$) 10-element linear antenna array, which is used as the benchmark architecture in the following discussion. The referenced array pattern is Taylor tapered with $SLL^{\text{ref}} = -20$ dB and $\bar{n} = 2$, it has the main lobe with the first nulls at $\theta = \pm 17^\circ$, and the highest SLLs (the first sidelobe at $\theta = \pm 19^\circ$) are approximately equal to -20 dB as designed. Some Monte Carlo simulations are also presented with $\sigma_a = 0.05$ and $\sigma_p = 5^\circ$, whose results show some performance degradation. The SLL is raised to as high as -16 dB for some cases.

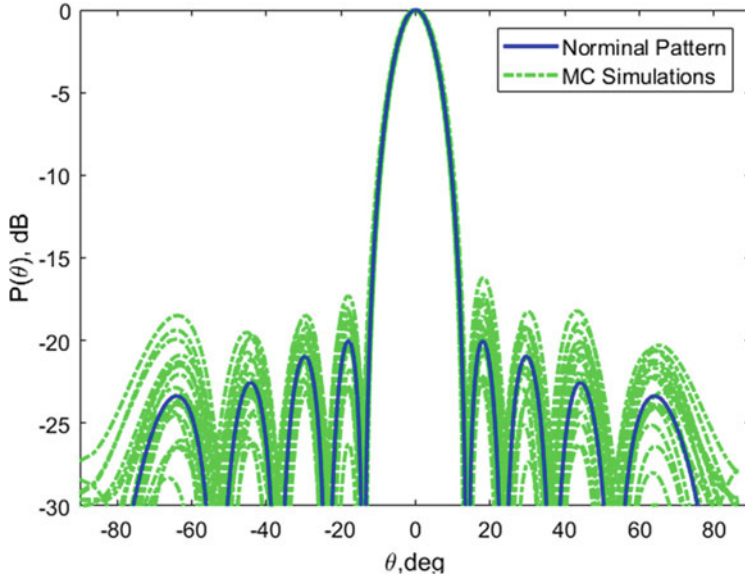


Fig. 3.1 The referenced array pattern of the 10-element Taylor tapered linear antenna array, with Monte Carlo simulations are presented with standard deviations of the amplitude error $\sigma_a = 0.05$ and phase error $\sigma_p = 5^\circ$

3.1.2 The Expectation of the Array Pattern

The expected array pattern aims to evaluate the average pattern performance and provide approximate formulas for the confidence boundaries. From Eqs. (3.1)–(3.3), the real array pattern $f^{\text{real}}(\theta, \phi)$ is rewritten as,

$$f^{\text{real}}(\theta, \phi) = \sum_{n=1}^N a_n (1 + \gamma_n) \exp[j(b_n + \delta_n)] \quad (3.5)$$

Generally, the array pattern is a complex value and can be splitted into the real and imaginary parts,

$$\begin{aligned} f_R^{\text{real}}(\theta, \phi) &= \sum_{n=1}^N a_n (1 + \gamma_n) \cos(b_n + \delta_n) \\ f_I^{\text{real}}(\theta, \phi) &= \sum_{n=1}^N a_n (1 + \gamma_n) \sin(b_n + \delta_n) \end{aligned} \quad (3.6)$$

From the Biggelaar's work [4], the expectations of these two parts, $u_R(\theta, \phi)$ and $u_I(\theta, \phi)$, can be derived as follows,

$$\begin{aligned}
u_R(\theta, \phi) &= E(f_R^{\text{real}}(\theta, \phi)) = \exp(-\sigma_p^2/2) \sum_{n=1}^N a_n \cos b_n \\
u_I(\theta, \phi) &= E(f_I^{\text{real}}(\theta, \phi)) = \exp(-\sigma_p^2/2) \sum_{n=1}^N a_n \sin b_n
\end{aligned} \tag{3.7}$$

And the variances of these two parts, $\sigma_R^2(\theta, \phi)$ and $\sigma_I^2(\theta, \phi)$, have the following expressions,

$$\begin{aligned}
\sigma_R^2(\theta, \phi) &= \frac{1}{2}(1 + \sigma_a^2) [1 - \exp(-2\sigma_p^2)] \sum_{n=1}^N a_n^2 \\
&\quad + [(1 + \sigma_a^2) \exp(-2\sigma_p^2) - \exp(-\sigma_p^2)] \sum_{n=1}^N a_n^2 \cos^2 b_n \\
\sigma_I^2(\theta, \phi) &= \frac{1}{2}(1 + \sigma_a^2) [1 - \exp(-2\sigma_p^2)] \sum_{n=1}^N a_n^2 \\
&\quad + [(1 + \sigma_a^2) \exp(-2\sigma_p^2) - \exp(-\sigma_p^2)] \sum_{n=1}^N a_n^2 \sin^2 b_n
\end{aligned} \tag{3.8}$$

Figure 3.2 illustrates the expectations and variances of the real and imaginary parts of the array pattern. As can be seen, the expectation for the real part is symmetric while the imaginary part is anti-symmetric. Their variances closely follow each other at most of the angles except at the boresight and the grating lobe. This is because the first term of their variances are the same in Eq. (3.8), while the latter term differs in their values especially in these two regions.

From Eqs. (3.7) and (3.8), the expectation of the array power pattern $P(\theta, \phi)$ can be derived as,

$$\begin{aligned}
E(P(\theta, \phi)) &= E((f_R^{\text{real}}(\theta, \phi))^2) + E((f_I^{\text{real}}(\theta, \phi))^2) \\
&= u_R^2(\theta, \phi) + \sigma_R^2(\theta, \phi) + u_I^2(\theta, \phi) + \sigma_I^2(\theta, \phi) \\
&= \exp(-\sigma_p^2) \left[\left(\sum_{n=1}^N a_n \cos b_n \right)^2 + \left(\sum_{n=1}^N a_n \sin b_n \right)^2 \right] \\
&\quad + (1 + \sigma_a^2) [1 - \exp(-2\sigma_p^2)] \sum_{n=1}^N a_n^2 \\
&\quad + [(1 + \sigma_a^2) \exp(-2\sigma_p^2) - \exp(-\sigma_p^2)] \sum_{n=1}^N a_n^2 \\
&= \exp(-\sigma_p^2) P^{\text{ref}}(\theta, \phi) + [1 + \sigma_a^2 - \exp(-\sigma_p^2)] \sum_{n=1}^N a_n^2
\end{aligned} \tag{3.9}$$

where $P^{\text{ref}}(\theta, \phi)$ is the normal/ideal array power pattern without excitation errors,

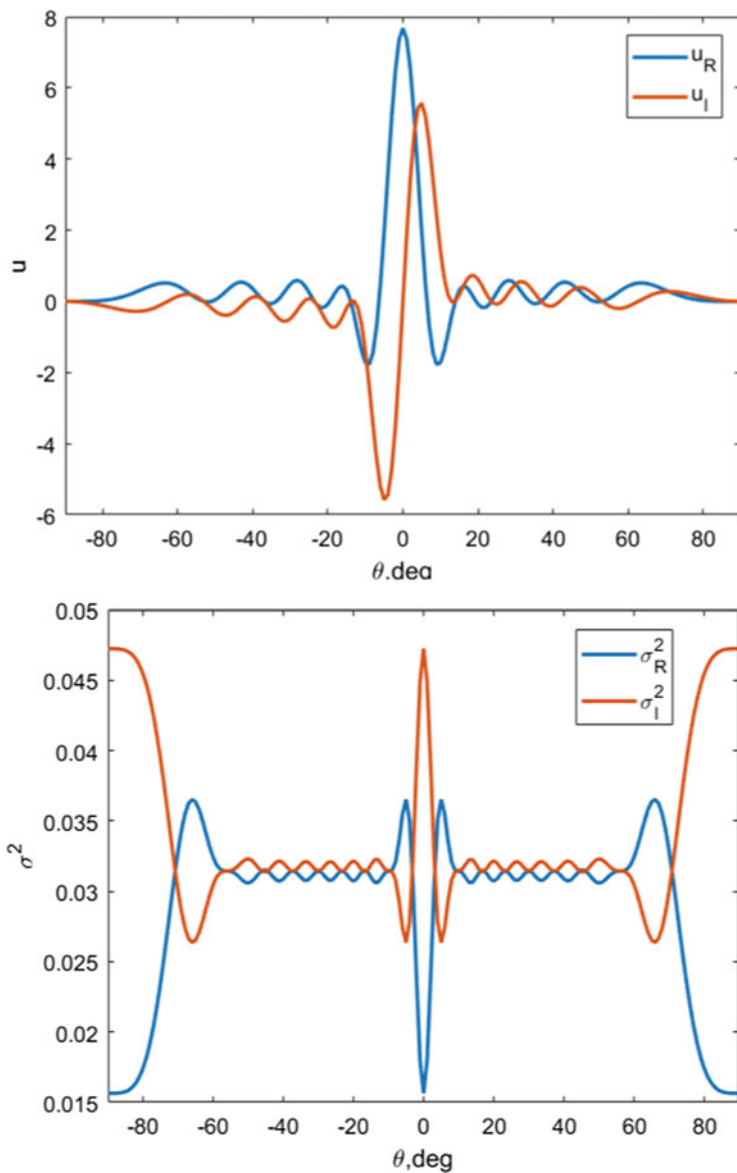


Fig. 3.2 The expectations and variances of the real and imaginary parts of the array pattern for the 10-element Taylor tapered linear antenna array

$$\begin{aligned}
P^{\text{ref}}(\theta, \phi) &= |f(\theta, \phi)|^2 = \left| \sum_{n=1}^N a_n \exp(jb_n) \right|^2 \\
&= \left[\left(\sum_{n=1}^N a_n \cos b_n \right)^2 + \left(\sum_{n=1}^N a_n \sin b_n \right)^2 \right]
\end{aligned} \tag{3.10}$$

This expression in Eq. (3.9) shows that the effect of random excitation errors produces the real array power pattern consisting of the ideal pattern reduced by a factor that accounts for phase error, plus another term that is a constant with no angular dependence. It is convenient to normalize the latter term to the peak of the resulting pattern, and the normalized sidelobe level is given by,

$$\begin{aligned}
SLL &= \frac{[1 + \sigma_a^2 - \exp(-\sigma_p^2)] \sum_{n=1}^N a_n^2}{\exp(-\sigma_p^2) \left(\sum_{n=1}^N a_n \right)^2} \\
&= [(1 + \sigma_a^2) \exp(\sigma_p^2) - 1] \sum_{n=1}^N a_n^2 / \left(\sum_{n=1}^N a_n \right)^2 \\
&\approx [(1 + \sigma_a^2)(1 + \sigma_p^2) - 1] \sum_{n=1}^N a_n^2 / \left(\sum_{n=1}^N a_n \right)^2 \\
&\approx (\sigma_a^2 + \sigma_p^2) / N\eta_A
\end{aligned} \tag{3.11}$$

where η_A is known as the array taper efficiency [5],

$$\eta_A = \left(\sum_{n=1}^N a_n \right)^2 / N \sum_{n=1}^N a_n^2 \tag{3.12}$$

The symmetrical form of the amplitude and phase errors in Eq. (3.11) suggests the convenience of converting them from one to the other. This helps in the trade-off to determine how much of the error variance to allot between the two. Figure 3.3 presents several curves of average sidelobe level due to the phase and amplitude errors for the linear array as presented in Fig. 3.1. Due to their interchangeability, each sidelobe level curve is like an arc.

3.1.3 The PDF of the Array Pattern

As can be seen in Fig. 3.2, for most regions except at the bore-sight, grating lobe, and other selective regions, it's reasonable to assume that the real and imaginary parts of the array factor have identical mean value and share the same variance. Hence, as sum of these two Gaussian distributed variances, the array pattern at these regions follows Rician distribution with the PDF defined as,

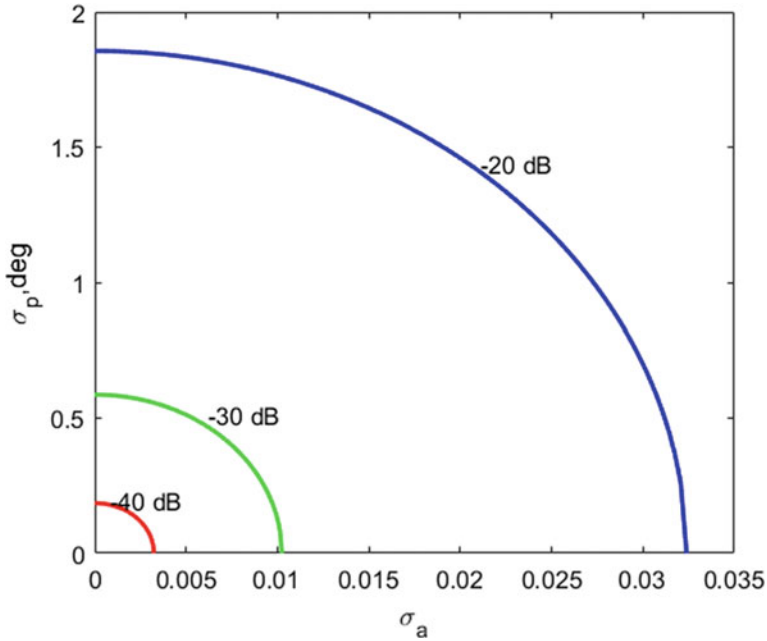


Fig. 3.3 Array average sidelobe level due to phase and amplitude errors for the 10-element Taylor tapered linear antenna array

$$prob(r(\theta, \phi)) = \frac{r(\theta, \phi)}{\sigma_r^2(\theta, \phi)} \exp\left(-\frac{r^2(\theta, \phi) + u_r^2(\theta, \phi)}{\sigma_r^2(\theta, \phi)}\right) \cdot I_0\left(\frac{ru_r(\theta, \phi)}{\sigma_r^2(\theta, \phi)}\right) \quad (3.13)$$

where $I_0()$ is the modified Bessel function of the first kind with order zero, and

$$\begin{aligned} u_r(\theta, \phi) &= \sqrt{u_R^2(\theta, \phi) + u_I^2(\theta, \phi)} \\ \sigma_r^2(\theta, \phi) &= \frac{1}{2} (\sigma_R^2(\theta, \phi) + \sigma_I^2(\theta, \phi)) \end{aligned} \quad (3.14)$$

Generally, the same variance assumption is good enough for most analyses. However, it has been pointed out that this assumption is not mandatory and a Beckman distribution can be derived [4]. For other special regions that have unidentical variances for the real and imaginary parts, such as the boresight ($u_I = 0$) and null locations ($u_R = u_I = 0$), detailed mathematical derivation can be found in Bhattacharyya's excellent book [6].

3.2 Interval Arithmetic Methods

For practical antenna array engineering, it's of great interest to evaluate the worst performance of the array pattern along with statistical characteristics. Recently, the IA technique was introduced by P. Rocca et al. [7, 8] to evaluate the effects of excitation errors and the corresponding upper and lower bounds of the array patterns. Thanks to the intrinsic of IA to deal with uncertainties, it's very efficient to determine bounds of the array pattern in a robust and reliable way. It needs no detailed information unless the knowledge of the intervals about the excitation errors, which are the only terms involved in the IA based mathematical operations. According to the geometry of the excitation error representation, three IA based methods have been proposed, i.e., the Cartesian IA method, the Circular IA method, and the Polygonal IA method. One more method call the Matrix IA method is also found be more effective when dealing with amplitude excitation error.

3.2.1 Interval Arithmetic Basics

A real-valued variable x has its interval which can be represented with the brackets $[x^{\text{inf}}, x^{\text{sup}}]$ defined by a lower bound x^{inf} and an upper bound x^{sup} . It can also be defined by a midpoint x^{mid} and a radius x^{rad} and represented with the brackets $< x^{\text{mid}}, x^{\text{rad}} >$. Assuming the actual value of x is the sum of the referenced (or expected) value x^{ref} and some deviation ε with the interval $[\varepsilon^{\text{inf}}, \varepsilon^{\text{sup}}]$, i.e., $x = x^{\text{ref}} + \varepsilon$. These parameters can be interchanged by the following relationships,

$$\begin{aligned} x^{\text{inf}} &= x^{\text{ref}} + \varepsilon^{\text{inf}} = x^{\text{mid}} - x^{\text{rad}} \\ x^{\text{sup}} &= x^{\text{ref}} + \varepsilon^{\text{sup}} = x^{\text{mid}} + x^{\text{rad}} \end{aligned} \quad (3.15)$$

$$\begin{aligned} x^{\text{mid}} &= \frac{x^{\text{inf}} + x^{\text{sup}}}{2} = x^{\text{ref}} + \frac{\varepsilon^{\text{inf}} + \varepsilon^{\text{sup}}}{2} \\ x^{\text{rad}} &= \frac{x^{\text{sup}} - x^{\text{inf}}}{2} = \frac{\varepsilon^{\text{sup}} - \varepsilon^{\text{inf}}}{2} \end{aligned} \quad (3.16)$$

Interval arithmetic have some very interesting properties [9, 10], here we list just some of them which will be used in the following discussion.

(1) Addition. Supposing x_1 and x_2 are two real-valued variables, and $y = x_1 + x_2$. Then the interval of y can be expressed as follows,

$$y^{\text{inf}} = x_1^{\text{inf}} + x_2^{\text{inf}}, \quad y^{\text{sup}} = x_1^{\text{sup}} + x_2^{\text{sup}} \quad (3.17)$$

(2) Scalar multiplication. Supposing x is a real-valued variable and k is a scalar, and $z = k \cdot x$. Then the interval of z can be expressed as follows,

$$\begin{aligned} z^{\text{inf}} &= k \cdot x^{\text{inf}}, \quad z^{\text{sup}} = k \cdot x^{\text{sup}} \quad \text{if } k \geq 0 \\ z^{\text{inf}} &= k \cdot x^{\text{sup}}, \quad z^{\text{sup}} = k \cdot x^{\text{inf}} \quad \text{if } k < 0 \end{aligned} \quad (3.18)$$

(3) Matrix multiplication. Supposing $\mathbf{X}$ is a $m \times n$ real-valued variable matrix and A is a $n \times k$ scalar matrix, and $\mathbf{B} = A\mathbf{X}$ is a $m \times k$ matrix to be determined. Then the interval of $\mathbf{B}$ can be expressed as follows,

$$\mathbf{B}^{\text{mid}} = A\mathbf{X}^{\text{mid}}, \quad \mathbf{B}^{\text{rad}} = |A| \mathbf{X}^{\text{rad}} \quad (3.19)$$

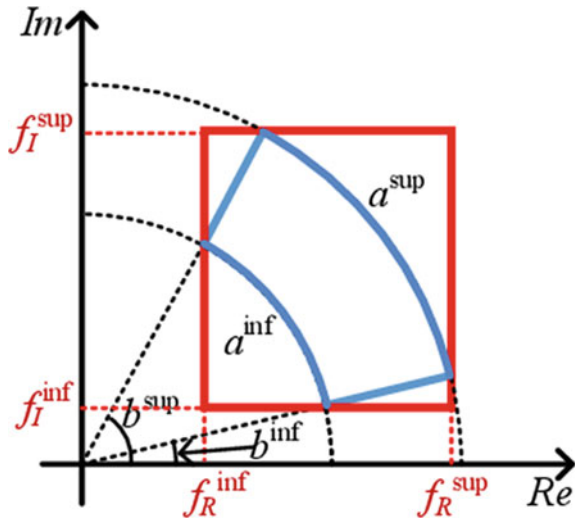
3.2.2 The Cartesian IA (rIA) Method

As shown in Fig. 3.4, a specified excitation error has a fan-like geometry defined by the boundaries of amplitude interval $[a^{\text{inf}}, a^{\text{sup}}]$ and phase interval $[b^{\text{inf}}, b^{\text{sup}}]$. Please keep in mind that b^{inf} and b^{sup} are also dependent on the observation direction (θ, ϕ) . For the Cartesian IA (rIA) method, the error geometry is expanded into the rectangular form for convenience [7, 8]. Again, the complex array pattern of equation (3.1) is rewritten with the real part $f_R(\theta, \phi)$ and the imaginary part $f_I(\theta, \phi)$,

$$\begin{aligned} f(\theta, \phi) &= f_R(\theta, \phi) + j \cdot f_I(\theta, \phi) \\ &= \sum_{n=1}^N a_n \cos b_n + j \cdot \sum_{n=1}^N a_n \sin b_n \end{aligned} \quad (3.20)$$

The array power pattern $P(\theta, \phi)$ can be obtained by first squaring these two parts then adding them together,

Fig. 3.4 Illustration of the Cartesian IA method. The boundary of each element is defined by the amplitude interval $[a^{\text{inf}}, a^{\text{sup}}]$ and phase interval $[b^{\text{inf}}, b^{\text{sup}}]$. $[a_R^{\text{inf}}, a_R^{\text{sup}}]$ and $[a_I^{\text{inf}}, a_I^{\text{sup}}]$ represent the interval of the real and imaginary parts, respectively



$$P(\theta, \phi) = P_R(\theta, \phi) + P_I(\theta, \phi) \\ = |f_R(\theta, \phi)|^2 + |f_I(\theta, \phi)|^2 \quad (3.21)$$

Since $a_n > 0$, the bound of the real part of the array pattern $[f_R^{\text{inf}}(\theta, \phi), f_R^{\text{sup}}(\theta, \phi)]$ can be derived as

$$f_R^{\text{inf}}(\theta, \phi) = \sum_{n=1}^N a_n^{\text{inf}} (\cos b_n)^{\text{inf}} \\ f_R^{\text{sup}}(\theta, \phi) = \sum_{n=1}^N a_n^{\text{sup}} (\cos b_n)^{\text{sup}} \quad (3.22)$$

where

$$(\cos b)^{\text{inf}} = \begin{cases} -1 & \text{if } (2k+1)\pi \in [b^{\text{inf}}, b^{\text{sup}}] \\ \min \{ \cos(b^{\text{inf}}), \cos(b^{\text{sup}}) \} & \text{otherwise} \end{cases} \quad (3.23)$$

$$(\cos b)^{\text{sup}} = \begin{cases} 1 & \text{if } 2k\pi \in [b^{\text{inf}}, b^{\text{sup}}] \\ \max \{ \cos(b^{\text{inf}}), \cos(b^{\text{sup}}) \} & \text{otherwise} \end{cases} \quad (3.24)$$

A similar expression can be obtained for the imaginary part, and the interval of the array power pattern is as,

$$P^{\text{inf}}(\theta, \phi) = P_R^{\text{inf}}(\theta, \phi) + P_I^{\text{inf}}(\theta, \phi) \\ P^{\text{sup}}(\theta, \phi) = P_R^{\text{sup}}(\theta, \phi) + P_I^{\text{sup}}(\theta, \phi) \quad (3.25)$$

where

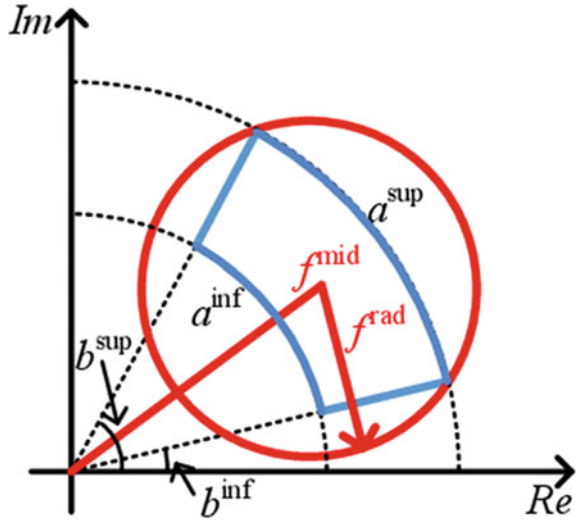
$$P_R^{\text{sup}} = \max \left\{ |f_R^{\text{inf}}|^2, |f_R^{\text{sup}}|^2 \right\} \\ P_R^{\text{inf}} = \begin{cases} 0 & \text{if } f_R^{\text{inf}} \leq 0 \leq f_R^{\text{sup}} \\ \min \left\{ |f_R^{\text{inf}}|^2, |f_R^{\text{sup}}|^2 \right\} & \text{otherwise} \end{cases} \quad (3.26)$$

It should be mentioned that, due to the intrinsic of the complex notation, the real and imaginary parts of the array pattern in Eq. (3.22) are dependent and highly correlated to each other. If an interval variable occurs several times in the expression, and each occurrence is taken independently, then this so-called dependency problem will cause expansion of the resulting intervals. In the above derivation, the array excitation error occurs in both the real and imaginary parts, this will cause an unwanted overestimation when adding them together.

3.2.3 The Circular IA (cIA) Method

For the circular IA (cIA) method, as shown in Fig. 3.5, the geometry of the excitation error is represented into the circular form. It's denoted as $\langle f^{\text{mid}}, f^{\text{rad}} \rangle$, where f^{mid}

Fig. 3.5 Illustration of the Circular IA method, where $\langle f^{\text{mid}}, f^{\text{rad}} \rangle$ denotes the center f^{mid} and the radius f^{rad}



is the center and f^{rad} is the radius covering all possible excitation error values [11]. The advantage of the circular IA method is the simplicity of IA calculation, the radius of the array pattern is the sum of the radius of the interval of each element, respectively,

$$f^{\text{rad}} = \sum_{n=1}^N f_n^{\text{rad}} \quad (3.27)$$

where f_n^{rad} is the radius covering the excitation error of the n -th element,

$$f_n^{\text{rad}} = \max \left\{ \begin{aligned} &|a_n \exp(jb_n) - a_n^{\text{inf}} \exp(jb_n^{\text{inf}})|, \\ &|a_n \exp(jb_n) - a_n^{\text{inf}} \exp(jb_n^{\text{sup}})|, \\ &|a_n \exp(jb_n) - a_n^{\text{sup}} \exp(jb_n^{\text{inf}})|, \\ &|a_n \exp(jb_n) - a_n^{\text{sup}} \exp(jb_n^{\text{sup}})| \end{aligned} \right\} \quad (3.28)$$

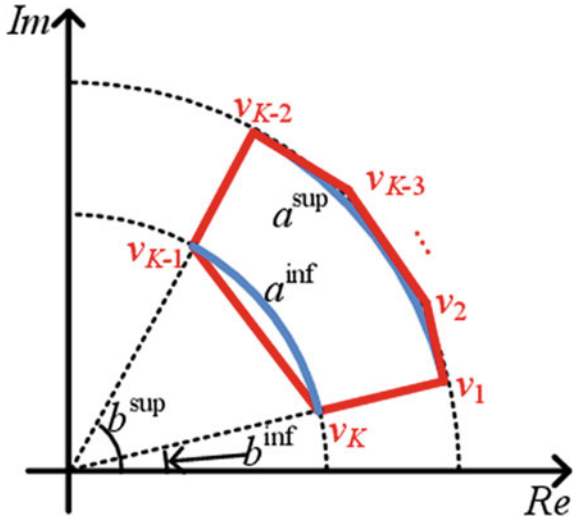
The bounds of the array power pattern are obtained by first calculating the bounds of the array pattern and then squaring it,

$$P^{\text{sup}} = (|f^{\text{mid}}| + f^{\text{rad}})^2$$

$$P^{\text{inf}} = \begin{cases} (|f^{\text{mid}}| - f^{\text{rad}})^2 & \text{if } |f^{\text{mid}}| > f^{\text{rad}} \\ 0 & \text{otherwise} \end{cases} \quad (3.29)$$

In this way, the excitation of the antenna element as an interval variable appears only once in Eq. (3.27). However, the radius is enlarged due to the expression of the circular IA form.

Fig. 3.6 Illustration of the Polygonal IA method. The boundary is a convex polygon with K vertexes $\{v_n\}$



3.2.4 The Polygonal IA (pIA) Method

As shown in Fig. 3.6, a more compact geometry is to expand the excitation error region as a convex polygon. The outer arc of the possible region is surrounded by several tangent lines, while the inner arc is surrounded by a line between the two vertices. In such a case, the computation procedure of the sum of two convex polygons is known as the Minkowski sum. Let V and W be two convex polygons representing possible regions of the excitation errors of two antenna elements with K vertices $\{v_1, \dots, v_K\}$ and M vertices $\{w_1, \dots, w_M\}$, respectively. Their Minkowski sum can be computed as follows [12],

Step 1: initial $k=1, m=1$;

Step 2: add $v_k + w_m$ as the vertex of the sum.

Step 3: if $\text{angle}(v_k v_{k+1}) \leq \text{angle}(w_m w_{m+1})$, then $k = k+1$;
 else if $\text{angle}(v_k v_{k+1}) \geq \text{angle}(w_m w_{m+1})$, then $m = m+1$;
 else $k = k+1, m = m+1$.

Step 4: repeat step 2 and step 3 until $k = K+1$ and $m = M+1$.

The above procedure runs in linear time because at each execution of the repeat loop either k or m is incremented. Moreover, the Minkowski sum is also a convex polygon with at most $K + M$ edges. Hence, by sequentially adding each element with $N-1$ times of Minkowski sum calculation, the interval of the array pattern can be obtained with improved bounds thus more reliable results. An extensive comparison of these three IA based methods with representative numerical results can be found in [13]. Generally speaking, the Polygonal IA method performs the best for most scenarios, while the Circular IA method performs the worst. An analytic method is also proposed to analyze and further dig the information from the shape of the

convex polygon of the array pattern, and map it into the probability distribution of the bound interval [14].

3.2.5 The Matrix IA (mIA) Method

If only weighted amplitude is used for antenna array synthesis, the above equations can be simplified. Recalling that an antenna array with N isotropic elements, each element located at the position $\{\vec{r}_n\}$ has the excitation amplitude $\{a_n\}$ and phase $\{p_n\}$, $n = 0, 1, \dots, N - 1$. The array radiation pattern or the array factor $f(\theta, \phi)$ can be mathematically expressed as follows,

$$f(\theta, \phi) = \sum_{n=1}^N a_n \exp(jk\vec{r}_0(\theta, \phi) \cdot \vec{r}_n + p_n) = \sum_{n=1}^N a_n \exp(jb_n) = \mathbf{a}E \quad (3.30)$$

where $\mathbf{a} = \{a_n\}$ and E is a $N \times 1$ vector with $E_n = \exp(jb_n)$. Let the interval of the excitation amplitude of the n -th element is $\langle a_n^{\text{mid}}, a_n^{\text{rad}} \rangle$, where a_n^{mid} and a_n^{rad} are the midpoint and the radius, respectively.

For the cIA method, the interval of the array factor $\langle f^{\text{mid}}, f^{\text{rad}} \rangle$ is calculated directly by summarizing the magnitudes of each element,

$$f^{\text{mid}} = \mathbf{a}^{\text{mid}} E, f^{\text{rad}} = \mathbf{a}^{\text{rad}} |E| = \sum_n a_n^{\text{rad}} \quad (3.31)$$

Then the antenna array pattern upper bound can be estimated as [8],

$$\begin{aligned} P_{cIA}^{\text{sup}} &= (|f^{\text{mid}}| + f^{\text{rad}})^2 = (|\mathbf{a}^{\text{mid}} E| + \mathbf{a}^{\text{rad}} |E|)^2 \\ &= \mathbf{a}^{\text{mid}} E E^* \mathbf{a}^{\text{mid}T} + 2 |\mathbf{a}^{\text{mid}} E| |E|^T \mathbf{a}^{\text{rad}T} + \mathbf{a}^{\text{rad}} |E| |E|^T \mathbf{a}^{\text{rad}T} \end{aligned} \quad (3.32)$$

Generally speaking, due to the intrinsic of the complex notation, the radius of the array factor interval is enlarged relatively too much by Eq. (3.30), which leads to a poor estimation for antenna array pattern bound estimation.

For the rIA algorithm, the array factor is split into the real and imaginary parts of the Cartesian form,

$$\begin{aligned} f(\theta, \phi) &= \sum_{n=1}^N a_n \cos b_n + j \sum_{n=1}^N a_n \sin b_n \\ &= f_R + f_I = \mathbf{a}C + j\mathbf{a}S \end{aligned} \quad (3.33)$$

where C and S are two $N \times 1$ vectors with $C_n = \cos b_n$ and $S_n = \sin b_n$. And f_R and f_I stand for the real and imaginary parts, respectively, their intervals $\langle f_R^{\text{mid}}, f_R^{\text{rad}} \rangle$ and $\langle f_I^{\text{mid}}, f_I^{\text{rad}} \rangle$ can be expressed as follows,

$$f_R^{\text{mid}} = \mathbf{a}^{\text{mid}} C, f_R^{\text{rad}} = \mathbf{a}^{\text{rad}} |C| \quad (3.34)$$

$$f_I^{\text{mid}} = \mathbf{a}^{\text{mid}} S, f_I^{\text{rad}} = \mathbf{a}^{\text{rad}} |S| \quad (3.35)$$

For the rIA algorithm, the antenna array pattern upper bound is the combination of these two parts,

$$\begin{aligned} P_{rIA}^{\text{sup}} &= P_R^{\text{sup}} + P_I^{\text{sup}} = (|f_R^{\text{mid}}| + f_R^{\text{rad}})^2 + (|f_I^{\text{mid}}| + f_I^{\text{rad}})^2 \\ &= (|\mathbf{a}^{\text{mid}} C| + \mathbf{a}^{\text{rad}} |C|)^2 + (|\mathbf{a}^{\text{mid}} S| + \mathbf{a}^{\text{rad}} |S|)^2 \\ &= \mathbf{a}^{\text{mid}} (CC^T + SS^T) \mathbf{a}^{\text{mid}T} + 2(|\mathbf{a}^{\text{mid}} C| |C|^T + |\mathbf{a}^{\text{mid}} S| |S|^T) \mathbf{a}^{\text{rad}T} \\ &\quad + \mathbf{a}^{\text{rad}} (|C| |C|^T + |S| |S|^T) \mathbf{a}^{\text{rad}T} \end{aligned} \quad (3.36)$$

It's noteworthy that, the real and imaginary parts are dependent and highly correlated to each other. Thus the array excitation amplitude occurred in both of these two parts would also cause overestimation of the array power pattern.

A novel matrix-based IA method (mIA) is proposed for array tolerance analysis with excitation amplitude errors in [15, 16]. Rather than in the Cartesian form or the complex form, the antenna array power pattern is expressed in the quadratic matrix form. This enable us to predict the array power pattern bounds in a very straightforward way. For the mIA algorithm, the antenna array power pattern is expressed in the quadratic matrix form,

$$\begin{aligned} P_{mIA} &= |f(\theta)|^2 = f(\theta) f^*(\theta) = \mathbf{a} E \mathbf{a} E^* = \mathbf{a} E E^{*T} \mathbf{a}^T \\ &= \mathbf{a} (C + jS) (C - jS)^T \mathbf{a}^T = \mathbf{a} (CC^T + SS^T) \mathbf{a}^T = \mathbf{a} \Theta \mathbf{a}^T \end{aligned} \quad (3.37)$$

where

$$\Theta = E E^{*T} = CC^T + SS^T \quad (3.38)$$

Hence, the antenna array upper bound derived by the mIA algorithm should be,

$$P_{mIA}^{\text{sup}} = \mathbf{a}^{\text{mid}} \Theta \mathbf{a}^{\text{mid}T} + 2 |\mathbf{a}^{\text{mid}} \Theta| \mathbf{a}^{\text{rad}T} + \mathbf{a}^{\text{rad}} |\Theta| \mathbf{a}^{\text{rad}T} \quad (3.39)$$

By comparing Eqs. (3.32), (3.36) and (3.39), one can see that the first items for these three algorithms are all the same. However, the latter two items denoting amplitude error introduced impacts are different. For the cIA and mIA algorithms,

$$|\mathbf{a}^{\text{mid}} \Theta| \mathbf{a}^{\text{rad}T} = |\mathbf{a}^{\text{mid}} E E^{*T}| \mathbf{a}^{\text{rad}T} \leq |\mathbf{a}^{\text{mid}} E| |E|^T \mathbf{a}^{\text{rad}T} \quad (3.40)$$

$$\mathbf{a}^{\text{rad}} |\Theta| \mathbf{a}^{\text{rad}T} = \mathbf{a}^{\text{rad}} |E E^{*T}| \mathbf{a}^{\text{rad}T} \leq \mathbf{a}^{\text{rad}} |E| |E|^T \mathbf{a}^{\text{rad}T} \quad (3.41)$$

And for the rIA and mIA algorithms,

$$\begin{aligned}
|\mathbf{a}^{\text{mid}} \Theta| \mathbf{a}^{\text{rad}T} &= |\mathbf{a}^{\text{mid}} (CC^T + SS^T)| \mathbf{a}^{\text{rad}T} \\
&\leq (|\mathbf{a}^{\text{mid}} CC^T| + |\mathbf{a}^{\text{mid}} SS^T|) \mathbf{a}^{\text{rad}T} \\
&\leq (|\mathbf{a}^{\text{mid}} C| |C|^T + |\mathbf{a}^{\text{mid}} S| |S|^T) \mathbf{a}^{\text{rad}T}
\end{aligned} \tag{3.42}$$

$$\begin{aligned}
\mathbf{a}^{\text{rad}} |\Theta| \mathbf{a}^{\text{mid}T} &= \mathbf{a}^{\text{rad}} |CC^T + SS^T| \mathbf{a}^{\text{mid}T} \\
&\leq \mathbf{a}^{\text{rad}} (|CC^T| + |SS^T|) \mathbf{a}^{\text{mid}T} \\
&\leq \mathbf{a}^{\text{rad}} (|C| |C|^T + |S| |S|^T) \mathbf{a}^{\text{mid}T}
\end{aligned} \tag{3.43}$$

Hence, we can conclude that the mIA algorithm has better performance than the rIA and cIA algorithms for antenna array pattern upper bound estimation.

3.2.6 Numerical Examples for Performance Comparison

In order to demonstrate the effectiveness of the proposed method, a half-wavelength equally spaced 10-element linear antenna array is used as the benchmark architecture in the following discussion. The referenced array pattern is Taylor tapered with $SLL^{\text{ref}} = -20$ dB and $\bar{n} = 2$, and the corresponding excitation amplitudes are $a_1 = a_{10} = 0.542$, $a_2 = a_9 = 0.629$, $a_3 = a_8 = 0.771$, $a_4 = a_7 = 0.913$, and $a_5 = a_6 = 1.000$. Without loss of generality, we assume the possible excitation amplitude deviates the same along its referenced value in both positive and negative directions, then we have $a_n^{\text{mid}} = a_n^{\text{ref}}$. In such case, the midpoint value of the array power pattern is equal to the referenced/expected array power pattern. We also assume the excitation amplitude error is proportional to the referenced value, i.e., $a_n^{\text{rad}} = \gamma_n * a_n^{\text{ref}}$, where γ_n is the tolerance ratio. Figure 3.7 illustrates the referenced array power pattern and 10000 Monte Carlo simulation results, along with the bounds estimated by the mIA method when the tolerance ratio $\gamma = 5\%$. During the Monte Carlo simulation, all γ_n are assumed to have uniform probability density and independent to each other. It clearly shows that, all the Monte Carlo simulations vary around the referenced array power pattern but none of them exceeds the lower and upper bounds. This verifies the effectiveness of the proposed mIA method.

We further investigate the statistical results of the normalized power level at $\theta = 5.7^\circ$, which corresponds to the half power level (-3 dB) of the referenced array power pattern. The bounds at $\theta = 5.7^\circ$ estimated by the rIA methods are $[-3.48$ dB, -2.52 dB] with the interval of 0.96 dB, while the bounds estimated by the cIA method are $[-3.63$ dB, -2.40 dB] with the interval of 1.23 dB. The mIA method shows the smallest estimation bounds of $[-3.45$ dB, -2.57 dB] and also the shortest interval of 0.88 dB. Figure 3.8 illustrates the histogram of the Monte Carlo simulation results, the width of each bin is set as 0.02 dB. It can be seen that, due to the central limit theorem, it has a nearly normal probability distribution with the center of -3.0 dB equal to its referenced value. Most of the Monte Carlo simulation results at $\theta = 5.7^\circ$ are in the range of $[-3.2$ dB, -2.8 dB], and have the minimum value of -3.34 dB

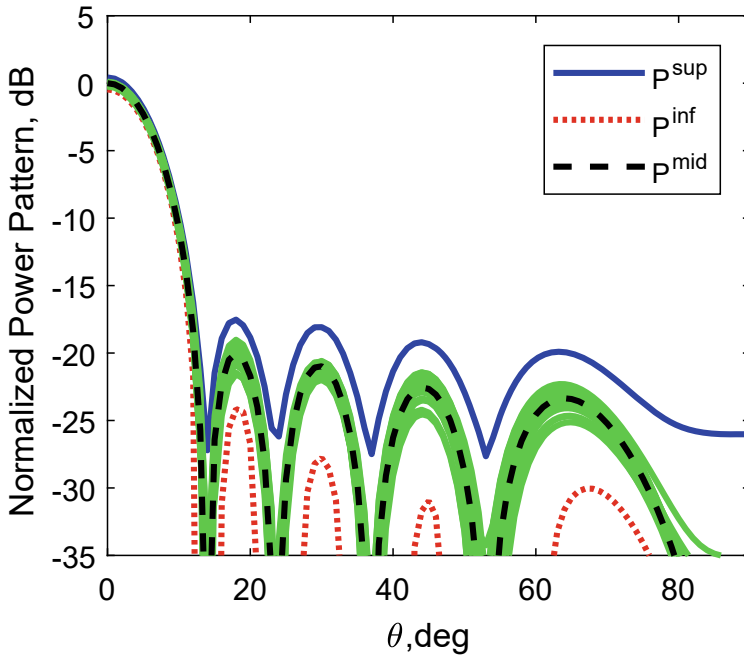


Fig. 3.7 The referenced array power pattern and the bounds estimated by the mIA method of the 10-element Taylor tapered linear array with Monte Carlo simulation results ($N = 10$, $d = \lambda/2$, $SLL^{\text{ref}} = -20$ dB, $\bar{n} = 2$, $\gamma = 5\%$)

and the maximum value of -2.73 dB, which fall in all the bounds predicted by these three methods.

Since the cIA method performs the worst than the other two methods for most scenarios, we present much fewer results of cIA in the following discussion for simplicity. Figure 3.9 illustrates the upper and lower bounds estimated by the rIA and mIA methods under various tolerance ratios. It indicates that a larger tolerance ratio leads to a higher upper bound and a smaller lower bound simultaneously, and both the upper and lower bounds become worse when the tolerance increases. To compare the performance of the rIA and mIA methods more comparatively, we define their difference as $\Delta P^{\text{sup}} = P_{mIA}^{\text{sup}} - P_{rIA}^{\text{sup}}$ and $\Delta P^{\text{inf}} = P_{mIA}^{\text{inf}} - P_{rIA}^{\text{inf}}$, and the results are illustrated in Fig. 3.10. It shows that ΔP^{sup} remains negative for all cases, thus the mIA method has a better performance for the upper bound estimation than the rIA method. The mean values of ΔP^{sup} are -0.33 dB, -0.53 dB and -0.61 dB for the tolerance ratio $\gamma = 1\%$, 3% and 5% , respectively. These values can be viewed as the average performance improvement for upper bound estimation by the mIA method. Moreover, it can also be seen that, the improvement at the array pattern nulls are much larger than other regions. For example, at the second pattern null $\theta = 24^\circ$, the P^{sup} estimated by the mIA and rIA methods are -35.1 dB and -34.2 dB for

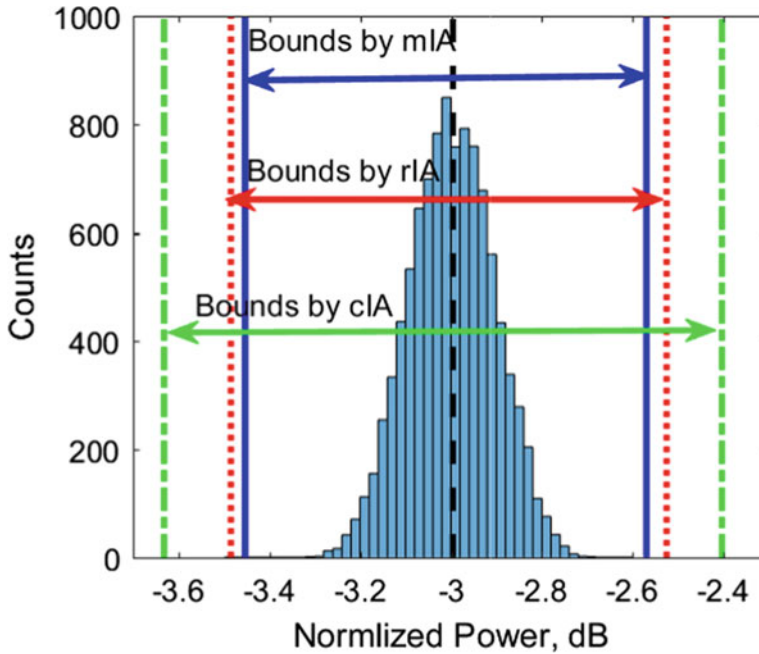


Fig. 3.8 The statistics of the normalized power level at $\theta = 5.7^\circ$ (half power level point) by 10000 Monte Carlo simulations

$\gamma = 1\%$, respectively, which means the improvement of 0.9 dB. For the lower bound estimation, these two methods show comparable performance. Sometimes one is better than the other, but this does not hold true for all the time.

3.3 Summary

Because of mechanical manufacturing imperfections, component aging, temperature variation, and circuit differences, the realistic antenna element excitation inevitably differs from their expected values in practice. Two kinds of impact analysis methods are introduced to evaluate impacts of array excitation errors. The probabilistic methods provide simple closed-form expressions of the features of the array pattern by exploiting the central limit theorem. On the contrary, the IA based methods introduce intervals to represent the element excitation errors and predict the array performance with its upper and lower bounds. Thanks to the inclusion property of IA to deal with uncertainties, the determined bounds of antenna array pattern are finite and inclusive thus reliable.

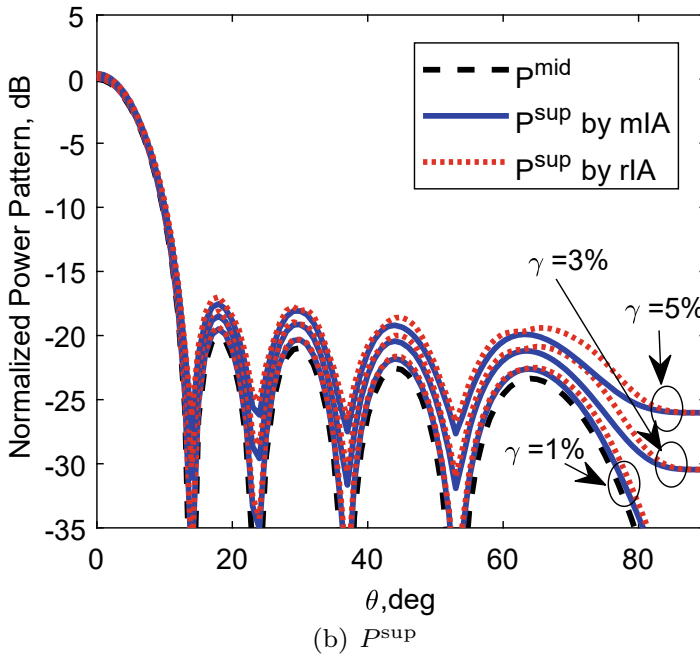
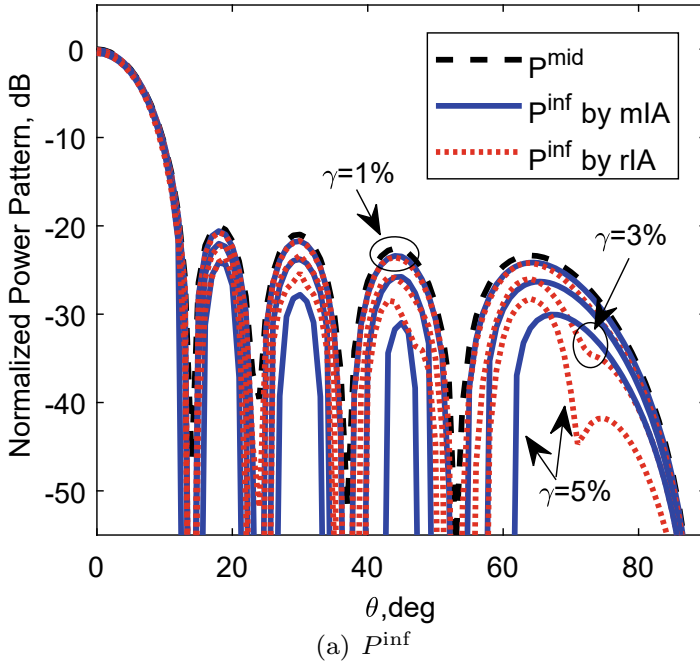


Fig. 3.9 The array power pattern bounds estimated by the rIA and mIA methods of the 10-element Taylor tapered linear array

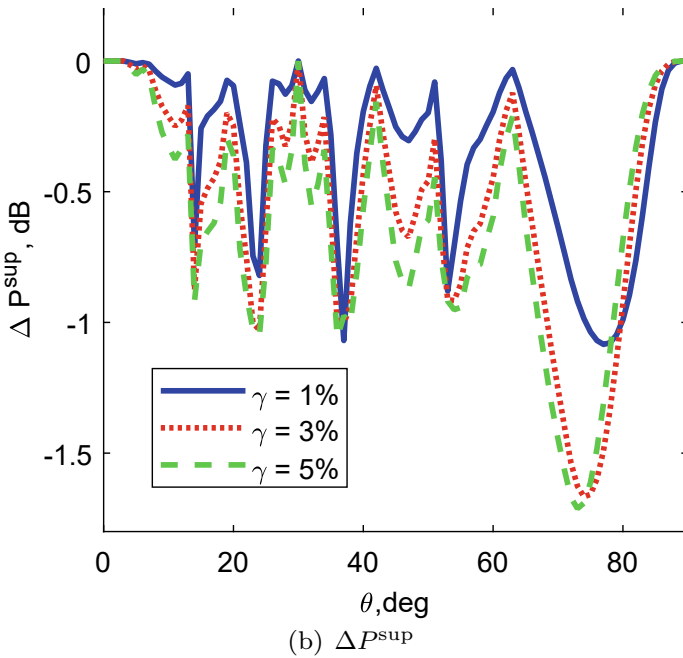
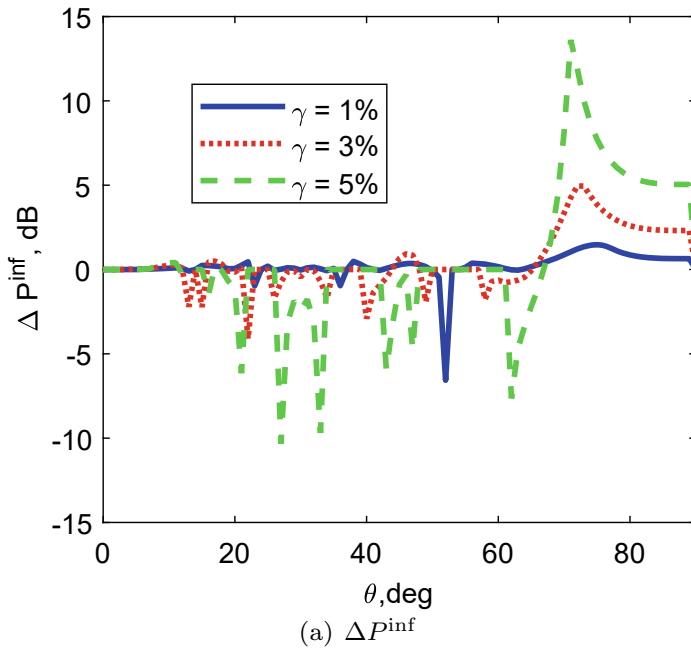


Fig. 3.10 The array power pattern bound differences estimated by the rIA and mIA methods of the 10-element Taylor tapered linear array

References

1. J. Ruze, The effect of aperture errors on the antenna radiation pattern. *Nuovo Cimento* **9**(3), 364–380 (1952)
2. J.K. Hsiao, Design of error tolerance of a phased array. *Electron. Lett.* **21**(19), 834–836 (1985)
3. J. Lee, Y. Lee, H. Kim, Decision of error tolerance in array element by the Monte Carlo algorithm. *IEEE Trans. Antennas Propag.* **53**(4), 1325–1331 (2005)
4. A.J. van den Biggelaar, Improved statistical model on the effect of random errors in the phase and amplitude of element excitations on the array radiation pattern. *IEEE Trans. Antennas Propag.* **66**(5), 2309–2317 (2018)
5. E.L. Holzman, A different perspective on taper efficiency for array antennas. *IEEE Trans. Antennas Propag.* **51**(10), 2963–2967 (2003)
6. A.K. Bhattacharyya, *Phased Array Antennas, Floquet Analysis, Synthesis, BFNs and Active Array Systems* (Wiley Inc., Hoboken, 2006)
7. P. Rocca, L. Manica, N. Anselmi, A. Massa, Analysis of the pattern tolerances in linear arrays with arbitrary amplitude errors. *IEEE Antennas Wirel. Propag. Lett.* **12**, 639–642 (2013)
8. N. Anselmi, L. Manica, P. Rocca, A. Massa, Tolerance analysis of antenna arrays through interval arithmetic. *IEEE Trans. Antennas Propag.* **61**(11), 5496–5507 (2013)
9. R.E. Moore, R.B. Kearfott, M.J. Cloud, *Introduction to Interval Analysis* (Society for Industrial and Applied Mathematics, Philadelphia, 2009)
10. S.M. Rump, INTLAB - INTerval LABoratory, in *Developments in Reliable Computing*, ed. by T. Csendes (Kluwer Academic Publishers, Dordrecht, 1999)
11. N. Anselmi, P. Rocca, M. Salucci, A. Massa, Power pattern sensitivity to calibration errors and mutual coupling in linear arrays through circular interval arithmetics. *Sensors* **16**(6), 1–14 (2016)
12. M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars, *Computational Geometry Algorithms and Applications*, 3rd ed. (Springer, Berlin, 2008), Chap. 13, p. 297
13. L. Tenuti, N. Anselmi, P. Rocca, M. Salucci, A. Massa, Minkowski sum algorithm for planar arrays sensitivity analysis with uncertain-but-bounded excitation tolerances. *IEEE Trans. Antennas Propag.* **65**(1), 167–177 (2017)
14. P. Rocca, N. Anselmi, A. Benoni, A. Massa, Probabilistic interval analysis for the analytic prediction of the pattern tolerance distribution in linear phased arrays with random excitation errors. *IEEE Trans. Antennas Propag.* **68**(12), 7866–7878 (2020)
15. G. He, X. Gao, H. Zhou, H. Zhu, Comparison of three interval arithmetic based algorithms for antenna array pattern upper bound estimation. *Electron. Lett.* **55**(14), 775–776 (2019)
16. G. He, X. Gao, H. Zhou, Matrix-based interval arithmetic for linear array tolerance analysis with excitation amplitude errors. *IEEE Trans. Antennas Propag.* **67**(5), 3516–3520 (2019)

Chapter 4

Array Excitation Error Calibration



4.1 Calibration Methods

The objective of phased array antenna calibration is to sample each antenna element in the array, and compare the obtained results to identify the differences among the elements. Phased array calibration for both transmit and receive are similar, the primary difference lies in the test-signal distribution at the input and the combination at the output. For receive calibration, a calibration source and/or a distribution network is required to inject the test signal into the input of each element, the array beam-former can be used as a test signal combiner. For transmit calibration, it requires a calibration source signal into each element and additional equipment or beam combiner for combined signal monitoring. Only receive calibration is presented here to keep the discussion focused, the extension for transmit calibration is straightforward. The calibration methods vary from different systems and applications, either suitable for factory test or for in-field test. These calibration methods can be generally divided into four main categories, i.e., the near-field scanning probe method, the peripheral fixed probe method, the mutual coupling method, and the built-in network method. The advantages and challenges of these calibration methods are shortly summarized in Table 4.1 and presented in the following subsections.

4.1.1 *The Near-Field Scanning Probe Method*

The traditional near-field scanning probe method, also known as “park and probe”, is one of the most widely used methods in the industry for phased array antenna calibration [1–3]. It’s considered to be very reliable and accurate with element-level pattern accuracy as high as 0.1 dB and 0.1 deg, respectively [3]. The procedure is straightforward, the tested element is excited to its default state, a robotic manipulator places a near-field scanning probe antenna at the boresight of the tested element, to directly measure the relative phase and amplitude by Vector Network Analyzer

Table 4.1 The advantages and challenges of phased array calibration methods

Method	Advantages	Challenges
The near-field scanning probe method	Reliable and accurate	Requiring a precise automated mechanical system Time-consuming Only suitable for in-factory test
The peripheral fixed probe method	Widely used in-field test Calibration can be done in parallel Calibration can be done in noncoherent	Requiring additional hardware Involving complicated signal processing
The mutual coupling method	No external equipment required Calibration can be done in parallel	Would encounter unwanted edge effect
The built-in network method	High accuracy and efficiency Calibration can be done in parallel Calibration can be done in noncoherent	Requiring dedicated coupling network

(VNA). Then it moves to test the next element under the same testing condition, and repeats this procedure until all elements of the antenna array are tested.

As depicted in Fig. 4.1, supposing g_p^t is the transmit gain of the calibration probe, C_{pn} is the coupling coefficient between the probe antenna and the n -th antenna element under test, and g_n^r is the receive gain of the n -th element to be calibrated. Then the signal received by the n -th antenna element R_{pn} can be formulated as,

$$R_{pn} = g_p^t C_{pn} g_n^r \quad (4.1)$$

Analogously, the signal received by the m -th antenna element R_{pm} can be formulated as,

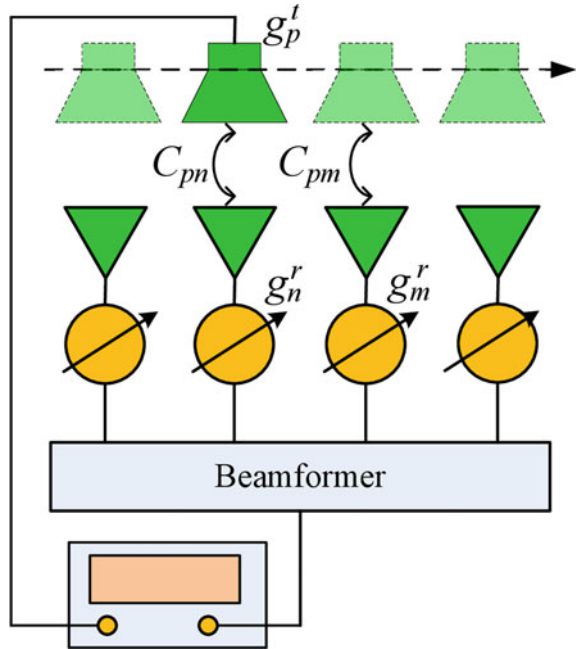
$$R_{pm} = g_p^t C_{pm} g_m^r \quad (4.2)$$

Since the technique scans and tests all the antenna elements under the same condition, it assumes that the coupling coefficient between the probe and the tested element is the same for all array elements, i.e., $C_{pn} = C_{pm}$. Hence, the relative excitation error between these two elements can be obtained as,

$$\frac{g_n^r}{g_m^r} = \frac{R_{pn}}{R_{pm}} \quad (4.3)$$

Usually, probe position error will result in some of the measurement error in near-field measurement [4]. However, for phased array near-field scanning probe calibration, it is not necessary to keep the probe antenna precisely at the boresight of

Fig. 4.1 Simplified diagram of the near-field scanning probe method



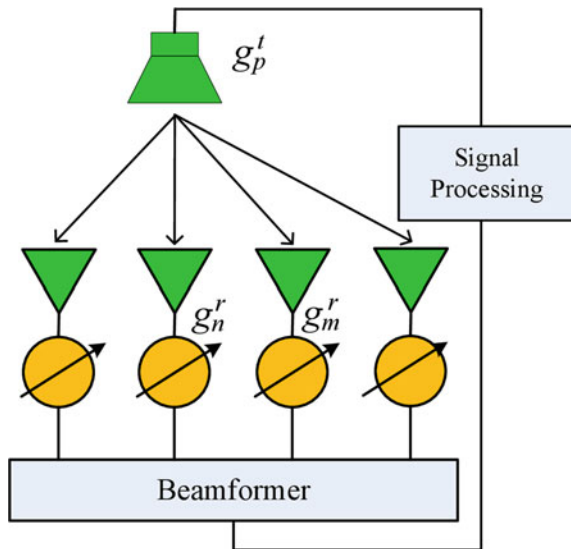
each antenna element under test. It is sufficient to just keep the relative position and angle between the probe antenna and the tested element unchanged. In that case, even if the probe is somewhat offset from its desired position for every antenna element, the amount of offsets and therefore couplings will be still the same for all tested antenna elements, making this technique still suitable for use [5].

The near-field scanning probe method usually requires a precise automated mechanical system including elements such as the axes controllers, actuators, motors for accurate probe movement. This makes the setup suitable to work only in the near-field environment, but too complicated to arrange it in the field test. Moreover, due to the mechanical movement of the probe antenna, the measurement process is very time-consuming, especially for large-scale phased array antennas. For example, for the THAAD (Theater High Altitude Area Defense) radar, each of the 25,344 T/R modules had to be near-field scanned individually, result in a very long and even unacceptable test time [6, 7]. Therefore, the near-field scanning probe method is most suitable for initial factory test rather than periodic in-field calibration.

4.1.2 The Peripheral Fixed Probe Method

As shown in Fig. 4.2, in some large phased array antennas, unlike the near-field scanning method with a moveable probe antenna, one or several fixed probe antennas

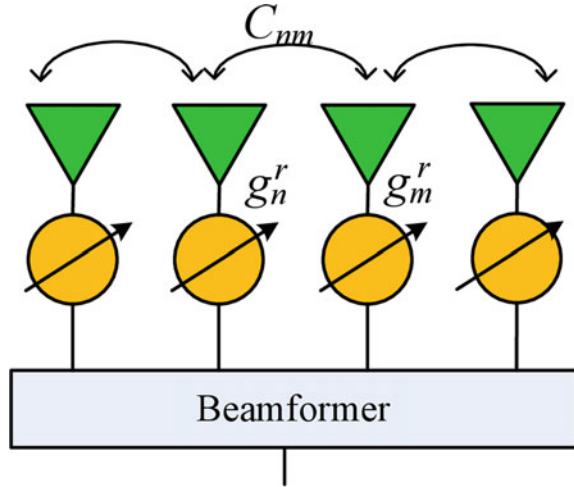
Fig. 4.2 Simplified diagram of the peripheral fixed probe method



are placed at the periphery of the array, or moving platform such as a balloon or unmanned aerial vehicle (UAV). In [8], calibration is made by external fixed probe antennas at certain locations near the array. A circle around the array center is chosen for symmetry reasons. It's proposed that one probe antenna is integrated at the side for a spaceborne phased array calibration [9]. High calibration accuracy can be achieved for actual satellite systems because this allows easy realization with high signal-to-noise ratio (SNR). The Space Fence, used to detect, track, and catalog small objects in Space Situational Awareness for the U.S. Air Force Space Surveillance Network, both the transmit and receive arrays are calibrated with horn antennas mounted on calibration towers in the near-field [10].

If the fixed probe is at the far-field of the array under test, then the effects of each antenna element can be viewed as the same. Otherwise, the coupling between the probe and each antenna element is going to be different. In this case, the peripheral fixed probe method can be applied only if the coupling amounts were previously measured or calibrated and moved out. This can be done by using a near-field scanning probe in factory test, then compared with the results from peripheral fixed probes made in-field to calculate the couplings. For example, the SAMPSON Multi-Function Radar has four fixed open waveguide auxiliary radiators used for calibration [11]. The signal received from an individual active channel is compared with a stored reference level obtained during the factory test of the phased array antenna. In this way, a replacement module or one whose characteristics have simply drifted with time may be reset to the original RF performance, thus restoring the original factory-fresh condition of the antenna. With the peripheral fixed probe method, some advanced signal processing techniques can be adopted to reduce the complexity of measurement hardware requirements, which is presented in detail in Sect. 4.2.

Fig. 4.3 Simplified diagram of the mutual coupling method



4.1.3 The Mutual Coupling Method

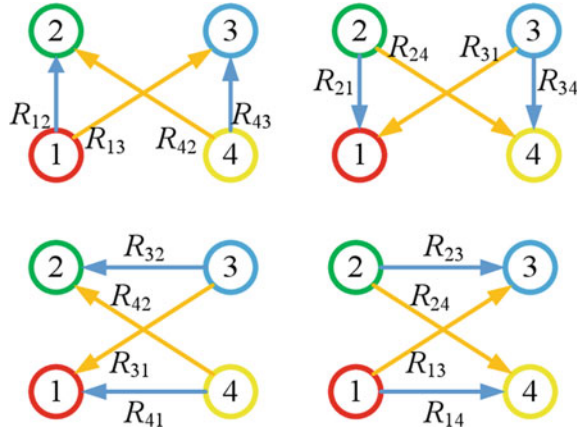
As shown in Fig. 4.3, the mutual coupling method was first proposed by Aumann et al. [12] on a linear antenna array, based on the idea that the inherent mutual coupling among the array elements can be used by transmitting from an element and receiving from another. The measured signals between all pairs of elements in the array allow a complete characterization of the relative amplitude and phase of each element in the array. Compared to the near-field scanning probe method and the peripheral fixed probe method, this method utilizes the inherent property of mutual coupling and requires no external hardware resources, resulting in a much faster calibration process and more suitable for in-field calibration.

The mutual coupling method was extended to two-dimensional (2D) antenna array and became more practically useful [13]. The detail of the mutual coupling method depends on the specific array geometry. Taking the most common rectangular array as an example, as shown in Fig. 4.4, there are four elements under test and two parameters need to be calibrated for each element. Let g_m^t denotes the transmit gain of the m -th element and g_n^r denotes the receive gain of the n -th element to be calibrated. A mutual coupling measurement consisting of a signal transmitted from the m -th element and received by the n -th element. When the #2 and #3 elements as receive and the #1 and #4 elements as transmit, two pairs of mutual coupling measurements are formulated as,

$$\begin{aligned} R_{12} &= g_1^t C_{12} g_2^r, & R_{13} &= g_1^t C_{13} g_3^r \\ R_{42} &= g_4^t C_{42} g_2^r, & R_{43} &= g_4^t C_{43} g_3^r \end{aligned} \quad (4.4)$$

From these four measurements, these elements can be calibrated relative to each other as,

Fig. 4.4 Coupling schemes for rectangular phased array calibration by performing the mutual coupling method



$$\frac{g_2^r}{g_3^r} = \left(\frac{R_{12}R_{42}}{R_{13}R_{43}} \right)^{1/2}, \quad \frac{g_1^t}{g_4^t} = \left(\frac{R_{12}R_{13}}{R_{42}R_{43}} \right)^{1/2} \quad (4.5)$$

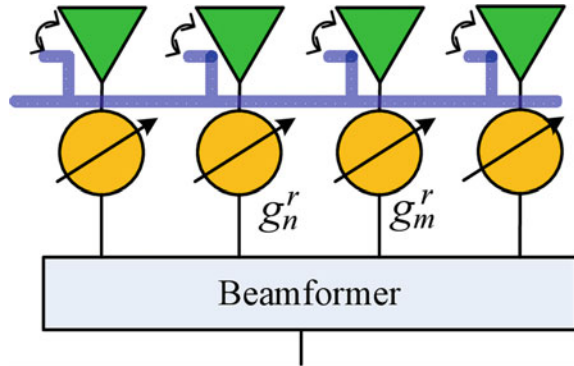
From the other three topographies with six more pairs of measurements, the relative transmit and receive gains of all four elements can be solved. Moreover, a mathematical framework by the least square method is proposed that using measurement results as many as possible [14].

There is also something to be mentioned for the mutual coupling method. It relies on the assumption that coupling between an antenna element and its neighbors are the same for all elements in the array. Nonetheless, it was revealed by using a prototype system, this assumption does not hold for small arrays due to unwanted edge effects [15]. A before/after approach is introduced to avoid errors introduced by edge effects, the ratio between the after and before status will quantify the changes suffered by the elements.

4.1.4 The Built-In Network Method

It's very common to use a built-in network connected to each antenna element for periodic in-field calibration of integrated phased array antennas, especially for Si-based RF and mmWave-band integrated phased array [16–18]. As shown in Fig. 4.5, this technique employs microstrip transmission line as test signal injector or weak signal coupler embedded under the element for calibration, with the coupling ratio between the transmission line and the radiating element usually several tens of dB. These calibration lines sample the signals received or transmitted by the antenna elements. The measured signals are then used to calculate the amplitude and phase differences among the elements. An early example of built-in performance

Fig. 4.5 Simplified diagram of the built-in network method



monitoring system was developed and tested in [19]. Test distribution networks can also be integrated with advanced signal processing techniques on-chip as described in Sect. 5.2.

To successfully calibrate a phased array using this method, phase shifts and amplitude losses caused by the transmission lines and the couplers that connect them should be equal or already known. A high-quality, embedded calibration network near the array face could source or receive nearly identical signal levels at each element. If the transmission lines are built using the same material and have the same radius and length, their effects would be very similar. For example, the ELTA/IAI Systems used microstrip line couplers near parallel plate waveguide radiators to individually probe T/R modules through the analog beamformers to supplement initial near-field calibration. The TerraSAR-X imaging satellite was one of the first SAR systems using embedded coupling to supplement the initial calibration once it is launched into space. For the S-band digital array testbed at Lincoln Laboratories, a coupler was used for injecting in front of the LNAs for receive-only calibration monitoring [20]. In [21], a coupled line network that is weakly coupled to the antenna array is designed to provide a calibration signal path without adding any switch after the antenna port.

4.2 Calibration Signal Processing Techniques

For the peripheral fixed probe method and the built-in network method for phased array calibration, it's very convenient to apply advanced signal processing techniques to save hardware or time requirements. As shown in Fig. 4.6, these techniques can be distinguished as serial or parallel by the number of elements to be calibrated each time. In serial calibration techniques, the antenna elements are characterized one by one, and the characterizations of different array elements are independent. Some techniques save time by handling several calibrations simultaneously, thus can be viewed as more efficient. These techniques can also be distinguished as coherent or

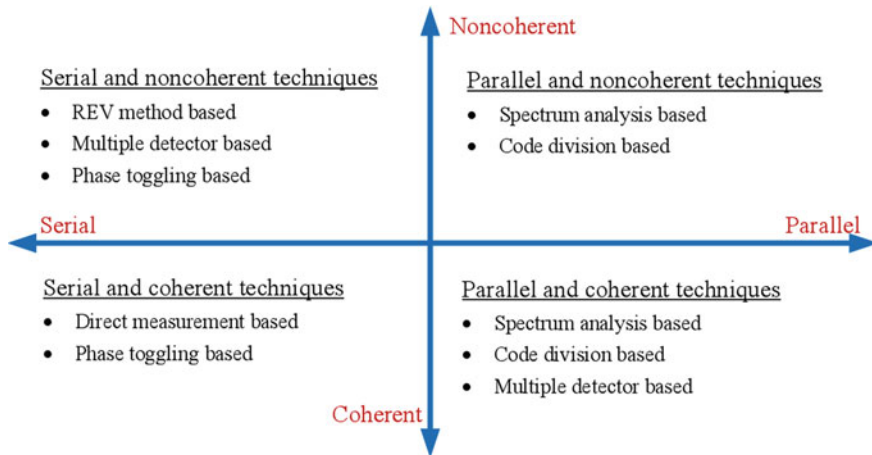


Fig. 4.6 Signal processing techniques for phased array calibration

noncoherent by the measurement requirements between the calibration source and the element under calibration. For coherent measurement, it's required that the calibration source and sink must be coherent to obtain the complex value containing both amplitude and phase information. On the contrary, noncoherent measurement requires no such synchronization, which can be done by amplitude or power only measurements. All these four kinds of techniques can be conducted with both the peripheral fixed probe method and the built-in network method, only slight differences in the measurement setup are required depending on practical applications.

4.2.1 Serial and Coherent Techniques

The serial and coherent techniques simply calibrate each element serially and require a coherent measurement to gather the element excitation information. The near-field scanning probe method can also be conducted as one of such kind of techniques. For some smart antenna applications and mobile satellite communication applications, these simple and direct calibration techniques are very common, especially with the element-level digital phased array. One of the main advantages of digital beamforming over every element is the ease that amplitude and phase errors among elements can be calculated in the digital domain.

There are also some measurement setups involving phase toggling. In [19], the phase of the element under calibration is switched between two states of 0 and π , while the other elements stay unchanged. Let $\vec{E}_n$ is the electric field of the n -th antenna element, and E_n and ϕ_n denote its amplitude and phase, $n = 1, 2, \dots, N$. The phased array electric field $\vec{E}_0$, the combination of the individual element electric fields, can be described as follows,

$$\begin{aligned}\vec{E}_0 &= \sum_{n=1}^N \vec{E}_n \\ &= \vec{E}_{\tilde{n}} + \vec{E}_n = E_{\tilde{n}} \exp(j\phi_{\tilde{n}}) + E_n \exp(j\phi_n)\end{aligned}\quad (4.6)$$

where $\vec{E}_{\tilde{n}} = E_{\tilde{n}} \exp(j\phi_{\tilde{n}})$ denotes the combined electric field of the array expect the n -th element. By shifting the phase π of the n -th antenna element,

$$\begin{aligned}\vec{E}_{\pi} &= \vec{E}_{\tilde{n}} + \vec{E}_n \exp(j\pi) = E_{\tilde{n}} \exp(j\phi_{\tilde{n}}) + E_n \exp(j\phi_n) \exp(j\pi) \\ &= E_{\tilde{n}} \exp(j\phi_{\tilde{n}}) - E_n \exp(j\phi_n)\end{aligned}\quad (4.7)$$

By comparing the combined array signal of these two measurements, the element excitation can be easily determined, while the other stationary element outputs can be canceled out,

$$\frac{\vec{E}_n}{\vec{E}_0} = \frac{1 - \vec{E}_{\pi} / \vec{E}_0}{2} \quad (4.8)$$

4.2.2 Serial and Noncoherent Techniques

Compared to coherent measurement methods, noncoherent techniques which involve amplitude or power only measurements are considered to be more convenient. In such a case, coherent measurements are no longer required with precise synchronization between the calibration source and the element under calibration. However, this is usually done with some additional costs. e.g., more power detectors or more times of measurements.

4.2.2.1 The Rotating Element Electric Field Vector Method

The rotating element electric field vector (REV) method is the most well-known power only measurement method with only one power detector. It measures the amplitude of the array combined signal by shifting the element phase from 0 to 2π continuously [22]. Recalling that the phased array electric field $\vec{E}_0$ is described as follows,

$$\vec{E}_0 = E_0 e^{j\phi_0} = \sum_{n=1}^N \vec{E}_n = \sum_{n=1}^N E_n e^{j\phi_n} \quad (4.9)$$

where E_0 and ϕ_0 are the amplitude and phase of the array electric field, respectively.

If the phase of the n -th element is shifted by φ_i , then the array electric field $\vec{E}_{0,i}$ becomes as

$$\begin{aligned}\vec{E}_{0,i} &= E_0 e^{j\phi_0} - E_n e^{j\phi_n} + E_n e^{j(\phi_n + \varphi_i)} \\ &= E_0 e^{j\phi_n} \left[e^{j(\phi_0 - \phi_n)} + \frac{E_n}{E_0} (e^{j\varphi_i} - 1) \right]\end{aligned}\quad (4.10)$$

Let k_n and ψ_n denote the relative amplitude and relative phase of the n -th element with the array electric field as,

$$\begin{aligned}k_n &= \frac{E_n}{E_0} \\ \psi_n &= \phi_0 - \phi_n\end{aligned}\quad (4.11)$$

Then Eq. (4.10) becomes as,

$$\begin{aligned}\vec{E}_{0,i} &= E_0 e^{j\phi_n} [e^{j\psi_n} + k_n (e^{j\varphi_i} - 1)] \\ &= E_0 e^{j\phi_n} [(\cos \psi_n - k_n + j \sin \psi_n) + k_n e^{j\varphi_i}] \\ &= E_0 e^{j\phi_n} (Y_n e^{j\psi_n} + k_n e^{j\varphi_i})\end{aligned}\quad (4.12)$$

where

$$\begin{aligned}Y_n^2 &= (\cos \psi_n - k_n)^2 + \sin^2 \psi_n \\ \tan \Phi_n &= \frac{\sin \psi_n}{\cos \psi_n - k_n}\end{aligned}\quad (4.13)$$

Then the relative power expression p_i become as,

$$p_i = \left| \vec{E}_{0,i} / \vec{E}_0 \right|^2 = Y_n^2 + k_n^2 + 2k_n Y_n \cos(\Phi_n + \varphi_i) \quad (4.14)$$

From Eq. (4.14), one can see that, the relative power changes in a cosine function when the element phase from 0 to 2π . Figure 4.7 illustrates one of such an example. The element excitation can be determined from statistics of three parameters, the maximum and minimum power of the array signal, and the element rotation phase to maximum power.

4.2.2.2 The Two Phase Shift Method

In [23], it was pointed out that measurement results at four orthogonal phase shifts are sufficient enough to obtain a maximum likelihood estimation for each array element. An even simpler expression was further derived in [24] that requires only two phase shifts of $\pi/2$ and π to yield the element complex excitation information. In order to calibrate the n -th array element complex excitation distortion, we can divide Equation (4.6) into two parts,

$$\begin{aligned}\vec{E}_0 &= \vec{E}_{\tilde{n}} + \vec{E}_n = E_{\tilde{n}} \exp(j\phi_{\tilde{n}}) + E_n \exp(j\phi_n) \\ &= (E_{\tilde{n}} \cos \phi_{\tilde{n}} + E_n \cos \phi_n) + j(E_{\tilde{n}} \sin \phi_{\tilde{n}} + E_n \sin \phi_n)\end{aligned}\quad (4.15)$$

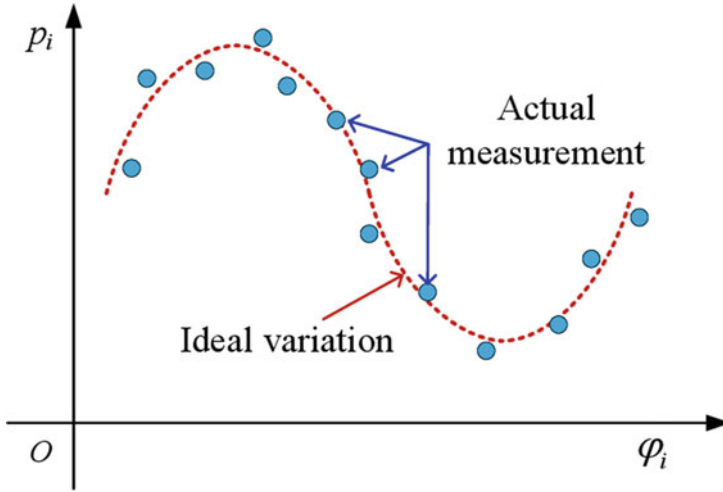


Fig. 4.7 Rower variation of the REV method with the phase shift

Then the array power P_0 can be expressed as,

$$\begin{aligned} P_0 &= E_0^2 = (E_{\tilde{n}} \cos \phi_{\tilde{n}} + E_n \cos \phi_n)^2 + (E_{\tilde{n}} \sin \phi_{\tilde{n}} + E_n \sin \phi_n)^2 \\ &= E_{\tilde{n}}^2 + E_n^2 + 2E_{\tilde{n}}E_n \cos(\phi_{\tilde{n}} - \phi_n) \end{aligned} \quad (4.16)$$

We shift the phase of the n-th antenna element by $\pi/2$, the corresponding electric filed $\vec{E}_{\pi/2}$ and the array power $P_{\pi/2}$ can be written as,

$$\begin{aligned} \vec{E}_{\pi/2} &= E_{\tilde{n}} \exp(j\phi_{\tilde{n}}) + E_n \exp(j\phi_n) \exp(j\pi/2) \\ &= (E_{\tilde{n}} \cos \phi_{\tilde{n}} - E_n \sin \phi_n) + j(E_{\tilde{n}} \sin \phi_{\tilde{n}} + E_n \cos \phi_n) \end{aligned} \quad (4.17)$$

$$\begin{aligned} P_{\pi/2} &= (E_{\tilde{n}} \cos \phi_{\tilde{n}} - E_n \sin \phi_n)^2 + (E_{\tilde{n}} \sin \phi_{\tilde{n}} + E_n \cos \phi_n)^2 \\ &= E_{\tilde{n}}^2 + E_n^2 + 2E_{\tilde{n}}E_n \sin(\phi_{\tilde{n}} - \phi_n) \end{aligned} \quad (4.18)$$

Similarly, by shifting the phase of the n-th antenna element by π , the corresponding electric filed $\vec{E}_{\pi}$ and the array power P_{π} can be written as,

$$\begin{aligned} \vec{E}_{\pi} &= E_{\tilde{n}} \exp(j\phi_{\tilde{n}}) + E_n \exp(j\phi_n) \exp(j\pi) \\ &= (E_{\tilde{n}} \cos \phi_{\tilde{n}} - E_n \cos \phi_n) + j(E_{\tilde{n}} \sin \phi_{\tilde{n}} - E_n \sin \phi_n) \end{aligned} \quad (4.19)$$

$$\begin{aligned} P_{\pi} &= (E_{\tilde{n}} \cos \phi_{\tilde{n}} - E_n \cos \phi_n)^2 + (E_{\tilde{n}} \sin \phi_{\tilde{n}} - E_n \sin \phi_n)^2 \\ &= E_{\tilde{n}}^2 + E_n^2 - 2E_{\tilde{n}}E_n \cos(\phi_{\tilde{n}} - \phi_n) \end{aligned} \quad (4.20)$$

From the above equations, we can derive that,

$$\begin{aligned} E_{\tilde{n}}^2 + E_n^2 &= \frac{P_0 + P_{\pi}}{2} \\ E_{\tilde{n}} E_n \cos(\phi_{\tilde{n}} - \phi_n) &= \frac{P_0 - P_{\pi}}{4} \\ E_{\tilde{n}} E_n \sin(\phi_{\tilde{n}} - \phi_n) &= \frac{2P_{\pi/2} - P_0 - P_{\pi}}{4} \end{aligned} \quad (4.21)$$

It was pointed out that [25], from the mathematical point of view, it's required at least two amplitude (power) measurements to determine the complex excitation of one array element.

4.2.3 Parallel and Coherent Techniques

Compared to serial techniques, parallel techniques that simultaneously calibrate a number of array elements are more efficient, which is especially suitable for large phased arrays with small calibration time slots. Some parallel techniques measure the array combined signals with several measurement probes at multiple positions. However, this requires some hardware burden. A more popular alternative technique allows simultaneously measuring excitations of multiple elements relying on code modulation. A method that controls the element phases based on time multiplexed orthogonal codes was proposed in [26, 27], thus the individual element excitation can be derived from the combined array signal. The pseudo-random (PN) code for phased array calibration has been successfully verified in a space born environment by TerraSAR-X [28]. In addition, a similar technique with Walsh codes was used for Sentinel-1 [29]. A recursive matrix-forming method for Hadamard matrix construction is presented in [30] for phased array calibration.

Besides these aforementioned methods, the spectrum analysis technique can also be adopted for phased array parallel calibration. In [31], a multi-element phase toggling method with Fast Fourier Transform (FFT) based spectrum analysis is proposed for parallel calibration. Each element under calibration is toggled with a particular odd time of fundamental step frequency. By applying FFT analysis of the combined complex signal, both the amplitude and phase of each element can be found in the corresponding spectrum.

4.2.4 Parallel and Noncoherent Techniques

The parallel and noncoherent techniques, with multiple element calibration each time and no coherent measurement required, should be the most preferred technique at the expense of more complex signal processing. In [32], the phases of multiple antenna elements are successively shifted with the specified phase intervals. The measured array power variation is expanded into Fourier series and the terms are rearranged to put them into the form of the conventional REV method. In [33],

another spectrum analysis based method is proposed by periodically modulating the element phase sequentially. The phase of the antenna element is shifted to π on one time slot while staying unchanged for the other time slots. By applying FFT analysis of the combined array signal, the element complex excitation is obtained from the corresponding modulation frequency harmonics.

In [34], a novel calibration technique that employs code-modulated interferometry is proposed for parallel measurements. Moreover, it requires only a simple power squaring detector without the need for a coherent receiver. First, the common test signal in the form of a single-frequency tone injected into each element, both the in- and quadrature-phase components of each element are encoding, then combined and squared using an on-chip power detector. Benefiting those orthogonal code products (OCPs) where the product of any two codes is another unique code, the squaring operation downconverts the combined signal to baseband and creates an interference pattern between all of the individual elemental responses. This pattern contains complex cross-correlations that are each modulated according to the OCP. Using each OCP, the correlation of interest can be demodulated, and then the full set of correlations can be used to extract amplitude and phase information for all elements.

4.2.5 *Machine Learning Based Techniques*

The rise and proliferation of artificial intelligence (AI) has found numerous applications including natural language processing, remote sensing, image recognition, and fraud detection. Machine learning (ML) approaches are scientific disciplines that build a mathematical model based on training data, to enable AI by improving an outcome through the experience without being explicitly programmed. Some popular ML techniques are radial basis functions (RBFs), support vector machines (SVMs), artificial neural networks (ANN), and deep neural networks (DNNs), just to name a few [35, 36]. Recently, it has also been gaining increasing popularity in the antenna and propagation community to solve complex electromagnetic problems [37–39]. For phased array calibration, a graph coloring theory based method is proposed in [40], which transforms the calibration problem into a coloring problem that aims at minimizing the number of used colors. In [41], the number of calibration measurements is minimized by using a compressed sensing (CS) approach. Sparsity is introduced into the combined signals by assigning binary delay vectors, which allows to recovery excitation of individual array elements in a computationally effective way. In [42], the ANN approach is applied to calibrate both transmit and receive channels for phased array antenna, which allows simple hardware implementation and requires little computational power. ML techniques are undoubtedly great analysis tools, particularly for problems with large and complex data sets. As ML techniques grow in sophistication and are still in boosting, more powerful ML based calibration techniques and practical applications are expected for phased array calibration and status monitoring.

4.3 Array Element Failure Diagnosis

The large antenna array consisting of thousands of elements has the possibility of failure of some of them. These element failures cause sharp variations in the field intensity across the array aperture, thus increasing both the side lobe and ripple level of the array radiation pattern. In such a case, the first thing that needs to know is the exact locations of the failed elements in the array. It may be too expensive to distribute a network of sensors integrated with the beamforming network for monitoring the array status in real time. Therefore, it is best to perform the antenna array diagnosis by measuring the array radiation field of a given number of spatial directions of the distorted radiation pattern. Several failure detection methods have been proposed for antenna array diagnosis and are shortly introduced in the this subsection.

4.3.1 The Backward Transform Method

The backward transformation method estimates the array element excitation coefficients by taking the Fourier transform relationship with the array radiation pattern [43]. For some applications, the array elements excitation coefficients are processed through Fourier transform to obtain the far-field pattern of the antenna under test. On the other hand, the array elements excitation coefficients can also be obtained through a backward transform to reconstruct the aperture field for diagnostic purposes. Consequently, this method is simple and straightforward.

However, to maintain an acceptable resolution, it is necessary to collect all the relevant energy radiated by the antenna as much as possible. Using the standard half-wavelength measurement step, the number of measurement points turns out to be very large, thus requiring an unaccepted data acquisition time. Moreover, it can only be used in a planar array limited by the Fourier transform.

4.3.2 The Pattern Mapping Method

A more effective way to reduce the set of data in antenna array diagnosis under a-priori knowledge of the shape of the antenna under test [44]. The failed elements can be detected by comparing the measured pattern with the nominal pattern. Suppose $f_p(\theta, \phi)$ is the array radiation pattern containing some failed elements,

$$f_p(\theta, \phi) = \sum_{n=1}^N p_n f_n(\theta, \phi) \quad (4.22)$$

where $f_n(\theta, \phi)$ is the element pattern and p_n is the status of the n -th element, which '1' denotes active and '0' denotes failed, respectively. With several measurements

y from M directions (θ_m, ϕ_m) , The problem becomes as choosing a set of p_n , to minimize the difference between the measurements and the nominal pattern,

$$\min \sum_{m=1}^M [y(\theta_m, \phi_m) - f_p(\theta_m, \phi_m)]^2 \quad (4.23)$$

Several evolutionary optimization techniques have been demonstrated to be effective to locate the failed array elements. The optimization process is to find a solution that minimize the difference between the measured array pattern and the supposed array pattern containing several failed elements. One of the most popular methods is the genetic algorithm (GA) for its natural advantage to use the gene to represent the discrete status of array elements [45]. A simple and fast technique is described in [46], which tabulates the nominal patterns radiated by the array with one failed element only. it can be applied together with other methods based on the evaluation of the array to improve the computational efficiency.

Some other stochastic approaches have also been developed in the last years. Neural networks and support vector machines (SVM) have been used but only within small arrays, where the number of possible failure combinations is also correspondingly small. In [47], a multilayer perceptron (MLP) was used to locate a maximum of three defective antennas from a 16-element array.

4.3.3 The Matrix Inversion Method

A different technique, applicable also to non-planar arrays, is based on matrix inversion computation [48]. The inverse problem consisting of processing measured fields to retrieve the source excitation. The relationship between measurement $Y = y_m$, $m = 1, \dots, M$ and element excitation $W = w_n$, $n = 1, \dots, N$ leads to a N -dimensional linear matrix equality. Considering a linear problem,

$$Y = \Phi W \quad (4.24)$$

where $\Phi \in C^{M \times N}$ is the matrix describing the interaction between the individual array elements and the measurement points. Usually, it requires that a number of measurements not smaller than the number of the array elements, i.e., $M > N$. Then the rank of Φ is full, it can be solved as,

$$W = \Phi^{-1} Y \quad (4.25)$$

However, due to the finite dimension of the measurement volume (the so-called truncation error [49]), this inverse problem may be ill-conditioned or over-determined. To solve this problem, several mathematical methods can be used, such as LU decomposition, Least Squares method, Singular Value Decomposition. A

parametric study and comparison of some matrix inversion methods can be found in [50].

4.3.4 The Compressed Sensing/Sparse Recovering Method

A more effective way to reduce the set of data in antenna array diagnosis is to utilize a-priori knowledge of the failure-free pattern of the antenna array under test. If the number of measurements is less than the array elements, i.e., $M < N$, then the rank of A is not full, thus there are many possible solutions for Eq. (4.24). However, if there are some more practical restrictions, it can also still be resolved. For antenna array diagnosis problem as an example, let ref and aut be the antenna array with all normal and with failed elements. Their array patterns are Y_{ref} and Y_{aut} , respectively, and the excitation are W_{ref} and W_{aut} . From these, we have,

$$\begin{aligned} Y^{ref} &= \Phi \cdot W^{ref} \\ Y^{aut} &= \Phi \cdot W^{aut} \end{aligned} \quad (4.26)$$

By differentiating these two equations, one can obtain that,

$$\Delta Y = \Phi \cdot \Delta W \quad (4.27)$$

where ΔY is the difference between the normal array pattern and the pattern with fault elements, and ΔW is their excitation difference, which contains only little non-zero elements.

The compressed sensing/sparse recovery (CS/SR) technique discusses the solution of equation (4.27) under the hypothesis that the unknown vector has a large number of null (or almost null) entries. It is a novel paradigm that allows representing sparse data in an efficient and accurate way based on nonlinear interpolation. Under the hypothesis that there are a large number of nulls (i.e., healthy antenna elements) in the solution, the required number of measurements is dramatically compressed, even fewer than the number of array elements [51]. To guarantee the singularity of the solution, it should satisfy the following condition [52],

$$Ax_1 - Ax_2 \neq 0, \forall x_1, x_2 \in S \quad (4.28)$$

Then Eq. (4.27) can be solved as,

$$\min \|x\|_{1,2} \text{ s.t. } \|Y - AX\|_{1,2} \leq \varepsilon \quad (4.29)$$

where ε is the measurement noise, $\|\cdot\|_{1,2}$ is the mixed norm $l1/l2$,

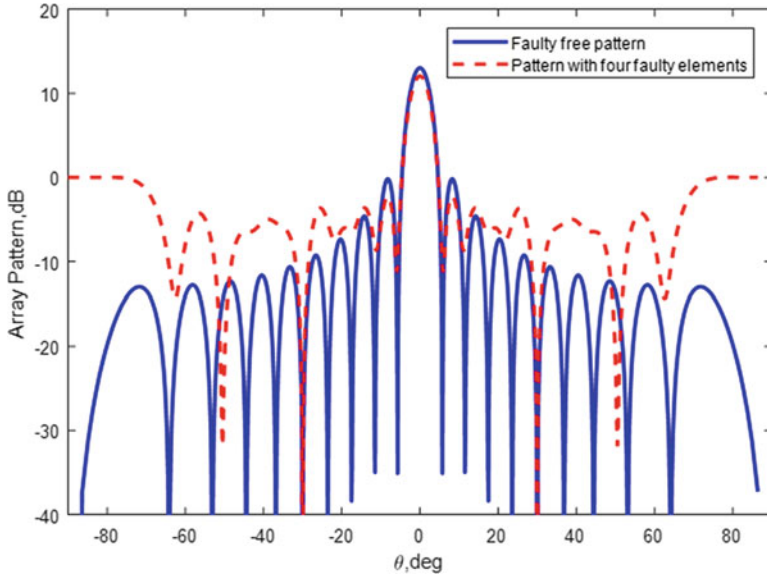


Fig. 4.8 The failed free array radiation pattern and the distorted pattern radiated with four failed elements

$$\|x\|_{1,2} = \sum_{g=1}^G \|x^g\|_2 = \sum_{g=1}^G \sqrt{|x_1^g|^2 + \dots + |x_{N_g}^g|^2} \quad (4.30)$$

To demonstrate the effectiveness of the CS/SR technique, a half-wavelength interspaced, 20 element linear array with uniform excitation is used as an example, i.e., $d = 0.5\lambda$, $w_n = 1$ for all elements. Figure 4.8 illustrates the array free pattern and array pattern containing four failed elements, $w_2 = w_4 = w_8 = w_{14} = 0$. As can be seen, the array pattern is distorted with much worse SLL. The number and the signal-to-noise ratio of far-field measurements are set as 12 and 40 dB, respectively. Figure 4.9 presents the exact and estimated excitation with some failed elements. It clearly shows that the failed elements have much lower weights than the normal ones. Figure 4.10 shows the estimated error of these two array patterns, such low error on the level of -30 dB demonstrates the high performance of the proposed CS method. Finally, the impacts of number of measurements on the performance are evaluated by the Monte Carlo simulations. With the increased number of measurements, the CS/SR method rapidly reaches a high success rate better than 80% (Fig. 4.11).

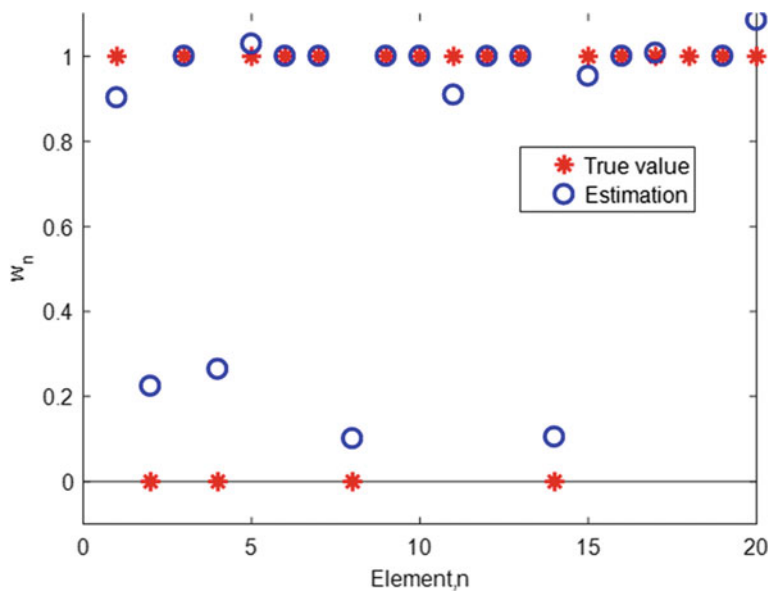


Fig. 4.9 The array element weights derived by the CS/SR technique

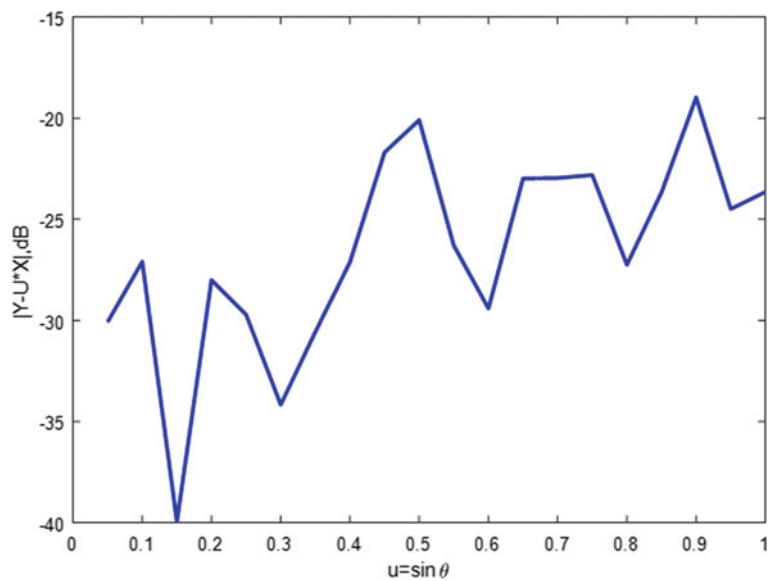


Fig. 4.10 The pattern error estimated by the the CS/SR technique

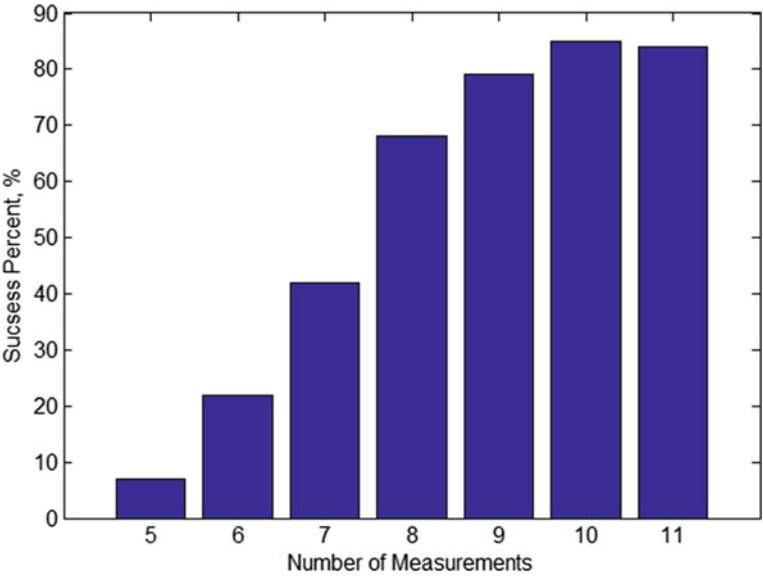


Fig. 4.11 The success rate of the CS/SR technique by Monte Carlo simulation

4.4 Summary

To meet the requirements and maintain acceptable performance, phased array antennas must be carefully calibrated to compensate the errors as low as possible. These calibration methods vary from different systems and applications, either suitable for factory test or for in-field test. The near-field scanning probe method that employs a robotic scanner is well-known and widely used as the standard for factory test. The peripheral fixed probe method requires external probes with calibration signal processing. In contrast, the mutual coupling method takes advantage of the inherent property of mutual coupling among elements of the array, and uses it to avoid the employment of external equipment. Some dedicated coupling network that using transmission lines connected to each antenna element for the periodic in-field calibration of integrated phased array antenna, especially for RF and mmWave-band integrated phased array. It's very convenient to apply advanced signal processing techniques to save hardware or time requirements. These techniques can be distinguished as serial or parallel by the number of elements to be calibrated each time, or distinguished as coherent or noncoherent by the measurement requirements between the calibration source and the element under calibration. Several failure detection methods for antenna array diagnosis are also shortly introduced, where the CS/SR technique is found to be very effective.

References

1. A. Stark, B. Rohrdantz, U. Johannsen, A.F. Jacob, In-situ probes for patch antenna array calibration. *Int. J. Microw. Wireless Tech.* **3**(3), 273–280 (2011)
2. K. Hassett, Phased array antenna calibration measurement techniques and methods, in *Proceedings of the European Conference on Antennas and Propagation* (2016), pp. 1–4
3. C. Fulton, Digital array radar calibration and performance monitoring techniques, including direct conversion and dual polarization architectures, Ph.D. dissertation (Purdue University, West Lafayette, 2011)
4. Y.V. Korotetskiy, A.M. Shitikov, V.V. Denisenko, Phased array antenna calibration with probe positioning errors. *IEEE Antennas Propag. Mag.* **58**(3), 65–80 (2016)
5. I. Seker, Calibration methods for phased array radars. *Proc. SPIE* **8714**, 1–15 (2013)
6. J. Mulcahey, M. Sarcione, Calibration and diagnostics of the THAAD solid state phased array in a planar near field facility, in 1996 *IEEE International Symposium Phased Array System Technology* (Boston, 1996), pp. 322–326
7. M. Sarcione, J. Mulcahey, D. Schmidt, K. Chang, M. Russell, P. Enzmann, P. Rawlinson, W. Guzak, R. Howard, M. Mitchell, The design, development and testing of the THAAD (Theater High Altitude Area Defense) solid state phased array (formerly Ground Based Radar), in *IEEE International Symposium Phased Array System Technology* (Boston, 1996), pp. 260–265
8. H. Pawlak, A.F. Jacob, An external calibration scheme for DBF antenna arrays. *IEEE Trans. Antennas Propag.* **58**(1), 59–67 (2010)
9. T. Takahashi, N. Nakamoto, M. Ohtsuka, T. Aoki, Y. Konishi, I. Chiba, M. Yajima, On-board calibration methods for mechanical distortions of satellite phased array antennas. *IEEE Trans. Antennas Propag.* **60**(3), 1362–1372 (2012)
10. J.A. Haimel, B. Hudson, G.P. Fonder, D.K. Lee, Overview of the large digital arrays of the space fence radar, in *IEEE International Symposium Phased Array System Technology* (Waltham, 2016), pp. 1–8
11. M. Scott, Sampson MFR active phased array antenna, in *IEEE International Symposium Phased Array System Technology* (Boston, 2003), pp. 119–123
12. H.M. Aumann, A.J. Fenn, F.G. Willwerth, Phased array antenna calibration and pattern prediction using mutual coupling measurements. *IEEE Trans. Antennas Propag.* **37**(7), 844–850 (1989)
13. C. Shipley, D. Woods, Mutual coupling-based calibration of phased array antennas, in *Proceedings of the IEEE International Symposium Phased Array System Technology*, Dana Point (USA, 2000), pp. 529–532
14. D. Bekers, R. van Dijk, F. van Vliet, Mutual-coupling based phased-array calibration: a robust and versatile approach, in *Proceedings of the IEEE International Symposium Phased Array System Technology* (Waltham, 2013), pp. 630–637
15. R.M. Lebron, P.S. Tsai, J.M. Emmett, C. Fulton, J.L. Salazar-Cerreno, Validation and testing of initial and in-situ mutual coupling-based calibration of a dual-polarized active phased array antenna. *IEEE Access* **8**, 78315–78329 (2020)
16. D. Kissinger, B. Laemmle, L. Maurer, R. Weigel, Integrated test for silicon front ends. *IEEE Microw. Mag.* **11**(3), 87–94 (2010)
17. O. Inac, D. Shin, G.M. Rebeiz, A phased array RFIC with built-in self-test capabilities. *IEEE Trans. Microw. Theory Techn.* **60**(1), 139–148 (2012)
18. O. Inac, F. Golcuk, T. Kanar, G. Rebeiz, A 90–100 GHz phased-array transmit/receive silicon RFIC module with built-in self-test. *IEEE Trans. Microw. Theory Techn.* **61**(10), 3774–3782 (2013)
19. K.M. Lee, R.S. Chu, S.C. Liu, A built-in performance monitoring/fault isolation and correction (PM/FIC) system for active phased-array antennas. *IEEE Trans. Antennas Propag.* **41**(11), 1530–1540 (1993)
20. D.J. Rabideau, R.J. Galejs, F.G. Willwerth, D.S. McQueen, An S-band digital array radar testbed, in *Proceedings of the IEEE International Symposium Phased Array System Technology* (Boston, 2003), pp. 113–118

21. S.C. Chae, H.W. Jo, J.I. Oh, G. Kim, J.W. Yu, Coupler integrated microstrip patch linear phased array for self-calibration. *IEEE Antennas Wirel. Propag. Lett.* **19**(9), 1615–1619 (2019)
22. S. Mano, T. Katagi, A method for measuring amplitude and phase of each radiating element of a phased array antenna. *Electron. Commun. Jpn.* **65**(5), 58–64 (1982)
23. R. Sorace, Phased array calibration. *IEEE Trans. Antennas Propag.* **49**(4), 517–525 (2001)
24. R. Long, J. Ouyang, F. Yang, W. Han, L. Zhou, Fast amplitude-only measurement method for phased array calibration. *IEEE Trans. Antennas Propag.* **65**(4), 1815–1822 (2017)
25. G. He, X. Gao, H. Zhou, Fast phased array calibration by power-only measurements twice for each antenna element. *Int. J. Antennas Propag.* **2019**, Article ID 6432149, 10 (2019)
26. S.D. Silverstein, Application of orthogonal codes to the calibration of active phased array antennas for communication satellites. *IEEE Trans. Signal Process.* **45**(1), 206–218 (1997)
27. E. Lier, M. Zemlyansky, D. Purdy, D. Farina, Phased array calibration and characterization based on orthogonal coding: theory and experimental validation, in *Proceedings of the IEEE International Symposium Phased Array System Technology* (Waltham, 2010), pp. 271–278
28. M. Bachmann, M. Schwerdt, B. Brautigam, TerraSAR-X antenna calibration and monitoring based on a precise antenna model. *IEEE Trans. Geosci. Remote Sens.* **48**(2), 690–701 (2010)
29. M. Schwerdt, K. Schmidt, N.T. Ramon, P. Klenk, N. Yague-Martinez, P. Prats-Iraola, M. Zink, D. Geudtner, Independent system calibration of sentinel-1B. *Remote Sens.* **9**(511), 1–34 (2017)
30. R. Long, J. Ouyang, F. Yang, W. Han, L. Zhou, Multi-element phased array calibration method by solving linear equations. *IEEE Trans. Antennas Propag.* **65**(4), 2931–2939 (2017)
31. A.B. Smolders, G.A. Hampson, Deterministic RF nulling in phased arrays for the next generation of radio telescopes. *IEEE Antennas Propag. Mag.* **44**(4), 13–22 (2002)
32. T. Takahashi, Y. Konishi, S. Makino, H. Ohmine, H. Nakaguro, Fast measurement technique for phased array calibration. *IEEE Trans. Antennas Propag.* **56**(7), 1888–1899 (2008)
33. C. He, X. Liang, J. Geng, R. Jin, Parallel calibration method for phased array with harmonic characteristic analysis. *IEEE Trans. Antennas Propag.* **62**(10), 5029–5036 (2014)
34. K. Greene, V. Chauhan, B. Floyd, Built-in test of phased arrays using code-modulated interferometry. *IEEE Trans. Microw. Theory Techn.* **66**(5), 2463–2479 (2018)
35. A. Massa, G. Oliveri, M. Salucci, N. Anselmi, P. Rocca, Learning-by-examples techniques as applied to electromagnetics. *J. Electromag. Waves and Appl.* **32**(4), 516–541 (2017)
36. L. Li, L. Wang, F.L. Teixeira, C. Liu, A. Nehorai, T. Cui, DeepNIS: deep neural network for nonlinear electromagnetic inverse scattering. *IEEE Trans. Antennas Propag.* **67**(3), 1819–1825 (2019)
37. A. Massa, D. Marcantonio, X. Chen, M. Li, M. Salucci, DNNs as applied to electromagnetics, antennas, and propagation—a review. *IEEE Antennas Wirel. Propag. Lett.* **18**(11), 2225–2229 (2019)
38. S.D. Campbell, R.P. Jenkins, P.J. O'Connor, D.H. Werner, The explosion of artificial intelligence in antennas and propagation. *IEEE Antennas Propag. Mag.* **63**(3), 16–27 (2021)
39. F. Zardi, P. Nayeri, P. Rocca, R.L. Haupt, Artificial intelligence for adaptive and reconfigurable antenna arrays. *IEEE Antennas Propag. Mag.* **63**(3), 28–38 (2021)
40. L. Yang, R. Dang, M. Li, K. Zhao, C. Song, Z. Xu, A fast calibration method for phased arrays by using the graph coloring theory. *Sensors* **18**(12), 1–19 (2018)
41. G. Baburs, D. Caratelli, A. Mirmanov, Phased array calibration by binary compressed sensing. *Prog. Electromagn. Res. M* **73**, 61–70 (2018)
42. C. Shan, X. Chen, H. Yin, W. Wang, G. Wei, Y. Zhang, Diagnosis of calibration state for massive antenna array via deep learning. *IEEE Wirel. Commun. Lett.* **8**(5), 1431–1434 (2019)
43. J.J. Lee, E.M. Ferren, D.P. Woollen, K.M. Lee, Near-field probe used as a diagnostic tool to locate defective elements in an array antenna. *IEEE Trans. Antennas Propag.* **36**(6), 884–889 (1988)
44. O.M. Bucci, A. Capozzoli, G. Delia, Diagnosis of array faults from far-field amplitude-only data. *IEEE Trans. Antennas Propag.* **48**(5), 647–652 (2000)
45. J.A. Rodríguez, F. Ares, H. Palacios, J. Vassallo, Finding defective elements in planar arrays using genetic algorithms. *Progr. Electromagn. Res.* **29**, 25–37 (2000)

46. J.A. Rodriguez, F. Ares, M. Fernandez-Delgado, R. Iglesias, S. Barro, Rapid method for finding faulty elements in antenna arrays using far field pattern samples. *IEEE Trans. Antennas Propag.* **57**(6), 1679–1683 (2009)
47. A. Patnaik, B. Choudhury, P. Pradhan, R.K. Mishra, C. Christodoulou, An ANN application for fault finding in antenna arrays. *IEEE Trans. Antennas Propag.* **55**(3), 775–777 (2007)
48. L. Gattoufi, D. Picard, Y.R. Samii, J.C. Bolomey, Regularized matrix method for near-field diagnostic techniques of phased array antennas, in *IEEE Antennas and Propagation Society International Symposium* (1997)
49. M. Bucci, M.D. Migliore, G. Panariello, G. Sgambato, Accurate diagnosis of conformal arrays from near-field data using the matrix method. *IEEE Trans. Antennas Propag.* **53**(3), 1114–1120 (2005)
50. J.A. Rodriguez, M. Fernandez-Delgado, J. Bregains, R. Iglesias, S. Barro, F. Ares, A comparison among several techniques for finding defective elements in antenna arrays, in *The Second European Conference on Antennas and Propagation, EuCAP 2007* (2007), pp. 1–8
51. D. Migliore, A compressed sensing approach for array diagnosis from a small set of near-field measurements. *IEEE Trans. Antennas Propag.* **59**(6), 2127–2133 (2011)
52. D. Migliore, A simple introduction to compressed sensing/sparse recovery with applications in antenna measurements. *IEEE Antennas Propag. Mag.* **56**(2), 14–26 (2014)

Chapter 5

Some Key Design Issues for Satellite Ground Station Application



5.1 Array Geometry Design for Hemispherical Coverage

As the satellite ground station, a phased array antenna should have the hemispherical coverage from the horizon to the zenith. In each pass, once the satellite comes above the horizon, the ground station sets a link with the satellite, and follows its trajectory until it moves below the horizon. The hemispherical coverage can be defined as the angular interval $\{(\theta, \phi) \mid \theta \in [\theta_1, \pi/2], \phi \in [0, 2\pi]\}$, where θ_1 is the minimum working elevation angle of the ground station, and $\theta_1 = 0$ for the horizontal coverage. Only a few array geometries have the potential to realize the hemispherical coverage, including pyramidal frustum, sphere, geodesic dome, and geodesic dome with cylinder base.

5.1.1 Pyramidal Frustum

In the past, multifaced planar arrays with pyramidal frustum geometry were the mainstream style for the hemispherical coverage [1–3]. Figure 5.1 illustrates a photo of AN/FPS-115 radar. It has two side faces with 3,600 active antenna elements, each face scans at azimuth angle of 120° , and elevation between 3° and 10° above horizontal. As shown in Fig. 5.2, it comprises several planar arrays mounted on the side faces and the top face of a pyramidal frustum. The hemispherical coverage is divided into several spaces for each face, thus reducing the maximum scan angle and scan losses relative to a single planar array. Sectoral coverage refers to using the n -th side face for $\{(\theta, \phi) \mid \theta \in [\theta_1, \theta_2], \phi \in [2\pi(n-1)/N, 2\pi n/N]\}$, while the top face is to cover the region of the space $\{(\theta, \phi) \mid \theta \in [\theta_2, \pi/2], \phi \in [0, 2\pi]\}$, where N is the number of side faces, and θ_2 is the elevation angle divided by the top face and the side faces.

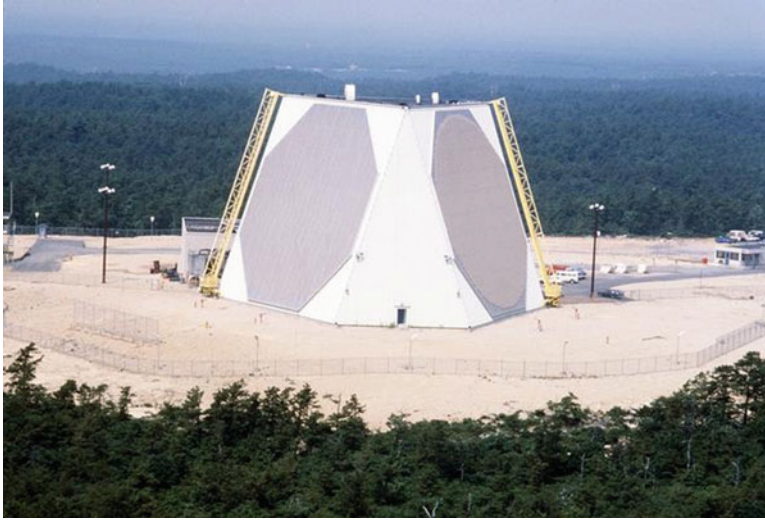


Fig. 5.1 The AN/FPS-115 radar. It has two side faces with 3,600 active antenna elements, each face scans at azimuth angle of 120° , and elevation between 3° and 10° above horizontal. Figure reproduced with permission from the National Electronics Museum

The planar array would suffer gain degradation and polarization loss when the beam shifts away from the boresight, usually the scan angle cannot scan beyond 60° . For the top face, the maximum scan angle is equal to $90^\circ - \theta_2$. While for the side faces, it is clear that the maximum scan angle occurs when the beam is scanned from the boresight direction of the face to either of the vertexes of the side faces. Figure 5.3 illustrates the relationship between the maximum scan angle θ_{max} and the side face elevation angle above the horizontal plane γ . It clearly shows that there is an optimal elevation angle for the side faces to minimize the maximum scan angle. Following the analysis in [4, 5], the optimal side face elevation angle γ_{opt} and the corresponding maximum scan angle θ_{max} are given by,

$$\gamma_{opt} = \tan^{-1} \frac{(\cos \theta_1 - \cos \theta_2)}{\cos \frac{\pi}{N} (\sin \theta_2 - \sin \theta_1)} \quad (5.1)$$

$$\theta_{max} = \cos^{-1} \left(\sin \gamma_{opt} \cos \frac{\pi}{N} \sin \theta_1 + \cos \gamma_{opt} \cos \theta_1 \right) \quad (5.2)$$

The maximum scan angle can be gradually reduced by increasing the number of side faces. However, only marginal benefit can be gained by doing this when the number of the pyramid faces is large. Figure 5.4 shows the maximum scan angle θ_{max} and the number of the pyramid faces N at different scan ranges. It shows that there is little reduction when increasing the number of side faces beyond six.

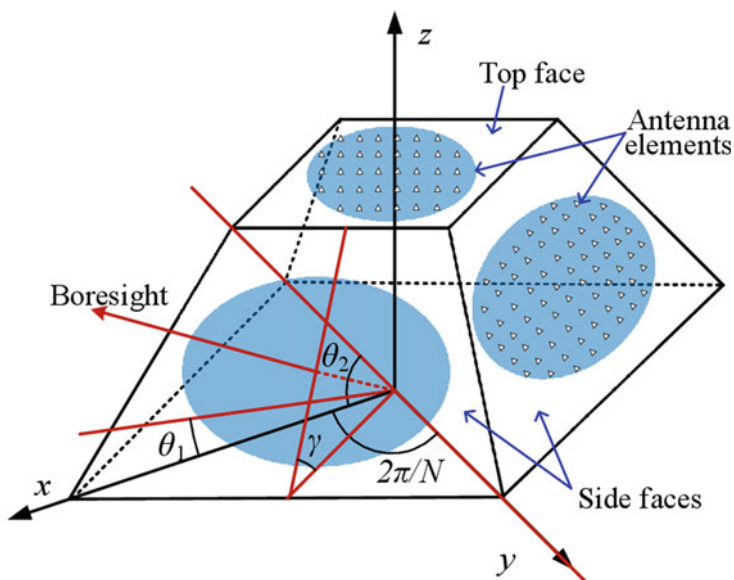


Fig. 5.2 The phased array antenna with pyramidal frustum geometry, the blue region means the planar antenna arrays at the side faces and the top face

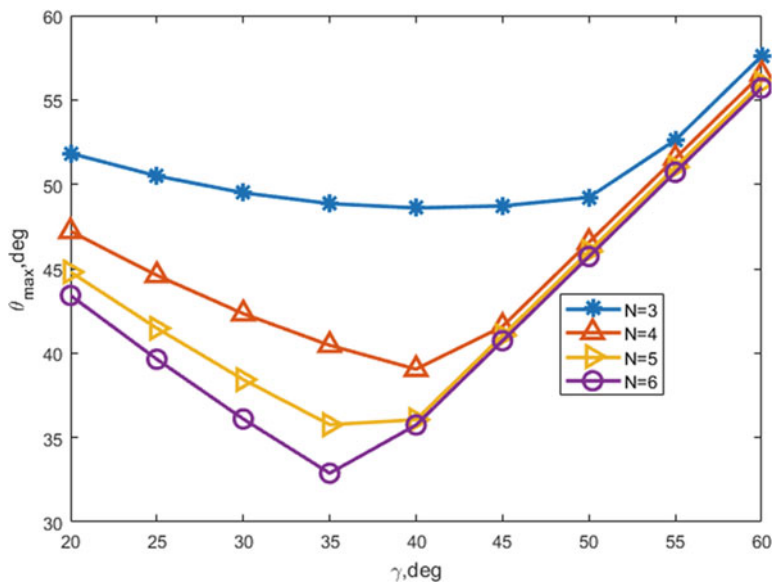


Fig. 5.3 The relationship between the maximum scan angle θ_{max} and the side face elevation angle γ for $\theta_1 = 5^\circ$, $\theta_2 = 60^\circ$

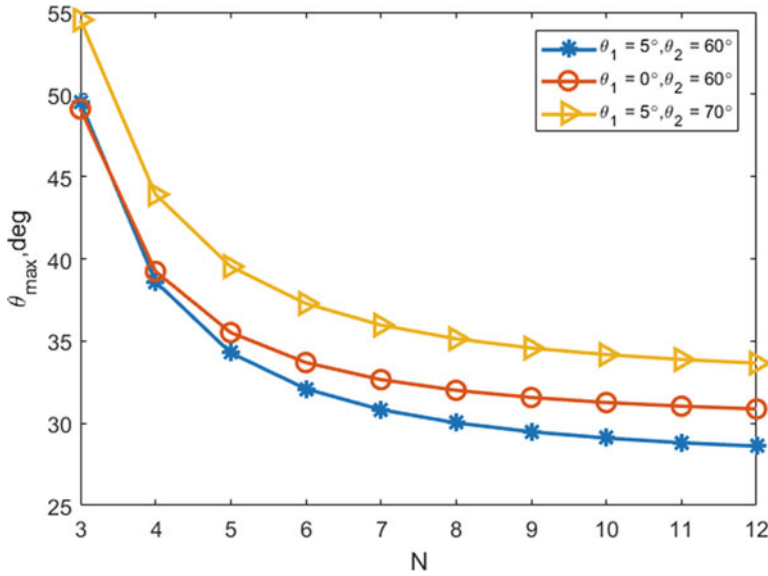


Fig. 5.4 The relationship between the maximum scan angle θ_{max} and the number of pyramid faces N at different scan ranges

5.1.2 Sphere

Ideally, the best geometry should be sphere for phased array antenna to realize the hemispherical coverage. As shown in Fig. 5.5, the beam direction is the axis of the cone, the antenna elements distributed on the surface of a sphere are symmetrically activated in the conical projection on the sphere surface. When the array beam moves, the active region also moves over the spherical surface, hence the antenna elements are dynamically adjusted to form the desired beam. Compared to the planar array having reduced gain for large scan angles, a spherical phased array can provide a practically identical beam that the pattern and gain remain almost unchanged over the entire hemisphere.

Moreover, the antenna elements required for the spherical array are also the least, which implies hardware saving and cost reduction. For the multifaced planar array, the minimum effective area A'_{ps} occurs at the maximum scan angle, for the side face,

$$A'_{ps} = A_{ps} \cos \theta_{max} \quad (5.3)$$

where A_{ps} is the physical area of the antenna array on the side face.

And for the top face,

$$A'_{pt} = A_{pt} \cos(90^\circ - \theta_2) = A_{pt} \sin \theta_2 \quad (5.4)$$

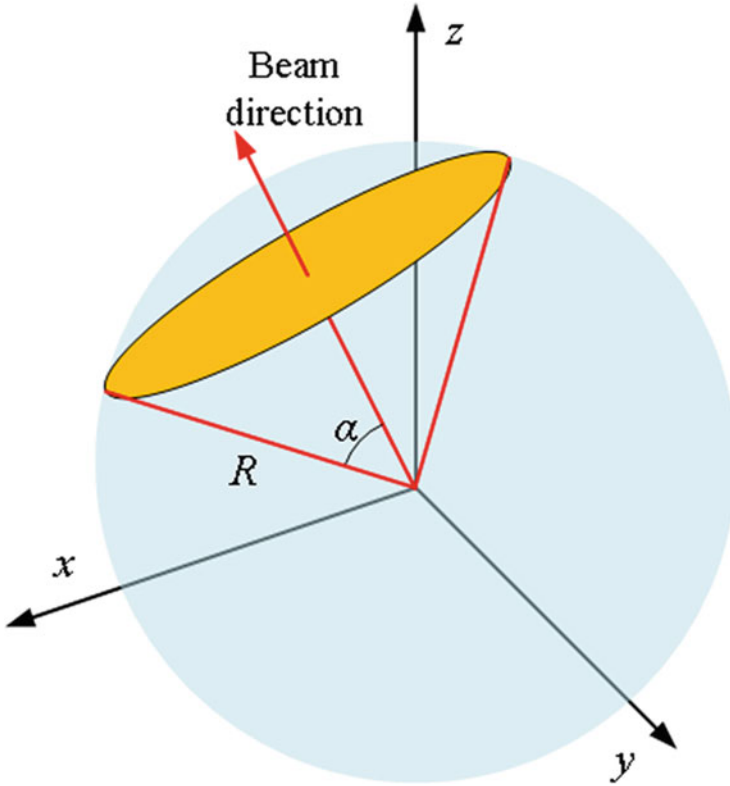


Fig. 5.5 The spherical phased array antenna with sphere geometry

where A'_{pt} and A_{pt} are the effective area and the physical area of the antenna array on the top face, respectively.

For the spherical array, the effective area A'_s is,

$$A'_s = \pi(R \cos \alpha)^2 = A_s \cos^2 \alpha \quad (5.5)$$

where R is the radius of the sphere, $A_s = \pi R^2$, and α is the cover cone angle. By applying equal gain criteria,

$$A'_{ps} = A'_{pt} = A'_s \quad (5.6)$$

Therefore, the total area (element number) of a spherical array normalized to its multifaced planar array is,

$$\eta = \frac{4A_s}{NA_{ps} + A_{pt}} = \frac{4/\cos^2 \alpha}{N/\cos \theta_{\max} + 1/\sin \theta_2} \quad (5.7)$$

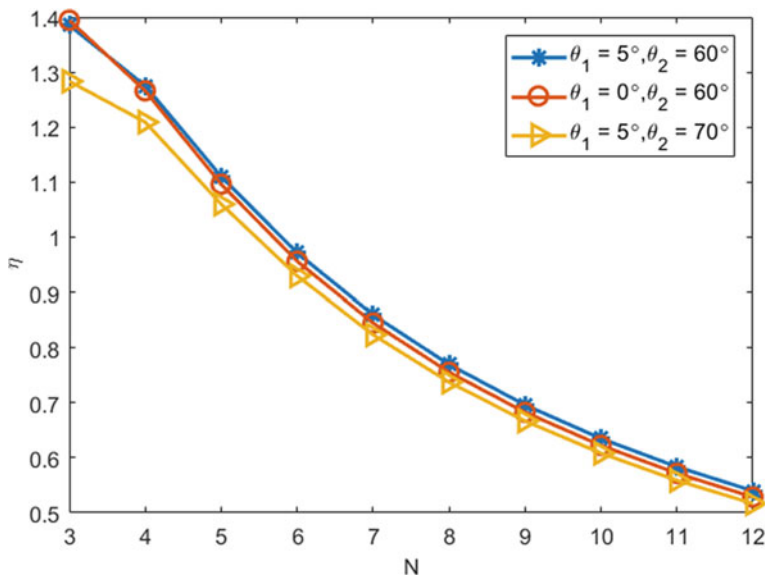


Fig. 5.6 Comparison of the total area of the spherical array and the multifaced planar array, assuming $\alpha = 45^\circ$

Figure 5.6 illustrates the ratio of the total area of the spherical array and the multifaced planar array. It shows that for a small number of side faces ($N \leq 6$), the antenna elements are comparable that are used in these two geometries. However, it dramatically becomes that the spherical array requires fewer antenna elements as the number of side faces increases.

Besides the no gain loss with beam steering, the spherical array has other advantages such as lower polarization loss and wider frequency bandwidth over planar array [6]. However, the spherical array has rarely been used in practice, only some cases can be found in the literature [7–9]. This is primarily because that the fabrication and assembly of antenna elements conformal to the spherical surface are much more difficult than the planar array.

5.1.3 Geodesic Dome

In order to overcome the fabrication and assembly difficulties of a spherical array, the geometry of a geodesic sphere or geodesic dome is proposed, which uses planar arrays to approximate a spherical shape. The geodesic dome geometry can be derived from any one of the five regular polyhedra or one of the 15 semi-regular polyhedra [9]. It preserves all the advantages of a spherical array while its fabrication is based

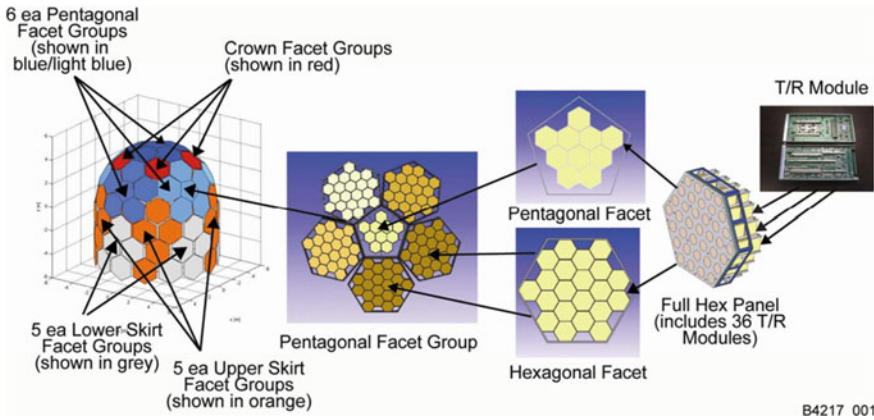


Fig. 5.7 The architecture of the GDPAA-ATD project. Figure reproduced with permission from [10], © 2010 IEEE

on the well developed, easily manufactured planar array. Moreover, a geodesic dome is mechanically very stable, no interior support structure or load-bearing walls is needed.

Figures 5.7 and 5.8 show the Geodesic Dome Phased Array Antenna (GDPAA) Advanced Technology Demonstration (ATD) antenna array built by Ball Aerospace & Technologies Corp. This array was developed to demonstrate multi-beam phased array capability for next generation Air Force Satellite Control Network station [10–12]. It consists of hexagonal and pentagonal panels in the geodesic dome, which contain several hexagonal planar arrays in them. Each hexagonal planar array has 37 elements in a triangular grid. Only 36 of the elements are active, the center element is devoted to calibration. Several panels and beam management software were built, integrated, and tested. Four beams were produced to support four satellites simultaneously, fully meet the goals of the demonstration program.

5.1.4 Geodesic Dome with Cylinder Base

A similar geometry that has geodesic dome with cylindrical base may be easier to construct. This geometry provides about 0.5 dB more gain at the horizon than full geodesic dome geometry [13]. This is favorable for satellite ground station with two main reason. First, it usually has the furthest distance between the ground station and the satellite for low elevation angle, cause the largest path loss. Second, the corresponding atmosphere attenuation becomes more severe due to multipath effects and tropospheric scintillation. The GEOdesic Dome Array (GEODA) project designed by Universidad Politecnica de Madrid falls into this type of geometry [14]. As shown in Fig. 5.9, the hexagonal panels of the geodesic dome are made of six triangular

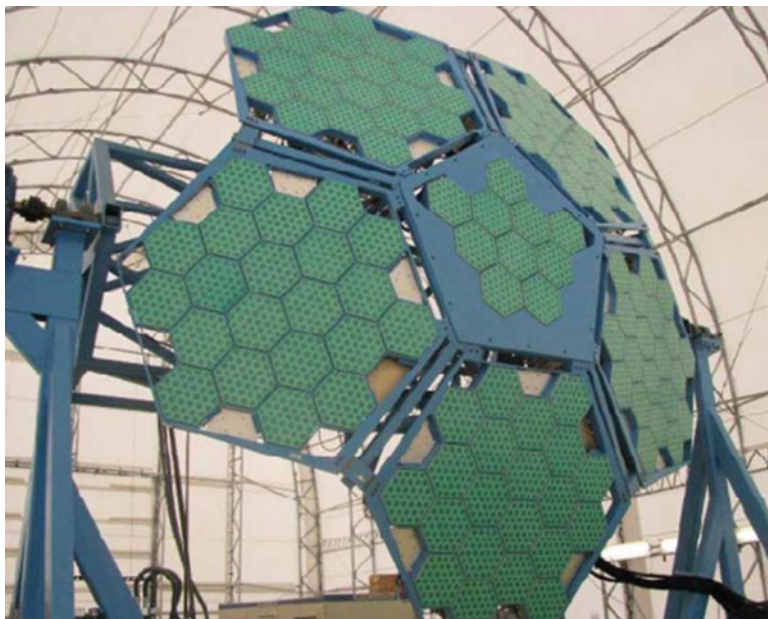


Fig. 5.8 The front view of the GDPAA-ATD project, with one pentagonal panel and five hexagonal panels. Figure reproduced with permission from [10], © 2010 IEEE

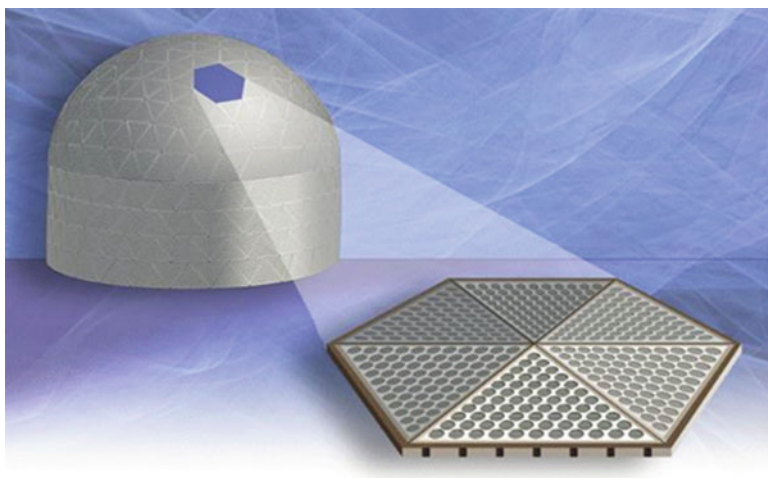


Fig. 5.9 The array geometry of the GEODA-SARAS project. Figure reproduced with permission from [14], © 2014 IEEE

arrays. Each triangular array consists of 45 radiating elements. The cylindrical base is also made by triangular array.

5.2 Multiple Beam Resource Management

5.2.1 The Traditional Work Flowchart of Reflector Antenna

Figure 5.10 illustrates a traditional work flowchart of the ground station to support a satellite pass. First, the ground station needs to make a plan ahead of time based on the satellite application tasks. With this plan, it configures channel parameters such as time duration, carrier frequency, modulation scheme, data rate, and uses satellite orbit prediction data to rotate the beam to the waiting point several minutes earlier. When the satellite enters into the antenna beam coverage of the ground station, the ground station emits an uplink carrier signal toward the satellite. The satellite transponder receives and locks into the carrier signal’s phase, and coherently forwards it back

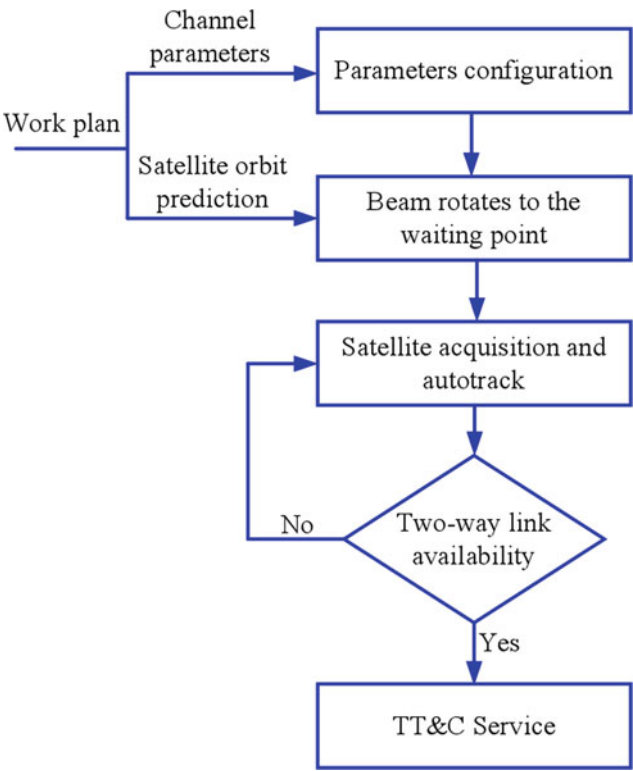


Fig. 5.10 A traditional work flowchart of the ground station to support a satellite pass

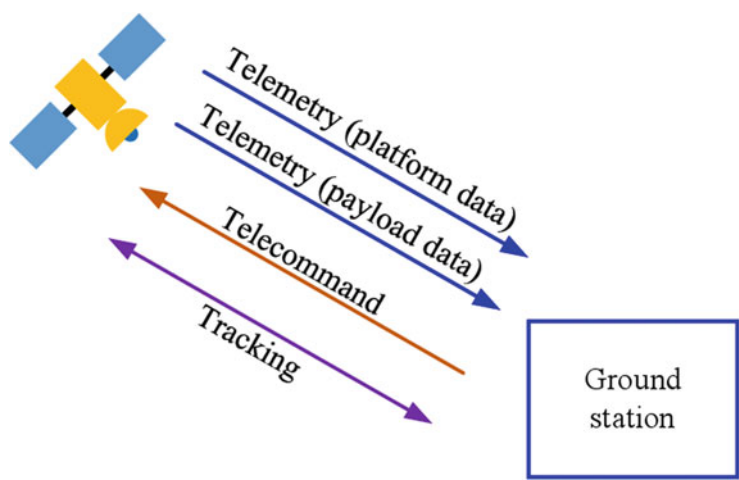


Fig. 5.11 TT&C services between satellite and ground station

to the ground station. The ground station also captures the satellite’s turnaround signal, indicating that the link for full duplex communication has been established successfully. Then the antenna rotates its beam by tracking the received satellite’s signal, and the TT&C service is carried out according to the plan as shown in Fig. 5.11.

In general, the working plan for each satellite pass needs to be formulated at least several hours in advance. Otherwise, even if the satellite is within the visible region of the ground station, the ground station cannot start TT&C service quickly. For some time-sensitive tasks, the whole procedure forbids the ground station to respond to the task requirement in time. Therefore, a more efficient way must be developed by transforming from the current ‘plan-driven’ manner to a new ‘demand-driven’ approach, which provides TT&C services on demand and achieves space-ground autonomous information exchange, thus significantly improving the efficiency of ground station resource utilization.

5.2.2 The Novel Work Flowchart of Phased Array Antenna

To support this new demand-driven approach mentioned above, the ground station needs to have two different kinds of antenna beams as shown in Fig. 5.12. The first one is a hemispheric coverage, receive only, low gain beam by a few sparsely arranged antenna elements, which is used for monitoring incoming signals from all directions with high reliability, 24 h a day, and 7 days a week. The other is traditional narrow and high gain beams formed by a large number of antenna elements to enable high data rate transmission. Table 5.1 compares some aspects of these two kinds of

Table 5.1 Comparison of ground station beams

Parameter	Hemispheric beam	Narrow beam
Beam width	Hemispheric coverage	$\leq 1^\circ$
Beam scan	No need	Yes
Number of antenna elements per beam	Tens	Thousands
Gain	Low	High
Capability	Receive only	Receive and/or transmit
Data rate	Low	High
Usage	Status monitoring	TT&C service

beams to show their differences. As illustrated in Fig. 5.13, with these beams in the ground station, the work flowchart of a satellite pass is totally renewed, which can be generally divided into five modes.

5.2.2.1 Satellite Arrival Registration

Nowadays, with the onboard high-sensitivity GNSS receiver, the satellite can simply achieve real-time precise orbit determination at the centimeter level [15, 16]. The position, velocity, and time of the satellite is obtained by GNSS broadcast ephemerides and single frequency psedorange observations. As an advantage without atmospheric effects, the spaceborne GNSS based technique has gradually become the primary method for precise orbit determination. Thus, with the pre-known ground station coordinates, the satellite calculates its relative position with the ground station in real-time. For already existing ground stations, their coordinates are loaded into the onboard navigation program prior to the launch of the satellite; while for newly built ground stations, their coordinates can also be updated through telecommand service. Once the satellite enters the visible region of the ground station, it sends an arrival registration message to the ground station. This message includes some basic information such as satellite ID, orbital elements, and important satellite platform status. Because multiple satellites send arrival registration messages to the ground station simultaneously, this should be done through code division multiplexing. After receiving and demodulating the arrival registration message and verifying its legitimacy, the ground station configures a narrow beam to reply by an acknowledgment message, informing the satellite that it has entered its service coverage. If the satellite does not receive a ground acknowledgment response after waiting for some time, it should attempt to send the arrival registration message again until the ground acknowledgment message is received.

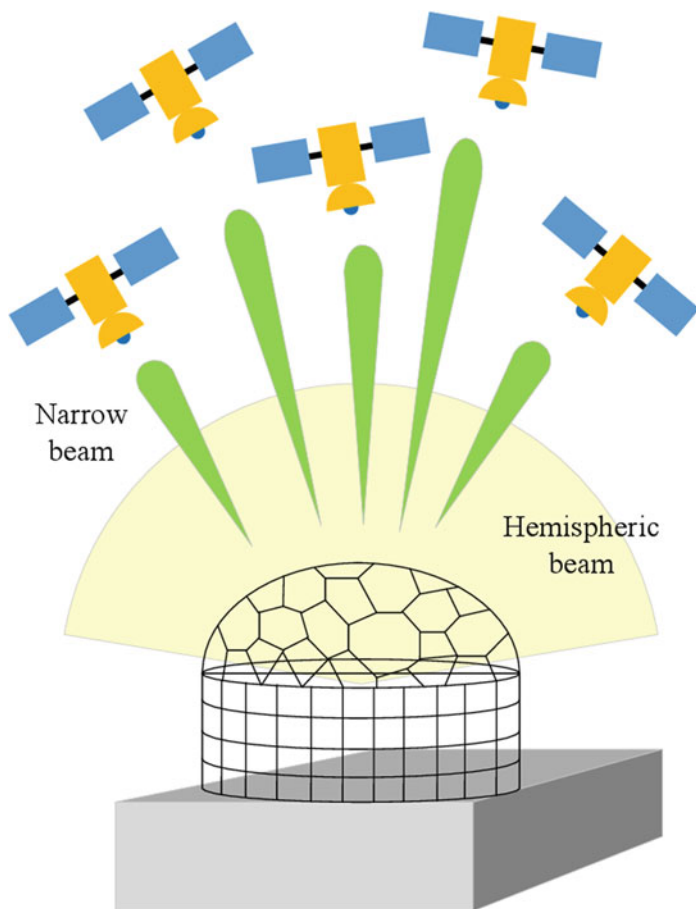


Fig. 5.12 Illustration of the ground station with a hemispheric beam and multiple narrow beams

5.2.2.2 Satellite Periodical Status Reporting

For an LEO satellite, its single pass in the visible range of the ground station is usually about several minutes. After successful arrival registration, the satellite decides which operating mode to adopt based on its own needs. If no telemetry information needs to transmit to the ground, a simple status reporting message is enough to inform the ground station of its current location and important platform status parameters. This status reporting message is sent within a certain period (typically one minute or other values as desired), and also received by the hemispheric beam of the ground station through code division multiplexing. After receiving and demodulating the status reporting message, the ground station stores the information in the local archive. To save the narrow beam resources, the ground station may not reply to the status reporting message. If the periodic status reporting message of a satellite has not been

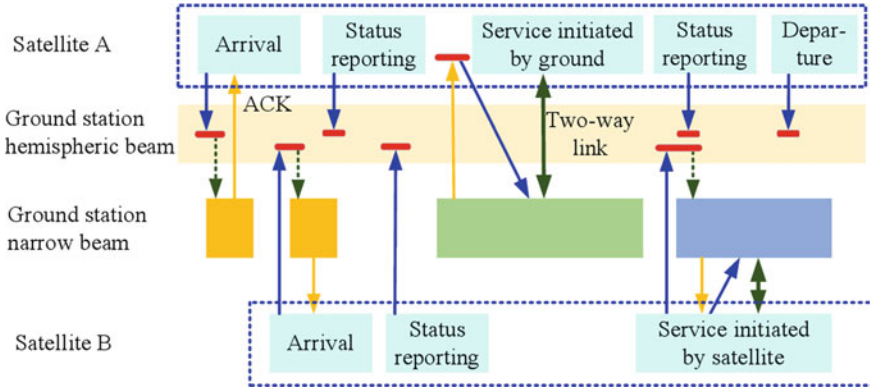


Fig. 5.13 The novel demand driven work flowchart for the ground station with multibeam phased array antennas

received for several periods, it can be considered that the satellite has left the service coverage of the ground station.

5.2.2.3 Ground Service Initiation

When the ground station wants to send telecommands to the satellite, or further inquire about detailed status parameters of the satellite, it needs to initiate a TT&C service from the ground. Based on the periodic status reporting messages of the satellite in the early stage, the ground station infers the current approximate location of the satellite, and configures a narrow beam to the inferred direction, sends a service initiation message to notify the satellite to synchronously start the corresponding TT&C service mode. This message also contains information about channel parameters, thus the satellite could set the parameters as needed. The following two-way carrier locking process is the same as in the traditional work process, the service link is established for full duplex communication. After the information transmission is completed, this link is automatically disconnected and the current narrow beam resource is released.

5.2.2.4 Satellite Service Initiation

The mode when the satellite wants to initiate a TT&C service is similar. The only difference is that the satellite first needs to apply for TT&C service from the ground station through the hemispheric beam. The ground station determines whether to respond based on current narrow beam resource usage. If all narrow beams have been occupied, the response to the current application must be temporarily postponed, otherwise an occupied narrow beam should be released in a mandatory manner. If

there are still free narrow beam resources available, a response can be made to satisfy the current TT&C service application. After this, the process of narrow beam configuration, two-way carrier locking, and information exchange will proceed the same.

5.2.2.5 Satellite Departure Forecasting

When the satellite is to leave the visible region of the ground station, it sends a departure forecast message to the ground station. The ground station also does not need to reply to this message, and the service for this current pass of the satellite ends. The service will be restarted when its arrival registration message is received again in the next pass.

5.3 Space-Ground Link Analysis

5.3.1 The Link Budget Equation

For evaluating the performance of a digital communication system, it's of great interest to focus on the signal-to-noise-ratio SNR or the bit energy per noise power spectral density E_b/N_0 , which directly shows the ability to transmit data in the presence of noise with a certain error probability. For constant envelope modulation schemes, the pre-detection SNR of the signal can be obtained as follows [17],

$$\frac{P_r}{N} = \frac{P_t G_t G_r / L_{path} L_{atm}}{k T_{sys} W} = \frac{EIRP_t \cdot G_r / T_{sys}}{k W L_{path} L_{atm}} \quad (5.8)$$

where P_r is the received signal power, N is the received noise power; P_t is the transmitter power, G_t is the transmitter gain, their product is denoted as $EIRP_t$ and called as the effective isotropic radiated power; G_r is the receiver gain, T_{sys} is the receiver system noise temperature, the grouping of G_r / T_{sys} is usually known as the receiver figure-of-merit; k is the Boltzmann constant and equal to 1.38×10^{-23} W/K-Hz or -228.6 dB/K-Hz, W is the signal bandwidth, L_{path} is the free path loss and equal to $(4\pi d/\lambda)^2$, d is the distance and λ is the signal wavelength, and L_{atm} is the atmospheric induced loss such as ionospheric absorption, ionospheric scintillation, and rain attenuation.

Assuming that all the receiving power P_r is the effective modulation signal, the expression of E_b/N_0 can be derived from Eq. (5.8),

$$\frac{E_b}{N_0} = \frac{P_r/R_b}{N/W} = \frac{P_r}{N} \cdot \frac{W}{R_b} = \frac{EIRP_t \cdot G_r / T_{sys}}{k L_{path} L_{atm} R_b} \quad (5.9)$$

where E_b is the bit energy, N_0 is the noise power spectral density, and R_b is the transmission data bit rate. The link margin M can be obtained as the difference between the received E_b/N_0 and the required E_b/N_0 to yield a specified error probability minus the demodulator implementation loss L_{imp} . It's convenient to express them in decibels as,

$$\begin{aligned} M(dB) &= (E_b/N_0)_r(dB) - (E_b/N_0)_{reqd}(dB) - L_{imp}(dB) \\ &= EIRP_t(dBW) + G_r/T_{sys}(dB/K) - k(dB/K-Hz) - L_{path}(dB) \quad (5.10) \\ &\quad - L_{atm}(dB) - R_b(dB-bps) - (E_b/N_0)_{reqd}(dB) - L_{imp}(dB) \end{aligned}$$

The above link budget equation is a very useful tool for allocating resources of a communication system. It's valid for both the ground-to-satellite link (uplink) and the satellite-to-ground link (downlink). Tradeoffs can be made between any items in the equation, the only limit is to maintain the link margin positive. For example, to increase the transmission data rate, one can enhance the transmitter power, or improve the receiver G/T value, or select efficient coding methods to reduce the required E_b/N_0 .

5.3.2 Case Study for Satellite-to-Ground Downlink

5.3.2.1 Downlink with the Hemispheric Beam

As stated in Sect. 5.2, a hemispheric beam is used for satellite status monitoring. A relatively low data rate (assuming 1 kbps) is required for these modes. For transmitting such low rate telemetry data, the satellite can use low power amplifier (1 W or 0 dBW) to save onboard battery energy. From the literature survey of cube-sat antenna design, 18 patch antenna designs are reviewed in [18]. These antennas operating in S-band can provide a total gain ranging from 4.3 to 8.5 dBi. Here a relatively conservative value of 5 dBi is taken. These two values result in the $EIRP$ of the satellite as 5 dBW. On the ground station, a small number of antenna elements are used to ensure the hemispheric coverage, hence the G/T value is set as low as -20 dB/K. The distance d between the satellite and the ground station varies and is determined by the orbital parameters of the satellite. A maximum distance 3000 km leads to a path loss of 168.8 dB when the signal carrier frequency $f_c = 2.2$ GHz. The atmospheric loss of 2 dB is enough for most case in S-band. By combining all these items and applying the link budget equation, the received E_b/N_0 is equal to 12.8 dB. Usually, some coding methods can be used to enhance the bit error performance and reduce the required E_b/N_0 . The Consultative Committee for Space Data Systems (CCSDS) recommends some of them, i.e., Reed-Solomon, Convolutional, Concatenated, TURBO, and LDPC codes [19]. Figure 5.14 illustrates their performance comparison with coding rate = 1/2. For example, the Convolutional (7,1/2) coding method requires the E_b/N_0 of 4.3 dB when $BER = 1e-5$, which leads to 5.4

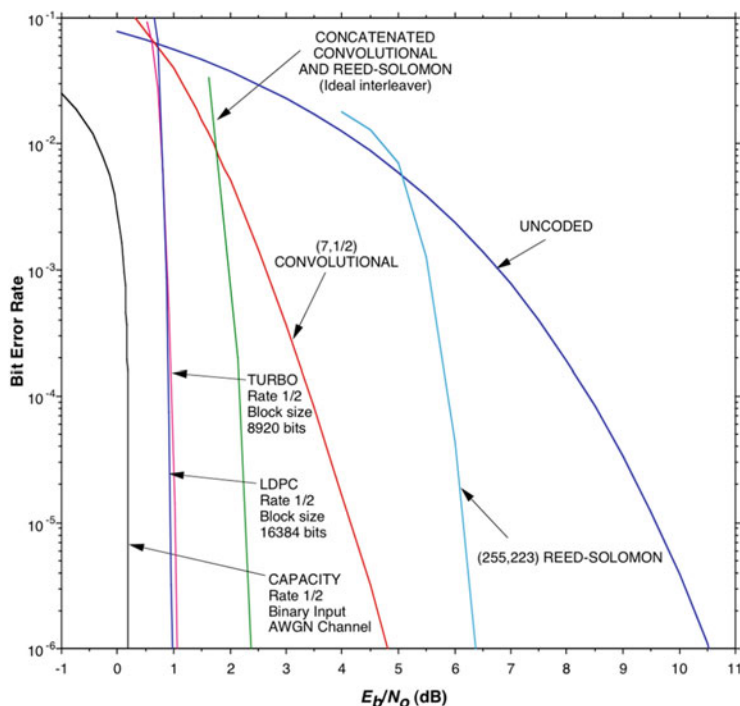


Fig. 5.14 Performance Comparison of CCSDS recommended channel coding techniques. The rightmost blue line is an uncoded scheme, and the lowest possible performance is the Shannon-limit of a 1/2 rate code. Figure reproduced with permission from [19], © CCSDS

dB more gain compared to the uncoded scheme. In this case, the link has a margin of 7.0 dB as shown in Table 5.2.

5.3.2.2 Downlink with the Narrow Beam

To transmit satellite payload data back to the ground, a much higher data rate is required. A data rate of 20 Mbps requires 43 more dB than the low status monitoring data rate of 1 kbps. Hence, items of the link budget equation should be improved to compensate for the new requirement. An onboard solid-state power amplifier of 10 W (10 dBW) is larger but still appropriate for power consumption. The G/T value of the ground station should be improved to 10 dB/K. This requires thousands of antenna elements to form such high performance. With these two improvements, the received E_b/N_0 is 9.8 dB. To meet this stringent link budget, more efficient coding methods can be used in the expense of decoding hardware cost. The link still has a margin of 7.0 dB if the LDPC coding method is chosen.

5.3.3 Case Study for Ground-to-Satellite Uplink

5.3.3.1 Uplink for the Normal Case

The link analysis for uplink is similar to downlink. The $EIRP$ of the ground station narrow beam is set as 45 dBW for the normal case. Limited by the onboard electronics, the G/T value and the demodulator implementation loss of the satellite receiver are usually worse than the ground station. The data rate of telecommand is assumed as 100 kbps. To save the onboard demodulator hardware, a simple RS(255,233) code is used to send telecommand toward the satellite, the required E_b/N_0 is 6.3 dB when $BER = 1e-5$. All the items result in an uplink margin of 9.9 dB for the normal case.

5.3.3.2 Uplink for the Emergency Case

For the emergency case, the satellite loses its attitude control and mis-points with the ground station. A telecommand must be injected to attempt to control the satellite with any orientation. The gain of the satellite antenna is very low under this situation, this leads to a poorer G/T value of the satellite receiver than the normal case. To overcome this drawback, the $EIRP$ of the ground station must be enhanced. Besides this, simple emergency telecommands are used with a much lower data rate, and an uncoded scheme must be used to avoid the case of onboard decoding program failure. As shown in Table 5.3, this results in a very high uplink margin for the emergency case.

Table 5.2 Examples of link budget for satellite-to-ground downlink

Parameter	Value (hemispheric beam)	Value (narrow beam)	Unit
P_{sat}	0	10	dBW
G_{sat}	5	5	dBi
$(EIRP)_{sat}$	5	15	dBW
$(G/T)_{gs}$	-20	10	dB/K
k	-228.6	-228.6	dB/K-Hz
d	3000	3000	km
f_{down}	2.2	2.2	GHz
L_{path}	168.8	168.8	dB
L_{atm}	2	2	dB
R_b	30	73	dB-bps
$(E_b/N_0)_r$	12.8	9.8	dB
$(E_b/N_0)_{reqd}$	4.3	1.3	dB
L_{imp}	1.5	1.5	dB
M	7.0	7.0	dB

Table 5.3 Examples of link budget for satellite-to-ground uplink

Parameter	Value (normal case)	Value (emergency case)	Unit
$(EIRP)_{gs}$	45	55	dBW
$(G/T)_{sat}$	-40	-50	dB/K
k	-228.6	-228.6	dB/K-Hz
d	3000	3000	km
f_{up}	2.1	2.1	GHz
L_{path}	168.4	168.4	dB
L_{atm}	2	2	dB
R_b	50	20	dB-bps
$(E_b/N_0)_r$	13.2	43.2	dB
$(E_b/N_0)_{reqd}$	6.3	9.7	dB
L_{imp}	2	2	dB
M	9.9	31.5	dB

5.3.4 Design Tradeoffs for the Phased Array Antenna

As indicated in the link budget equation, the $EIRP$ and the G/T value of the ground station are vital for the overall performance. However, these two parameters behave somewhat differently. Assuming all the antenna elements have the same receiving performance, the G/T value of the antenna array improves N times by capturing signals N times larger,

$$\left(\frac{G}{T}\right)_{array} = N \cdot \left(\frac{G}{T}\right)_{ant} \quad (5.11)$$

Hence, for a required array G/T value, the number of antenna elements is inversely proportional to the G/T value of the antenna element. The higher the antenna performance, the fewer elements are required. The choice of practical use of types of antenna elements comes down to engineering trade-offs of many aspects, including size, weight, bandwidth, gain, efficiency, and power handling capability. Figure 5.15 illustrates the relationships between the G/T value and the number of the antenna array for ground station use. For the hemispheric beam, the required G/T value is relatively low, thus only a few high performance antenna elements are needed. For the narrow beam, the antenna with a relatively simple structure and low G/T value is preferred. The reason is although the required antenna number is large, it is still the best choice for low cost and easy to integrate with the T/R module.

For signal transmission, the phased array provides a very efficient way to combine power in space. There is an N^2 improvement of the $EIRP$ of the antenna array with

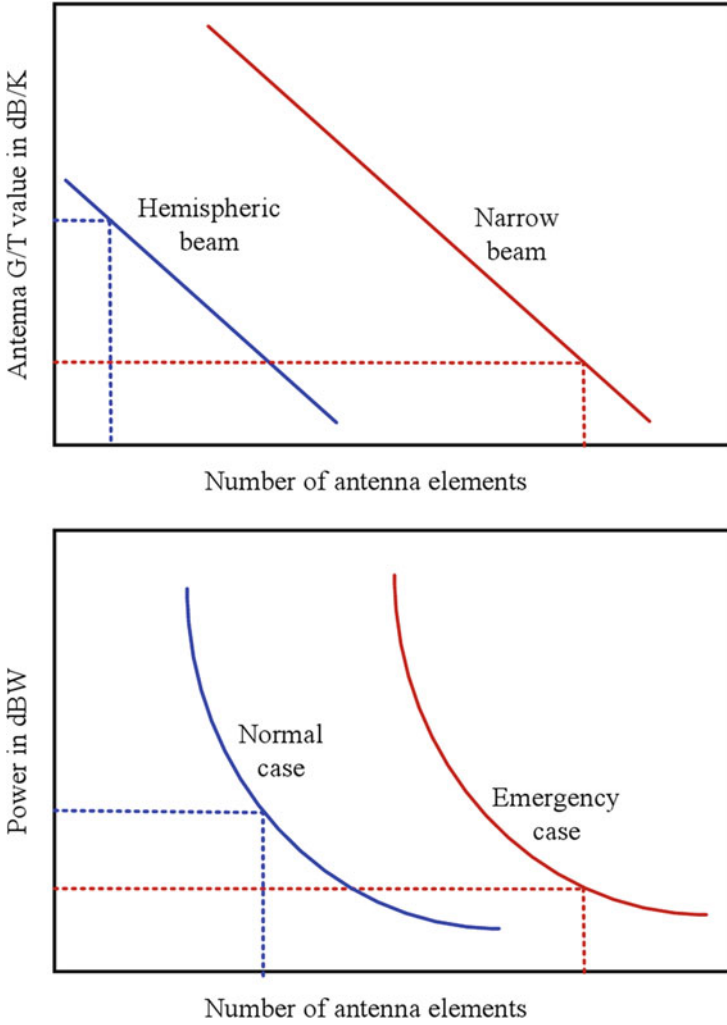


Fig. 5.15 Relationship of the antenna array performance, **a** the G/T value and **b** the EIRP

N antenna elements and N times of coherently compiled power,

$$EIRP_{array} = (NG_{ant}) \cdot (NP_{ant}) = N^2 G_{ant} P_{ant} = N^2 EIRP_{ant} \quad (5.12)$$

As the antenna gain is determined by the choice of antenna, the power amplifier is the remainder issue to realize a desired EIRP performance. Currently, there are a lot of available technologies of power amplifiers, such as Silicon, Silicon Carbide (SiC), Gallium Arsenide (GaAs), Gallium Nitride (GaN), Indium Phosphide (InP), and Silicon Germanium (SiGe). These technologies and capacities for power amplifiers

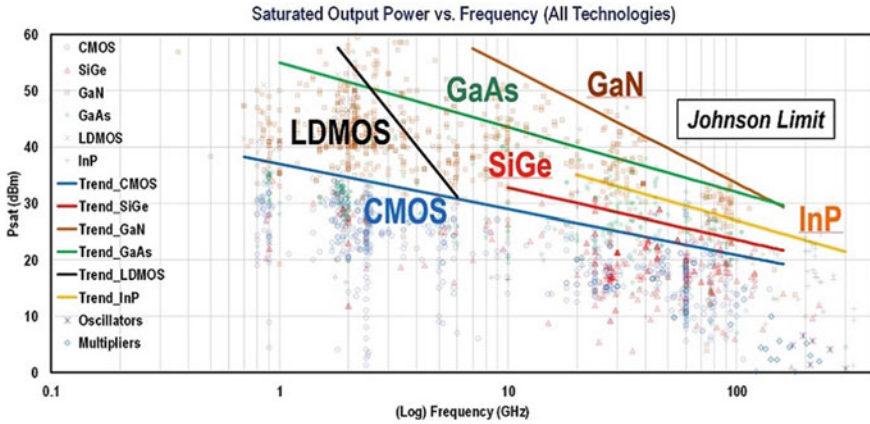


Fig. 5.16 Technologies and capacities for power amplifier. Figure reproduced with permission from [20], © 2021 IEEE

have been presented in a recent review paper [20] and reproduced as in Fig. 5.16. The practical selection of power amplifiers is based on the combined consideration of power generating capability, efficiency, and cost of the selected technology. When the required power is less than 0.1 W, Si-based materials are preferred for high density integration. Meanwhile, GaAs and GaN are usually used for power requirement of about 1 W. Moreover, the selected power level must fall in the power handling capability of the antenna element, to avoid saturation or even breakdown of it.

5.4 Summary

Some key design issues for satellite ground station application are discussed in this section. Only a few array geometries have the potential to realize the hemispherical coverage. The multifaced planar array with pyramidal frustum geometry is the simplest form, while the spherical array has no gain loss with beam steering, requires the least antenna elements, but difficult to fabrication and assembly. The geodesic dome geometry preserves all the advantages of a spherical array while its fabrication is based on the well developed, easily manufactured planar array. A similar geometry that has geodesic dome with a cylindrical base provides about 0.5 dB more gain at the horizon. To support the demand-driven approach, the ground station should have a hemispheric beam and multiple narrow beams, the work flowchart of a satellite pass is totally renewed and can be generally divided into five modes. Link budgets for satellite-to-ground downlink and ground-to-satellite uplink are also discussed, along with trade offs of phased array antenna of the ground station.

References

1. J.L. Kmetzo, An analytical approach to the coverage of a hemisphere by N planar phased arrays. *IEEE Trans. Antennas Propag.* **15**(3), 367–371 (1967)
2. L.E. Corey, A graphical technique for determining optimal array antenna geometry. *IEEE Trans. Antennas Propag.* **33**(7), 719–726 (1985)
3. A. Jablon, A. Agrawal, Optimal number of array faces for active phased array radars. *IEEE Trans. Aerosp. Electron. Syst.* **42**(1), 351–360 (2006)
4. I. Khalifa, R. Vaughan, Geometric design of pyramidal antenna arrays for hemispherical scan coverage. *IEEE Trans. Antennas Propag.* **55**(2), 468–471 (2007)
5. I. Khalifa, R. Vaughan, Geometric design and comparison of multifaceted antenna arrays for hemispherical coverage. *IEEE Trans. Antennas Propag.* **57**(9), 2608–2614 (2009)
6. C. Kumar, B.P. Kumar, V.S. Kumar, V.V. Srinivasan, Dual circularly polarized spherical phased-array antenna for spacecraft application. *IEEE Trans. Antennas Propag.* **61**(2), 598–605 (2013)
7. B. Pavan Kumar, C. Kumar, V. Senthil Kumar, V.V. Srinivasan, Active spherical phased array design for satellite payload data transmission. *IEEE Trans. Antennas Propag.* **63**(11), 4783–4791 (2015)
8. B.P. Kumar, C. Kumar, V.S. Kumar, V.V. Srinivasan, Reliability considerations of spherical phased array antenna for satellites. *IEEE Trans. Aerosp. Electron. Syst.* **54**(3), 1381–1391 (2018)
9. R.L. Haupt, *Antenna Arrays: A Computational Approach* (Wiley, New York, 2010)
10. M. Henderson, M.B. Davis, M. Huisjen, GDPAA advanced technology demonstration overview and results, in *IEEE International Symposium on Phased Array Systems and Technology* (Boston, 2010), pp. 140–143
11. H. Ahn, B. Tomasic, S. Liu, Digital beamforming in a large conformal phased array antenna for satellite operations support - Architecture, design, and development, in *IEEE International Symposium on Phased Array Systems and Technology* (Boston, 2010), pp. 423–431
12. R. Ahn, M. Weiner, B. Tomasic, Digital beamforming in a large conformal phased array antenna for satellite TT&C - Demonstration, in *2013 IEEE International Symposium on Phased Array Systems and Technology* (Waltham, 2013), pp. 807–814
13. B. Tomasic, J. Turtle, S. Liu, The geodesic sphere phased array antenna for satellite communication and air/space surveillance part I, in *AFRL in-House Technical Report*, AFRL-SN-HS-TR-2004-031 (2004)
14. J.M.F. Gonzalez, J.M.I. Alonso, J.G. Trujillo, A.N.S. de Toca, M.S. Perez, GEODA-SARAS: multi-phased array planar antenna design and measurements, in *The 8th European Conference on Antennas and Propagation (EuCAP 2014)* (2014), pp. 423–425
15. X. Li, K. Zhang, X. Meng, W. Zhang, Q. Zhang, X. Zhang, X. Li, Precise orbit determination for the FY-3C satellite using onboard BDS and GPS observations from 2013, 2015, and 2017. *Engineering* **6**(8), 904–912 (2020)
16. A. Hauschild, O. Montenbruck, Precise real-time navigation of LEO satellites using GNSS broadcast ephemerides. *Navigation* **68**(2), 419–432 (2021)
17. B. Sklar, *Digital Communications Fundamentals and Applications*, 2nd edn. (Prentice Hall, Upper Saddle River, 2001)
18. S. Abulgasem, F. Tubbal, R. Raad, P.I. Theoharis, S. Lu, S. Iranmanesh, Antenna designs for CubeSats: a review. *IEEE Access* **9**, 45289–45324 (2021)
19. Consultative Committee for Space Data Systems (CCSDS), TM Synchronization and Channel Coding, Summary of Concept and Rationale, CCSDS 131.0-G-2. Green Book. Issue 2. Washington, D.C., November 2012
20. V. Camarchia, R. Quaglia, A. Piacibello, D.P. Nguyen, H. Wang, A. Pham, A review of technologies and design techniques of millimeter-wave power amplifiers. *IEEE Trans. Microw. Theory Techn.* **68**(7), 2957–2983 (2020)