The Design of A Matrix Completion Signal Recovery Method for Array Processing

by

Jie Fan

A Thesis Presented in Partial Fulfillment of the Requirements for the Degree Master of Science

Approved October 2016 by the Graduate Supervisory Committee:

Andreas Spanias, Chair
Cihan Tepedelenlioglu
Konstantinos Tsakalis

ARIZONA STATE UNIVERSITY

December 2016
ABSTRACT

For a sensor array, part of its elements may fail to work due to hardware failures. Then the missing data may distort in the beam pattern or decrease the accuracy of direction-of-arrival (DOA) estimation. Therefore, considerable research has been conducted to develop algorithms that can estimate the missing signal information. On the other hand, through those algorithms, array elements can also be selectively turned off while the missed information can be successfully recovered, which will save power consumption and hardware cost.

Conventional approaches focusing on array element failures are mainly based on interpolation or sequential learning algorithm. Both of them rely heavily on some prior knowledge such as the information of the failures or a training dataset without missing data. In addition, since most of the existing approaches are developed for DOA estimation, their recovery target is usually the co-variance matrix but not the signal matrix.

In this thesis, a new signal recovery method based on matrix completion (MC) theory is introduced. It aims to directly refill the absent entries in the signal matrix without any prior knowledge. We proposed a novel overlapping reshaping method to satisfy the applying conditions of MC algorithms. Compared to other existing MC based approaches, our proposed method can provide us higher probability of successful recovery. The thesis describes the principle of the algorithms and analyzes the performance of this method. A few application examples with simulation results are also provided.
Dedicated to my workmates, friends and families.
ACKNOWLEDGMENTS

First of all, I’d like to express my sincere thanks to my advisor, Dr. Spanias, who has given me so much helpful advices and great inspirations on my study and research, and has tried his best to improve my paper. It is his patient instruction, insightful criticism and expert guidance that lead me to explore on the journey to knowledge.

Secondly, I want to thank those workmates and professors that I have worked with. During my research process, they may provide me advices and comments, or give me instruction for my study, or share me their research experience. Without their help, I cannot go this far in my research area.

Then, I need to show my gratitude to ASU and SenSIP Laboratory. Thank them to provide me useful information and precious resources for my living, study and research.

Last but not the least, I also want to take this opportunity to express my deep appreciation to my parents for their loving considerations and great confidence in me all through these years. Studying abroad alone is not an easy matter. It is not only for me, but also for my parents. Thank them again for their unselfish support, material as well as spiritual.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF FIGURES</th>
<th>vi</th>
</tr>
</thead>
</table>

## CHAPTER

1. INTRODUCTION ................................................................. 1
   1.1 Array Processing ...................................................... 1
   1.2 Problem Statement .................................................... 4
   1.3 Conventional Methods .................................................. 5
   1.4 Proposed Method ....................................................... 6

2. BACKGROUND REVIEW .......................................................... 9
   2.1 The Radiation Characteristics of Antenna Arrays .................. 9
   2.2 The Data Model of Array Signal Processing ......................... 11
   2.3 Common Algorithms ..................................................... 13
   2.4 Array Figuration ......................................................... 21
   2.5 Array Processing in The Missing Data Case ......................... 24

3. PROPOSED METHOD ............................................................. 34
   3.1 Introduction of Matrix Completion Theory .......................... 34
   3.2 Feasibility of Matrix Completion Theory ............................ 37
   3.3 Simple Reshaping Method .............................................. 39
   3.4 Proposed Improved Overlapping Reshaping Method ................. 42

4. SIMULATION RESULTS .......................................................... 46
   4.1 Success Rate Analysis ................................................... 46
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2 Application: Spatial Filtering</td>
<td>52</td>
</tr>
<tr>
<td>4.3 Application: DOA Estimation</td>
<td>57</td>
</tr>
<tr>
<td>5. CONCLUSION AND FUTURE WORK</td>
<td>62</td>
</tr>
<tr>
<td>5.1 Summary</td>
<td>62</td>
</tr>
<tr>
<td>5.2 Future Work</td>
<td>63</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>65</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Milwaukee’s Oldest Radio Array with Its Transmitter</td>
<td>2</td>
</tr>
<tr>
<td>2.1</td>
<td>The Radial Element and Coordinate Relation</td>
<td>10</td>
</tr>
<tr>
<td>2.2</td>
<td>A Typical Uniform Linear Array Structure</td>
<td>12</td>
</tr>
<tr>
<td>2.3</td>
<td>The Beamforming Model</td>
<td>14</td>
</tr>
<tr>
<td>2.4</td>
<td>Design of Ultrawideband Aperiodic Arrays</td>
<td>22</td>
</tr>
<tr>
<td>2.5</td>
<td>A Typical Artificial Neural Network</td>
<td>30</td>
</tr>
<tr>
<td>3.1</td>
<td>The Feasibility of Matrix Completion</td>
<td>39</td>
</tr>
<tr>
<td>3.2</td>
<td>The Simple Reshaping Method</td>
<td>40</td>
</tr>
<tr>
<td>3.3</td>
<td>The Overlapping Reshaping Method</td>
<td>43</td>
</tr>
<tr>
<td>4.1</td>
<td>Success Rate of A 36-Element ULA with 2 Input Signals</td>
<td>47</td>
</tr>
<tr>
<td>4.2</td>
<td>Success Rate of A 64-Element ULA with 2 Input Signals</td>
<td>48</td>
</tr>
<tr>
<td>4.3</td>
<td>Success Rate of A 100-Element ULA with 2 Input Signals</td>
<td>49</td>
</tr>
<tr>
<td>4.4</td>
<td>Success Rate of A 100-Element ULA with 4 Input Signals</td>
<td>50</td>
</tr>
<tr>
<td>4.5</td>
<td>Spatial Filtering (a)</td>
<td>52</td>
</tr>
<tr>
<td>4.6</td>
<td>Spatial Filtering (b)</td>
<td>54</td>
</tr>
<tr>
<td>4.7</td>
<td>Spatial Filtering (c)</td>
<td>55</td>
</tr>
<tr>
<td>4.8</td>
<td>Spatial Filtering (d)</td>
<td>56</td>
</tr>
<tr>
<td>4.9</td>
<td>Spatial Filtering (e)</td>
<td>57</td>
</tr>
<tr>
<td>4.10</td>
<td>DOA Estimation (a)</td>
<td>58</td>
</tr>
<tr>
<td>4.11</td>
<td>DOA Estimation (b)</td>
<td>60</td>
</tr>
<tr>
<td>4.12</td>
<td>DOA Estimation (c)</td>
<td>61</td>
</tr>
</tbody>
</table>
Chapter 1
INTRODUCTION

1.1 Array Processing

The basic goal of signal processing is to extract, recover and utilize the useful information from signal features. Therefore, it is particularly important to provide effective detection and accurate estimation for signal parameters in complicated electromagnetic environments. During the course of its development, the research of signal processing has extended from the simplest one-dimensional signal processing to a more complex field of multi-dimensional signals. For example, Milwaukee’s oldest radio station is shown in Figure 1.1. The array consisted of four Blaw-Knox self-supporting towers in a rectangle. Notice the lack of fencing, warning signs and the like around the towers. Since Wiener filtering theory was applied on array processing [1] in the 1960’s, by using sensor arrays or antenna arrays, researchers would be able to transfer time-domain-sampled signals to space-domain-sampled signals and deal with some spatial problems, for example estimating the direction of arrivals. Thus, various theoretical results in time domain could be extended to space domain and a new field for studying array signal processing was opened up. Until the 1990’s, antenna arrays had not been used for beamforming and spatial diversity. Then researchers found that through array systems, they could use multipath effects to improve the performance of communication systems. In 1994, A. J. Paulraj and T. Kailath patented a technique that allows us to use multi-input-multi-output (MIMO) systems to improve the capacity of wireless broadcast systems [70]. The first time-space
coding architecture was proposed in 1998 [71], which improved the data rate and the reliability of communications over fading channels. So far, array processing techniques and MIMO systems have played significant roles in large-scale wireless applications such as wireless local area and third generation networks [72].

![Figure 1.1 Milwaukee’s oldest radio array (source: Engineering Radio Blog).](image)

As the name suggest, array processing is a technique that places multiple sensors at different places in the space domain thereby forming a so-called sensor array, and then the sensor array can receive spatial signals. Array processing is regarded as an important branch in the signal-processing field and is widely used for civilian and military programs such as communication, radar, sonar, navigation, geological exploration, mechatronics measurement, biomedicine, and radio astronomy [2-14]. Specific processing methods will be adopted to deal with the received signals that can enhance the signal-of-interest (SOI), suppress the signal-not-of-interest (SNOI), extract useful signal features and perceive the inner information. Compared to traditional one-dimensional signal processing with a single-sensor-system, array processing can provide more flexible beam control, larger signal gain, stronger inference suppression and better spatial discrimination.

The main research directions of array processing, in general, are adaptive beamforming [15-18] and super resolution direction-of-arrival (DOA) estimation [19-23].
Adaptive beamforming technology is also known as adaptive spatial filtering technique. Essentially, it enhances or suppresses signals from different directions and finally extracts useful information from the received signals. This technology has been well applied in a variety of fields, such as radar, communication and sonar. Super resolution DOA estimation is a technology that estimates the DOAs of the array based on super resolution spectrum estimate methods, accordingly realizing the purpose of exact location. This technology helps people to implement the localizing function for many applications, for example radar, sonar, geological exploration and radio astronomy.

Adaptive beamforming is the process of adding weighting factors on each array element for spatial filtering. It enhances useful signals and suppresses interference signals. The weighting factors are adjusted adaptively due to a changing signal environment. Although adaptive beamforming methods can provide us an optimal signal-to-interference-and-noise-ratio (SINR) under a desired circumstance, there are always errors in a realistic environment including covariance matrix estimation errors due to a limited adaptive training set, steering errors due to the steering vector and system errors such as amplitude-phase errors, element location errors, mutual coupling among array elements and channel frequency response mismatch. At this point, the performance of adaptive beamforming will be greatly decreased, or even fail to work.

Since the famous multiple-signal-classification (MUSIC) algorithm introduced in 1979 [19], super resolution DOA estimation technologies have developed rapidly over the past thirty years. However, research on super resolution DOA estimation is based on a few desired assumptions, for example the ambient noise should be white Gaussian noise.
(WGN), or spatial stationarity. If these assumptions are invalid, there will be a significant performance deterioration of DOA estimation methods.

Although a lot of research has been conducted for the above two areas and a variety of algorithms have been introduced to provide better performance, there are still some difficulties in achieving engineering completion [23]. On one hand, the computational cost of these algorithms is very large; and on the other hand, these algorithms are not robust enough, which means they have a high requirement on system environment and signal environment. Therefore, research on robust array processing algorithms and engineering completion methods possess important theory significance and practical value.

1.2 Problem Statement

As mentioned before, the robustness of our array system has become a research priority. We should notice that hardware failure is an inevitable problem. If part of the elements in the array fail to work properly, the performance of our array processing algorithms will be heavily depressed. Specifically, for adaptive beamforming, when hardware failure happens and some signals are lost, the main-lobe will be expanded and the side-lobe power level will increase, causing the suppression of interference and system noise will be weakened and SINR will be lower. When we come to super resolution DOA estimation, the missing signals will distort the covariance matrix, which will increase the side-lobe dramatically. This may affect the accuracy of the DOA estimation.

Under the above circumstances, we may need to find out a way to help the array system recover the missing signals and overcome those problems. At this point, if we recognize hardware failure as we “passively” turning off these broken down sensors, we
will naturally find that this group of methods will also be feasible and practical for an “active” situation. Based on the Fourier transform theory [24], the main-lobe width is positively associated with the reciprocal of the array aperture. If we want to increase the array aperture, we will need a larger number of elements with more power consumption and memory requirement. Now suppose we “actively” turn off several sensors and apply a method to recover the missing signals, we could make our system more inexpensive and low-powered.

Therefore, by using methods that can help us to reconstruct the missing signals, we will be able to make our array system more robust and efficient. With such methods, it will be possible to achieve a certain performance with fewer sensors than we previously needed.

1.3 Conventional Methods

The existing approaches to resolve such issues as we discussed before can be generally divided into two main types, interpolation based algorithms [25-27] and neural network based algorithms [28]. Both of them focus on recovering the covariance matrix of the array signals and they are mainly applied on super resolution DOA estimation.

For interpolation algorithms, we first need to know the element and the time at which the failure occurs. Then, various estimation algorithms could be utilized to handle failures. It could be a relatively simple ad-hoc estimate algorithm or a maximum-likelihood (ML) estimate algorithm or a genetic algorithm (GA) with digital beamforming.
And in all the conventional algorithms from the cited papers [25-29], the missing signals will be replaced by data obtained by the interpolation of the available data.

In neural network based approaches introduced in the citation [28], a minimum resource allocation network (MRAN) [28, 30] could be used as a sequential learning algorithm for minimum radial basis function (RBF) neural network. This MRAN system should be first trained under no failure cases and the trained network can then be used for DOA estimation even under failures. In most other existing neural network solutions for this problem, the network size has to be previously fixed [29]. However, with an MRAN, the size can be automatically decided by the training data. Then, through the memory process in MRAN, the missed information in the covariance matrix can be successfully estimated.

1.4 Proposed Method

Though existing approaches can handle the failures in array processing, all of them strongly rely on prior knowledge, either some failure information or a training dataset without failure. And since those algorithms are mainly developed for DOA estimation, their recovery targets are usually the covariance matrix. That means they cannot recover the original receiving signals, which will make them infeasible for spatial filtering area.

In this thesis, a new approach based on matrix completion (MC) theory is introduced. It does not rely on any prior knowledge and it can directly provide us the recovered receiving signals instead of their covariance matrix. This would suggest that our new algorithm could be applied to both adaptive beamforming and super resolution DOA estimation.
To explain the principle of this method, first we need to briefly introduce the MC theory. Matrix completion theory is a remarkable new research field. For now, it can be effectively applied to many science and engineering areas, such as collaborative filtering, image inpainting, machine learning, system control, computer vision and predicting missing data in sensor networks. MC theory is intended to complete the matrix that contains absent entries from a few observed entries. Based on a large amount of actual situations, researchers established the theoretical model of MC theory. When a low-rank matrix or an approximately low-rank matrix satisfies some appropriate conditions [31-32], by solving a convex optimization problem, it can be accurately or approximately accurately completed in a high probability. This convex optimization problem is the smallest nuclear norm or a least square problem with a nuclear norm regularization.

We should notice that the feasibility of MC theory depends on some appropriate conditions and one of them is that the matrix has to follow the strong coherent property. Unfortunately, in the array processing with failure recovery elements problem mentioned before, when several array elements fail to work, we will lose their whole relative rows in the signal matrix, which dissatisfies the strong coherent property. Then the MC theory is infeasible for the array processing failure recovery problems.

To overcome the above barrier, a new algorithm is introduced in this thesis report. We utilize the inner relationship among array elements and apply a reshaping process to the array signal on each snapshot. This reshaping process will allocate the absent entries in a new matrix to satisfy the strong coherent property. And considering about improving the performance of this approach, we make the reshaping step as an adjustable transform
process. By changing the parameters in this process, we will be able to improve the probability of successfully recovering the signal matrix.
Chapter 2

BACKGROUND REVIEW

2.1 The Radiation Characteristics of Antenna Arrays

The radiation characteristics of antenna arrays follow the far-field characteristics, which require the distance between the observation point and the source to be much larger than the overall size of antenna arrays and the operating wavelength. The antenna array radiation characteristics depend on the polarization of array elements, the quantity of array elements, array figuration, the spacing between elements, excitation amplitude and phase. We can change the characteristics by controlling these factors. Besides that, we should also consider the influence on the general characteristics of arrays from the single element characteristics.

In the field of antenna design, the far-field pattern refers to the directional dependence of the strength of the radio waves from antennas or other sources. It contains the radiation intensity, field strength, phase and polarization. Specially, we usually call the change rules of the received power level and filed strength at the far-filed antennas with varying orientation coordinates, respectively, as the power pattern and the filed pattern. If we write them as spatial orientation coordinate functions, we will call them pattern functions. They are just the functions of orientation and have nothing to do with the radial distance $r$. 

In a free space, if there is a source $S_n$ at the point $P_n(r_n, \theta_n, \varphi_n)$, as shown in Figure 2.1, the far-field function at observation point $P(r, \theta, \varphi)$ can be written as

$$E_n(\theta, \varphi) = f_n(\theta, \varphi) I_n e^{i(k r_n \cdot e_r + \beta_n)} = f_n(\theta, \varphi) I_n e^{i(k r_n \cos \psi_n + \beta_n)}, \quad (1)$$

and we have:

$$\cos \psi_n = e_{r_n} \cdot e_r = \cos \theta \cos \theta_n + \sin \theta \sin \theta_n \cos(\varphi - \varphi_n), \quad (2)$$

where $f_n(\theta, \varphi)$ represents the pattern function of the array element at the coordinate origin, $k = 2\pi / \lambda$ is a phase constant and $\lambda$ is the free space wavelength. $I_n$ and $\beta_n$, where $n = 1, 2, 3, \cdots N$, are excitation amplitude and phase, respectively. We should notice that $E_n(\theta, \varphi)$ in equation (1) does not entirely equal to the actual field strength. It makes an approximation and simplification of the term $e^{(-jk|\mathbf{r}-\mathbf{r}_n|+j\omega t)/|\mathbf{r}-\mathbf{r}_n|}$ from the actual field strength, where

$$\mathbf{r} = e_x x + e_y y + e_z z, \quad (3)$$

$$\mathbf{r}_n = e_x x_n + e_y y_n + e_z z_n, \quad (4)$$
\[ r - r_n = e_x(x - x_n) + e_y(y - y_n) + e_z(z - z_n), \]  
\[ |r - r_n| = \sqrt{(x - x_n)^2 + (y - y_n)^2 + (z - z_n)^2}, \]  
\[ |r - r_n| = \begin{cases} \frac{r}{r_n \cos \psi_n} & \text{for amplitude} \\ r - r_n & \text{for phase} \end{cases} \]

For an antenna array with multiple elements, the far-field strength will be the vector sum of radiation pattern functions from all array elements and it is given by,

\[ E(\theta, \phi) = \sum_{n=1}^{N} E_n(\theta, \phi). \]  

If all of the elements have the same polarized direction, the total far-field strength could derive from the algebraic sum:

\[ E(\theta, \phi) = \sum_{n=1}^{N} E_n(\theta, \phi). \]  

When all array elements are same, we will have \( f_n(\theta, \phi) = f(\theta, \phi) \), and equation (9) becomes:

\[ E(\theta, \phi) = f(\theta, \phi) \sum_{n=1}^{N} I_n e^{j(kr_n \cos \psi_n + \beta_n)} = f(\theta, \phi) \cdot AF(\theta), \]  

where

\[ AF(\theta) = \sum_{n=1}^{N} I_n e^{j(kr_n \cos \psi_n + \beta_n)}. \]

We usually call \( AF(\theta) \) as the array factor or the space factor. When the excitation of each element is fixed, the array factor will only depend on the distribution of radial elements in space.

### 2.2 The Data Model of Array Signal Processing

Similarly, for a sensor array to receive spatial signals, the wave functions at each array elements that represent signals from a same source can derive from the product of the source pattern function and a steering factor.
The function of a spatial plane radial wave is four-dimensional:

\[ g(t, r) = A e^{j[2\pi(f t - k^T r)]}, \]  

where \( r \) represents where the particle is, \( t \) represents time, \( k \) is the wave vector, \( A \) and \( f \) are amplitude and frequency, respectively.

To simplify the wave function, we need to make a narrowband assumption. Suppose all elements sample the signals at the same time and their received signals have a same complex envelope. That means we only need to consider the phase shift among array elements, which only depends on the array figuration.

![Figure 2.2 A typical uniform linear array structure.](image)

The most basic array structure is uniform linear array (ULA). As shown in Figure 2.2, it arranges identical sensor elements along a line in space with uniform distance. For a ULA of \( N \) sensors with an interelement spacing \( d \), the spatial narrowband signal function as mentioned above will be:

\[ s(t, r) = s(t) e^{j(\omega t - r^T a)}. \]
If we select element 1 as our reference point, the received signals of each element will be given by

\[
\begin{align*}
    x_1(t) &= s(t)e^{j\omega t} \\
    x_2(t) &= s(t)e^{j\omega t}e^{2\pi d \sin \theta} \\
    &\vdots \\
    x_N(t) &= s(t)e^{j\omega t}e^{2\pi d (N-1) \sin \theta},
\end{align*}
\]

where the vector format can be written as:

\[
X(t) = \begin{bmatrix}
x_1(t) \\
x_2(t) \\
\vdots \\
x_N(t)
\end{bmatrix} = s(t)e^{j\omega t} \begin{bmatrix}
1 \\
e^{2\pi d \sin \theta} \\
\vdots \\
e^{2\pi d (N-1) \sin \theta}
\end{bmatrix} = s(t)\mathbf{a}(\theta).
\]

We call \(\mathbf{a}(\theta)\) as the steering vector. In narrowband situation, it only depends on the configuration of the array elements, which is usually known and the DOAs, which is usually unknown.

### 2.3 Common Algorithms

As our previous briefing explains, there are two main research directions in array processing: spatial filtering and high resolution DOA estimation. In this section, a few common algorithms applied in these areas will be introduced.
Spatial filtering is similar to frequency domain filtering. It adopts beamforming technology, as shown in Figure 2.3, to process weighting summation for the sampled data:

\[ y(t) = W^H X(t) = s(t)W^H a(\theta), \]  

(16)

where \( W \) is the weighting vector. By this way, the signals in certain spatial areas can be enhanced or weaken. Now, we recognize

\[ P_W(\theta) = W^H a(\theta), \]  

(17)

as the beam pattern. When \( W \) completes an in-phase stacking for the signal from a certain direction \( \theta_0 \), the module value of \( P_W(\theta_0) \) will become the maximum value in our beam pattern. The received signal \( X(t) \) is actually the spatial sampled signal and beamforming allows for selecting the direction angle, which completes spatial filtering. We can compare this with the frequency selection in the frequency domain filtering.
Specially, for a ULA, when we want to steer the beam to a certain direction angle \( \theta_0 \), we can set \( W = a(\theta_0) \) and then the beamforming process will be:

\[
P(\theta) = W^H a(\theta) = a(\theta_0)^H a(\theta)
\]

\[
= \sum_{i=1}^{N} e^{j \frac{2\pi d (i-1)}{\lambda} (\sin \theta - \sin \theta_0)}
\]

\[
= \frac{1 - e^{j \frac{2\pi d N (\sin \theta - \sin \theta_0)}}}{1 - e^{j \frac{2\pi d (\sin \theta - \sin \theta_0)}}},
\] (18)

Thus, its module value can be expressed as:

\[
|P(\theta)| = \left| \frac{\sin[N(\phi - \phi_0)/2]}{\sin[(\phi - \phi_0)/2]} \right|,
\] (19)

where \( \phi = \pi \sin \theta \) and \( \phi_0 = \pi \sin \theta_0 \). \( |P(\theta)| \) is the beam pattern of our array.

We call this method normal beamforming. Received signals are coherently accumulated in the direction angle range of the main lobe. It is actually a simple matched filter that is easy to be implemented. However, the normal beamforming just relies on the geometric structure of the array and DOAs of the received signals. For a certain selected direction angle, the weighting vector is fixed and will not adapt to different signal environments. It has a poor ability of restraining interferences. Therefore, under white noise environment, normal beamforming may be an optimal method but under colored noise environment, we need to find an approach that is similar to the Wiener filtering theory [33-34].

Adaptive beamforming is a technique that applies Wiener filtering theory on spatial filtering. The weighting vector of an adaptive beamforming algorithm relies on the signal environment. Based on equation (14), for a stable random received signal, the output power is:
\[
E[|y(t)|^2] = E \left[ W^H X(t)(W^H X(t))^H \right]
= E[W^H X(t)X(t)^H W]
= W^H E[X(t)X(t)^H]W
= W^H R_x W, \quad (20)
\]
where \( R_x = [X(t)X(t)^H] \) is the covariance matrix of the array signals. It contains all statistical information from our received signals. At this point, when we try to obtain the optimal beamforming result, we would need to complete following optimization problems:

\[
\left\{ \begin{array}{l}
\min \limits_W W^H R_x W \\
\text{s.t. } f(W) = 0
\end{array} \right. \quad (21)
\]
where \( f(W) = 0 \) represents the optimal filtering criteria. There are usually three kinds of criteria: maximum signal to noise ratio (SNR) criterion, minimum mean square error (MSE) criterion and linearly constrained minimum variance (LCMV) criterion.

Specifically, for maximum SNR criterion, suppose the received array signals are given by

\[
X(t) = X_s(t) + X_n(t). \quad (22)
\]
The signal component \( X_s(t) \) is statistically uncorrelated with noise component \( X_n(t) \) and their covariance matrices are known as

\[
R_s(t) = E[X_s(t)X_s(t)^H], \quad (23)
\]
\[
R_n(t) = E[X_n(t)X_n(t)^H]. \quad (24)
\]
Then the output signal is

\[
y(t) = W^H X(t) = W^H X_s(t) + W^H X_n(t), \quad (25)
\]
and the output power can be expressed as
\[ E[|y(t)|^2] = W^H R_s(t)W + W^H R_n(t)W, \]

where \( W^H R_s(t)W \) is the signal power and \( W^H R_n(t)W \) is the noise power. Based on maximum SNR principle, we need to solve the following problem:

\[
\max_W W^H R_s(t)W \quad W^H R_n(t)W.
\]

By Rayleigh entropy, we know that

\[
\lambda_{\min}(R) \leq \frac{X^H R_s X}{X^H X} \leq \lambda_{\max}(R),
\]

then the above optimization problem can be solved to

\[
R_s W_{\text{opt}} = \lambda_{\max} R_n W_{\text{opt}}.
\]

By generalized eigenvalue decomposition (GEVD), the optimal weighting vector \( W_{\text{opt}} \) is the corresponding eigenvector of the greatest eigenvalue from matrix \((R_s, R_n)\).

When we want to use minimum MSE criterion, we have a condition to be satisfied, which requires a reference signal \( d(t) \). Suppose the MSE is expressed as \( \sigma(W) \) and our optimization target is given by \( \min_W \sigma(W) \). Then we have

\[
\sigma(W) = E\left[(W^H X(t) - d(t))^H W - d(t)\right] = W^H R_X W + E[|d(t)|^2] - W^H r_{xd} - r_{xd}^H W,
\]

where \( r_{xd} = E[X(t)d^*(t)] \) is the correlation vector and \( R_{xd} = E[X(t)X(t)^H] \) is the covariance matrix. The optimal weighting vector can be obtained by a derivative method for compound function, which is given by

\[
W_{\text{opt}} = R_X^{-1} r_{xd}.
\]

From the above equation, we can see that the minimum MSE criterion method requires the correlation vector of the received array signals and the desired reference signal.
When it comes to the LCMV criterion, the steering vector \( \mathbf{a}(\theta_0) \) of our desired signal is necessary. Then the received array signals can be expressed as

\[
\mathbf{X}(t) = s(t)\mathbf{a}(\theta_0) + \mathbf{N}.
\]  
(32)

Thus the output signal is given by

\[
y(t) = \mathbf{W}^H \mathbf{X}(t) = s(t)\mathbf{W}^H \mathbf{a}(\theta_0) + \mathbf{W}^H \mathbf{N}.
\]  
(33)

At this point, we can firstly set \( \mathbf{W}^H \mathbf{a}(\theta_0) = 1 \) to fix the desired signal component and then to minimize the covariance, which holds the equivalent of minimizing the noise component \( \mathbf{W}^H \mathbf{N} \). Therefore the optimization problem can be given by

\[
\begin{aligned}
\min_{\mathbf{W}} & \quad \mathbf{W}^H \mathbf{R}_x \mathbf{W} \\
\text{s.t.} & \quad \mathbf{W}^H \mathbf{a}(\theta_0) = 1.
\end{aligned}
\]  
(34)

One thing to note here is that \( \mathbf{W}^H \mathbf{a}(\theta_0) \) can be set to any nonzero constant value besides 1 because our purpose is to fix the desired signal component. The optimal weighting vector is given by

\[
\mathbf{W}_{\text{opt}} = \mu \mathbf{R}_x^{-1} \mathbf{a}(\theta_0),
\]  
(35)

and if we set \( \mathbf{W}^H \mathbf{a}(\theta_0) = 1 \), the coefficient \( \mu \) will be

\[
\mu = \frac{1}{\mathbf{a}^H(\theta_0) \mathbf{R}_x^{-1} \mathbf{a}(\theta_0)}.
\]  
(36)

The value of \( \mu \) wouldn’t affect SNR and beam pattern.

The above three adaptive beamforming criteria can be applied on various adaptive array-processing algorithms, either full adaptive algorithms that utilize all available degree of freedom or partial adaptive algorithms that use part of available degree of freedom [35-39].
For the other main research direction, high resolution DOA estimation, there are traditional direction-finding techniques, such as phase comparison method [40] and beam scanning method [41]. Moreover, the introduction of MUSIC algorithm marked the spatial spectrum estimation entering a prosperous period. After thirty-year development, this kind of techniques has already become mature.

First, for phase comparison method, there are phase differences among the received signals of different array elements due to different propagation distance between the signal sources to each array element. If we ignore the inconsistency of reception channel, in a single source situation, we can use a two-element array to obtain DOAs.

In a narrowband case, the received signal is given by

$$
\begin{bmatrix}
x_1(t) \\
x_2(t)
\end{bmatrix} = s(t) \begin{bmatrix}
1 \\
e^{j \frac{2\pi d}{\lambda} \sin \theta}
\end{bmatrix}.
$$

Therefore,

$$E[x_1(t)x_2^*(t)] = \sigma_s^2 e^{-j \frac{2\pi d}{\lambda} \sin \theta}.$$  

From equation (38), we can notice that we can obtain the DOA by comparing the phase of each element. In practical engineering applications, we can fast compute the DOA by table lookup.

However, this method also has an applying position. When there is no double-value fuzzy problem, which means

$$\frac{2\pi d}{\lambda} \sin \theta \ll 2\pi \theta,$$

the phase comparison method will be feasible. We can increase the interelement spacing $d$ to increase the accuracy.
Beam scanning method is based on beam forming approaches. For instance, we consider the simplest normal beam forming method, which is similar to the matched filter. In a narrowband case, the data model is given by,

$$y(t) = W^H X(t).$$ \hspace{1cm} (40)

If we set \( W = a(\theta_s) \), we will have

$$y(t) = a^H(\theta_s)X(t)$$

$$= a^H(\theta_s)s(t)a(\theta_0)$$

$$= s(t)a^H(\theta_s)a(\theta_0).$$ \hspace{1cm} (41)

where \( \theta_s \in [-90^\circ, 90^\circ] \) is our scanning range and \( \theta_0 \) is the DOA of the received array signals. When we plot the beam power pattern, we could see the DOA.

However, this method also has weakness. Although the beam scanning method can solve multiple-signal-source situations, when the included angle between two signals is smaller than the width of a beam, we will not be able to distinguish them.

Unlike above two methods, multiple-signal-classification (MUSIC) algorithm is based on matrix eigenvalue decomposition. From a geometric perspective, the observation space of the processed signal can be decomposed into two parts, signal subspace and noise subspace. Obviously, these two parts are orthogonal. The signal subspace can be formed from the eigenvectors related to the correlation matrix of received signal data. The noise subspace is formed from the eigenvectors related to the minimum eigenvalue in the correlation matrix of received signal data.

MUSIC algorithm is the cornerstone of space spectral estimation direction-finding theory. The principle of MUSIC algorithm is as follows:
(1) From the received array signal $X(t)$ we can estimate the covariance matrix by

$$R = E[XX^H] \approx \frac{1}{M} \sum_{i=1}^{M} X(t_i)X^H(t_i) = \hat{R}.$$  \hspace{1cm} (42)

(2) Do eigenvalue decomposition of estimated correlation matrix $\hat{R}$.

(3) Form $S_N^P$ from the eigenvectors related to $P$ relatively large eigenvalue. Or form $N_{N-P}^N$ from the eigenvectors related to $P$ relatively small eigenvalue.

(4) Project the steering vector $a(\theta)$ on $N_{N-P}^N$ and it is given by

$$P_n a(\theta) = (\sum_{i=P+1}^{N} v_i v_i^H)a(\theta).$$  \hspace{1cm} (43)

(5) The spectral peak can be given by

$$S(\theta) = \frac{1}{\|P_n a(\theta)\|} = \frac{1}{\sum_{i=P+1}^{N} |a^H(\theta)v_i|^2}.$$  \hspace{1cm} (44)

At this moment, the signals with different DOAs can be shown as spectral peaks. However, the spectrum is irrelevant to the signal strength. It only reflects the orthogonality between $a(\theta)$ and $N_{N-P}^N$.

2.4 Array Figuration

Besides ULA, we also have some other kinds of uniform spacing arrays, or periodical arrays, such as uniform rectangle planner array and uniform circular array. Their interelement spacing shall not exceed half of the wavelength. Since forties of last century, uniformly spaced arrays have been extensively and deeply studied [42]. When we require the array to have a high angular resolution, considering the main-lobe width of the beam pattern is positively associated with the physical dimension of our array, we need to increase the array aperture. At this point, to avoid grating lobe phenomenon, we have to use more elements in a traditional uniform spaced array. This sort of thinking may increase the
manufacturing cost. Therefore, researchers adopt nonuniform spaced array to avoid above shortcoming. Specifically, from Figure 2.4 we could see an example of designing broadband planner array [50]. On the left, there is a portion of a Penrose aperiodic tiling and its corresponding antenna array element locations. On the right, a wideband planar array design is given, which was generated by optimizing a perturbed Penrose tiling array.

![Figure 2.4 Design of ultrawideband aperiodic arrays.](image)

When we applied sparse spacing technique instead of the traditional uniform one, the array aperture becomes larger, the scanning beam becomes narrower, the directivity becomes stronger and the spatial resolution becomes higher. Unfortunately, the sparse array may result in a lower gain than a uniform array with the same aperture. To improve the practical applicability of sparse arrays, we need not only narrow main lobes in the beam pattern, but also beamforming algorithms that are effective and efficient. Those algorithms will be able to provide us enough gains with low side lobes. Thereby, we can improve their ability to detect targets and suppress clutter. For now, sparse array technique has been applied on military areas, for example high frequency groundwork phased radar antennas and synthetic impulse and aperture radar (SIAR) [43-44].
Theoretically, a single sensor can receive or transmit signals. However, in practice, we usually require sensors to have strong directivity, a high gain, a specific physical structure and sometimes the beam scanning ability. At this point, multiple sensors will be applied and the will be arranged under a certain rule to form an array. We call sensors arranged in a line as a linear array, call sensors arranged in a plane as a planner array and call sensors arranged in a certain carrier surface as a conformal array. In recent years, various kinds of advanced aircrafts, for instance supersonic airplanes, cruise missiles and artificial satellites, are increasingly keen to have their electronic equipment, such as radar antennas, installed on their surface areas to acquire higher aerodynamic performance and better weapon performance. That demands the sensors to be arranged in a conformal array for coinciding with the surface of the aircraft.

Since the 1960s researchers has extended their work from uniform spacing array to sparse array. By using nonuniform spacing method and randomly arrangement to avoid grating lobe phenomenon, a high side lobe is unluckily inevitable. Therefore, one of the main research directions in array processing is to find an optimal figuration for elements allocation, which can balance the main lobe performance and the side lobe performance [45-49].

2.5 Array Processing in Missing Data Case

As we mentioned in the introduction section, sensor failure is an inevitable problem in array processing. In the missing data case, for traditional analog beamforming methods, the array hardware has to be pulled out due to an unacceptable side lobe level. When it
comes to the digital beamforming methods, instead of replacing the sensors, we should invoke the relative algorithm again to recalculate the weighting vector and form a new beam pattern that is close to the original one [27]. Either hardware replacement or weighting parameters recalculation will present us a severe test of time and efficiency, especially for some military applications and strategy resource facilities, such as radar, sonar, wireless communication system, etc. In consequence, how to design a more robust sensor array system has become a significant research focus. Furthermore, suppose we could find an efficient approach with low consumption to handle sensor failures, we would also be able to proactively turned off several sensors or directly use a sparse array instead of a uniform one. That will help us to save cost of hardware and operational support.

To deal with the missing data case in DOA estimation applications, researchers developed two primary approaches, methods based on interpolation and methods based on neural network. Specifically, for interpolation-based methods, the basic ideal is replacing the missing data with interpolation values generated through the available signal data, which could complete the estimation of missing data. This kind of methods requires the elements and time where failure happened as the prior knowledge.

Before introducing the mathematical principle, we should first find out what is interpolation. The spatial signal data gathered by various methods are often observed according to user request determination. That is, the received data sets are comprised of random points in range of interest (ROI) or observations of regular grid points. However, sometimes users may require data from non-observed points. Then the spatial distribution of the observed data makes it possible for us to estimate the unknown data. Interpolation is
a process that is able to estimate the non-observed data based on the received data. For instance, interpolation can be applied on estimating almost all kinds of unknown geographical values, such as elevation, precipitation, chemical concentration, noise level, etc.

As for DOA estimation applications with missing array data, an interpolation-based approach introduced in [25] is presented as follows. We firstly assume a far-field narrowband data model. The received signal of an $N_1$-element array is given by

$$x(t) = A(\theta)s(t) + n(t), \quad (45)$$

where $\theta = [\theta_1, \theta_2, \ldots, \theta_k]$ contains the DOAs of the $k$ received signals, the matrix $A(\theta) = [a(\theta_1), a(\theta_2), \ldots, a(\theta_k)]$ contains the array steering vectors $a(\theta)$ in different DOAs, $s(t) = [s_1(t), s_2(t), \ldots, s_k(t)]^T$ is our received signal vector at time $t$, and $n(t) = [N_1(t), N_2(t), \ldots, N_k(t)]^T$ represents the noise.

By this stage, suppose some of the elements fail to work after a certain time. Assume the spatial data of our $N_1$-element array was collected for $t = 1, 2, \ldots, T_1$. After a certain time step $T_1$, a few elements in the array fail to work and the signal data is collected for $t = T_1 + 1, T_1 + 2, \ldots, T_1 + T_2$ with $N_2$ working elements remaining. Then for $t > T_1 + T_2$, a few more element fail to work and $N_3$ elements remain. Now let $p - 1$ represents the number of times that element failure occurred and the total measurement time is $t = N_1 + N_2 + \cdots + N_p$. We should note that $N_1 > N_2 > \cdots > N_p$.

As the required prior knowledge mentioned above, we assume $\{N_1, N_2, \ldots, N_p\}$ and $\{T_1, T_2, \ldots, T_p\}$ are known. Without loss of generality, we further assume that sensors are numbered. For instance, the sensors numbered $N_1 - N_2 + 1, \ldots, N_1$ work for $t = T_1 +$
1, \(T_1 + 2, \cdots, T_1 + T_2\), the sensors numbered \(N_2 - N_3 + 1, \cdots, N_2\) work for \(t = T_1 + T_2 + 1, T_1 + T_2 + 2, \cdots, T_1 + T_2 + T_3\) and so on.

At this point, we assume
\[
x_i(t) = [\mathbf{O}_{N_{i}(N_1 - N_i)} \quad \mathbf{I}_{N_{i}}]x(t),
\]
(46)
as the last \(N_i\) elements of \(X(t)\), where \(\mathbf{O}_{N_{i}(N_1 - N_i)}\) denotes an \(N_i \times (N_1 - N_i)\) matrix of zeros and define the following snapshot matrix \(X_i\) is given by
\[
X_i = [x_i(N_1 + \cdots + N_{i-1} + 1) \cdots x_i(N_1 + \cdots + N_i)],
\]
(47)
where \(i \in [1, p]\). The snapshot matrix \(X_i\) contains all available data during the time interval that \(N_i\) sensors were properly working in.

In DOA estimation applications, most of the existing algorithms are based on covariance matrix \(R\), which is defined as
\[
R = E[x(t)x^H(t)] = ASA^H + \sigma^2 \mathbf{I},
\]
(48)
where for any \(t_1\) and \(t_2\), we have
\[
E[s(t_1)s^H(t_2)] = \delta(t_1 - t_2)S,
\]
(49)
\[
E[n(t_1)n^H(t_2)] = \delta(t_1 - t_2)\sigma^2 \mathbf{I}_{N_i}.
\]
(50)
Therefore, in measurement period \(p\), the covariance matrix \(R_i\), as the \(N_i \times N_i\) lower right corner of \(R\), is given by
\[
R_i = E[x_i(t_1)x_i^H(t_2)] = [\mathbf{0} \quad \mathbf{I}_{N_i}]R[\mathbf{0} \quad \mathbf{I}_{N_i}]^H.
\]
(51)
At last, we define the sample covariance matrix \(\hat{R}_i\) from the measurement period \(p\) is given by
\[
\hat{R}_i = \frac{1}{N_i}X_iX_i^H.
\]
(52)
In order to overcome the missing data case in DOA estimation, the simplest idea is to estimate the covariance matrix $\mathbf{R}$ based on the data from the first measurement period. So mathematically speaking, we first let $i = 1$ and then the sample covariance matrix is 

$$\tilde{\mathbf{R}}_1 = \frac{1}{N_1} \mathbf{X}_1 \mathbf{X}_1^H.$$  

Unfortunately, it is expected to be a poor estimation owing to obvious reasons [25]. In the following, we present two estimation algorithm algorithms, a relatively simple ad-hoc estimate method and the Maximum-Likelihood (ML) estimate of the covariance matrix.

First for ad-hoc estimate algorithm, the estimation $\tilde{\mathbf{R}}_{ah}$ is the unstructured estimate of the covariance of array element $a$ and $b$ based on data from the measurement period that both of the elements were working properly. The estimate matrix is given by

$$\tilde{\mathbf{R}}_{ah}^{(a,b)} = \frac{1}{|\Omega_{a,b}|} \sum_{t \in \Omega_{a,b}} \mathbf{x}^{(a)}(t) \mathbf{x}^{(b)H}(t),$$  \hspace{1cm} (53)

where $\mathbf{x}^{(a)}(t)$ is the $a$th element of $\mathbf{x}(t)$, $\Omega_{a,b}$ is the set of time instants at which both sensor $a$ and $b$ were working properly, and $|\Omega_{a,b}|$ represents the number of elements in set $\Omega_{a,b}$.

Although $\tilde{\mathbf{R}}_{ah}$ is guaranteed to be Hermitian, it is easy to find express where $\tilde{\mathbf{R}}_{ah}$ becomes indefinite, which is an undesired property.

Besides the basic ad-hoc estimate algorithm, we also have the ML estimate algorithm [50], whose estimate covariance matrix $\tilde{\mathbf{R}}_{ml} = (\mathbf{H}^H)^{-1} (\mathbf{D})^{-1} (\mathbf{H})^{-1}$, where $\mathbf{H}$ and $\mathbf{D}$ are computed as follows:

(1) According to equation (52), we can compute $\tilde{\mathbf{R}}_i$ for $i = 1, \cdots, p$.

(2) For $b = 1, \cdots, N_l$ and $i = 1, \cdots, p$, we could compute $\mathbf{s}_{i,b}$ by
\[ s_{i,b} = \sum_{k=1}^{i} \Gamma_{k,b}, \]  

where \( \Gamma_{k,b} = N_i [\mathbf{0} \quad I_b] \tilde{R}_k [\mathbf{0} \quad I_b] \).

(3) We can compute \( \hat{g}_b \) for \( b = N_{i+1} + 1, \ldots, N_i \) and \( i = 1, \ldots, p \) by

\[
\hat{g}_b = \frac{(s_{i,b})^{-1} u}{u^H(s_{i,b})^{-1} u},
\]

where \( u = [1,0,\ldots,0]^T \).

(4) For \( b = 1, \ldots, N_1 \), we can compute \( \hat{d}_b \) according to

\[
\hat{d}_b^{-1} = \frac{\hat{g}_b^{-1}s_{b,b}\hat{g}_b}{N_1 + N_2 + \cdots + N_i}.
\]

(5) Now we define \( \hat{D} = \text{diag}\{\hat{d}_{N_1}, \ldots, \hat{d}_1\} \).

(6) Finally, \( \hat{H} \) can be constructed to a lower triangular matrix by using

\[
\hat{g}_b = \left[ \begin{array}{c} 1 \\ h_j \end{array} \right].
\]

When we use ML estimate algorithm to generate the estimate covariance matrix \( \hat{R}_{ml} \), it is positive semi-definite by construction [25]. Since \( \hat{d}_b > 0 \) holds a probability equal to one, \( \hat{R}_{ml} \) is positive definite with probability one. Therefore, considering that ad-hoc estimate method could be indefinite, the ML estimate algorithm has a certain advantage. In spite of the two algorithms introduced above, some other interpolation-based algorithms, such as genetic algorithm (GA) approach and Inverse Free Krylov Subspace algorithms (IFKSA) based techniques, are proposed in [26, 27, 73].

Besides the interpolation-based approach, another widely applied method is neural-network-based approach. Since 1980s, artificial neural network (ANN) has become a hotspot of intelligence research. It abstracts the human cranial nerve into a signal process
The neural network is a computation model composed of a large sum of interconnected artificial neurons. As shown in Figure 2.5, each neural node has a certain output function, which is called as activation function. The connection between two nodes contains a weighting term to the transmitted signal. The output of the network rely on the selected connection modes, weighting terms and activation functions. The ANN itself is an approximation to an algorithm or a function in nature and it might be an express of a kind of logical strategies.

Figure 2.5 A typical artificial neural network.

The ANN-based algorithms have been developed for real-time DOA estimation applications. A comprehensive summary of neural network based methods for array processing is presented in [51]. In array processing, the most outstanding advantage of the neural network based algorithms is that they have a better performance on computational speed and accuracy than the conventional linear algebra based methods [28]. In addition,
ANN based method has some ability in learning and is adaptive to the changing signal environment. Therefore, it is feasible for array sensor failure in array processing.

As an illustration, a typical ANN based algorithm from [28] is introduced as following. Similar to interpolation-based method, we firstly assume a far-field narrowband data model. According to equation (45), the received signal of an \( N \)-element array is expressed as \( \mathbf{x}(t) = \mathbf{A}(\theta)\mathbf{s}(t) + \mathbf{n}(t) \). From equation (48), the covariance matrix is given by \( \mathbf{R} = E[\mathbf{x}(t)\mathbf{x}^H(t)] = \mathbf{A}\mathbf{S}\mathbf{A}^H + \sigma^2 \mathbf{I} \). At this point, with a given snapshots number \( T \), we use the array correlation matrix to approximate the covariance matrix, which is given by

\[
\mathbf{R} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}(t)\mathbf{x}^H(t).
\] (58)

When sensor failures occur in the array, it is assumed that there is no signal comes from the sensors under failure but only the noise. If a certain sensor \( i \) encounters a failure just after a certain snapshot \( p \) and it does not recover from the failure in the last \( T - p \) snapshots. With this, the array signal after the failure is give by

\[
\mathbf{x}_f(t) = [x_1(t), x_2(t), \ldots, x_{i-1}(t), n_i(t), x_{i+1}(t), \ldots, x_N(t)]^T,
\] (59)

where \( p < t \leq T \). The array correlation matrix under failure can be computed as

\[
\mathbf{R}_f(t) = \frac{1}{T} [\sum_{t=1}^{p}[\mathbf{x}(t)\mathbf{x}^H(t)] + \sum_{t=p+1}^{T}[\mathbf{x}_f(t)\mathbf{x}_f^H(t)]].
\] (60)

In above equation, the two terms respectively represent the correlation before and after the array element failure.

In all conventional DOA estimation algorithms from [25-27] that can overcome the failures to a certain extent, the missing data term \( n_i(t) \) is replaced by an interpolated data
from the rest available data. For ANN based algorithms, the network should be firstly trained with a failure-free training data set. Different from above conventional algorithms, ANN based methods do not require to know when and where the failure occurred as it automatically generalized the missing data. In order to automatically select the network size, a minimum resource allocation network (MRAN) [30] is applied, which is a sequential learning algorithm for minimum radial basis function (RBF) neural network. The output of an RBF network with \( h \) hidden neurons can be expressed as

\[
    f(x_k) = \alpha_0 + \sum_{n=1}^{k} \alpha_n \phi_n(x_k),
\]

where \( \phi_n(x_k) \) is the response of the \( n \)th hidden neuron to the input \( x_k \), \( \alpha_n \) is the weight parameter connecting the \( n \)th neuron to its output and \( \alpha_0 \) is the bias parameter. The response function \( \phi_n(x_k) \) is a Gaussian function, which is given by

\[
    \phi_n(x_k) = \exp \left( -\frac{1}{(\sigma_n)^2} \|x_k - \mu_n\|^2 \right),
\]

where \( \mu_n \) is the center and \( \sigma_n \) is the width of the Gaussian function. At the beginning, the MRAN has no hidden units. A brief outline of the built-up process based on the training data \( (x_k, d_k) \) of an MRAN is introduced as following steps.

1. Use the array signal \( x_k \) to compute the network output \( y_k \).
2. When the following conditions are satisfied, a new RBF center can be created.
   i. The error \( \|y_k - d_k\| \) exceeds a threshold value.
   ii. The mean square root of the above error computed over a window exceeds a threshold value.
   iii. The new input is far enough from the centers of the existing neurons.
(3) If a normalized hidden neuron’s contribution to the output with consecutive inputs is below a certain threshold, this hidden neuron should be pruned.

(4) Adjust the centers, widths and weights of the network using an extended Kalman filter (EKF).

(5) Increase $k$ to $k + 1$ and go back to step (1).

To generate training data sets, concrete to DOA estimation applications in array processing, the entries of the array correlation matrix $\mathbf{R}$ are selected as the training data set instead of the array signal $\mathbf{x}(t)$. This handling removes redundant or irrelevant information and reduces the size of the network, which requires less signal parameter space [28].

As mentioned in [52], the training data set can be generated by complex elements in the correlation matrix $\mathbf{R}$. However, the shortcoming of this method is that it gives a considerably high error when the DOAs of the sources are near $\pm 90^\circ$ within a scanning sector $[-90^\circ, +90^\circ]$. To overcome this shortcoming, a summary of improved pre-processing scheme, which adopts both its magnitude and phase angle, is introduced as following [28].

1. Obtain $m$ snapshots of the array output vector.
2. Estimate the correlation matrix $\mathbf{R}$.
3. Extract the lower triangular half of $\hat{\mathbf{R}}$.
4. Appending the columns of $\hat{\mathbf{R}}$ into a single columns $\mathbf{b}$ in order of their column number.
(5) Form the training vector \( t \) whose elements are the magnitude and phase of the corresponding elements in \( b \).

(6) By generating above training vector \( t \) corresponding to different DOAs, the training data set can be formed of such vectors, which are presented in random order.

When signal sources locate in a sufficiently small range, magnitude values of the array covariance matrix \( R \) is approximately constant. Therefore, under this circumstance, we can obtain the training data only by using the phase angles in \( R \).
3.1 Introduction of Matrix Completion Theory

To illustrate matrix completion theory, we can first examine several interesting practical problems. For example, suppose \( m \) audiences watched and graded some of \( n \) movies from a movie rental system. Then the audiences and the movie grades form a two-dimensional data set \((i,j) \in E \subseteq [n_1] \times [n_2]\), which means for any set of \((i,j) \in E\), there is a grade \( M_{ij} \in \mathbb{R} \). The movie rental company wants to recommend movies to customers for gaining more profits by recovering and estimating the missing grades of all \( n \) movies. This is the famous Netflix problem [53]. For another, suppose there is a low-power and wireless sensor network that is randomly distributed in a particular area. Assume each sensor can only generalize the distance estimations based on the signal intensity of its nearest sensor. By these noised distance estimations, we can form a distance matrix that has only part of the entries. Then we can try to estimate the correct matrix. If the sensors are in a plane, the rank of the matrix will be two; if the sensors distribute in a three-dimensional space, the rank of the matrix will be three [54, 55]. In this situation, we only need to know the partial distance of each node, which contains enough information to rebuild the object.

Although the descriptions of the above two examples are very brief, we can still find certain commonalities between them, and they are also the general questions that we need to answer. First, under some conditions, the known data can provide enough
information to estimate the unknown data. Second, this estimation can be effectively processed.

In recent years, for the first question, researchers assume the original data matrix is low-rank or approximately low-rank. This assumption is reasonable on various occasions, for example recommender systems, because in such matrices, there will be coherence among some rows or columns. Therefore, on the basis of actual situation and compressive sensing theory [56,57], Candès and Recht [31] proposed an optimization problem that can be presented as:

\[
\text{minimize } \text{rank}(X) \quad \text{subject to } \quad X_{ij} = M_{ij} \quad (i,j) \in \Omega, \quad (63)
\]

where $\Omega$ is a set of observed entries in matrix $X$, which means if $M_{ij}$ is observed, there will be $(i,j) \in \Omega$.

Unfortunately, the optimization problem in equation (63) is NP-hard [60] and all existing methods that can provide accurate optimizing results will take an exponential amount of time. In compressive sensing, researchers use convex optimization to solve vector problems and, thereby, we adopt a convex approximation to transform the optimization problem in equation (63) into

\[
\text{minimize } \|X\|_* \quad \text{subject to } \quad X_{ij} = M_{ij} \quad (i,j) \in \Omega, \quad (64)
\]

where $\|X\|_* = \sum_{j=1}^{n} \sigma_j(X)$ and $\sigma_j(X)$ is the $j^{th}$ largest singular value of $X$. Meanwhile they have proved that there exists a positive integer constant $C$, and when the number of observed entries $m$ satisfy

\[
m \geq Cn^{1.2-r} \log n, \quad (65)
\]
the low-rank data matrix can be successfully completed in a high probability. In inequality (65), \( n \) is the maximum value of matrix dimensions and \( r \) is the rank of the matrix. After this, a few improvements on the lower bound of \( m \) were proposed in [32, 58, 59].

When it comes to the matrix completion optimization algorithms, there are abundant research results. Fazel [60, 61] proposed an interior-point algorithm to solve the optimization problem in equation (63) by semi-definite programs. However these interior-point based programs are not suitable for the solution of large-scale matrix completion. Toh and Yun [62] introduced an accelerated proximal gradient (APG) algorithm with an iteration complexity \( O\left(\frac{1}{\sqrt{\varepsilon}}\right) \). To enhance the convergence rate, they adopt linesearch-like technique, continuation technique and truncation technique. Ma, Goldfarb and Chen [63] applied approximate singular value decomposition (SVD) technology to the fixed-point continuation (FPC). They also proposed the fixed-point continuation with approximate SVD (FPCA). At almost the same time, Cai, Candès and Shen [64] proposed the singular value thresholding (SVT) algorithm, which keeps the sparse character of the matrix to be recovered. This optimizing algorithm largely reduces the memory requirements and can effectively treat large-scale matrix completion problems. However, the selection of threshold values is still an ongoing topic of research. After this, Lin, Chen and Ma [65] proposed the augmented Lagrange multiplier (ALM) algorithm and used numerical experiments to show that ALM has a better convergence rate than SVT and APG. Besides that, other research is developed for matrix completion with noise [66], fast singular value thresholding without singular value decomposition [67] and accelerated singular value thresholding for matrix completion [68].
3.2 Feasibility of Matrix Completion Theory

It’s important to note that matrix completion theory is not feasible for all situations. First, a heavily disorganized matrix cannot be reconstructed. As shown in equation (65), to successfully complete a size-fixed matrix, there will be a lower bound for the number of observed entries $m$ and an upper bound for the matrix rank $r$. Besides that, matrix completion theory also requires the matrix that we try to recover to follow strong incoherence property [66]. For a given matrix $M_{n_1 \times n_2}$, the singular value decomposition is given by

$$M_{n_1 \times n_2} = \sum_{1 \leq k \leq r} \sigma_k u_k v_k^* = U \Sigma V,$$  \tag{66}

where $\Sigma = diag(\sigma_1, \cdots, \sigma_r)$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. And if there are two positive number $\mu_1, \mu_2$ and matrix $U$, $\Sigma$ and $V$ satisfy:

$$\left\{ \begin{array}{l}
|\langle e_i, \tilde{P} e_i \rangle - \frac{r}{n_1} 1_{i = i'}| \leq \mu_1 \frac{\sqrt{r}}{n_1} , \\
|\langle e_j, \tilde{P} e_j \rangle - \frac{r}{n_2} 1_{j = j'}| \leq \mu_2 \frac{\sqrt{r}}{n_2} ,
\end{array} \right. \forall (i, i') \in [n_1] \times [n_2], \forall (j, j') \in [n_1] \times [n_2], \tag{67}$$

and

$$|E_{i,j}| \leq \mu_2 \frac{\sqrt{r}}{\sqrt{n_1 n_2}}, \forall (i, j) \in [n_1] \times [n_2], \tag{68}$$

where $E_{i,j}$ is an entry of matrix $E = \sum_{i \in [r]} u_i v_i^*$ with a coordinate $(i, j)$, we will say matrix $M$ follows strong incoherence property. Specific to the matrix completion problem, following strong incoherence property means when most of the entries are absent in a row or column, the matrix will be hard to complete and when a row or column is entirely missing, the matrix cannot be recovered by matrix completion theory. Therefore, in our matrix completion model, the missing entries should be randomly distributed in the matrix.
To further explore the feasibility of matrix completion theory, we conduct a few Monte Carlo simulation tests to look for the relationship between the feasibility and matrix parameters. Suppose we have a matrix $X_{n_1 \times n_2}$ with $N$ entries and $n = \max(n_1, n_2)$. Now we define two parameters: the proportion of absent entries $\rho_s$ and the relative value of the rank $\rho_r$, which is given by

$$\rho_s = \frac{D}{N}, \quad \rho_r = \frac{r}{n},$$

(69)

where $D$ is the number of absent entries and $r$ is the rank of the matrix. In each round of our Monte Carlo simulations, we will randomly set $D$ entries to 0 and obtain a matrix $M$ with missing data. Then we use SVT algorithm to complete the matrix. Finally we will compute the error rate, which is given by

$$ER = \frac{\|X - M\|_F}{\|M\|_F},$$

(70)

where $\|\cdot\|_F$ is the Frobenius norm. In our numerical simulation, we would declare matrix $M$ was successfully completed if the error rate is less than 0.02. The simulation results are shown in Figure 3.1. The white zone means the matrix was successfully recovered in all Monte Carlo tests. The black zone means the completion process failed in all simulations. Gray points represent that the matrix could be recovered with a probability between 0 and 1. For each set of $\rho_s$ and $\rho_r$, 50 repetitions of Monte tests were made.
Figure 3.1 The feasibility of matrix completion.

From Figure 3.1 we find that the matrix will be successfully recovered in a high probability if both $\rho_s$ and $\rho_r$ are less than their upper bounds. That means with a size-fixed matrix, we need the number of observed entries in the matrix greater than a certain value and the matrix rank less than a certain value.

3.3 Simple Reshaping Method

Considering the inner relationship among the array figuration, the arrival signals have the features of sparsity in spatial domain, which means some information in the data matrix is redundant. Therefore, within certain limits, when some entries are absent in the data matrix, we can use the remaining entries to estimate those missing ones. It seems like a compressive sensing process in spatial domain and it’s also our basic idea to recover the data matrix with element failures.
Suppose we have an N-element array and the sampling time range is from \( t_0 \) to \( t_L \).

At time step \( t_s \), the \( i^{th} \) element stops working. We assume that when an element stops working, the received in the data matrix will be 0 and the data matrix is given by

\[
M(t) = \begin{bmatrix}
x_1(t_0) & x_1(t_1) & \ldots & x_1(t_{s-1}) & x_1(t_s) & \ldots & x_1(t_{L-1}) & x_1(t_L) \\
x_2(t_0) & x_2(t_1) & \ldots & x_2(t_{s-1}) & x_2(t_s) & \ldots & x_2(t_{L-1}) & x_2(t_L) \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
x_i(t_0) & x_i(t_1) & \ldots & x_i(t_{s-1}) & 0 & \ldots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
x_N(t_0) & x_N(t_1) & \ldots & x_N(t_{s-1}) & x_N(t_s) & \ldots & x_N(t_{L-1}) & x_N(t_L)
\end{bmatrix}
\] (71)

From (71) we can see that the absent entries are distributed on the same row that corresponds to when the element stops working at a certain time. When \( t_s \) is very small or even equals to \( t_0 \), the matrix will not follow the strong incoherence property. Besides that, the matrix \( M(t) \) is not a low-rank or approximately low-rank matrix. At this point, this received array signal matrix seems impossible to be recovered by matrix completion algorithms. Therefore, in [74], researchers assume the absent entries are randomly distributed in the matrix, which makes the problem easy to solve but not practical to apply.

Fortunately, a simple reshaping method [69] was found to overcome these barriers.

In Figure 3.2, an example shows us that the second element stopped working since the first time step \( t_0 \) and the received signal matrix is unable to be recovered by matrix
completion algorithms. By this time, we first select a certain time step and take out the signal vector under this snapshot as our recovering target in this round. Then we reshape the signal vector into a matrix. We orderly take out the entries from the vector, from top to bottom, to form the rows of the new matrix. For each row, we take out a sequence of entries with a length $p$ and all sequences have no overlapping with each other. That finally provides us a new matrix with $p$ columns and $q$ rows, which requires $N = p \times q$.

With this reshaping process, the absent entries will be reallocated in the matrix. Although the problem that the matrix might not follow strong incoherence property still remains, the entries would not be distributed in same rows or columns in a high probability, especially when the number of absent entries is very small.

From (14) and (15) we can see that in a uniform linear array, when there is only one input signal, the received signals from array elements form a geometric sequence and when there are $K$ input signals, the received signals form a sequence that is summed by $K$ geometric sequence. If we apply the above reshape method to this matrix and the matrix rank $K$ is much less than the minimum dimension of the new matrix, this new matrix at a certain time step will be a low-rank matrix or an approximately low-rank matrix and the matrix rank will be $K$.

At this point, the new matrix after the reshape processing can satisfy the condition of applying matrix completion algorithm and now the missing data would be recovered. Considered the feasibility of matrix completion algorithm, we prefer smaller $\rho_s$ and $\rho_r$. Since the number of entries $N$ and number of absent entries $D$ cannot be changed, the proportion of absent entries $\rho_s = \frac{D}{N}$ cannot be changed either. With a fixed signal
environment, the matrix rank $r$ is also fixed. Therefore, to obtain a smaller $\rho_r = \frac{r}{n}$, we need to make the maximum dimension of the matrix as small as possible. By this point, during our reshaping process, we set $n$ to be the smallest divisor of $N$ that is not less than $\sqrt{N}$.

However, this simple reshape method has some limits. First it requires the element number is reducible. If it’s not reducible, we may need to add zero point in the new matrix and that will increase the matrix rank, which is a negative effect for matrix completion algorithms. To improve this method, a new improved overlapping is developed which is able to tackle such disadvantage and improve the success rate of applying matrix completion algorithms at the same time.

3.4 Proposed Improved Overlapping Reshaping Method

As we discussed before, though the simple reshape method allows us to use matrix completion algorithms to recover the array signal matrix with failures, it has a major disadvantage. Therefore, we developed an overlapping reshape method that not only can overcome the shortcoming but also improve the performance of the applied matrix completion algorithms.

In overlapping reshaping method, we still use same-length sequences of entries from the signal vector $X(t_i)$ to form the new matrix. However, different from the simple reshaping method, there is an overlapping portion $o$ between every two adjacent sequences. With this overlapping portion, the dimensions of the matrix are increased to larger values $\bar{n}_1$ and $\bar{n}_2$. Therefore, we obtain larger value of $\bar{n} = \max(\bar{n}_1 \times \bar{n}_2)$ the relative value of the rank $\bar{\rho}_r = \frac{r}{\bar{n}}$ becomes smaller. But when it comes to the proportion of
absent entries, the situation becomes more complicate. Although the number of matrix entries $\tilde{N} = \tilde{n}_1 \times \tilde{n}_2$ increased, the number of absent entries $\tilde{D}$ may also be increased, which makes it difficult to judge whether the proportion of absent entries $\rho_s$ is increased or decreased.

\begin{align*}
&\begin{bmatrix}
  x_1(t)
  x_2(t) = 0 \\
  \vdots \\
  x_{p-w}(t) \\
  \vdots \\
  x_{p-w+1}(t) \\
  \vdots \\
  x_{N-p+1}(t) \\
  \vdots \\
  x_N(t)
\end{bmatrix} \\
&\begin{bmatrix}
  x_{1-2}(t) \\
  \vdots \\
  x_{(p-2)(p-1)+1}(t) \\
  \vdots \\
  x_{N-(p-1)(p-1)+1}(t) \\
  \vdots \\
  x_{N-(p-1)(p-1)+2}(t) \\
  \vdots \\
  x_N(t)
\end{bmatrix}
\end{align*}

Figure 3.3 The overlapping reshaping method.

The new number of absent entries is related with the value of overlapping portion and also has certain randomness. From Figure 3.3 we can see that if an absent entry is distributed in the overlapping portion, it will be copied in more than one row in the reshaped matrix. Since the absent value is randomly generated, the value of the overlapping portion will determine the probability of copying the absent entries. If the overlapping portion is more than half of the length of the sequence $p$, the absent entry in the overlapping portion may not be just doubled in the new matrix. It could be tripled, quadrupled or even more. Due to this reason, we need to limit the value of overlapping portion $\phi$ to avoid a substantial growth of it. Usually, we should limit the overlapping value to the half of the sequence length $p$.  

43
Now, here is an example to explain why we need to limit the overlapping portion. Suppose for the same signal matrix, the simple reshaping method generates an $n \times n$ matrix for a certain time step with $D$ absent entries. The proportion of absent entries is given by

$$\rho_s = \frac{D}{N} = \frac{D}{n^2}. \quad (72)$$

For the overlapping reshaping method, $\tilde{n} = n + o$ and there are $\tilde{D}$ absent entries. Under this circumstance, the new proportion of absent entries is given by

$$\tilde{\rho}_s = \frac{\tilde{D}}{\tilde{N}} = \frac{\tilde{D}}{(n+o)^2}. \quad (73)$$

When $o = n/2$ and $\tilde{D} = 2D$, the equation (73) can be reformulated as

$$\tilde{\rho}_s = \frac{\tilde{D}}{(n+o)^2} = \frac{2D}{2.25n^2} = \frac{8}{9} \rho_s. \quad (74)$$

That means though the number of absent entries achieve the maximum value, the new proportion of absent entries can still be decreased by this overlapping method. From Figure 3.1, we can see that when the relative value of the rank $\rho_r$ is very small, the acceptable range of $\rho_s$ become very large. Since the $\rho_r$ is surely with decreased overlapping method, even an increase of $\rho_s$ that is not too huge could be acceptable. Besides that, considering the computational cost, we may use a smaller overlapping portion to save the space complexity.

Generally speaking, the reshaping method is a bridge over array signal recovery and matrix completion theory. Therefore, its complexity mainly depends on the complexity of the selected matrix completion algorithm. It uses the redundant information of the array signal matrix to estimate the missing data. The proposed overlapping reshaping method
creates more redundant information based on the simple reshaping method, which can obtain better estimation performance. To more intuitively show the feasibility and advantage of this method, we present series of simulation tests results with analysis in the next chapter.
Chapter 4

SIMULATION RESULTS

4.1 Success Rate Analysis

Since we noticed that the matrix completion algorithms are not always feasible, in this section, we are going to analyze the success rate of our matrix completion based approach.

We first generated several narrowband input signals with different amplitudes and frequencies from different directions within an angle range \([-90^\circ, 90^\circ]\). Then three ULAs with 36, 64 and 100 elements were used for receiving the input signals. For an array with \(N\) elements, we randomly selected \(D\) elements and set their received signals to zero. The observing rate is defined as

\[
\rho_o = \frac{(N-D)}{N}.
\]  

(75)

We used \(X_{des}\) to represent the reshaped matrix without any absent entries and used \(M\) to represent the reshaped matrix recovered from missing data case. We tried both the simple reshaping method and the proposed overlapping reshaping method. Then the basic singular value thresholding (SVT) algorithm is applied as our matrix completion algorithm. We would declare the matrix \(M\) is successfully completed if the error rate satisfies

\[
ER = \frac{\|X_{des}-M\|_F}{\|M\|_F} < 0.01,
\]

(76)

where \(\|\cdot\|_F\) is the Frobenius norm. To estimate the success rate, we applied \(T\) rounds of Monte Carlo tests for each sampled observing rate value. We randomly selected a certain number of failed elements for each round, while the array size and signal environment were...
fixed. If there are $T_s$ rounds of simulations are resulted in a success completion, the success rate will be given by

$$r_s = \frac{T_s}{T}.$$  \hfill (77)

In Figure 4.1, we used a 36-element ULA to receive 2 narrowband input signals. In the simple reshaping method, we got a $6 \times 6$ reshaped matrix. In the overlapping reshaping method, we set the overlapping portion $\sigma = 2$ and the size of the reshaped matrix became $7 \times 7$. 200 rounds of Monte Carlo tests are processed.

![Figure 4.1 Success rate of a 36-element ULA with 2 narrowband input signals.](image)

In Figure 4.2, we used a 64-element ULA to receive 2 narrowband input signals. In the simple reshaping method, we got an $8 \times 8$ reshaped matrix. In the overlapping
reshaping method, we set the overlapping portion $o = 2$ and the size of the reshaped matrix became $9 \times 9$. 200 rounds of Monte Carlo tests are processed.

Figure 4.2 Success rate of a 64-element ULA with 2 narrowband input signals.

In Figure 4.3, we used a 100-element ULA to receive 2 narrowband input signals. In the simple reshaping method, we got a $10 \times 10$ reshaped matrix. In the overlapping reshaping method, we set the overlapping portion $o = 4$ and the size of the reshaped matrix became $12 \times 12$. 200 rounds of Monte Carlo tests are processed.
Figure 4.3 Success rate of a 100-element ULA with 2 narrowband input signals.

In Figure 4.4, we still used the 100-element ULA but generated 4 narrowband input signals. Similar to the reshaping process for the previous group of simulation, we got a $10 \times 10$ reshaped matrix in the simple reshaping method. In the overlapping reshaping method, we set the overlapping portion $o = 4$ and the size of the reshaped matrix became $12 \times 12$. 200 rounds of Monte Carlo tests are processed.
From any of these four figures, we can see that the success rate grows with the observing rate. This meets the feasibility analysis that we made in the previous chapter. By comparing the results shown in Figure 4.1 and Figure 4.2, we can find that a larger size of array can provide larger range of observing rate with a success rate close to 1. That is because under a same signal environment, a larger size array can provide more redundant information, which means it will be more robust in a missing data case.

From both Figure 4.1 and 4.2, the performance of the overlapping reshaping method, the blue curve, is obviously better than the simple reshaping method, which is represented by the red curve. However, in Figure 4.3, the performance of two reshaping
method is very close to each other. The overlapping reshaping method is no longer privilege in this case and strictly speaking, its performance is even worse than the performance of the simple reshaping method. This is because if the size of array grows larger, the dimensions of the reshaped matrix will also grow larger. Therefore, the increasing effect of reshaped matrix size by the overlapping reshaping method is diluted. For example, in the simulation shown in Figure 4.3, a $12 \times 12$ will have very limit advantage over a $10 \times 10$ matrix with 2 narrowband input signals. Because the relative values of the rank of those reshaping methods are very similar, which is given by

$$\rho_r = \frac{2}{10} = \frac{1}{5} \approx \tilde{\rho}_r = \frac{2}{12} = \frac{1}{6}. \quad (78)$$

Since the overlapping reshaping method may increase the number of absent entries, when the increasing effect is very limit, the overlapping process would exert negative influence and that is why there is some range that the simple reshaping method provides better performance than the overlapping reshaping method does.

With the same 100-element array, if we make the signal environment more complicated, just like we did in the simulation shown in Figure 4.4, the advantage of the overlapping reshaping method will appear again. In that group of simulations, we increased the number of input signals form 2 to 4, which equals to rank of the reshaped matrix. That means the numerator of the relative value of the rank $\rho_r$ became larger. By this point, the increasing effect of the reshaped matrix size became important again, which makes the overlapping reshaping method to gain obvious advantage over the simple reshaping method.
4.2 Application: Spatial Filtering

Figure 4.5 Spatial filtering (a): beam pattern of a 64-element ULA with 32% of the elements out of work and DOA=10.

From the review of content about spatial filtering provided in chapter 2, we know that the main goal of spatial filtering is to enhance or restrain the signal power in certain spatial regions. To achieve this goal, we usually add different weighting terms to each array elements and obtain the output by summing those weighted signals. Therefore, in missing data case, if a received signal turns to zero, corresponding weighted signal will be zero, which equals to set the corresponding weighting term to zero. That will undermine the original beam pattern and make negative effect on the output performance. An example is
shown in Figure 4.5. We applied the minimum mean square error (MMSE) as our beamforming criterion, and the optimal weighing vector is given by,

\[ W_{opt} = (R_x)^{-1}r_{xd}, \]  

where \( R_x = E[x(t)x(t)^H] \) and \( r_{xd} = E[x(t)d^*(t)] \). \( d(t) \) is a desired reference output signal.

The traditional non-adaptive algorithms are helpless in the face of the missing data case. For those algorithms, people have to generate the weighting vector only with the remaining elements. However, with fewer elements, the output performance will be impaired, which is mainly shown by higher side-lobe and wider main-lobe. For adaptive approaches, no matter what kind of iterative algorithm is used, a recalculation process for the weighting vector is indispensable. Though those algorithms have some ways, such as using reference terms, to recover the distorted pattern to its original state as much as they can, the distortion is still inevitable. However, by the matrix completion signal recovery method, there will be no need to recalculate the weighting term, because the signals matrix with missing data will be directly completed to its original state.

Several groups of simulations with different SNRs are made to show the performance of matrix completion signal recovery method. Both the simple reshaping method and the overlapping reshaping method were applied. Same as the simulation shown in Figure 4.5, we used a 64-element ULA with 32% of elements out of work and the DOA equals to 10°. The SVT algorithm was selected as the matrix completion algorithm.
Figure 4.6 Spatial filtering (b): beam pattern of a 64-element ULA with 32% of the elements out of work and $SNR = 10\, \text{dB}$.

In Figure 4.6, either the simple reshaping method or the overlapping reshaping method could reduce the side-lobe level and form the main-lobe in $10^\circ$. However, both of them were unable to provide us a pattern that is similar enough to the desired beam pattern. That is because the SNR is too low for the SVT algorithm to successfully complete the matrix. Actually speaking, the sensitivity to the noise of our proposed method mainly depends on the selected matrix completion algorithm.
Figure 4.7 Spatial filtering (c): beam pattern of a 64-element ULA with 32% of the elements out of work and $SNR = 20$ dB.

Then after we increased the SNR from 10 dB to 20 dB, the performances of those two reshaping methods were significantly improved. From Figure 4.7 we can see that both of them could generate a pattern that is very close to the original one. However we should also notice that the overlapping reshaping method made no advantage over the simple reshaping method. This is because in this simulation, only one input signal was introduced. That means the rank of the reshaped matrix equals to one, and if we generate an $8 \times 8$ reshaped matrix by the simple reshaping method and a $9 \times 9$ reshaped matrix by the overlapping reshaping method, their relative values of the rank $\rho_r$ will be approximately
equal. As we mentioned in the previous section, by this point, the matrix size increasing effect become very limit, which makes the overlapping reshaping method have no advantage of the simple reshaping method.

![Spatial Filtering with MMSE Criterion](image)

Figure 4.8 Spatial filtering (d): Beam pattern of a 64-element ULA with 32% of the elements out of work and $SNR = 30$dB.

As we kept increasing the SNR to 30dB, the performances of the both two reshaping methods are acceptable. With this noise power level, the matrix completion signal recovery method can generate a beam pattern that is greatly close to the desired pattern.
Figure 4.9 Spatial filtering (e): beam pattern of a 64-element ULA with 32% of the elements out of work and $SNR = 40$ dB.

In Figure 4.9, when we finally increased the SNR to 40dB, the recovered beam patterns were almost coincided with the desired pattern, which provides a nearly perfect performance.

### 4.3 Application: DOA Estimation

Besides spatial filtering, we also applied our proposed matrix completion signal recovery method on DOA estimation. By this point, we used the famous MUSIC algorithm as our DOA estimation algorithm. In the missing data case, the orthogonality between signal subspace and noise subspace will be reduced, which leads to a worse performance.
In our simulations, we used a 36-element ULA with 30% of the elements were out of work. When it comes to the signal environment, we tried the worst situation. Two strong coherence narrowband input signals were introduced and their incident angles were very close. Both the simple reshaping method and the overlapping reshaping method were applied with different noise power level. The SVT algorithm was selected as our matrix completion algorithm.

Figure 4.10 DOA estimation (a): beam pattern of a 36-element ULA with 30% of the elements out of work and $SNR = 5$dB.

We started the simulation with a high noise power level. From Figure 4.10 we can see that both two reshaping methods can effectively improve the performance. However,
neither of them could generate a pattern that is similar to the desired one. Besides that, in this situation, our proposed overlapping reshaping method seems to have no advantage over the simple reshaping method. That is because with an SNR equals to 5dB, it is very hard for the SVT algorithm to successfully recovered missing entries in the reshaped matrix, either the simply reshaped one or the one provided by the overlapping reshaping method. Actually, with such high noise power level, the matrix size increasing effect in the overlapping reshaping method becomes pointless. Compared to the simple reshaping method, its performance could be better, could be worse, or could be nearly the same just like what is shown in Figure 4.10.

Then as we increased the SNR to 20dB, the performances from the recovery methods were also improved. In Figure 4.11, the side-lobe of the patterns generated by the recovered signals were at a new low. We are happy to see that the overlapping reshaping method started to show its advantage over the simple reshaping method. That is because since the noise power went down, the SVT algorithm could start to complete the absent entries in the reshaped matrices effectively.

However, with different observed entries, the difference between the performances of those two reshaping methods may be different. Therefore, with a certain observing rate, we can make success rate analysis like we did in the section 4.1 and find out whether the overlapping reshaping holds a better probability to successfully recover the signal. Specific to our simulation, the observing rate $\rho_o = 70\%$ may not be a value that grants the overlapping reshaping method a considerable advantage over the simple reshaping method.
But with a certain set of absent entries in different noise situations, the advantage will remain if it is established at the beginning.

Figure 4.11 DOA estimation (b): beam pattern of a 36-element ULA with 30% of the elements out of work and $SNR = 20\text{dB}$.

At this point, we can image that if we keep increasing the SNR, we will obtain better signal recovery performance. And from Figure 4.12 we can see that the side-lobe was further reduced, which was more and more close to the desired one. Under this noise power level, the SVT algorithm essentially fully implemented the role in the matrix completion signal recovery method.
Figure 4.12 DOA estimation (c): beam pattern of a 36-element ULA with 30% of the elements out of work and $SNR = 30$dB.
Chapter 5
CONCLUSION AND FUTURE WORK

5.1 Summary

In this paper, a matrix completion signal recovery method for array processing is proposed. We first made a comprehensive literature review, including the history of array processing, the mathematical model of array processing, classical and popular array processing algorithms. Besides that, to explain the matrix completion theory, we also presented a brief introduction of matrix completion principles and a basic mathematical model of its optimization problems and applying conditions.

To explain our motivation, we provided an extraordinarily detailed picture of what is the missing data case in array processing. We described why this case would occur and what kind of influence would be generated when it occurred in the array. Before the introduction of our proposed method, several existing algorithms are introduced as comparisons. We gave a brief study of their logic thoughts, mathematical derivation, applicable conditions, advantages and disadvantages.

Different from the existing algorithms, our proposed method does not rely on any prior knowledge such as when and where the failure occurs or a training dataset without any missing data. Actually, the matrix completion signal recovery method utilizes the inner relationship among the array figuration. That means we required the array elements to be regularly distributed and in our simulation tests, we adopted ULA as our selected figuration. When it comes to the core, the matrix reshaping method, it is essentially a bridge over
matrix completion algorithms and missing data case in array processing. As we have elaborated in the previous chapters, there are a few barriers between matrix completion theory and missing data case in array processing. It is the matrix reshaping method that helps us break through those barriers and apply various matrix completion algorithms to recovering the missing data in array processing. In addition, our proposed overlapping reshaping method not only can be used for an array whose element number is not reducible, but also improve the success rate of the applied matrix completion algorithms.

Based on the structure of our method introduced above, the matrix reshaping process needs to be used for all snapshots. Therefore, the complexity of our method depends on the number of snapshots and the complexity of the selected matrix completion algorithm. As a tradeoff, when we applied the overlapping reshaping method, there is a matrix size increasing effect, which may cost more computation in each round of matrix completion. Considering of the randomness in absent entries distribution, we would better made success rate analysis for the specific signal environment before we decide the overlapping value.

5.2 Future Work

As we mentioned above, our proposed method needs the matrix figuration to be regular and we used several ULA in our simulation test. ULA is the most basic array model but is not very practical. Therefore, to make our method have more practical significance, we should study for more kinds array figurations, such as uniform circular array (UCA), uniform rectangular array (URA) or even some conformal arrays. Since our reshaping method can cooperate with various matrix completion algorithms, we also need to test
different algorithms under different circumstances and give a conclusion about what we should apply for different applications or in different situations. In this paper, some conclusion were draw from some qualitative analysis and simulation results, more detailed and sophisticated mathematical derivation may need to be provided in the future.
REFERENCES


