TARGET DETECTION IN ULTRA-WIDEBAND NOISE RADAR SYSTEMS

A Dissertation in

Electrical Engineering

by

Yangsoo Kwon

© 2012 Yangsoo Kwon

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

December 2012
The dissertation of Yangsoo Kwon was reviewed and approved* by the following:

Ram M. Narayanan
Professor of Electrical Engineering
Dissertation Advisor, Chair of Committee

William E. Higgins
Distinguished Professor of Electrical Engineering

Vishal Monga
Assistant Professor of Electrical Engineering

Mark Levi
Professor of Mathematics

Kultegin Aydin
Professor of Electrical Engineering
Head of the Department of Electrical Engineering

*Signatures are on file in the Graduate School.
Ultra-wideband (UWB) noise radar has been widely considered as a promising technique for covert high-resolution detection of multiple targets due to several advantages such as excellent immunity from jamming and interference, low probability of detection and interception, and relatively simple hardware architectures. The most important advantages of noise radar is immunity of interception by an adversary, since the transmitted noise-waveform is constantly varying and never repeats exactly. In addition, UWB noise radar obtains high-resolution detection of multiple targets due to its high instantaneous bandwidth. Based to the above advantages, UWB noise radar systems are garnering more and more attention recently. In this dissertation, two issues concerning detection and estimation based on information theory and compressive sensing applied to UWB noise radar are investigated in detail.

This dissertation explores a target detection method using the total correlation formalism based on information theory which enables the detection of multiple targets at
intermediate and low signal-to-noise ratio (SNR) regimes. This approach uses the largest eigenvalue of the sample covariance matrix to extract information from the transmitted signal replica, and outperforms the conventional total correlation detector when reflected signals have intermediate or low SNR values. Additionally, in order to avoid ambiguous target occurrence, an adaptive threshold is proposed, which guarantees the detection performance with the same receiving antenna elements for a given false alarm probability. The threshold is computed from the largest and smallest eigenvalue distributions based on random matrix theory. Numerical simulations show this detection method can be used for a wide range of SNR environments, and the threshold provides definitive target detection.

This dissertation also explores an application of compressive sensing for multiple-input multiple-output (MIMO) UWB noise radar imaging. Two schemes to improve the system performance of a sample selection and an adaptive weighting allocation are investigated. The sample selection is based on comparing the norm values of candidates among the received signal, and selecting the largest $M$ samples among $N$ per antenna to obtain selection diversity. Moreover, an adaptive weighting allocation which improves reconstruction accuracy of compressive sensing by maximizing the mutual information between target echoes and the transmitted signals is investigated. Further, this weighting scheme is applicable to both sample selection schemes, a conventional random sampling and the proposed selection. Simulations show that this selection method enhances the multiple target detection probability and reduces the normalized mean square error.
This dissertation also develops the adaptive weighting allocation of compressive sensing for MIMO UWB noise radar imaging from a practical perspective. For an adaptive weighting allocation scheme in the previous chapter, however, perfect knowledge of the target scene is not available in practice due to the existence of noise at receivers or lack of measurements. Thus, the recovery error provides inaccurate weighting values which can degrade reconstruction accuracies and target detection probabilities. In order to mitigate this problem, an adaptive weighting allocation with reconstruction error by using knowledge of recovery error variances is investigated. Numerical simulations for various scenarios show that this practical scheme improves NMSE and the target detection probability.
# Table of Contents

List of Figures ......................................................... ix
List of Tables .......................................................... xi
List of Symbols .......................................................... xii
Acknowledgments .......................................................... xiv

Chapter 1
Introduction ................................................................. 1
1.1 Overview .............................................................. 1
1.2 Contributions of the dissertation .................................. 3
1.3 Structure of the dissertation ....................................... 4

Chapter 2
UWB Random Noise Radar and Compressive Sensing .............. 6
2.1 UWB random noise radar system .................................. 6
  2.1.1 Ultra-wideband (UWB) radar ................................ 6
  2.1.2 Random noise radar ........................................... 7
  2.1.3 UWB random noise radar ...................................... 8
2.2 Compressive sensing and radar .................................... 11
  2.2.1 Classical signal processing ................................... 11
  2.2.2 Compressive sensing .......................................... 14
  2.2.3 Compressive UWB noise radar ............................... 16
Chapter 3
Multi-Target Detection using Total Correlation for Noise Radar Systems 18
3.1 Introduction ............................................ 18
3.2 Detection Approach Based on Mutual Information .................. 21
  3.2.1 Signal Model ..................................... 21
  3.2.2 Detection Based on Mutual Information ..................... 23
3.3 Modified Multiple Target Detector based on Total Correlation ....... 27
  3.3.1 Modified Total Correlation Detector ...................... 27
  3.3.2 Performance Analysis ................................ 29
3.4 Multiple Target Threshold based on Random matrix theory ......... 31
3.5 Simulations ............................................. 36
  3.5.1 Detection performances ................................ 36
  3.5.2 Adaptive Thresholding ................................ 38

Chapter 4
Target Detection and Reconstruction for Compressive MIMO UWB Noise Radar Imaging 46
4.1 Introduction ............................................. 46
4.2 System model .......................................... 49
  4.2.1 System model of typical radar systems ...................... 49
  4.2.2 Compressive radar system ................................ 52
4.3 Sample selection ........................................ 53
  4.3.1 Proposed sample selection strategy ....................... 53
  4.3.2 Complexity analysis .................................. 55
4.4 An adaptive weighting allocation ................................ 56
4.5 Simulations ............................................. 60
  4.5.1 Reconstructed images .................................. 61
  4.5.2 Reconstruction accuracy ................................ 63

Chapter 5
Adaptive Weighting Allocation with the Knowledge of Recovery Error for Compressive Sensing MIMO Noise Radar 77
5.1 Introduction ............................................. 77
5.2 Reconstruction error ..................................... 78
  5.2.1 System model ....................................... 78

vii
5.2.2 Adaptive weighting allocation with the knowledge of recovery error 79
5.3 Simulations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 81

Chapter 6
Conclusions and Future Research 89
  6.1 Conclusions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 89
  6.2 Future research . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 91

Appendix A
  A.1 Proof of Lemma 1. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 92

Bibliography 96
List of Figures

2.1 Simplified block diagram of a UWB random noise radar. .................. 9

3.1 Three target detection using TC at 20-dB SNR. ......................... 37
3.2 Three target detection using MTC at 20-dB SNR. ....................... 38
3.3 Three closely located target detection using TC at 20-dB SNR. ....... 39
3.4 Three closely located target detection using MTC at 20-dB SNR. ..... 40
3.5 Three target detection performance comparison at -10-dB SNR. ....... 42
3.6 Probability of detection with 5 receiver elements for $P_f = 0.02$. .... 43
3.7 Probability of detection with 10 receivers elements for $P_f = 0.02$. ... 44
3.8 The receiver operating curves for different values of signal to noise ratio with 5 receiver elements. ............................. 44
3.9 Receiver operating characteristics curves for different thresholds. ....... 45

4.1 MIMO noise radar system model. ........................................ 49
4.2 A block diagram of the proposed adaptive weighting allocation ...... 57
4.3 Comparison of reconstructed images by RSM and LSM at 20-dB SNR with $N = 100, M = 30, K = 3$, and $M_T = M_R = 1$ or 2. (a) Original target scene. (b) RSM with $M_T = M_R = 1$. (c) LSM with $M_T = M_R = 1$. (d) RSM with $M_T = M_R = 2$. (e) LSM with $M_T = M_R = 2$. ................. 62
4.4 Comparison of reconstructed images by RSM and LSM at 5-dB SNR with $N = 100, M = 30, K = 3$ and $M_T = M_R = 1$ or 2. (a) Original target scene. (b) RSM with $M_T = M_R = 1$. (c) LSM with $M_T = M_R = 1$. (d) RSM with $M_T = M_R = 2$. (e) LSM with $M_T = M_R = 2$. ................... 64
4.5 Comparison of NMSE with $N = 100, K = 3$, $M_T = M_R = 1$ and $M = 10, 30$ or $50$. (a) $M = 10$. (b) $M = 30$. (c) $M = 50$. ............... 65
4.6 Comparison of NMSE with $N = 100, K = 3$, $M_T = M_R = 2$ and $M = 10, 30$ or $50$. (a) $M = 10$. (b) $M = 30$. (c) $M = 50$. ............... 66
4.7 Comparison of $P_d$ with $N = 100$, $K = 3$, $M_T = M_R = 1$ and $M = 10$, 30 and 50. (a) $M = 10$. (b) $M = 30$. (c) $M = 50$. ................................. 68

4.8 Comparison of $P_d$ with $N = 100$, $K = 3$, $M_T = M_R = 2$ and $M = 10$, 30 and 50. (a) $M = 10$. (b) $M = 30$. (c) $M = 50$. ................................. 69

4.9 Comparison of MI with $N = 100$, $K = 3$, $M_T = M_R = 1$ and $M = 10$, 30 and 50. (a) $M = 10$. (b) $M = 30$. (c) $M = 50$. ................................. 71

4.10 Comparison of MI with $N = 100$, $K = 3$, $M_T = M_R = 2$ and $M = 10$, 30 and 50. (a) $M = 10$. (b) $M = 30$. (c) $M = 50$. ................................. 72

4.11 NMSE, $P_d$ and MI according to $K$ at -10-dB SNR with $N = 100$, $M = 30$, and $M_T = M_R = 1$. (a) NMSE. (b) $P_D$. (c) MI. ................................. 73

4.12 NMSE, $P_d$ and MI according to $K$ at -10-dB SNR with $N = 100$, $M = 30$, and $M_T = M_R = 2$. (a) NMSE. (b) $P_D$. (c) MI. ................................. 74

4.13 $P_d$, MI and NMSE according to $M_T = M_R$ at -10 dB SNR with $N = 100$, $M = 30$ and $K = 3$. (a) $P_d$. (b) MI. (c) NMSE. ................................. 76

5.1 A concept of the adaptive weighting allocation with recovery error . . . 78

5.2 Comparison of MI for several reconstruction errors, $\sigma_e^2$ with $N = 100$, $M = 30$, $K = 3$, and $M_T = M_R = 1$ or 2 (a) $M_T = M_R = 1$. (b) $M_T = M_R = 2$. ................................. 83

5.3 Comparison of MI for several reconstruction errors, $\sigma_e^2$ with $N = 100$, $M = 50$, $K = 3$, and $M_T = M_R = 1$ or 2 (a) $M_T = M_R = 1$. (b) $M_T = M_R = 2$. ................................. 84

5.4 Comparison of NMSE for several reconstruction errors, $\sigma_e^2$ with $N = 100$, $M = 30$, $K = 3$, and $M_T = M_R = 1$ or 2 (a) $M_T = M_R = 1$. (b) $M_T = M_R = 2$. ................................. 85

5.5 Comparison of NMSE for several reconstruction errors, $\sigma_e^2$ with $N = 100$, $M = 50$, $K = 3$, and $M_T = M_R = 1$ or 2 (a) $M_T = M_R = 1$. (b) $M_T = M_R = 2$. ................................. 86

5.6 Comparison of $P_d$ for several reconstruction errors, $\sigma_e^2$ with $N = 100$, $M = 30$, $K = 3$, and $M_T = M_R = 1$ or 2 (a) $M_T = M_R = 1$. (b) $M_T = M_R = 2$. ................................. 87

5.7 Comparison of $P_d$ for several reconstruction errors, $\sigma_e^2$ with $N = 100$, $M = 50$, $K = 3$, and $M_T = M_R = 1$ or 2 (a) $M_T = M_R = 1$. (b) $M_T = M_R = 2$. ................................. 88
List of Tables

4.1 Proposed largest selection method (LSM) and the complexity corresponding to each line of the algorithm. . . . . . . . . . . . . . . . . 55
List of Symbols

$\Delta R$  Range resolution
$c$  Speed of light
$f_c$  Center frequency
$\mathcal{N}(0, \sigma^2)$  Normal distribution with zero mean and $\sigma^2$ variance
$\mathcal{CN}(0, \sigma^2)$  Complex normal distribution with zero mean and $\sigma^2$ variance
$\phi$  Phase
$\cdot^*$  Complex conjugate
$T_p$  Pulse duration
$T_s$  Sampling time
$A^\dagger$  Moore-Penrose pseudoinverse
$\| \cdot \|_p$  $p$ norm
$\Omega$  Measurement matrix
$\Psi$  Basis matrix
$\langle \cdot, \cdot \rangle$  Inner product
$I(\cdot; \cdot)$  Mutual information
$K$  Number of targets

$\Sigma$  Covariance matrix

$(\cdot)^H$  Hermitian or conjugate transpose

$h(\cdot)$  Entropy

$h(\cdot|\cdot)$  Conditional entropy

$h(\cdot, \cdot)$  Joint entropy

$f(\cdot)$  Probability of density function

$(\cdot)^T$  Transpose

$|\cdot|$  Absolute value

$(\cdot)^{-1}$  Inverse matrix

$P_{fa}$  False alarm probability

$P_d$  Detection probability

$M_T, M_R$  Number of transmit and receive antennas, respectively
Acknowledgments

First and foremost, I would like to thank my advisor, mentor, and my role model, Dr. Ram M. Narayanan for his guidance, support and patience during my study at Penn State. He has also highly influenced me with his optimistic nature, scientific and critical way of thinking. It was a great honor for me to study this most fascinating and rewarding field under his supervision.

I would like to thank my committee members, Dr. William E. Higgins, Dr. Vishal Monga, and Dr. Mark Levi for their time, thoughtful comments, suggestions and in-depth review of my dissertation. I would also like to thank all professors who have imparted to knowledge in my study. I am also very much grateful and thankful to my former advisor at Inha University, for giving me the strong foundation of engineering, over which, this dissertation has been built. I also express my thanks to Dr. Muralidhar Rangaswamy of the Air Force Research Laboratory (AFRL) for his constant support, encouragement, and his insightful suggestions and comments on my work.

I am grateful for funding support from the Air Force Office of Scientific Research.
(AFOSR) Grant #FA9550-09-1-0605, and I would also like to thank Dr. Jon Sjogren of AFOSR for his valuable comments and suggestions during the development of this work. Furthermore, I would like to thank my fellow graduate students, especially Mahesh Shastry, for extensive discussions and exchange of ideas during this dissertation work.

Last but not the least, I would like to dedicate this dissertation to my family. My father, mother, brother, his’s wife, and two sons have always prayed sincerely and support me that I would be successful. Without their deep concern and encouragement, I would never have made it through some of the hard times of my journey towards completing my Ph.D. degree.
Introduction

1.1 Overview

Ultra-wideband (UWB) noise radar is widely considered a promising technique for the convert detection of multiple targets due to several advantages, such as excellent electronic countermeasure (ECM), low probability of interception (LPI), counter-electronic support measure (CESM), and relatively simple hardware architectures [1, 2, 3]. One of the most important advantages of noise radar is LPI, since the transmitted noise-waveform is constantly varying and never repeats exactly [4, 5]. Further, UWB noise radar achieves high-resolution target detection by employing an ultra-wide bandwidth. Due to these advantages, research in UWB noise radar has garnered much attention in the past decades [6], and UWB technology is being used in many radar applications, such as ground penetrating radar (GPR), through-wall radar (TWR), synthetic aperture radar (SAR), and imaging. In this dissertation, the main research is focused on target detection for UWB noise radar, which is one of the most challenging issues in radar
signal processing.

Typically, target detection is accomplished in noise radar by correlating the returned signal from the target and the transmitted noise waveform, and identifying the peak to recognize the target range. The correlation can be done using a matched filter maximizing the SNR. However, if the transmitted signal suffers non-linear distortion owing to multi-paths or multiple targets, the correlation detector yields degraded detection performance. In order to overcome such weakness, mutual information can be utilized instead of the correlator [7]. However, the mutual information detector has a strict limitation of system infrastructure, such as feasibility with only one receiver. In order to extend its use for multiple random variables, as a general form of MI, Watanabe proposed the total correlation that is the amount of shared information among the variables in the set [8]. In [9], total correlation is applied for multiple target detection using multiple receivers (single-input multi-output, SIMO), which outperforms the conventional correlator with one receiver. However, the total correlation detector yields sub-optimal detection performance due to high noise variance. To overcome this problem, it is imperative to develop the conventional total correlation detector.

In spite of advances in analog-to-digital converter (ADC) technologies, a sampling process for UWB radar would place a high burden on the ADC component due to the high sampling rate. As a strong candidate to solve problem, the compressive sensing (CS) technique was introduced by Donoho et al. [10]. According to [10], CS allows us to acquire significantly fewer samples than that given by the Nyquist sampling theorem, and the original signal can be reconstructed from the undersampled measurements, assuming the original signal is sparse in a representation. Thereafter, several papers developed
theoretical underpinnings of CS [11, 12, 13]. Due to the major advantage of its ability to recover the undersampled signal, it is quite efficiently applied to UWB noise radar. More specifically, utilizing CS for UWB noise radar imaging, it is possible to require fewer samples or measurements to reconstruct the original target scene. The suitability of random Toeplitz or circulant matrices is derived [14]. Further, the feasibility of CS for UWB noise radar imaging is shown based on Toeplitz matrices [15]. Inspired by the proposed method in [15], in this dissertation, we aim to develop and improve these theories from a practical perspective.

1.2 Contributions of the dissertation

The main contributions of the current dissertation can be summarized as follows:

- A target detector based on total correlation for noise radar systems is proposed. More specifically, this detector utilizes the largest eigenvalue property of covariance matrices generated by received signals at multiple receivers with one transmitter. The proposed detector achieves the improved detection performance of conventional detectors.

- Based on random matrix theory, an adaptive thresholding approach is proposed that accommodates a false alarm probability, the length of the transmitted waveform, and the number of receiving elements. The limits of the maximum and minimum eigenvalues are employed to compute the threshold of a certain target scene.

- UWB noise radar imaging for multi-target detection utilizing compressive sensing
is extended to MIMO scenarios. Furthermore, the reconstruction accuracies of SISO and MIMO are compared through numerical results obtained by Monte Carlo simulations.

• A sample selection strategy for compressive MIMO noise radar systems is proposed based on comparing received energies of candidate samples. The sorting and selection process leads to selection diversity gain, which improves reconstruction accuracy and detection probability. Moreover, computational complexity of the strategy is investigated and compared with the random sample selection method.

• An adaptive weighting allocation for compressive MIMO noise radar is proposed. This is achieved by formulating and solving a convex optimization problem with the constraint. Mutual information is utilized as the optimization metric, which provides a minimization of the error bounds of compressive sensing.

• From a practical perspective, an adaptive weighting allocation with recovery error for compressive MIMO noise radar is proposed. This scheme is a modified weighting allocation considering reconstruction error by a CS solver. The recovery accuracies and target detection probabilities of a conventional scheme and the modified scheme are compared through simulation results.

1.3 Structure of the dissertation

The rest of the dissertation is organized as follows:

Chapter 2 introduces the background and basic properties of UWB random noise radar systems and CS technique. It also describes how compressive sensing can be
efficiently applied to UWB noise radar systems for target detection and estimation.

Chapter 3 presents a multi-target detector based on the concept of total correlation, which is a modified version of the conventional total correlation detector for noise radar systems. Further, an adaptive thresholding approach based on random matrix theory is presented to differentiate between target and non-target pixels.

Chapter 4 presents an effective sample selection strategy for compressive sensing-based UWB noise radar systems, which enhances target detection and reconstruction performances. In addition, an adaptive weighting allocation for each measurement at the receiver is introduced based on convex optimization. Furthermore, these two schemes are extended to MIMO scenarios to obtain diversity and array gain.

Chapter 5 presents an adaptive weighting allocation considering reconstruction error for CS MIMO UWB noise radar, which mitigates the inherent error of inaccurate recovery by a CS solver.

Chapter 6 concludes the whole dissertation by briefly summarizing the main points of the dissertation, pointing to several problems that should be explored further in order to consider practical implementations.
Chapter 2

UWB Random Noise Radar and Compressive Sensing

2.1 UWB random noise radar system

2.1.1 Ultra-wideband (UWB) radar

UWB radar signals have been investigated by researchers to get as much information as possible from radar targets [16, 17, 18, 19]. While a narrow-band signal utilizes less than 10% of the fractional bandwidth, UWB utilizes the fractional bandwidth at greater than 25%. Since the waveform of UWB has very high fractional bandwidth, the determination of a carrier frequency is generally impossible; a situation known as carrier-free transmission techniques. The fractional bandwidth is defined as

$$\text{FBW} = \frac{2(f_H - f_L)}{f_H + f_L} = \frac{f_H - f_L}{f_a}, \quad (2.1)$$
where \( f_H \) and \( f_L \) are the upper and lower band edges of the system, and \( f_a \) is the average of \( f_H \) and \( f_L \).

The main advantage of UWB radar is using either short-duration pulses or waveforms over a wide spectral range, enhancing the detection probability of metrics such as range measurement accuracy and range resolution [20], which is given by

\[
\Delta R = \frac{c}{2\beta},
\]

where \( c \) is the speed of light and \( \beta \) is the bandwidth. Using low transmission power or below can be considered another advantage of UWB over narrow-band radar systems, and such a property leads UWB radar to share spectrums with other radar systems without causing any interference. By sharing low-frequency bands with other radar or communication systems, the spectrum efficiency increases and waveforms effectively propagate through materials such as bricks and other dense media.

### 2.1.2 Random noise radar

Research work has been conducted on the effective usage of random or pseudo-random signals since the 1950s [21, 22, 23]. While both types developed relatively slowly at first, the solid state microwave, wide-band communications, and digital signal processing progression of the last few years has boosted the theory and applications. The only difference between the random noise radar and traditional radars is that the random noise radar, also known as random signal radar (RSR), utilizes a random or random-like noise waveform as the transmit signal. The designs and classifications of noise radar
systems can be divided into two major categories: truly random noise radars and pseudo-random noise (PRN) radars. There are several advantages of noise radar systems. Using truly random waveforms, the random noise radar obtains high ECM capability, high range resolution, LPI, CESM capability, and relatively simple hardware [24]. Hence, noise radar systems provide great advantages in military applications.

2.1.3 UWB random noise radar

The concept of UWB random noise has been significantly developed in the last decade [25, 26, 27]. In comparison to conventional radar, which transmits predefined pulses or frequency-modulated waveforms, a UWB noise radar transmits an ultra-wideband random noise waveform towards the target scene. In general, target detection processes are identical to the traditional radar detector, which utilizes a matched filer. The returned signal reflected by random targets is cross-correlated with the transmit signal using the same pseudo-random number generator at the transmitter and receiver. Because the transmit waveform in a certain time frame is not deterministic in noise radar systems, the detector requires a correlator. It is mathematically well known that the matched filter and the correlator operate alike and yield the same output signal-to-noise ratio (SNR), maximizing the output SNR and determining the target locations and ranges.

A random noise radar signal can be written as

$$x(t) = x_I(t) \cos(2\pi f_c t) - x_Q(t) \sin(2\pi f_c t),$$  \hspace{1cm} (2.3)

where $x_I(t)$ and $x_Q(t)$ are random variables drawn from $\mathcal{N}(0, \sigma^2)$, and $f_c$ is the carrier
As a simple form, (2.3) can be rewritten as

$$x(t) = a(t) \cos [2\pi f_c t + \phi(t)],$$  \hspace{1cm} (2.4)$$

where $a(t)$ is the Rayleigh distributed envelope function, and $\phi(t) = \tan^{-1}\left[\frac{x_Q(t)}{x_I(t)}\right]$ is a uniformly distributed phase function over $[-\pi, +\pi]$. (4.14) can be expressed in a complex analytic form as

$$x(t) = \frac{1}{2} [x_C(t) \exp(j2\pi f_c t) + x_C^*(t) \exp(-j2\pi f_c t)],$$  \hspace{1cm} (2.5)$$

where $x_C(t) = x_I(t) + jx_Q(t)$, and $(\cdot)^*$ is a complex conjugate. A simplified block diagram of a UWB random noise radar system using the transmit signal model, (4.15) is shown in Fig. 4.11.

Assuming that the transmit waveform is (4.15), the returned signal by a point target in a discrete domain can be written as

$$y(t) = \sum_{i=-L/2}^{L/2} r(t)x(t-i) + n(t),$$  \hspace{1cm} (2.6)$$
where \( r(t) \) and \( n(t) \) are the point target reflectivity drawn from Rayleigh distribution and additive white Gaussian noise, \( \mathcal{N}(0, \sigma^2) \), respectively. \( L \) denotes the time duration of the extended target in the down range. Thus, the received signal \( y(t) \) is also stationary and ergodic. Therefore, the cross-correlation in time domain between the returned signal and the time delayed \( x(t) \), i.e., \( x^*(t - \tau) \), can be used to estimate the target reflectivity as

\[
 z(t) = \sum_{j=-T/2}^{T/2} \sum_{i=-L/2}^{-L/2} r(t)x(t - i)x^*(t - j) + \sum_{j=-T/2}^{T/2} n(t)x^*(t - j), \tag{2.7}
\]

where \( T \) is the correlation summation period. Statistically, the second term of the right-hand side is zero where \( T \to \infty \), since \( n(t) \) is independent to \( x^*(t - j) \). Therefore, the received target response \( g(\tau) \) is

\[
g(\tau) = E[z(t)] = E \left[ \sum_{j=-T/2}^{T/2} \sum_{i=-L/2}^{-L/2} r(t)x(t - i)x^*(t - j) \right]. \tag{2.8}
\]

As one can see, \( g(\tau) \) is a function of the delayed time \( \tau \), and that the target reflectivity can be extracted from (4.18). Hence, the radar system can detect a target range and reflectivity. For example, if a UWB waveform is transmitted over a 500 MHz bandwidth, the receiver achieves an excellent resolution of 0.3 meters. Moreover, since the transmitted signal \( x(t) \) is unknown due to random nature of the waveform, this system is robust against jamming signals, \( j(t) \), i.e.,

\[
\sum_{j=-T/2}^{T/2} j(t)x^*(t - j) \approx 0. \tag{2.9}
\]

UWB noise radar signals are not only difficult to detected, but also robust against
jamming by hostile objects, since the transmit signal is never repeated.

2.2 Compressive sensing and radar

“Everyone now knows that most of the data we acquire “can be thrown away” with almost no perceptual loss. Witness the broad success of lossy compression formats for sounds, images, and specialized technical data. The phenomenon of ubiquitous compressibility raises very natural questions: why go to so much effort to acquire all the data when most of what we get will be thrown away? Can we not just directly measure the part that will not end up being thrown away?”

- D.L. Donoho in 2006 -

CS is a new paradigm in signal processing that trades sampling frequencies for computing power and allows the accurate reconstruction of signals sampled at rates less than the conventional Nyquist frequency. Due to such benefits, CS has received considerable attention and has been successfully applied in various fields, such as radar and magnetic resonance imaging (MRI). In this section, we describe a brief history of the developments in signal processing techniques leading to compressed sensing.

2.2.1 Classical signal processing

We start with the following Shannon-Nyquist sampling theorem. The theorem states that any band-limited continuous signals can be represented by uniformly spaced samples taken at a rate greater than twice the maximum frequency of the signal. Shannon’s 1949 version of the theorem is given below.
Theorem 1. [28] If a function $f(t)$ contains no frequencies higher than $W$ hertz, it is completely determined by giving its ordinates at a series of points spaced $1/2W$ seconds apart.

$$f(t) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2W}\right) \sin\left(\frac{2\pi W \left[t - \frac{n}{2W}\right]}{2\pi W}\right). \quad (2.10)$$

In plain language, Theorem 1 states that any band-limited signals (the bandwidth $W$) can be represented with the uniform spaced samples taken at a rate greater than twice the maximum frequency of the signal (sampling rate), $f_s = 2W$. For an arbitrary band-limited signal, $f_s$ is necessary, otherwise the original signal cannot be reconstructed from the samples.

Now, consider taking samples of a frequency spectrum at every $f_p$ in inside the band $f_s \triangleq [-W,W]$. Let $T_p$ and $T_s$ denote the sampling time duration and $1/W$. Suppose that $N$ samples are collected during $T_p$, i.e., $T_p/T_s = T_pW = N$, in each domain. Then there are $N$ distinct samples in each domain. One sample of each domain can be obtained using a discrete fourier transform (DFT) or inverse DFT. For example, one sample, $b(n)$, in a time domain is composed of $N$ samples frequency domain, $x(k)$, as

$$b(n) = \frac{1}{N} \sum_{k=0}^{N-1} x(k) \exp\left(\frac{j2\pi nk}{N}\right). \quad (2.11)$$

Then we can write $b(n), n = 0, \cdots, N - 1$ samples in the matrix form as

$$b = Ax, \quad (2.12)$$

where $A$ is the $N \times N$ IDFT (inverse DFT) matrix (linear operator) and $x$ is the $N \times 1$
vector. Thus, $b$ is composed of a linear combination of $x$, which means that $x$ can be obtained with a vector $b$, i.e., an inverse problem. In other words, we seek an estimate $\hat{x}$ of $x$.

Suppose the linear operator $A \in \mathbb{R}^{M \times N}$. Then, there are two possible cases of finite dimension of $A$ to solve inverse problems, i.e., $M \geq N$ (over-determined) and $M < N$ (under-determined). Consider a case in which $M \geq N$, so the problem is over-determined. The popular approach, also known as the maximum likelihood (ML) estimator, is to set $\hat{x} = \arg \min_x \| b - Ax \|_2^2$, which is the least square (LS) estimator. LS can be implemented as $\hat{x} = A^\dagger b$ where $A^\dagger$ is the Moore-Penrose pseudoinverse of $A$. If $A$ is ill-conditioned (a singular or large-condition number), the Tikhonov regularization recovers as

$$\hat{x} = \arg \min_x \| Ax - b \|_2^2 + \| \Gamma x \|_2^2,$$

(2.13)

where $\Gamma x$ is the properly chosen Tikhonov matrix. Note that ML, LS and Tikhonov regularization can successfully recover the original signal of over-determined problems, and are based on $l_2$ norm minimization.

Similar to $l_2$ solvers, $l_1$ optimizations were well studied and developed by many researchers. Most $l_1$ minimizations are related to sparse signal reconstruction. It began in 1965 by Logan [29]. The dissertation proved that continuous and band-limited signals corrupted by noise (very sparse support in time domain) can be perfectly reconstructed by finding minimal $l_1$ norm projection. In the discrete domain, the $l_1$ norm was presented in [30]. In [30], the reconstruction of $K$ sparse vector is possible with just $M$ samples using LASSO [31], even though the original signal has $N$ elements. The work
of [30] was developed by [32, 33]. In [34], Rudin, Osher, and Fatemi demonstrated the total-variation minimization (TV) usage, which is connected to $l_1$ and CS.

### 2.2.2 Compressive sensing

As a first step of CS, Donoho and Huo initialized $l_1$ minimization as the sparsest solution of the CS theory [10]. Thereafter, several papers developed the theoretical underpinnings of CS [11, 12, 13, 35, 36, 37, 38]. Those articles developed concrete theories for CS, and suggested that CS has strong potential for various application areas.

This section provides background on CS. Suppose a discrete signal vector $x \in \mathbb{C}^N$. Then, the measurement vector $b$ is

$$ b = \Phi \Psi \alpha = \Phi x, \quad (2.14) $$

where $\Phi \in \mathbb{C}^{M \times N}$ is a measurement matrix, and $\Psi \in \mathbb{C}^{N \times N}$ denotes the basis matrix that spans the vector space of $\alpha$. We assume that $\Phi$ and $\Psi$ are orthogonal. Due to the under-determined case $M \ll N$, classical linear system solvers such as LS and ML would have the infinite number of solutions. However, if $x$ is a $K$ sparse vector, the solution can be obtained using the exhaustive search ($l_0$ minimization) as

$$ \min \|x\|_0 \quad \text{subject to } \Phi x = b. \quad (2.15) $$

Using $l_0$ minimization, (2.15) can be solved with a strict assumption that the solution is unique [39]. However, uniqueness is not guaranteed in general, and (2.15) is computationally intractable. Instead of $l_0$, alternative tractable approaches have been proposed
using $l_1$ [40, 12, 41], which is written as

$$\min \|x\|_1 \quad \text{subject to } \Phi x = b,$$  \hspace{1cm} (2.16)

which is a convex optimization problem and can be seen as a convex relaxation of (2.15).

In (2.16), the measurement matrix $\Phi$ can be described as a linear transformer. There are three popular types of matrix: (1) randomly selected rows of the Fourier transform matrices, (2) i.i.d. Gaussian matrix $\mathcal{N}(1, 1/M)$, and (3) i.i.d. Bernoulli $-1, +1$. These matrices satisfy two important properties of CS, incoherence and restricted isometry property (RIP), which describe the full recovery properties of CS.

The coherence is a classical way of analyzing the recovery abilities of $\Phi$ and $\Psi$ [42, 12], defined as,

$$1 \leq \mu(\Phi, \Psi) \triangleq \sqrt{N} \max_{1 \leq k,j \leq N} |\langle \phi_j, \psi_k \rangle| \leq \sqrt{N},$$  \hspace{1cm} (2.17)

where $\langle \cdot, \cdot \rangle$ is the inner product. The coherence of the matrices $\Phi$ and $\Psi$ is a metric to select a proper measurement matrix. Thus, $\Phi$ is usually constructed with the random Gaussian matrix since the statistical property of $\Phi \Psi$ remains the same as that of $\Phi$ when $\Psi$ is a unitary (orthogonal) matrix.

Another important and convenient tool of CS is the restricted isometry property (RIP), defined below.

**Definition 1.** (RIP, Restricted Isometry Property). The restricted isometry constant
\( \delta_k \) of order \( k \) for matrix \( \Phi \) is the smallest number such that

\[
(1 - \delta_k) \leq \frac{\| \Phi x \|_2^2}{\| x \|_2^2} \leq (1 + \delta_k)
\]

(2.18)

holds for all \( k \) sparse vectors \( x \).

The best RIP constant \( \delta_k = 0 \) for any \( k \) that can be obtained with \( \Phi \) is an orthogonal matrix. More generally, if a small \( \delta_k < 1 \) exists for a class of \( \Phi \), then \( \Phi x \) should behave like a unitary transformation. Thus, \( b \) and \( x \) are one-to-one. On occasion, adding more rows to a measurement matrix changes RIP constants, generally improving \( \delta_k \). Thus, \( \delta_{2k} \) is utilized to verify the perfect recovery property of \( \Phi \). For example, if \( \delta_{2k} < 1 \), then the \( l_0 \) solution is unique. On the other hand, if \( \delta_{2k} < \sqrt{2} - 1 \), the \( l_1 \) solution attains the \( l_0 \) solution. Furthermore, the RIP provides for robust noisy signal recovery [13, 43].

### 2.2.3 Compressive UWB noise radar

A typical radar system transmits a certain type of waveforms, and then uses a matched filter to correlate the signal received based on the information of transmitted waveforms. The receiver utilizes a pulse compression system with a high-rate ADC for signal processing. This traditional approach increases the complexity and cost of hardware architecture design, as well as limits the resolution of the radar. Such a weakness becomes worse for a traditional UWB radar due to its wide bandwidth. Utilizing a pseudo-random waveform generated by a noise waveform generator is fortunately well-suited to CS. The transmit waveform can be constructed as a Toeplitz matrix, and satisfies the RIP condition. Suppose that the transmitted waveform is generated from a Gaussian distribution, and
the reflected signal from targets measured with a restriction matrix. In this case, each possible target scene can be treated as a vector or matrix. If the number of targets is small enough, then the grids with targets will be sparse. CS techniques thus allow us to estimate, detect multiple targets, and reconstruct the target scene with a significantly smaller number of samples than that of the traditional UWB noise radar systems [44].
Chapter 3

Multi-Target Detection using Total Correlation for Noise Radar Systems

3.1 Introduction

A radar system which transmits white Gaussian noise, usually amplified thermal noise, is known as a noise radar. Noise radar has been widely considered as a promising technique for high-resolution detection of multiple targets [1, 2, 3]. One of the most important advantages of noise radar is their low probability of intercept since the transmitted noise-waveform is constantly varying and never repeats exactly [4, 5]. In typical noise radar systems, target detection is accomplished by correlating the reflected signal from the target with a delayed replica of the transmitted random waveform, and identifying the peak in the correlation value. However, if other factors such as clutter reflections, propagation factors or multipath effects cause non-linear distortions on the transmitted signal, the correlation detector may not always successfully detect targets.
In order to deal with this limitation of correlators, the mutual information (MI) metric has been proposed as an alternative to the correlation [7]. MI, $I(X;Y)$, is a measure of dependence between two variables $X$ and $Y$. If the two are correlated or related, MI has a non-zero value. This implies that MI can be utilized to extract dependent information between two variables. Moreover, one of interesting properties of MI is its insensitivity to non-linearities. Even though $X$ and $Y$ are separately (and non-linearly) transformed, MI remains the same as that of the original variables, which means that MI is robust in quantifying the relationship between any two random variables. As a first step towards applying information theory for radar systems, a waveform optimization approach which maximizes MI between a random target impulse response and the returned signal based on the knowledge of target impulse response was proposed [45]. In [46], an effective method is proposed to minimize minimum mean-square error (MMSE) in an additive white Gaussian noise environment. Waveform design approaches for various environments to optimize MI between reflected signals and target reflectivity based on information of transmitted waveforms have been proposed [47, 48].

Thus, MI serves as an alternate metric to measure the dependence between the reflected signal and the delayed transmit replica in noise radar systems. However, MI is limited to only two random variables. In order to extend its use for multiple random variables, as a general form of MI, Watanabe proposed the total correlation which is the amount of shared information among the variables in the set [8]. However, the theory proposed in [8] is expressed by probability distributions which is difficult to implement in practice, since the length of received signal is a finite number which might not yield the exact distributions in some cases. Considering a Gaussian environment in which all
random variables are Gaussian distributed, MI can be formulated as a simple form as a function of variances of random variables [49]. In [50], the estimated mutual information is known to be inherently biased with a finite number of samples. A number of approaches to obtain more accurate MI from the limited samples have been presented by many researchers. However, MI in this chapter is utilized as a metric to qualify the dependency of two random variables even though the estimated value is positively biased. However, that sample estimation is expected to be more accurate than the estimation of the distribution itself which means that the sample based MI is still informative [51]. In radar systems, if the received signal is composed of random variables drawn from a multivariate distribution, MI can be computed from the sample variances and the correlation coefficients of the signals. In the same manner, the total correlation can be reformulated via a simple equation, since the total correlation is an ad hoc extension of MI. This concept can be directly applied to the noise radar systems with multiple receiving antennas for multiple target detection since noise radar waveforms are generated from a set of waveforms based on orthogonal bases [9]. The proposed multiple target detection method outperforms the conventional correlation methods in SAR (synthetic aperture radar) imaging. However, this detection method may not achieve optimal performance at low SNR regimes, since receiving antennas suffer high noise variance which leads to a low total correlation value. Thus, at low SNRs the performance of total correlation detector would be deteriorated.

In this chapter, we propose a modified total correlation target detector based on the largest eigenvalue. The proposed method achieves better performance for multiple targets compared to the conventional total correlation method which is at intermediate
SNR regimes. At low SNRs, the proposed detection method outperforms the conventional method. Moreover, in order to identify target pixels at low SNRs, we propose an adaptive thresholding method based on random matrix theory. Simulation results verify that the proposed method can judge the existence of the targets in a certain area.

The remainder of this chapter is organized as follows. The system model considered and the conventional total correlation detector using MI are presented in Section 2.2. Section 2.3 introduces and discusses our detection scheme. In Section 2.4, an adaptive threshold method is presented which identifies target pixels or non-target pixels (target existence or absence). Section 2.5 illustrates simulations to evaluate the proposed method for intermediate and low SNR regimes. In addition, we compare the conventional and the modified multi-target detector proposed. Conclusions are presented in Section 2.6.

3.2 Detection Approach Based on Mutual Information

3.2.1 Signal Model

Consider a single input multiple output (SIMO) system with one transmitting antenna and \( N \) receiving antennas. The number of targets in the scene is assumed to be \( K \). Denote the transmitted sequence (waveform) at time instant \( n \) by \( x(n) \). Then the total received signal at the \( i \)th receiving antenna element is given as

\[
y_i(n) = \sum_{k=1}^{K} r_{ki} x(n - \tau_{ki}(k)) + \eta_i(n),
\]

\( i = 1, 2, \cdots, N, \quad k = 1, 2, \cdots, K, \)
where $r_{ki}$ is the $k$th target response at the $i$th receiver. Assuming a Rayleigh distributed fluctuating amplitude and a uniformly distributed random phase, the individual target response is modeled as zero mean complex Gaussian random variables with variance $\sigma^2$ [52, 53]. Denote the duration of the transmitted waveform $x(k)$ as $M$. Since each component of the waveform is a random variable drawn from a Gaussian distribution, the samples are independent and identically distributed (i.i.d.). In (4.21), $\tau_{ki}$ represents the time delay for the received signal from the $k$th target at the $i$th receiver, and $\eta_i$ denotes complex white Gaussian noise assumed to be i.i.d. with zero mean and variance $\sigma^2_\eta$ at the $i$th receiver. As shown in (4.21), the total received signal $y_i(n)$ at each receiver element are the superposition of the reflected signals from multiple $K$ targets. In this chapter, we assume that the receiving antenna elements are a uniform linear array (ULA) that each element is sufficiently separated for received signal decorrelation. Thus, the spatial correlations among components of target response vectors can be ignored. Further, we assume that the noise vectors are independent.

The received signal model given in (4.21) can be cast into a vector form by considering $M$ consecutive samples as follows,

$$r_i(n) = [r_{1i} \ r_{2i} \ \cdots \ r_{ki}]^T,$$  

(3.2)

$$x(n) = [x(n) \ x(n-1) \ \cdots \ x(n-M+1)]^T,$$  

(3.3)

$$\eta_i(n) = [\eta_i(n) \ \eta_i(n-1) \ \cdots \ \eta_i(n-M+1)]^T,$$  

(3.4)

$$y_i(n) = [y_i(n) \ y_i(n-1) \ \cdots \ y_i(n-M+1)]^T.$$  

(3.5)
\[
\mathbf{X} = \begin{bmatrix}
  x(n) & x(n-1) & \cdots & x(n-M+1) & 0 & \cdots & 0 \\
  0 & x(n) & \cdots & x(n-M) & x(n-M+1) & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & 0 & 0 & \cdots & x(n-M+1)
\end{bmatrix}.
\]

(3.7)

Then we have the received signal at the \(i\)th antenna given by

\[
y_i(n) = r_i^T \mathbf{X}_i^{i+k} + \eta_i(n),
\]

(3.6)

where \(\mathbf{X}_i^{i+k}\) is the sub-matrix of transmitted waveform matrix, and \(\mathbf{X}\) defined by (4.22).

The superscript and subscript of \(\mathbf{X}\) in (4.27) are column indices for generating the sub-matrix. For instance, \(\mathbf{X}_i^{i+k}\) is a \(K \times M\) matrix generated by \(\mathbf{X}\) from the \(i\)th to the \(i+k\)th column. Now, we define a matrix having \(M \times (N+1)\) elements composed of \(\mathbf{x}\) and \(\mathbf{y}_i\)s as

\[
\Phi = \begin{bmatrix}
  \mathbf{x} & \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_N
\end{bmatrix}^T,
\]

(3.8)

and the sample covariance matrix,

\[
\Sigma = \frac{1}{N+1} \Phi \Phi^H = \frac{1}{N+1} \sum_{n=1}^{N+1} \phi_n \phi_n^H,
\]

(3.9)

where \(\phi_n\) is the \(n\)th column of \(\Phi\).

### 3.2.2 Detection Based on Mutual Information

Suppose that there are two continuous random variables \(\mathbf{x} \in \{x_i\}, \mathbf{y} \in \{y_i\}, i \in \{1, 2, \cdots, M\}\). The logarithm is to base 2 throughout this chapter. When we consider
two random variables, \( x \) and \( y \), MI is defined as [54]

\[
I(x; y) = h(x) - h(x|y),
\]

(3.10)

where \( h(x) \) is the entropy of \( x \) and \( h(x|y) \) is the conditional entropy of \( x \) given \( y \). The differential entropy of variable \( x \) is defined as [49]

\[
h(x) = -\int_{\mathbb{R}^N} f(x) \log f(x) \, dx.
\]

(3.11)

where \( f(x) \) represents the probability density function of \( x \).

Then MI can be written as

\[
I(x; y) = h(x) + h(y) - h(x, y)
\]

\[
= \int_{\mathbb{R}^M} \int_{\mathbb{R}^M} f(x, y) \log \left( \frac{f(x, y)}{f(x) f(y)} \right) \, dx \, dy,
\]

(3.12)

where \( h(x, y) \) and \( f(x, y) \) are the joint entropy and the joint distribution of \( x \) and \( y \) [54, 49]. If the variables are drawn from a Gaussian random distribution, MI is obtained as

\[
I(x; y) = -\frac{1}{2} \log \left( \frac{|\Sigma|}{\sigma_x^2 \sigma_y^2} \right),
\]

(3.13)

where \( |\cdot| \) is determinant of a square matrix. \( \Sigma \) is the covariance matrix of two variables, and \( \sigma_x^2 \) and \( \sigma_y^2 \) are variance of \( x \) and \( y \), respectively [55].

As mentioned earlier, MI measures the dependence between two random variables.
Thus MI can be utilized as an alternate metric which is justified by the following Lemma.

**Lemma 1.** Let \( s_{ii} \) be the cross correlation coefficient between two unit variance normal random variables corresponding to \( x \) and \( y \). Then, as \( s_{ii} \to -1 \) or \(+1\), \( I(x; y) \to \infty \).

**Proof.** See Appendix A. \( \Box \)

Therefore, MI can be utilized for multiple target detection. Moreover, if the length of the variable is longer, MI obtains much larger value as shown in (A.12).

Now suppose that there is a set of discrete random variables \( S = \{x_i\}_{i=1}^{N} \) and \( Q = \{y_i\}_{i=1}^{N} \) where \( N \) is the number of variables. As an extension of mutual information, the general form of total correlation introduced by Watanabe [8] is given by (3.26) where \( p(x) \) and \( p(x, y) \) are the probability mass function of \( x \) and the joint probability mass function of \( x \) and \( y \). From a practical perspective, (3.26) would not be able to measure the exact MI value of target response, since characterization of the exact distribution of the reflected signals requires a large number of samples. Assuming that all variables are Gaussian random variables, the total correlation can be written as the extended version of MI,

\[
C(x_1, x_2, \cdots, x_N) = \sum_{i=1}^{n} H(x_i) - H(x_1, x_2, \cdots, x_n) = \sum_{x_1 \in X_1} \cdots \sum_{x_n \in X_n} p(x_1, x_2, \cdots, x_N) \log_2 \left( \frac{p(x_1, x_2, \cdots, x_N)}{p(x_1)p(x_2) \cdots p(x_n)} \right).
\] (3.26)
Fortunately, for practical noise radar systems under the Gaussian environments, the transmitted and received signals can be considered to be random variables drawn from a multivariate normal distribution. Utilizing the properties of the total correlation, Davydov introduced a detector scheme which is able to successfully detect multiple targets [9]. The total correlation detector with \( N \) receiving antennas for a certain pixel is given as

\[
C(x, y_1, y_2, \cdots, y_N) = -\frac{1}{2} \log \left( \frac{|\Sigma|}{\sigma_x^2 \prod_{i=1}^{N} \sigma_{y_i}^2} \right),
\]

(3.15)

where \( x \) is the transmitted signal and \( y_i \) is the received signal vector at the \( i \)th receiving antenna. Note that total correlation is equivalent to the mutual information in (A.12) and (3.14) with only two random variables. The total correlation detector provides ranging and SAR imaging for multiple targets, and outperforms the cross-correlation detector [9].

However, for measured signals at receiving antennas at low SNRs, \( \sigma_{y_i}^2 \) will be large. \( \sigma_{y_i}^2 \) is determined by amplitudes of reflected signal plus additive noise at the \( i \)th receivers, and the noise variance is much larger than that of the reflected signal at low SNRs. Thus, the total correlation detector, (4.29) would have relatively high values over the whole given area, and the target detection performance of the detector would be degraded.
3.3 Modified Multiple Target Detector based on Total Correlation

In order to overcome the performance degradation of the total correlation method at low SNR environments, we adopt the largest eigenvalue property from random matrix theory. The details are as follows.

3.3.1 Modified Total Correlation Detector

Covariance matrices can be considered as functions of matrix arguments such as eigenvalues, specifically the largest eigenvalues attaining the original characteristics of the covariance matrix. From the property of the largest eigenvalue, we can extract information of the covariance matrix ($\Sigma$) using spectral decomposition techniques, e.g., SVD. In this sense, (3.15) can be reformulated as

$$C'(x, y_1, y_2, \cdots, y_N) = -\frac{1}{2} \log \left( \frac{\lambda_{\text{max}}(\Sigma)}{\frac{\sigma_x^2}{N} \prod_{i=1}^{N} \rho(x, y_i)} \right), \quad (3.16)$$

where $\rho(x, y_i)$ is the correlation coefficient between $x$ and $y_i$. Since the receivers of noise radar systems have full knowledge of the predefined transmitted signal, $x$, $\Sigma$ can be readily obtained. If the transmitter and receivers share the same pseudo-code generator such as the Walsh-Hadamard code, it can be practically implemented.

Consider a case of one transmitter and one receivers for simplicity. In this case, the received signal can be written as $y = rx + \eta$, where $y$, $x$ and $\eta$ are $M \times 1$ vectors, where $M$ is the length of the transmitted signal. Scalar $r$ is a the target response. Let us define
statistical covariance matrices as

\[ R_y = E[yy^H], \quad (3.17) \]
\[ R_{rx} = r^2E[xx^H]. \quad (3.18) \]

Then, we have

\[ R_y = R_{rx} + \sigma^2 \mathbf{I}_M, \quad (3.19) \]

where \( \sigma^2 \) is variance of noise.

\( R_y \) and \( R_{rx} \) can be eigendecomposed as \( R_y = U \Lambda U^H \) and \( R_{rx} = \Lambda \Gamma \Lambda^H \), respectively, where \( U \) and \( \Lambda \) are unitary matrices. \( \Lambda \) and \( \Gamma \) are diagonal matrices whose components are the ordered eigenvalues of \( R_y \) and \( R_{rx} \) as \( \lambda_1, \lambda_2, \ldots, \lambda_M \) and \( \gamma_1, \gamma_2, \ldots, \gamma_M \), respectively, where the first eigenvalues, \( \lambda_1 \) and \( \gamma_1 \) are the largest eigenvalues. If there is no target in a certain pixel, the receiver will not measure the returned signal, which leads to \( R_{rx} = 0 \), i.e., \( \lambda_1 = \sigma^2 \). On the other hand, in the case of target existence, \( R_{rx} \neq 0 \), which leads to \( \lambda_1 = \gamma_1 + \sigma^2 \). In (3.19), the largest eigenvalue is much greater than the noise variance. Thus, we can verify \( \lambda_1 > \sigma^2 \), and considering only the largest eigenvalue, we can extract the original information of target. Moreover, since \( 0 \leq \| \rho(x, y_i) \| \leq 1 \), \( \prod_i \| \rho(x, y_i) \| \) is very small which boosts the numerator in (3.16). Therefore, the ratio between the minimum and maximum value of the modified total correlation is less than the conventional total correlation value.
3.3.2 Performance Analysis

In order to mathematically verify that the proposed method achieves better performance, we derive the ratio of the conventional total correlation to the modified method’s ratio between the normalized minimum and peak values. For simplicity, we assume that the received signal at each receiving antenna contains the entire transmitted signals.

Let $R_{pro}$ and $R_{con}$ denote the ratio of minimum to the maximum value of the proposed and the conventional total correlation method, respectively. With this setting, we can say that the proposed method achieves better performance if $R_{pro}/R_{con} > 1$, since it implies that the proposed scheme has the large value for the ratio.

For each ratio, we need two different pixels which contain a target presence and absence. In the case of target absence for the conventional detector, the total correlation value $C_{abs}$ can be obtained by

$$C_{abs} = -\frac{1}{2} \log (|\Sigma|) + \frac{1}{2} \log \left( \sigma_{x}^{2} \prod_{i=1}^{N} \sigma_{y_{i}}^{2} \right)$$

$$= -\frac{N}{2} \log (\sigma_{n}^{2}) + \frac{1}{2} \log (\sigma_{s}^{2} \sigma_{n}^{2N})$$

$$= -\frac{N}{2} \log (\sigma_{n}^{2}) + \frac{1}{2} \log (\sigma_{s}^{2}) + \frac{N}{2} \log (\sigma_{n}^{2})$$

$$= \frac{1}{2} \log (\sigma_{s}^{2}) , \quad (3.20)$$

where $|\Sigma|$ is equal to the product of all eigenvalues of the sample covariance matrix, $\prod_{i} \lambda_{i}$, and if there is no target, $\lambda_{\text{max}} = \lambda_{2} = \cdots = \lambda_{N+1} = \sigma_{n}^{2}$. Each $\sigma_{y_{i}}^{2}$ is equal to $\sigma_{n}^{2}$ due to the received signal absence. For the case that a target is present in a pixel, the
conventional method achieves the peak value $C_{\text{pre}}$ as

$$C_{\text{pre}} = -\frac{1}{2} \log \left( \prod_{i} \lambda_i (\Sigma) \right) + \frac{1}{2} \log \left( \sigma_x^2 (\sigma_y^2 + \sigma_{\eta}^2)^N \right)$$

$$= \frac{1}{2} \log \left( \frac{\sigma_x^2 (\sigma_y^2 + \sigma_{\eta}^2)^N}{\prod_{i} \lambda_i (\Sigma)} \right).$$  \hspace{1cm} (3.21)

In (3.21), since there is a target at a specific pixel, the product of all eigenvalues is not equal to $\sigma_{\eta}^2 (N+1)$ due to the transmitted signal replica. The target is present in this case, which implies $\sigma_x^2$ is $(\sigma_y^2 + \sigma_{\eta}^2)$. Hence, $\prod_{i} \lambda_i (\Sigma)$ will not be able to be expanded in terms of the signal variance.

In the same manner as above, we obtain the $C'_{\text{abs}}$ and $C'_{\text{pre}}$ for the proposed method as follows,

$$C'_{\text{abs}} = -\frac{1}{2} \log \left( \frac{\lambda_{\text{max}} (\Sigma)}{\sigma_x^2 \prod_{i=1}^{N} \| \rho_{\eta} (x, y_i) \|} \right)$$

$$= -\frac{1}{2} \log \left( \frac{\sigma_{\eta}^2}{\sigma_x^2 \prod_{i=1}^{N} \| \rho_{\eta} (x, y_i) \|} \right), \hspace{1cm} (3.22)$$

$$C'_{\text{pre}} = -\frac{1}{2} \log \left( \frac{\lambda_{\text{max}} (\Sigma)}{\sigma_x^2 \prod_{i=1}^{N} \| \rho_{\eta} (x, y_i) \|} \right)$$
\begin{align*}
&= -\frac{1}{2} \log \left( \frac{\lambda_{\text{max}}(\Sigma)}{\sigma_s^2 \prod_{i=1}^{N} \|\rho_s(x, y_i)\|} \right). \quad (3.23)
\end{align*}

Then, we finally compute $R_{\text{pro}}/R_{\text{con}}$ given by,

\begin{align*}
R_{\text{pro}} &= \frac{\prod_{i=2}^{N+1} \lambda_i(\Sigma) \prod_{i=1}^{N} \|\rho_s(x, y_i)\|}{(\sigma_s^2 + \sigma_n^2)^N \prod_{i=1}^{N} \|\rho_n(x, y_i)\|}.
\end{align*} \quad (3.24)

Since the product of correlation coefficient $\prod_{i=1}^{N} \rho_n(x, y_i)$ with no signal is close to zero (non-zero due to the finite number of receiving antennas), $R_{\text{con}}/R_{\text{pro}}$ is always larger than unity, which means the proposed method achieves better performance than the conventional method.

### 3.4 Multiple Target Threshold based on Random matrix theory

The elements of $\eta_i$ are drawn from independent Gaussian distribution. We define the SNR of the received signal of the $i$th receiver element as

\begin{equation*}
\text{SNR}_i = 10 \log_{10} \left( \frac{E\left(\|\mathbf{y}_i - \eta_i\|^2\right)}{E\left(\|\eta_i\|^2\right)} \right) \text{ dB},
\end{equation*} \quad (3.25)

\begin{itemize}
  \item $i = 1, 2, \cdots, N$.
\end{itemize}

We assume that $\eta_i$ has the same variance $\sigma^2_{\eta_i}$. Thus, $\text{SNR}_i$ can be generalized as SNR
over the receiving antenna elements.

At low SNR regimes, the correlation value becomes very small, since the noise is i.i.d. This causes the ratio of maximum to minimum value of (3.16) to have a relatively smaller value. In order to improve this performance degradation, we propose a threshold to obtain target existence or absence at a certain pixel.

Suppose there is no target in a certain pixel. In this case, only noise is measured by the $N$ receivers since there is no reflected signal from targets. Then, the covariance matrix can be obtained by

$$\Sigma_{\eta} = \frac{1}{N+1} \Phi \Phi^H.$$  \hfill (3.26)

From random matrix theory, the covariance matrix of the received signals, $\Sigma_{\eta}$ has a Wishart distribution [56] when there is no target in an area, since each element of the sample covariance matrix is the sum of random variables drawn from i.i.d. normal distribution. We restate as Theorem 1, the result of [57] as follows.

**Theorem 2.** Suppose $A$ is a $n \times p$ matrix, each element of each row is drawn from $p$-variate IID normal distribution with zero mean and covariance $R$, $N_p(0, R)$. If we define $B = A^T A$, $B$ is said to be a Wishart matrix which has Wishart distribution with $n$ degrees of freedom, $\mathcal{W}_p(R, n)$. The joint probability density function (PDF) of $B$ is given as

$$\frac{|B|^{(n-p-1)/2}}{2^{pn/2} \Gamma(n/2) |R|^{n/2}} \exp \left( -\frac{1}{2} tr \left( R^{-1}B \right) \right),$$  \hfill (3.27)

where $\Gamma(x)$ is the gamma function of $x$, and notation $tr$ is trace.
Following in Wishart’s footsteps, the density function of eigenvalues (ordered) of Wishart matrix was developed by James in [58]. However, it is complicated to be expressed numerically and there is no known closed form solution. Thus, instead of Wishart matrix analysis, the principal component analysis (PCA) can be utilized for obtaining an adaptive threshold for a pixel by pixel detection. In this chapter, we focus on the convergence properties of largest and smallest eigenvalues of a sample covariance matrix.

The asymptotic distribution of the largest eigenvalue in Wishart matrix have been center and scaling constants [59] 60, and the distribution converges to Tracy-Widom distribution. For the Tracy-Widom distribution, $F_\beta$ stands for the limiting cumulative distributions of the largest eigenvalues in three types of ensembles, viz Gaussian orthogonal ensemble, Gaussian unitary ensemble, and Gaussian symplectic ensemble [61, 62]. In the complex signal case, the value of $\beta$ is two for the Gaussian unitary ensemble (GUE) case which means that the Tracy-Widom distribution with order 2 is defined as

$$F_2(s) = \exp \left( - \int_s^\infty (x-s) q^2(x) \, dx \right). \tag{3.28}$$

In (3.28), $q(s)$ is the unique solution to the Painlevé II equation given by $\frac{d^2q}{dx^2} = sq + 2q^3$ with boundary condition $q(x) \sim Ai(x)$ as $x \to \infty$ where $Ai(x)$ denotes the Airy function [61, 62].

**Theorem 3.** Define the normalized sample covariance as $\tilde{\Sigma} = \alpha \Sigma_n$ with a scaling factor $M/\sigma_n^2$. Assume $\lim_{M \to \infty} \frac{N+1}{M} = \alpha$, where a constant $\alpha$ is defined as $0 < \alpha < 1$. Also, let $\mu = (\zeta + \xi)^2$ and $\omega = (\zeta + \xi) \left( \frac{1}{\zeta} + \frac{1}{\xi} \right)^{1/3}$ where $\zeta$ and $\xi$ denote $\sqrt{M}$ and $\sqrt{N+1}$, respectively. This setting is to apply Tracy-Widom distribution, following which $\left( \lambda_{\max}(\tilde{\Sigma}) - \mu \right)/\omega$
converges to Tracy-Widom distribution of order 2 \cite{59, 60}.

This implies that the largest eigenvalue of $\Sigma$ is $\frac{\sigma_n^2}{\zeta}(\zeta + \xi)^2$ as $M \to \infty$.

In \cite{63}, Bai and Yin found the limit of the smallest eigenvalue of a sample covariance matrix as the following theorem.

**Theorem 4.** Assume $\lim_{N \to \infty, M \to \infty} \frac{N+1}{M} = \alpha$, where a constant $\alpha$ is defined as $0 < \alpha < 1$.

With the above condition, the smallest eigenvalue converges to a certain value, that is

$$\lim_{N \to \infty, M \to \infty} \lambda_{\min}(\tilde{\Sigma}) \to \sigma_n^2(1 - \sqrt{\alpha})^2 \ [63].$$

All eigenvalues of the sample covariance matrix $\Sigma$ from $\mathbf{x}$ and $\{y_i\}_{i=1}^N$ have the same value as $\sigma_n^2$, i.e., $\lambda_1 = \lambda_2 = \cdots = \lambda_{N+1} = \sigma_n^2$. When targets are present, $\lambda_1 > \lambda_{N+1} = 1$.

Hence, we can determine the target existence in a pixel. To compute a threshold, we start with the predefined false alarm probability $P_{fa}(P_{fa} = P(\lambda_{\max} > \gamma_{thr}\lambda_{\min}))$. From this criterion, the threshold $\gamma_{thr}$ can be obtained based on these two properties of the eigenvalue convergence. The pre-defined false alarm probability can be reformulated as

$$P_{fa} = P(\lambda_{\max} > \gamma_{thr}\lambda_{\min})$$

$$= P\left(\frac{\sigma_n^2}{\zeta^2}\lambda_{\max}(\tilde{\Sigma}) > \gamma_{thr}\lambda_{\min}(\tilde{\Sigma})\right)$$

$$= P\left(\frac{\sigma_n^2}{\zeta^2}\lambda_{\max}(\tilde{\Sigma}) > \gamma_{thr}\sigma_n^2\left(1 - \frac{\xi}{\zeta}\right)^2\right)$$

$$= P\left(\frac{\lambda_{\max}(\tilde{\Sigma}) - \mu}{\omega} > \gamma_{thr}(\zeta - \xi)^2 - \mu\right)$$

(3.29)

where $\mu$ and $\omega$ are defined in Theorem 2.

In (3.29), since the left hand side of the inequality can be expressed using the com-
plementary CDF of Tracy-Widom distribution of order 2, we have

\[ P_{fa} = 1 - F_2 \left( \frac{\gamma_{thr}(\zeta - \xi)^2 - \mu}{\omega} \right). \]  

(3.30)

Taking the inverse function of Tracy-Widom distribution, we have

\[ F_2^{-1}(1 - P_{fa}) = \frac{\gamma_{thr}(\zeta - \xi)^2 - \mu}{\omega}. \]  

(3.31)

Substituting \( \mu \) and \( \omega \) into (3.31) and simplifying, we obtain the threshold as

\[ \gamma_{thr} = \left( \frac{(\zeta + \xi)^2}{(\zeta - \xi)^2} + \frac{F_2^{-1}(1 - P_{fa})(\zeta + \xi)^{-2/3}}{(\zeta \xi)^{1/3}(\zeta - \xi)^2} \right). \]  

(3.32)

Utilizing the proposed total correlation method, each pixel in a certain area has positive values for the total correlation. Then, applying the above threshold, target existence can be determined. Note that the proposed threshold is a function of \( M, N, \) and \( P_{fa}. \)

It is of interest to determine a threshold value below which target detection becomes unreliable, especially at low SNRs. However, determining such a threshold value is cumbersome and computationally intensive. In addition, this is scenario-specific, as stated above.
3.5 Simulations

3.5.1 Detection performances

In order to evaluate the proposed detection algorithm, we conduct simulations for multiple target detection at intermediate and low SNR environments. We assume that three point targets are randomly distributed in an area and there are one transmitter and 10 receiver antennas \((N)\) as shown. The transmitted signal is a 1000 pseudo Gaussian random sequence \((M)\) (i.i.d.), and \(P_f\) is set 0.02. The closed-form representation of the inverse of the Tracy-Widom distribution has not been found. However, [59] provides tables for the functions which are various orders of Tracy-Widom distribution. Thus, we utilize the value of the inverse Tracy-Widom distribution to verify detection performance of the proposed scheme.

To test whether the two competing detection methods can detect multiple targets, we assume that SNR is 20 dB, which of course is a high value. With this setting, the conventional total correlation (TC) can successfully detect three targets shown in Figure 3.1 with the normalized total correlation values. The x-axis and y-axis units are scenario-specific and context-specific, and could denote target locations in 2-D space in meters. On the other hand, the proposed method, the modified total correlation (MTC) attains better target detection performance than the conventional method shown in Figure 3.2, since maximum to minimum ratio is larger than that of the conventional total correlation value.

Further, in order to compare the detection performance between two methods for closely located targets, two targets are manually located at a close distance between grid
Figure 3.1: Three target detection using TC at 20-dB SNR.

line $x = 32$ and $x = 33$. Figure 3.3 and 3.4 show the performance utilizing TC and MTC, respectively. All other assumptions are the same as that of the above simulations. In Figure 3.3, the two targets appear as a single large target when utilizing the conventional TC method. MTC however can unambiguously detect two targets as shown in Figure 3.4. Thus, we observe that the proposed detection scheme provides better resolution performance compared to that of TC.

For the low SNR scenario, we assume that SNR is -10 dB with the same system model as the high SNR case. As shown in Figure 3.6a, the conventional detector has relatively high value at the around targets due to the high noise variance. However, the proposed method outperforms the conventional method in Figure 3.6b, since maximum
to minimum ratio is larger than the conventional total correlation value. This implies that the proposed scheme can separate closely located targets.

### 3.5.2 Adaptive Thresholding

In section IV, we proposed a threshold for multiple target, $\gamma_{thr}$. In order to verify the detection performance, the threshold is applied to the result of Figure 3.6b to determine target existence in a certain pixel. In this simulation, one and zero for a location indicate the target presence and absence, respectively. As shown in Figure 3.5c shows that the proposed threshold achieves the exact positions of multiple targets. Three target locations have unity and others are zeros. Thus exact target locations can be obtained.
Figure 3.3: Three closely located target detection using TC at 20-dB SNR.

by the proposed threshold at low SNR environment.

Simulations so far are for instantaneous detection performance. To test the overall impact of SNR, we vary the SNR from -20 to +20 dB to obtain how the proposed system operates at the SNR range in terms of probability of detection. The number of targets in a certain area increases from one to six to verify probability of detection for different number of number of receivers. In this case, simulations are performed for, five and 10 receivers (locations of the transmitter and the receivers are fixed). The false alarm probability is a constant, 0.02. Based on this value, the threshold is computed to determine of target existence at a certain location. All the results are averaged over 10,000 Monte Carlo simulations. For each iteration, reflectivity, location of targets, and
random noise signal from the transmitter are generated.

Figures 3.6 and 3.7 show the probability of detection with five and 10 receiver antennas for detecting from one to six targets, respectively. It is seen that the probability of detection increases as the received SNR increases for both scenarios, since performance degradation of target detection is caused by higher noise variance. As we can easily expect, detection with 10 antenna elements is always much better than the five antenna case since detection performance degrades with a lower number of receivers for the same environment due to loss of diversity gain. Interestingly, 10 receiver systems applied with the proposed thresholding method achieves detection probability of close to unity for a typical radar operating range from 10 to 13 dB [64].

Figure 3.4: Three closely located target detection using MTC at 20-dB SNR.
(a) TC

(b) MTC
To evaluate the target detection performance of the proposed thresholding method, the receiver operating characteristics (ROC) relating its detection probability to the false alarm probability is depicted in Figure 3.8. To generate the ROC curve, we consider three different SNRs environments: -10 dB, +5 dB, and +20 dB. We assume that the receiver antenna has five elements. Moreover, in order to compare the impact of the transmitted signal length $M$ to the detection performance, we also consider two different cases for $M$, namely 1000 and 500. Moreover, the ROC curve by the conventional target detector is plotted for comparing detection performances (black diamond line). Generally, the curve moves upward with increasing SNR. This means that the lower the noise variance is the higher the target detection performance. Furthermore, the proposed detection
scheme obtains the better detection performance with the longer $M$.

In order to test the effectiveness of the threshold value given by (3.32), we compare the detection probability with manually varied threshold values with coefficients ranging 0.5 to 5. The threshold from (3.32), $\gamma_{thr}$, is 1.8117 with parameters, $M=500$, SNR=20 dB and $N=5$. We see from Figure 3.9 that as expected, detection probabilities with lower values (coefficients 0.5 and 0.25) are almost identical to the case of $\gamma_{thr}$. On the other hand, in the cases of higher thresholds (coefficients 1.5, 3 and 5), the detection probabilities decrease since the system misses targets which have fluctuating target reflection coefficients. Thus, the threshold obtained using (3.32) yields an upper bound above which the detection probability is degraded.
Figure 3.7: Probability of detection with 10 receivers elements for $P_f = 0.02$.

Figure 3.8: The receiver operating curves for different values of signal to noise ratio with 5 receiver elements.
Figure 3.9: Receiver operating characteristics curves for different thresholds.
Target Detection and
Reconstruction for Compressive
MIMO UWB Noise Radar Imaging

4.1 Introduction

UWB noise radar systems offer high range resolution compared to a narrow band radar, since the spectrum of UWB noise radar covers more than 500 MHz bandwidth with equivalently short pulse duration. Moreover, UWB noise radar systems utilize stochastically random waveforms which are non-repeatedly transmitted. Hence, UWB noise radar waveforms are frequently construed as noise by an adversary. This leads to one of the most important advantages of noise radar systems in that the transmit signal is immune from detection, interference, and jamming [65]. Due to these two advantages, UWB noise radar systems have been considered as a promising technique for high-resolution
target detection and estimation \[4, 2\].

In spite of advance in ADC technologies, a sampling process for the wide information
bandwidth would place a high burden in the ADC circuit due to the high sampling rate.
As a strong candidate for this problem, CS technique has been introduced by Donoho et.
al. \[10\] Thereafter, several papers developed theoretical understanding of CS \[11, 12, 13\].
Those papers made concrete theories of CS, and suggested that CS has strong potential
for various signal processing areas. CS allows us to acquire significantly fewer samples
than that given by the Nyquist sampling theorem, and the original signal can be re-
constructed from the under sampled measurements assuming the signal is sparse in a
representation and information bandwidth is less than the total system bandwidth. Due
to this major advantage of CS, it can be efficiently applied to UWB radar systems. In
particular, UWB noise radar is suitable for CS since the transmit signal is randomly
generated from bell shaped distributions, such as a Gaussian distribution. The suitabil-
ity is derived and shown for UWB noise radar imaging based on Toeplitz matrices\[15\].
In the paper \[15\], Toeplitz matrices constructed as a transmit signal matrix satisfy the
restricted isometry property (RIP) and UWB noise radar waveforms obtain almost iden-
tical reconstruction performance to general CS theories employing random measurement
matrices using phase transition diagram \[66\]. Based on the CS noise radar theory, the
CS system measures only \(M\) samples among \(N\) candidates utilizing a restriction matrix
which has only \(M\) entries with 1s utilized for a naive sample selection, where \(N\) and \(M\)
are the length of the transmit signal length and the number of measured samples at the
receiver, respectively.

The performance bounds of CS are analyzed by a number of researchers. In par-
ticular, reconstruction bounds of CS are shown using information theory [67, 68]. In Ref. [10], two possible models of CS called output and input noise model are considered. Between those two models, the lower bound of the error probability of output noise model is shown in terms of MI between the measurement matrix and the received signal [67]. The relationship between CS and MI is established in Ref. [11]. In the paper of Kirachaiwanich et. al., the information theoretic lower bound of error probability is shown using the eigenvalue properties of covariance matrices. Both papers stated that the error bound of compressive sensing can be expressed as a function of MI between two known vectors with the given information [68].

In this chapter, we extend the conventional compressive noise radar imaging by Shastry et. al.[15] to MIMO scenarios. Moreover, we propose a new sample selection approach. The conventional CS randomly takes samples using a $M \times N$ matrix (random selection method, RSM), while our selection approach compares the values of candidates and selects those $M$ samples that offer selection diversity gain (largest selection method, LSM). Further we propose an adaptive weighting allocation for minimizing the normalized mean square error (NMSE). Our weighting allocation method is based on maximizing MI between the known information such as the transmitted waveform, the received signal, and the target response. This is achieved by formulating a convex optimization problem with the constraint. The solution of this problem is applicable to RSM and LSM, and the weighting allocation with RSM and LSM are hereafter called RSMW and LSMW, respectively.

The chapter is organized as follows. In Section 2, we describe a system model of compressive noise radar imaging and extend it to MIMO scenarios. In Section 3, we
proposed a sample selection strategy and analyze computational complexity. In Section 4, we propose an adaptive weighting allocation scheme for compressive MIMO noise radar. Simulation results illustrate the performance comparisons and improvement of our approaches in Section 5, and the conclusions are presented in Section 6.

4.2 System model

4.2.1 System model of typical radar systems

In typical radar systems, target detection is accomplished by processing the reflected signal from targets with the transmitted waveforms, wherein the returned waveforms at the receiver are represented the transmitted signal convolved with the target scene and suitably delayed to account for the round-trip travel. Compressive sensing (CS) can be efficiently applied to these UWB noise radar systems to detect and estimate multiple target impulse response with the limited number of samples at the receiver side [15, 69].

A MIMO noise radar system we considered is depicted in Figure 4.1. We assume that the transmitter and receiver have $M_T$ and $M_R$ antenna elements, respectively. The $i$th transmitter transmits $N$ length pseudo-noise waveforms. Each transmitter sends
independent and identically distributed (i.i.d.) waveform drawn from a complex Gaussian distribution, \( \mathcal{CN}(0, 1/M_T) \). Let \( s_{ij} \) denote the impulse response corresponding to the target between the \( i \)th transmitter and the \( j \)th receiver, where each \( s_{ij} \) follows a complex Gaussian distribution, \( \mathcal{CN}(0, \sigma_s^2) \) The number of targets in the scene is assumed to be \( K \) which is much less than \( N \), i.e., \( K \ll N \). Then the received waveform at the \( j \)th receive element is given by
\[
y_j(n) = \sum_{i=1}^{M_T} \sum_{k=1}^{K} s_{ij}(k)x_i(n-k) + w_j(n),
\]
where \( x_i(n) \) is the transmitted waveform from the \( i \)th transmit element at time instant \( n \), and \( w_j(n) \) is the additive complex Gaussian noise at the \( j \)th element, drawn from \( \mathcal{CN}(0, \sigma_w^2) \). Since we assume that the length of waveform is \( N \), (4.1) can be written using matrix notations as
\[
y_j = \sum_{i=1}^{M_T} X_i s_{ij} + w_j,
\]
where
\[
Y_j = \begin{bmatrix} y_j(n) y_j(n+1) \cdots y_j(n+N-1) \end{bmatrix}^T, \quad (4.3)
\]
\[
s_{ij} = \begin{bmatrix} s_{ij}(n) s_{ij}(n+1) \cdots s_{ij}(n+N-1) \end{bmatrix}^T, \quad (4.4)
\]
\[
w_j = \begin{bmatrix} w_j(n) w_j(n+1) \cdots w_j(n+N-1) \end{bmatrix}^T. \quad (4.5)
\]
In (4.2), the transmitted noise waveform matrix from the \( i \)th transmitter, \( X_i \) is
modeled as a $N \times N$ partial Toeplitz matrix given as

$$\mathbf{X}_i = \begin{bmatrix}
  x_i(n) & \cdots & x_i(n - N + 1) \\
  \vdots & \ddots & \vdots \\
  x_i(n + N - 1) & \cdots & x_i(n)
\end{bmatrix}.$$  \hfill (4.6)

Considering $M_T$ transmitters, (4.2) can be rewritten as

$$\mathbf{y}_j = \mathbf{Xs}_j + \mathbf{w}_j,$$  \hfill (4.7)

where

$$\mathbf{X} = \left[ \mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_{M_T} \right],$$  \hfill (4.8)

$$\mathbf{s}_j = \left[ \mathbf{s}_1^T \mathbf{s}_2^T \cdots \mathbf{s}_{M_T}^T \right]^T.$$  \hfill (4.9)

Stacking the received signals at $M_R$ receivers, the $N M_R$ received waveform vector from the all antennas is given by

$$\mathbf{y} = \tilde{\mathbf{X}} \mathbf{s} + \mathbf{w},$$  \hfill (4.10)

where

$$\mathbf{y}_{(N M_R) \times (1)} = \left[ \mathbf{y}_1^T \mathbf{y}_2^T \cdots \mathbf{y}_{M_T}^T \right],$$  \hfill (4.11)

$$\mathbf{s}_{(N M_R M_T) \times (1)} = \left[ \mathbf{s}_1^T \mathbf{s}_2^T \cdots \mathbf{s}_{M_T}^T \right]^T.$$  \hfill (4.12)
\[
\begin{aligned}
\mathbf{w}_{(NM_R) \times (1)} &= \left[ \mathbf{w}_1^T \mathbf{w}_2^T \cdots \mathbf{w}_{M_R}^T \right]^T, \\
\end{aligned}
\]

and \( \mathbf{X}_{(NM_R) \times (NM_R M_T)} = \mathbf{I}_{M_R} \otimes \mathbf{X} \), where \( \otimes \) is the Kronecker product. Note that the \( NM_R M_T \times 1 \) target impulse response vector, \( \mathbf{s} \), has only \( K M_T M_R \) non-zero elements.

### 4.2.2 Compressive radar system

Based on the theory of CS, each receiver measures only \( M \) samples among \( N \) candidates, assuming \( M \ll N \) [15, 69]. In order to select \( M \) samples per antenna, a restriction matrix \( \mathbf{R} \) can be utilized which is defined as

\[
\mathbf{R}_{(MM_R) \times (NM_R)} = \text{diag} \left( \left[ \mathbf{R}_1 \mathbf{R}_2 \cdots \mathbf{R}_{M_R} \right] \right),
\]

where a sub-matrix \( \mathbf{R}_j \) is utilized to pick samples at the \( j \)-th antenna. The sub-matrix \( \mathbf{R}_j \) is given by

\[
\mathbf{R}_j = \left[ \mathbf{r}_{1j}^T \mathbf{r}_{2j}^T \cdots \mathbf{r}_{Mj}^T \right]^T, j = 1, 2, \cdots, M_R.
\]

where each row of \( \mathbf{R}_j \) is a \( N \) length vector which has only one non-zero entry at a random index, otherwise zeros. For instance, consider a compressive noise radar system with \( M_T = M_R = 1 \). If the receiver measures the first \( M \) samples among \( N \), the restriction matrix is \( \mathbf{R} = \left[ \mathbf{I}_M \mathbf{0}_{M \times (N-M)} \right] \), where \( \mathbf{I}_M \) is a \( M \times M \) identity matrix. Applying \( \mathbf{R} \), the \( MM_R \times 1 \) received signal vector \( \tilde{\mathbf{y}} \) can be written as

\[
\tilde{\mathbf{y}} = \mathbf{R} \tilde{\mathbf{X}} \mathbf{s} + \tilde{\mathbf{w}},
\]
where $\tilde{w}$ is a $MMR \times 1$ noise vector.

Our goal is to reconstruct the target scene vector $s$ from (4.16) and to detect multiple targets. The ill-posed signal recovery problem can be solved using the various theories and algorithms of CS. In order to obtain the estimated target vector $\tilde{s}$, we utilize the basis pursuit de-noising (BPDN) algorithm first proposed by Chen, et. al [40, 70]. BPDN involves solving the convex optimization problem as

$$\hat{s} = \arg \min_s \frac{1}{2} \left( \| \tilde{y} - RXs \|_2^2 \right) + \lambda \|s\|_1,$$

(4.17)

where $\lambda$ is a penalizing parameter.

4.3 Sample selection

4.3.1 Proposed sample selection strategy

According to the theories of CS radar, the samples are randomly selected at the receiver to reconstruct the target impulse response [37, 71, 72]. Stochastic noise radar waveforms generated from random distributions satisfy the restricted isometry property (RIP) [37, 15]. However, there exists a finite non-zero probability that the instantaneous transmitted energy of a truly random waveform may at times be quite low, resulting in a correspondingly low reflected signal energy. If the CS receiver measures low energy samples, the recovery accuracy is generally degraded. In order to select the best $M$ samples per antenna in an optimal manner, the reconstruction accuracy has to be computed for $\binom{NMR}{MMR}$ possible combinations assuming $N$ maximum measurements at each receiver.
However, computing the combinations is infeasible in practice, since \( N \) is typically not a small number. In this section, in an effort to address this problem, we propose a sample selection method, LSM.

Let us denote an index set called RSM and another index set called LSM by \( C_R \) and \( C_L \), respectively, with both sets containing \( MM_R \) components. As mentioned previously, \( C_R \) is composed of \( MM_R \) random indices from a set of \( \{1, 2, \cdots, NM_R\} \) as general compressive radars acquire samples by random selection. On the other hand, after finding the set of LSM, \( C_L \), the proposed sample selection for CS MIMO UWB noise radar starts with first defining a temporal set, \( T = \{1, 2, \cdots NM_R\} \), and then removing one sample index per step to this set. In each step, one sample is removed from \( T \), which increases the average energy of the received signal vector. Then, \( C_L \) is finally obtained by excluding the remained indices of \( T \). Assuming perfect knowledge of \( s \) and \( \tilde{X} \) at the receiver, the proposed LSM computes the values of \( \beta_i \) as

\[
\beta_i = \left| \tilde{X}_i s \right|, \quad i = 1, 2, \cdots, NM_R, \quad (4.18)
\]

where \( \tilde{X}_i \) is the \( i \)th row of \( \tilde{X} \). Define an index variable \( I \). Finding \( I \) is equivalent to obtaining

\[
I = \arg \max_{i \in T} \beta_i, \quad (4.19)
\]

\[
T = T - \{I\}, \quad (4.20)
\]

Then, by repeating (4.19)-(4.20) \( MM_R \) times, \( T \) has \( (N - M)M_R \) elements, and we
Table 4.1: Proposed largest selection method (LSM) and the complexity corresponding to each line of the algorithm.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ( \mathcal{T} := {1,2,\ldots,NM_R} )</td>
<td>( \mathcal{O}(N^2M_TM_R^2) )</td>
</tr>
<tr>
<td>2 ( \text{for } i := 1 \text{ to } NM_R )</td>
<td></td>
</tr>
<tr>
<td>3 ( \beta_i :=</td>
<td>\bar{X}_i s</td>
</tr>
<tr>
<td>4 end</td>
<td></td>
</tr>
<tr>
<td>5 ( \text{for } i := 1 \text{ to } MM_R )</td>
<td>( \mathcal{O}(NM_M^2) )</td>
</tr>
<tr>
<td>6 ( \mathcal{I} := \arg \max_{i \in \mathcal{T}} \beta_i )</td>
<td></td>
</tr>
<tr>
<td>7 ( \mathcal{T} := \mathcal{T} - {\mathcal{I}} )</td>
<td></td>
</tr>
<tr>
<td>8 end</td>
<td></td>
</tr>
<tr>
<td>9 ( \text{return } \mathcal{C}_L := {1,2,\ldots,NM_R} - \mathcal{T} )</td>
<td></td>
</tr>
</tbody>
</table>

finally obtain \( \mathcal{C}_L \) as

\[
\mathcal{C}_L = \{1,2,\ldots,NM_R\} - \mathcal{T}. \tag{4.21}
\]

4.3.2 Complexity analysis

The proposed LSM is summarized in Table 1 with the complexity corresponding to each step of the algorithm. In order to compare the computational complexity between the two sample selection methods, RSM and LSM, we only consider the complexities of two approaches, not the entire recovery process of CS, since the complexity for recovering \( \hat{s} \) is identical for both cases.

Without loss of generality, we assume that the length of original signal \( N \) is much larger than \( M \) based on CS theories. Moreover, since the number of measurements at each receiver is greater than the number of transmitters or receivers in practice, we assume that \( M \gg M_T \) or \( M_R \). Thus, \( N \gg M \gg M_T \) or \( M_R \). With these assumptions, it is clear that \( \mathcal{C}_R \) is obtained without computations due to random selection. On the other hand, LSM is based on sorting candidate samples and selecting \( MM_R \) samples which
leads to additional complexity. As shown in Table 1, LSM computes values of $NM_R$ samples, and each multiplication needs $O(NM_TM_R)$. Thus, the complexity of (4.18) is $O(N^2MTM_R^2)$. In (4.19)-(4.20), $NM_R$ samples are compared and $MM_R$ samples are selected with the complexity $O(NMM_R^2)$. Thus, the overall complexity of LSM is $O(N^2MTM_R^2)$. Therefore, LSM clearly requires higher complexity than that of RSM. However, LSM obtains selection diversity gain which improves the recovery accuracy and target detection probability. Performance improvements will be shown in Section 5.

4.4 An adaptive weighting allocation

A major advantage of compressive sensing noise radar is that multiple targets can be detected with significantly less measurements than that given by the Nyquist sampling theorem. So far, $M$ indices per antenna can be selected based on RSM or LSM for MIMO noise radar systems. The selected samples are governed by $R$ that has exactly one entry of 1 in each row and 0s elsewhere. However, there is an additional degree of freedom to allocate proper weighting values with respect to the selected samples at each receive element. In this section, we formulate the problem of weighting allocation for $M$ samples as a constrained convex optimization problem that can be efficiently solved based on a criterion of MI [73].

The probability of error bound of CS can be expressed as a function of MI [67, 68]. From the reconstruction performance limits shown in the papers [67, 68], maximizing MI results in minimizing the error probability of CS. In particular, since the transmitted signals are drawn from a Gaussian distribution in compressive MIMO noise radar
Figure 4.2: A block diagram of the proposed adaptive weighting allocation systems, we can improve the minimum error probability.

A block diagram of adaptive weighting is shown in Figure 4.2 assuming that the transmitted waveform matrix and the reconstructed target impulse response for weighting calculation are perfectly known at the receiver. Note that the samples are selected by RSM or LSM, and the selected indices are utilized for the proposed weighting allocation.

Let us define $k = \tilde{X}_s$. Then, the well known MI between $\tilde{y}$ and $k$ given $R$ is given as

$$I (\tilde{y}; k|R) = h (\tilde{y}|R) - h (\tilde{y}|k, R)$$

$$= h (\tilde{y}|R) - h (\tilde{w}) , \quad (4.22)$$

where $h(\cdot)$ and $h(\cdot|\cdot)$ denote the differential entropy and the conditional entropy, respectively.

Since the measurements at the receivers $\tilde{y}$ given $R$ are Gaussian distributed, i.e., $\mathcal{CN}(0, R\Sigma_k R^H + \sigma^2 I_{MMR})$ where the covariance matrix $\Sigma_k = kk^T$, (4.22) can be rewrit-
\[
I(\mathbf{y}; \mathbf{k} | \mathbf{R}) = \frac{1}{2} \log_2 \left| \mathbf{R} \Sigma_k \mathbf{R}^H + \sigma_w^2 \mathbf{I}_{NM_R} \right| - \frac{1}{2} \log_2 \left| \sigma_w^2 \mathbf{I}_{NM_R} \right|
\]

\[
= \frac{1}{2} \log_2 \left| \mathbf{R} \Sigma_k \mathbf{R}^H \sigma_w^{-2} + \mathbf{I}_{NM_R} \right|, \quad (4.23)
\]

where \(| \cdot |\) is the determinant.

Using Sylvester’s theorem, i.e., \(| \mathbf{I}_m + \mathbf{A} \mathbf{B} | = | \mathbf{I}_n + \mathbf{B} \mathbf{A} |\), (4.23) can be rewritten as

\[
I(\mathbf{y}; \mathbf{k} | \mathbf{R}) = \frac{1}{2} \log_2 \left| \Sigma_k \mathbf{R}^H \mathbf{R} \sigma_w^{-2} + \mathbf{I}_{NM_R} \right|, \quad (4.24)
\]

where \(\mathbf{R}^H \mathbf{R}\) is the \(N M_R \times N M_R\) diagonal matrix with diagonal entries with 0 or 1.

The each receive antenna measures only \(M\) samples corresponding to \(C_R\) or \(C_L\). In (4.24), \(\mathbf{R}^H \mathbf{R}\) has \((N - M)M_R\) zero entries which do not affect to the MI. Using the property of \(\mathbf{R}^H \mathbf{R}\), i.e., \(\text{tr}(\mathbf{R}^H \mathbf{R}) = M M_R\), we introduce a new diagonal matrix, \(\mathbf{\tilde{R}} = \text{diag}([r_1, r_2, \cdots, r_{MM_R}])\) which satisfies \(\text{tr}(\mathbf{\tilde{R}}) = \sum_{i=1}^{MM_R} r_i = M M_R\), where \(\text{tr}(\cdot)\) denotes the trace operator. In the same manner, we define the effective covariance matrix \(\mathbf{\tilde{\Sigma}}_k = \mathbf{\tilde{k}} \mathbf{\tilde{k}}^H\), where we stack entries of \(\mathbf{k}\) by \(\mathbf{C} = [c_1, c_2, \cdots, c_{MM_R}]\) corresponding to \(C_R\) or \(C_L\) into a vector \(\mathbf{\tilde{k}}\), i.e., \(\mathbf{\tilde{k}} = [k_{c_1}, k_{c_2}, \cdots, k_{c_{MM_R}}]^T\).

Then, we obtain

\[
I(\mathbf{\tilde{y}}; \mathbf{\tilde{k}} | \mathbf{\tilde{R}}) = \frac{1}{2} \log_2 \left| \mathbf{\tilde{R}} \mathbf{\tilde{\Sigma}}_k \sigma_w^{-2} + \mathbf{I}_{MM_R} \right|. \quad (4.25)
\]

Typical compressive MIMO radar systems randomly pick \(M\) samples per antenna using \(\mathbf{R}\), i.e., \(r_i = 1, i = 1, 2, \cdots, M\). In this chapter, we propose a method to adaptively
compute the optimum weighting values of $R$ which improves NMSE of compressive MIMO noise radar. The optimization problem can be formulated to find $\tilde{R}$ such that which maximizes the MI of (4.25), as

$$\max I \left( \tilde{y}; \tilde{k} | \tilde{R} \right)$$

subject to $\tilde{R}$ is a diagonal matrix,

$$r_i \geq 0, \quad i = 1, \cdots, MM_R,$$

$$\text{tr}(\tilde{R}) = MM_R. \quad (4.26)$$

The maximization of the MI in (4.26) is a constrained optimization problem, and this can be solved by applying of Lagrange multipliers. In order to solve the problem, we firstly form

$$L(\tilde{R}, \alpha, \mu) = -\log_2 \left| \tilde{R} \tilde{\Sigma}_k \sigma_w^{-2} + I_{MM_R} \right| - \text{tr}(\tilde{R} D_\alpha) + \mu(\text{tr}(\tilde{R}) - MM_R),$$

$$\alpha_i \geq 0, \mu \geq 0, D_\alpha = \text{diag}(\alpha_i). \quad (4.27)$$

Since the objective function is a convex function on $\tilde{R}$, (4.27) is a convex optimization problem. Deriving partial derivatives with respect to $\tilde{R}$, and applying the Karush-Kuhn-Tucker (KKT) conditions for the globally optimum solutions on $\tilde{R}$, we have

$$\frac{\partial L(\tilde{R}, \alpha, \mu)}{\partial \tilde{R}} = 0,$$

$$\alpha_i \tilde{R}_i = 0, \quad i = 1, 2, \cdots, MM_R. \quad (4.28)$$
\[ \mu \left( \text{tr}(\tilde{R}) - MM_R \right) = 0. \quad (4.30) \]

Using \( \frac{\partial}{\partial x} \log |A + BC| = C(A + BC)^{-1}B \), (4.28) is given by

\[ \tilde{\Sigma}_k (\tilde{R} \tilde{\Sigma}_k \sigma_w^{-2} + I_{MM_R})^{-1} - D_\alpha + \mu I_{MM_R} = 0. \quad (4.31) \]

Using (4.29)-(4.31), we obtain the following solution,

\[ r_i = \left( \mu - \frac{\sigma_w^2}{\tilde{\Sigma}_{kii}} \right)^+, \quad (4.32) \]

where \( \mu \) is chosen according to \( \text{tr}(\tilde{R}) = MM_R \) and \( (x)^+ = \max[0, x] \). The obtained weighting values from (4.32) are applied to (4.16) taking the square root of each element.

However, although we have the \( MM_R \) non-zero entries for \( \tilde{R} \), there is no certainty that (4.32) provides \( MM_R \) optimal solutions. In other words, if some entries are below a certain level, (4.32) allocates zeros to them which discards the samples measured for recovery. According to Ref. [8], the required samples for accurate reconstruction is at least \( O(S \log N) \) for \( S \) sparsity scenarios [15]. Therefore, the number of \( r_i > 0 \) for \( i = 1, 2, \cdots, MM_R \) should be more than \( O(KM_T M_R \log N M_T M_R) \).

### 4.5 Simulations

In order to evaluate the performance of the proposed sample selection and weighting allocation methods, several simulations are presented for different scenarios. For various scenarios, we assume that three point targets are randomly distributed in an area and that there are \( M_T \) transmitters and \( M_R \) receivers. For simulations, we consider \( M_T = \)
$M_R = 1$ or $2$. The transmitted noise radar waveform is a 100 pseudo i.i.d. Gaussian random sequence ($N$), and the number of measured samples per antenna ($M$) varies between 10 and 50. The number of sparse targets in the scene ($K$) is three with target impulse responses drawn from a complex normal distribution, $\mathcal{CN}(0,1)$. All entries of $s$, $\bar{X}$ and $\bar{w}$ are independently generated, and simulations are conducted using Monte Carlo method averaged over 3,000 independent repetitions. We define SNR in dB scale for the compressive MIMO UWB noise radar systems as

$$\text{SNR} = 10\log_{10}\left(\frac{E\|\bar{X}\|^2}{\sigma_w^2}\right) \text{dB}. \quad (4.33)$$

### 4.5.1 Reconstructed images

In order to visualize the original and reconstructed target scene by RSM and LSM with $M_T = M_R = 1$ or $2$ at 20-dB SNR, we assume that three targets with random target impulse responses are randomly located in a $10 \times 10$ grids target scene. The x-axis and y-axis units are scenario-specific and context-specific, and could denote target locations in 2-D space in meters. With this setting, Figure 4.3a shows the original target scene, while the reconstructed images by RSM and LSM with $M_T = M_R = 1$ are shown in Figure 4.3b and 4.3c, respectively. It is seen that the conventional RSM and LSM can reconstruct the original scene with only 30% of samples. Note that the sparsity $K$ of $s$ is 3 for the single antenna case, while that of $M_T = M_R = 2$ is $KM_T M_R = 12$, and the length of $\tilde{y}$ for $M_T = M_R = 2$ is $MMR = 60$. Even though the sparsity of $M_T = M_R = 2$ is larger than the single antenna case, $s$ can be successfully recovered as shown in Figure
Figure 4.3: Comparison of reconstructed images by RSM and LSM at 20-dB SNR with $N = 100, M = 30, K = 3$, and $M_T = M_R = 1$ or 2. (a) Original target scene. (b) RSM with $M_T = M_R = 1$. (c) LSM with $M_T = M_R = 1$. (d) RSM with $M_T = M_R = 2$. (e) LSM with $M_T = M_R = 2$. 
4.3d and 4.3e. Note that the values of target pixels with $M_T = M_R = 2$ are higher than those of $M_T = M_R = 1$ which implies that multiple antennas provide clear reconstructed images compared to the single antenna case.

For the low SNR scenario, we assume that SNR is 5 dB with the same system model as the high SNR case. As shown in Figure 4.4b and 4.4c, RSM and LSM with $M_T = M_R = 1$ can successfully recover two targets with relatively high reflection coefficients at $x = 7, y = 3$ and $x = 9, y = 7$. However, the target pixel at $x = 2, y = 7$ with a relatively small impulse response value may not be discriminated with from other non-target pixels. Using LSM, the recovered value is higher than that of RSM which implies that LSM attains better reconstruction performance. Figure 4.4d and 4.4e show the recovered images by RSM and LSM with $M_T = M_R = 2$, respectively. We can see that $M_T = M_R = 2$ can unambiguously detect multiple targets in the low SNR regime which provides better resolution performance compared to that of $M_T = M_R = 1$.

4.5.2 Reconstruction accuracy

To compare the reconstruction accuracy of four schemes, RSM, LSM, RSMW and LSMW according to $M$, we utilize the normalized mean square error (NMSE) defined as

$$\text{NMSE} = \mathbb{E} \left[ \frac{\| s - \hat{s} \|^2}{\| s \|^2} \right].$$

(4.34)

Figure 4.5 depicts NMSE results of four schemes with $N = 100$, $K = 3$, and $M_T = M_R = 1$, and we vary the number of samples at each receiver, $M$ to 10, 30 and 50. From Figure 4.5a, 4.5b and 4.5c, LSM achieves more accurately recovered signal compared to
Figure 4.4: Comparison of reconstructed images by RSM and LSM at 5-dB SNR with $N = 100, M = 30, K = 3$ and $M_T = M_R = 1$ or 2. (a) Original target scene. (b) RSM with $M_T = M_R = 1$. (c) LSM with $M_T = M_R = 1$. (d) RSM with $M_T = M_R = 2$. (e) LSM with $M_T = M_R = 2$. 
Figure 4.5: Comparison of NMSE with $N = 100, K = 3, M_T = M_R = 1$ and $M = 10, 30$ or 50. (a) $M = 10$. (b) $M = 30$. (c) $M = 50$. 
Figure 4.6: Comparison of NMSE with $N = 100$, $K = 3$, $M_T = M_R = 2$ and $M = 10, 30$ or $50$. (a) $M = 10$. (b) $M = 30$. (c) $M = 50$. 
RSM throughout the entire SNR range. Moreover, the proposed weighting allocation
method improves the NMSEs of RSM and LSM, and the reconstruction enhancement is
larger at low SNR regimes. Interestingly, we observe that the smaller number of samples
at the receiver provides the lower NMSE at low SNR regimes than that of high SNR.
This is because each sample suffers high noise variance at low SNRs which leads to
recovery performance degradation. However, as shown Figure 4.5a, the small number
of samples brings high NMSE at high SNR due to lack of samples, while the larger M
provides more accurate reconstruction at high SNR since the reconstruction accuracy
relies on M. As shown in Figure 4.6, the NMSE trends of $M_T = M_R = 2$ are the same
as $M_T = M_R = 1$ case. However, at low SNR, NMSEs are lower than those of the single
antenna, while convergence speeds of four schemes are slower. Even though the length
of the original target scene vector increases with the order of $\mathcal{O}(NM_T M_R)$, the length
of measurements is $\mathcal{O}(MM_R)$. In other words, $M_T = M_R = 2$ requires more samples
rather than $\mathcal{O}(MM_R)$ to attain the same NMSE of $M_T = M_R = 1$.

Figure 4.7 shows the detection probability ($P_d$) with $M_T = M_R = 1$, and we vary the
number of samples at each receiver, $M$ to 10, 30 and 50. In each Figure, the black solid
line with cross marker is the $P_d$ with $M = N$ as a reference. As can be noted, $P_d$ of RSM,
LSM, RSMW and LSMW increase as SNR increases in all cases. In the case of $M = 10$
shown in Figure 4.7a, LSM obtains higher $P_d$ than that of RSM over the whole SNR range
due to sample selection diversity. Moreover, the proposed adaptive weighting allocation
in (4.32) provides the additional SNR gain for RSM and LSM, and the obtained gains
are higher at low SNR. As we can easily expect, $P_d$ increases with more samples at the
fixed SNR for all cases. As shown in Figure 4.7a-4.7c, we see that the $P_d$ improvement
Figure 4.7: Comparison of $P_d$ with $N = 100$, $K = 3$, $M_T = M_R = 1$ and $M = 10, 30$ and 50. (a) $M = 10$. (b) $M = 30$. (c) $M = 50$. 

(a)

(b)

(c)
Figure 4.8: Comparison of $P_d$ with $N = 100$, $K = 3$, $M_T = M_R = 2$ and $M = 10, 30$ and 50. (a) $M = 10$. (b) $M = 30$. (c) $M = 50$. 
of the proposed LSM with \( M = 10 \) is higher than those of \( M = 30 \) or \( M = 50 \). In other words, the SNR gap between RSM(W) and LSM(W) decreases, which means that the selection diversity decreases as \( M \to N \). In Figure 4.8, \( P_d \) with \( M_T = M_R = 2 \) is shown with the same \( M \) cases. We can see that more antennas achieve higher \( P_d \) for each case.

However, \( P_d \) by RSM or RSMW with \( M = 10 \) and \( M_T = M_R = 2 \) is lower than those of \( M_T = M_R = 1 \), since the ratio of \( \mathcal{O}(MM_R)/\mathcal{O}(NM_TM_R) \) of \( M_T = M_R = 2 \) is smaller than that of \( M_T = M_R = 1 \). Note that the proposed LSMW outperforms RSMW due to selection diversity gain, and obtains higher \( P_d \) with \( M = 30 \) or \( M = 50 \) even below \(-10\)-dB SNR than \( P_d \) with \( M = N \), respectively.

Figure 4.9 and 4.10 demonstrate the estimated MI of the four schemes in (4.25) as a function of SNR with \( M_T = M_R = 1 \) and \( M_T = M_R = 2 \), respectively. As shown in the figures, MI increases as the SNR increases, and as \( M_T = M_R \) increases for all schemes. Moreover, the estimated MI increases with more samples per antenna. As we can expect, LSM has higher MI than RSM, and the proposed weighting allocation improves MI of RSM and LSM which enhances NMSE and \( P_d \). Further, the estimated MI is to be close to the MI of \( M = N \) as \( M \) increases, and the proposed weighting method attains higher MI than \( M = N \) at SNR below \(-5\) dB with \( M = 50 \) using \( M_T = M_R = 1 \) or \( M_T = M_R = 2 \).

To test the overall impact of the number of targets (\( K \)) in the certain area, we vary the \( K \) from three to 15 to ascertain how the proposed four methods perform over the \( K \) range in terms of NMSE, \( P_d \) and MI at \(-10\)-dB SNR with \( N = 100, M = 30 \). With those parameters, Figure 4.11 and 4.12 depict NMSE, \( P_d \) and MI with \( M_T = M_R = 1 \) and 2, respectively. As the sparsity increases, MI increases due to the smaller number of
Figure 4.9: Comparison of MI with $N = 100$, $K = 3$, $M_T = M_R = 1$ and $M = 10$, $30$ and $50$. (a) $M = 10$. (b) $M = 30$. (c) $M = 50$. 

(a) 

(b) 

(c)
Figure 4.10: Comparison of MI with $N = 100$, $K = 3$, $M_T = M_R = 2$ and $M = 10, 30$ and 50. (a) $M = 10$. (b) $M = 30$. (c) $M = 50$. 
Figure 4.11: NMSE, $P_d$ and MI according to $K$ at -10-dB SNR with $N = 100$, $M = 30$, and $M_T = M_R = 1$. (a) NMSE. (b) $P_D$. (c) MI.
Figure 4.12: NMSE, $P_d$ and MI according to $K$ at -10-dB SNR with $N = 100, M = 30,$ and $M_T = M_R = 2$. (a) NMSE. (b) $P_d$. (c) MI.
non-zero elements in $s$ as shown Figure 4.11c and 4.12c. Additionally, higher MI leads to lower NMSE with $K$ in 4.11a and 4.12a. However, as we seen in 4.11a and 4.12a, NMSE of each case converges above a certain $K$. It can be inferred that a higher value of $K$ leads to recovery accuracy degradation with a fixed $M$ [40]. Further, in Figure 4.11b and 4.12b, $P_d$ decreases as $K$ increases for both scenarios, and converges with the similar trends of NMSE.

Figure 4.13a and 4.13b depict $P_d$ and MI as a function of $M_T = M_R$ varying one to five. We assume that $N = 100$, $M = 30$ and $K = 3$ at $-10$-dB SNR. As we can easily expect, the more antennas, the better $P_d$ shown in Figure 4.13a. The slopes of LSM and LSMW are steeper than those of RSM and RSMW due to selection diversity. In Figure 4.13b, RSMW and LSMW achieve higher MI than those of RSM and LSM. Fig. 4.13c exhibits the comparison of NMSE. It is clear that applying the proposed weighting allocation method attains lower NMSE than those of RSM and LSM. However, NMSEs of RSMW and LSMW turn to increase from $M_T = M_R = 3$. This is related to the length of original target vector $s$ and the number of measurements. $s$ increases much faster with the order of $O(NM_T M_R)$ than that of the measurements at the receiver with the order $O(M M_R)$. Hence, in order to obtain the more accurately recovered signal, each receiver requires more samples for MIMO.
Figure 4.13: \( P_d \), MI and NMSE according to \( M_T = M_R \) at -10 dB SNR with \( N = 100, M = 30 \) and \( K = 3 \). (a) \( P_d \). (b) MI. (c) NMSE.
Chapter 5

Adaptive Weighting Allocation with the Knowledge of Recovery Error for Compressive Sensing MIMO Noise Radar

5.1 Introduction

In Chapter 4, the target scene information must be known to obtain the optimized restriction matrix, assuming that the CS solver perfectly recovers the original target scene, \( s \), i.e., \( s = \hat{s} \). However, the recovered target scene invariably contains some reconstruction error due to the presence of noise or lack of measurements from a practical perspective, as shown in Figure 5.1. Thus, the restriction matrix obtained by (4.32) with the inaccurate \( \hat{s} \) would not be optimized to \( \hat{X} s \), thereby leading to performance degradations of
the NMSE and target detection. In order to mitigate performance degradation by the recovery error, we propose an adaptive weighting allocation for a reconstruction error environment.

### 5.2 Reconstruction error

#### 5.2.1 System model

In Chapter 4, we assumed that the original signal can be reconstructed by a CS solver. However, it is difficult to recover $s$ from noisy measurement at the intermediate SNR. Therefore, the system model should be modified with the reconstruction error. Considering recovery error, the received signal model is given by

$$\tilde{y} = R\tilde{X}s + \tilde{w} = R\tilde{X}s + R\tilde{X}e + \tilde{w}, \quad (5.1)$$

where $\tilde{s}$ is the recovery of $s$ with zero mean and variance $\sigma^2_s$, and $e$ is the reconstruction error vector with i.i.d. entries distributed according to $CN(0,\sigma^2_e)$. $\sigma^2_s$ is assumed the original target scene with $CN(0,\sigma^2_s - \sigma^2_e)$.
5.2.2 Adaptive weighting allocation with the knowledge of recovery error

As shown in (5.1), $e$ would not contribute to recover $s$ and compute the optimized restriction matrix. Thus, $R\tilde{X} e$ can be considered as an additional noise to $\tilde{w}$. Now, we define a new noise term as

$$\tilde{w}_t = R\tilde{X} e + \tilde{w}. \quad (5.2)$$

We also define the effective covariance matrix of $\tilde{k}_s$ as $\tilde{\Sigma}_{k_s}$, where $\tilde{k}_s = \tilde{X} s$. The covariance matrix of $\tilde{w}_t$ is given by

$$\Sigma_{\tilde{w}_t} = E \left[ (R\tilde{X} e + \tilde{w}) (R\tilde{X} e + \tilde{w})^H \right]$$

$$= RE \left[ \tilde{X} e e^H \tilde{X} e^H \right] + \sigma_w^2 I_{MMR}$$

$$=(\sigma_e^2 \sigma_e^2 NZM T + \sigma_w^2) I_{MMR}$$

$$=(\sigma_e^2 NMR + \sigma_w^2) I_{MMR}. \quad (5.3)$$

Using (5.3), MI with recovery error can be written as

$$I \left( \tilde{y}; \tilde{k}_s | \tilde{R}_e \right) = \frac{1}{2} \log_2 \left| \tilde{R}_e \tilde{\Sigma}_{k_s} \Sigma_{\tilde{w}_t}^{-1} + I_{MMR} \right|, \quad (5.4)$$

where $I \left( \tilde{y}; \tilde{k}_s | \tilde{R}_e \right)$ is a convex function on a new restriction matrix $\tilde{R}_e$. Thus, the maximization problem using (5.4) can be formulated as

$$\max \ I \left( \tilde{y}; \tilde{k}_s | \tilde{R}_e \right) \quad (5.5)$$
subject to \( \mathbf{R}_e \) is a diagonal matrix,

\[
\begin{align*}
\mathbf{r}_{e_{ii}} & \geq 0, \quad i = 1, \ldots, MM_R, \\
\text{tr}(\mathbf{R}_e) &= MM_R.
\end{align*}
\]

Similar to (4.27), we form a Lagrange function defined by

\[
\mathcal{L} \left( \mathbf{R}_e, \alpha, \mu \right) = -\log \left| \mathbf{R}_e \mathbf{\Sigma}_k \mathbf{\Sigma}_w^{-1} + \mathbf{I}_{MM_R} \right| - \text{tr} \left( \mathbf{R}_e \mathbf{D}_\alpha \right) + \mu \left( \text{tr} \left( \mathbf{R}_e \right) - MM_R \right),
\]

\[
\alpha_i \geq 0, \quad \mu \geq 0, \quad \mathbf{D}_\alpha = \text{diag} \left( \alpha_i \right).
\]  \hspace{1cm} (5.6)

Since the objective function is a convex function on \( \mathbf{R}_e \), (5.6) is a convex optimization problem. Deriving partial derivatives with respect to \( \mathbf{R}_e \), and applying the Karush-Kuhn-Tucker (KKT) conditions for the globally optimum solutions on \( \mathbf{R}_e \), we have

\[
\frac{\partial \mathcal{L} \left( \mathbf{R}_e, \alpha, \mu \right)}{\partial \mathbf{R}} = 0,
\]  \hspace{1cm} (5.7)

\[
\alpha_i \mathbf{R}_{e_{ii}} = 0, \quad i = 1, 2, \ldots, MM_R,
\]  \hspace{1cm} (5.8)

\[
\mu \left( \text{tr}(\mathbf{R}_e) - MM_R \right) = 0.
\]  \hspace{1cm} (5.9)

Using \( \frac{\partial}{\partial \mathbf{R}} \log |\mathbf{A} + \mathbf{B} \mathbf{X} \mathbf{C}| = \mathbf{C} (\mathbf{A} + \mathbf{B} \mathbf{X} \mathbf{C})^{-1} \mathbf{B} \), (5.7) is given by

\[
\mathbf{\Sigma}_k \left( \mathbf{R}_e \mathbf{\Sigma}_k \mathbf{\Sigma}_w^{-1} + \mathbf{I}_{MM_R} \right)^{-1} - \mathbf{D}_\alpha + \mu \mathbf{I}_{MM_R} = 0.
\]  \hspace{1cm} (5.10)
Using (5.8)-(5.10), we obtain the following solution as

$$r_{ei} = \left( \mu - \frac{\Sigma \bar{w}_i}{\Sigma k_{si}} \right)^+, \quad (5.11)$$

where $\mu$ is chosen according to $\text{tr}(\bar{R}_e) = MM_R$ and $(x)^+ = \max[0, x]$.

As shown in (5.11), the optimized restriction matrix with the knowledge of $\sigma_e^2$ uses the covariance matrix of the new error instead of $\sigma_w^2$. Therefore, each weighting value adequately reflects the error variance. However, we can easily deduce some performance degradation by the recovery error, which leads to a higher NMSE and lower probability of target detection than those cases with a perfect recovery of $s$.

### 5.3 Simulations

In order to evaluate the performance of the weighting allocation with the $e$, simulations are presented for different scenarios with with $\sigma_e^2 = 0.01$, 0.05, and 0.1. For the various scenarios, we assume that three point targets are randomly distributed in an area and that there are $M_T$ transmitters and $M_R$ receivers. For the simulations, we consider $M_T = M_R = 1$ or 2. The transmitted noise radar waveform is a 100 pseudo i.i.d. Gaussian random sequence ($N$), and the number of measured samples per antenna ($M$) varies between 30 and 50. The number of sparse targets in the scene ($K$) is three with target impulse responses drawn from a complex Gaussian distribution, $CN(0, 1)$. Entries of $s$, $\bar{X}$, and $\bar{w}$ are independently generated, and simulations are conducted using Monte Carlo method averaged over 5,000 repetitions.

In order to test the overall impact of the recovery error, we compare the estimated MI
with the perfect recovery, (4.32), and (5.11). Figure 5.2 depicts the MI comparisons of three schemes with \( N = 100, M = 30, K = 3, \) and \( M_T = M_R = 1 \) or 2. In each figure, as the upper bound, the blue solid line with circles is the MI of RSMW without the recovery error, (4.32), as described in Chapter 4. The MI results of (4.32) with \( \sigma_e^2 \) (dash-dot line) are compared with those of (5.11) for three error variances. First, we observe that the estimated MIs are sensitive to \( \sigma_e^2 \) at higher SNR or utilizing higher number of antennas. This is because \( e \) is overwhelmed by high noise variance at low SNRs. At a high SNR regimes, \( \sigma_w^2 \) is close to zero, which implies that \( e \) only affects MI. Irreducible MI floors verify this observation. As one might expect, higher \( \sigma_e^2 \) causes a lower MI over all SNR ranges in both Figure 5.2a and 5.2b, despite a small amount of \( \sigma_e^2 \). MI degradation is worse with \( M_T = M_R = 2 \), as shown in Figure 5.2b, by using (4.32). However, (5.11) obtains slightly higher MIs than those obtained with (4.32), since the recovery error is considered to compute the optimal weighting values. In other words, the solution, (5.11) treats \( \sigma_e^2 \) as an additional noise. Moreover, the MI gain utilizing (5.11) is higher with more antennas. Figures 5.3a and 5.3b depict the estimated MI with \( M = 50 \). The trends are the same as the results of \( M = 30, \) and the MIs are relatively higher than those in Figure 5.2 due to increased selection diversity. In both figures, we observe that \( \sigma_e \) causes higher MI degradation than those obtained with \( M = 30 \). However, (5.11) achieves higher MI gains than those of \( M = 30 \). In other words, the proposed weighting allocation with \( e \) can compensate for the effect of recovery error.

Figures 5.4a and 5.4b demonstrate NMSE comparisons between (4.32) and (5.11) with \( \sigma_e^2 = 0.01, 0.05, \) and 0.1. The NMSE results of \( N = 100, M = 30, K = 3, M_T = M_R = 1, \) and 2 are compared. As the lower bound of NMSE, RSMW with
Figure 5.2: Comparison of MI for several reconstruction errors, $\sigma_e^2$ with $N = 100$, $M = 30$, $K = 3$, and $M_T = M_R = 1$ or 2 (a) $M_T = M_R = 1$. (b) $M_T = M_R = 2$.

the perfect recovery of $s$ is also compared. First, for all cases, we observe that NMSE increases as $\sigma_e^2$ increases. Moreover, the reconstruction error increases as SNR increases, which implies that $\sigma_e^2$ becomes dominant for recovery accuracies at high SNRs. Also, the
reconstruction error produces an irreducible NMSE floor. Compared to (4.32) without considering $\sigma^2_e$, (5.11) reduces NMSEs over the entire SNR range. Moreover, NMSE gain exploiting (5.11) increases at higher SNR, as shown in Figures 5.4a and 5.4b. Figures
Figure 5.4: Comparison of NMSE for several reconstruction errors, $\sigma_e^2$ with $N = 100$, $M = 30$, $K = 3$, and $M_T = M_R = 1$ or 2 (a) $M_T = M_R = 1$. (b) $M_T = M_R = 2$.

5.5a and 5.5b demonstrate the NMSE results with $M = 50$ at each receive element. The trends are the same as those of $M = 30$, and NMSEs are relatively lower than those in Figure 5.4 due to more measurements. As shown in Figures 5.4a and 5.4b, (5.11)
Figure 5.5: Comparison of NMSE for several reconstruction errors, $\sigma_e^2$ with $N = 100$, $M = 50$, $K = 3$, and $M_T = M_R = 1$ or 2 (a) $M_T = M_R = 1$. (b) $M_T = M_R = 2$.

provides NMSE enhancements for all $\sigma_e^2$ scenarios. However, the recovery performances exploiting (5.11) are slightly higher than those of $M = 30$.

In order to verify $P_d$ enhancement using (5.11) for these same scenarios, we plot
Figure 5.6: Comparison of $P_d$ for several reconstruction errors, $\sigma_e^2$ with $N = 100$, $M = 30$, $K = 3$, and $M_T = M_R = 1$ or 2 (a) $M_T = M_R = 1$. (b) $M_T = M_R = 2$.

$P_d$ vs. SNR in Figures 5.6 and 5.7 using $M = 30$ and $M = 50$, respectively. As one might expect, (5.11) gives higher $P_d$ performances than those obtained with (4.32). Also note that $P_d$ using the proposed method, i.e., with the knowledge of recovery error
Figure 5.7: Comparison of $P_d$ for several reconstruction errors, $\sigma_e^2$ with $N = 100$, $M = 50$, $K = 3$, and $M_T = M_R = 1$ or 2 (a) $M_T = M_R = 1$. (b) $M_T = M_R = 2$.

variances, continuously increases throughout the SNR range, while the results of (4.32) almost saturate over 15-dB SNR for $M_T = M_R = 1$ and $M = 30$, or 10-dB SNR for $M_T = M_R = 2$ and $M = 50$. 
Conclusions and Future Research

This dissertation has been concerned with target detection for UWB noise radar using information theory and compressive sensing. This concluding chapter gives a summary of the results from previous chapters of this dissertation. Furthermore, some suggestions for possible future directions of the current research will be discussed.

6.1 Conclusions

In Chapter 2, we discussed a general signal model of UWB noise radar systems, and described several different types of UWB noise radars developed for multiple target-detection applications. Further, CS as a new paradigm of signal processing has been reviewed. Then, we have discussed the relationship between CS and UWB noise radar systems to develop compressive UWB noise radar systems.

In Chapter 3, a detector based on MI for noise radar systems has been presented. The proposed detection algorithm obtains enhanced discrimination results for multiple targets compared to the total correlation detector at high- and low-SNR environments.
without knowledge of target reflectivity. Further, the proposed scheme using a threshold achieves better detection results compared to the conventional method for the same number of receivers. Random matrix theory has been utilized to set a threshold for multiple target detection. In order to show the performance of the proposed thresholding scheme, simulation results have been done for randomly generated targets. Moreover, the probabilities of target detection and ROC curves based on the proposed thresholding method have been shown by numerical simulations.

In Chapter 4, a sample selection for compressive MIMO noise radar imaging has been proposed. The proposed selection method, based on comparing norm values, obtains the improved $P_d$ of multiple targets for compressive MIMO noise radar. Further, the proposed weighting allocation by formulating and solving a constrained convex optimization problem improves the reconstruction accuracy of CS. In order to show the performance of the proposed selection and weighting method, simulation results have shown that the proposed schemes obtain the improved $P_d$ and NMSE for various compressive MIMO noise radar scenarios.

In Chapter 5, an adaptive weighting allocation with recovery error for compressive MIMO noise radar is proposed. By maximizing the MI of our system, the proposed weighting allocation enhances the recovery accuracies and target detection probabilities considering the recovery error as an addition noise. Simulation results have shown performance improvements, and compared the recovery accuracies and target detection probabilities of a conventional scheme and the modified method.
6.2 Future research

This dissertation presents target detection methodologies of a UWB random noise radar systems utilizing information theory and compressive sensing. Despite this initial success, deeper work should be done to improve the target detection performance in a practical perspective.

In Chapter 3 we have proposed a multi-target detector that uses a concept of total correlation. This method is effective and provides high resolutions for point targets. However, if we consider an extended target, the received signals at multiple receivers are spatially correlated by the target impulse response containing high values of correlation coefficients. To obtain robust detection performance in more practical scenarios, it is important to investigate performance degradation and a possible modification of this detector.

Further, the proposed detector utilizes one transmitter, which uses noise waveform. It is well known that multiple transmitters obtain additional diversity gain, which leads to robust detection performance. Moreover, there is no guarantee that the transmitted noise waveform at a certain time instance has the intended signal energy due to randomness in the properties of noise waveform generators. This topic is worthy of further investigation.

In Chapter 4 we have considered an adaptive weighting allocation scheme that minimizes NMSE of the recovered target impulse response, assuming that the target vector for computing the optimized restriction matrix is perfectly recovered. Based on the theory of CS, the perfect recovery of sparse signals is possible by acquiring as few as $O(K \log(N))$ projections of a signal into an appropriate measurement basis.
Appendix A

A.1 Proof of Lemma 1.

Let \( z = [x^T y^T]^T \) denote zero mean normal random vector \( \mathcal{N}(0, R_z) \), where the covariance matrix \( R_z \) is given by

\[
R_z = E \{zz^H\} = \begin{bmatrix}
R_{xx} & R_{xy} \\
R_{yx} & R_{yy}
\end{bmatrix}.
\]  

(A.1)

where \( R_{xx} = E [xx^H] \) and \( R_{yy} = E [yy^H] \). Moreover \( R_{xy} = E [xy^H] \) and \( R_{yx} = E [yx^H] \).

We have the differential entropy of \( z \) as,

\[
h(z) = h(x, y) = -E \{\ln p(z)\}
\]

\[
= M \log (2\pi e) + \frac{1}{2} \log |R_z| + M.
\]

(A.2)

Moreover, MI can be written as

\[
I(x; y) = h(x) + h(y) - h(x, y)
\]
\[
\frac{1}{2} \log |R_{xx}| + \frac{1}{2} \log |R_{yy}| - \frac{1}{2} \log |R_z|.
\] (A.3)

Now, let \( A \) denote a unitary matrix so that

\[
A^H z = w = \begin{pmatrix} (x - \hat{x}) \\ y \end{pmatrix},
\] (A.4)

where \( \hat{x} = R_{xy} R_{yy}^{-1} y \). Thus, the covariance matrix of \( w \) can be written as

\[
R_w = E \left[ ww^H \right] = A^H E \left[ zz^H \right] A
= A^H R_z A.
\] (A.5)

In (A.5), the determinant of \( R_w \) equals to that of \( R_z \), since \( A \) is a unitary matrix.

Moreover, (A.5) can be written as

\[
R_w = \begin{bmatrix}
E \left[ (x - \hat{x}) (x - \hat{x})^H \right] & E \left[ (x - \hat{x}) y^H \right] \\
E \left[ y(x - \hat{x})^H \right] & E \left[ yy^H \right]
\end{bmatrix}
= \begin{bmatrix}
R_{xx} - R_{xy} R_{yy}^{-1} R_{yx} & 0 \\
0 & R_{yy}
\end{bmatrix}.
\] (A.6)

The off-diagonal terms of (A.6) are zero, since estimation error \( x - \hat{x} \) is orthogonal to \( y \) and vice versa.

Now, let \( P = R_{xx} - R_{xy} R_{yy}^{-1} R_{yx} \). Then, \( P \) can be expanded as

\[
P = R_{xx}^{1/2} \left( I_M - R_{xx}^{-1/2} R_{xy} R_{yy}^{-1/2} R_{yy}^{-1/2} R_{yx} R_{xx}^{-1/2} \right) R_{xx}^{1/2},
\] (A.7)
where $\mathbf{I}_M$ is an identity matrix with $M$ diagonal elements.

Let us define $\tilde{x} = \mathbf{R}_{xx}^{-1/2} x$ and $\tilde{y} = \mathbf{R}_{yy}^{-1/2} y$. The cross correlation matrix between $\tilde{x}$ and $\tilde{y}$ is given by

$$
\mathbf{R}_{\tilde{x}\tilde{y}} = \mathbb{E} [\tilde{x}\tilde{y}^H] = \mathbf{R}_{xx}^{-1/2} \mathbb{E} [xy^H] \mathbf{R}_{yy}^{-1/2}
$$

$$
= \mathbf{R}_{xx}^{-1/2} \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1/2}.
$$

(A.8)

Applying the Singular Value Decomposition (SVD), we can compute $\mathbf{R}_{\tilde{x}\tilde{y}} = \mathbf{USV}^H$.

Using $\mathbf{R}_{\tilde{x}\tilde{y}}$, (A.7) can be written as

$$
\mathbf{P} = \mathbf{R}_{xx}^{1/2} (\mathbf{I} - \mathbf{R}_{\tilde{x}\tilde{y}} \mathbf{R}_{\tilde{x}\tilde{y}}^H) \mathbf{R}_{xx}^{1/2}
$$

$$
= \mathbf{R}_{xx}^{1/2} (\mathbf{I} - \mathbf{US}^2 \mathbf{U}^H) \mathbf{R}_{xx}^{1/2}.
$$

(A.9)

Taking the determinant of (A.9), we have

$$
|\mathbf{P}| = |\mathbf{R}_{xx}^{1/2} (\mathbf{I} - \mathbf{US}^2 \mathbf{U}^H) \mathbf{R}_{xx}^{1/2}|
$$

$$
= |\mathbf{R}_{xx}| |\mathbf{I} - \mathbf{US}^2 \mathbf{U}^H|
$$

$$
= |\mathbf{R}_{xx}| \prod_{i=1}^{M} (1 - s_{ii}^2).
$$

(A.10)

Plugging (A.10) into (A.6) and using the distributive property of determinant, $|\mathbf{R}_w|$ can be written as

$$
|\mathbf{R}_w| = |\mathbf{R}_z| = |\mathbf{R}_{xx}| |\mathbf{R}_{yy}| \prod_{i=1}^{M} (1 - s_{ii}^2).
$$

(A.11)
Finally, (A.3) can be rewritten as

\[ I(x; y) = -\frac{1}{2} \sum_{i=1}^{M} \log (1 - s_{ii}^2) \]  \hspace{1cm} (A.12)

Since \(-1 \leq s_{ii} \leq 1\), we can say that as \(s_{ii} \to -1\) or \(+1\), \(I(x; y) \to \infty\).
Bibliography


Vita
Yangsoo Kwon

Research Interests
Signal processing techniques for radar and wireless communications, signal recovery and sampling for compressive sensing, numerical optimization, detection and estimation.

Education
- The Inha University, M.S. in Electronics Engineering, Feb. 2008.
- The Inha University, B.S. in Electronics Engineering, Feb. 2006.

Publications