Signal Processing Systems
Mekelweg 4,
2628 CD Delft
The Netherlands
https://sps.ewi.tudelft.nl/

SPS-2025-5939526

M.Sc. Thesis

Wideband Ambiguity Function Shaping and Sonar Imaging

Xingyuan Ma B.Sc.

Abstract

Synthetic Aperture Sonar (SAS) is an advanced sonar imaging technique that uses multiple pulses or high duty cycle waveforms from a moving surface vehicle to create a synthetic aperture array for imaging. One of the key challenges of SAS systems is the design of continuous waveforms that enable high-quality imaging under practical constraints (e.g., unimodular constraint). This thesis addresses the design of such waveforms and observes that waveform design for a SAS system over a short time window can be approximated as a wideband ambiguity function shaping problem. To tackle this, we formulate the wideband ambiguity function shaping problem as a non-convex optimization problem and propose four methods to solve it. Among the proposed methods, the wideband gradient descent method is proven to be the most efficient and effective in minimizing the average sidelobe energy in the region of interest of the wideband ambiguity function. Simulation and field trial results show that, although waveform design for SAS systems is not strictly equivalent to wideband ambiguity function shaping, the waveform obtained through this approach still yields good SAS imaging performance compared to conventional sonar waveforms, such as the random Binary Phase-Shift Keying (BPSK) waveform. These findings provide a direction for the waveform design of future SAS systems. Beyond SAS, the proposed wideband ambiguity function shaping methods also show potential for other wideband applications, such as waveform design for underwater target tracking.

Wideband Ambiguity Function Shaping and Sonar Imaging

Thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Science

in

Electrical Engineering

by

Xingyuan Ma B.Sc.
born in Nanchang, China

This work was performed in:

Signal Processing Systems Group
Department of Microelectronics
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

Delft University of Technology

Delft University of Technology
Department of
Microelectronics

The undersigned hereby certify that they have read and recommend to the Faculty of Electrical Engineering, Mathematics and Computer Science for acceptance a thesis entitled **"Wideband Ambiguity Function Shaping and Sonar Imaging"** by **Xingyuan Ma B.Sc.** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: July 30th, 2025

Chairman: _____
prof.dr.ir. Geert Leus

Advisor: _____
prof.dr.ir. Geert Leus

Committee Members: _____
dr.ir. Yanki Aslan

_____
dr.ir. Pim van der Meulen

Abstract

Synthetic Aperture Sonar (SAS) is an advanced sonar imaging technique that uses multiple pulses or high duty cycle waveforms from a moving surface vehicle to create a synthetic aperture array for imaging. One of the key challenges of SAS systems is the design of continuous waveforms that enable high-quality imaging under practical constraints (e.g., unimodular constraint). This thesis addresses the design of such waveforms and observes that waveform design for a SAS system over a short time window can be approximated as a wideband ambiguity function shaping problem. To tackle this, we formulate the wideband ambiguity function shaping problem as a non-convex optimization problem and propose four methods to solve it. Among the proposed methods, the wideband gradient descent method is proven to be the most efficient and effective in minimizing the average sidelobe energy in the region of interest of the wideband ambiguity function. Simulation and field trial results show that, although waveform design for SAS systems is not strictly equivalent to wideband ambiguity function shaping, the waveform obtained through this approach still yields good SAS imaging performance compared to conventional sonar waveforms, such as the random Binary Phase-Shift Keying (BPSK) waveform. These findings provide a direction for the waveform design of future SAS systems. Beyond SAS, the proposed wideband ambiguity function shaping methods also show potential for other wideband applications, such as waveform design for underwater target tracking.

Acknowledgments

This thesis was carried out at Fugro Innovation & Technology B.V. First of all, I would like to express my sincere gratitude to prof.dr.ir. Geert Leus for introducing me to this project and facilitating the connection with Fugro.

I would also like to thank my supervisor at Fugro, dr.ir. Pim van der Meulen, for his invaluable guidance and support throughout my thesis work. I am thankful as well to Kaan Demir for the generous help and assistance during the project. I also appreciate the support provided by dr.ir. Pim van der Meulen and Kaan Demir in processing the field trial data.

Finally, I would like to thank prof.dr.ir. Geert Leus, dr.ir. Pim van der Meulen, and Kaan Demir once again for their thoughtful feedback and suggestions on my thesis, which greatly contributed to its completion.

Xingyuan Ma B.Sc.
Delft, The Netherlands
July 30th, 2025

Contents

List of Figures

List of Tables

Notation and Nomenclature

$(\cdot)^*$	Conjugate of a scalar/vector/matrix
$(\cdot)^H$	Conjugate transpose of a vector/matrix
$(\cdot)^T$	Transpose of a vector/matrix
$\Im(\cdot)$	Imaginary part of a scalar/vector/matrix
$\Re(\cdot)$	Real part of a scalar/vector/matrix
$\|\cdot\|_p$	l_p norm of a vector
$\lambda_i(\cdot)$	The ith largest eigenvalue of a matrix
$\mathrm{Tr}(\cdot)$	Trace of a matrix
$\|\cdot\|$	Absolute value of a scalar/vector/matrix
$(g * h)(t)$	Convolution of $g(t)$ and $h(t)$
$\mathcal{F}(\cdot)$	Fourier transform of a signal
$\langle \mathbf{A}, \mathbf{B} \rangle$	Inner product of matrices $\mathbf{A}$ and $\mathbf{B}$, $\langle \mathbf{A}, \mathbf{B} \rangle = \mathrm{Tr}(\mathbf{B}\mathbf{A}^H)$
$\mathbf{0}$	all zero vector
$\mathbf{0}_n$	$n \times n$ all zero matrix
$\mathbf{I}_n$	$n \times n$ identity matrix
$\mathbf{a} \cdot \mathbf{b}$	Inner product of vector $\mathbf{a}$ and $\mathbf{b}$
$\mathbf{a} \odot \mathbf{b}$	Elementwise product of vector $\mathbf{a}$ and $\mathbf{b}$
B	Bandwidth of a signal
c	Propagation speed of a signal in the medium
f_c	Carrier frequency of a signal

1 Introduction

Active sonar detects and tracks underwater targets, by transmitting continuous[1] or pulsed acoustic signals and listening to the reflections of the transmitted signals. The distance and velocity of the target can be determined by measuring the delay and the Doppler effect of the received signal using the matched filter. The principle of active sonar is similar to active radar. However, due to the high absorption and high attenuation of electromagnetic waves in water, radar systems are not suitable for detection and tracking of underwater targets.

Sonar systems can also be used for imaging targets, such as cables, pipelines, shipwrecks, and boulders. If the target is exposed in the incident field of an underwater acoustic source, it can be imaged by analyzing the scatter field. Typically, sonar imaging systems can be categorized into three categories [2]: sectorscan sonar, sidescan sonar, and Synthetic Aperture Sonar (SAS). The schematic diagram of the three categories is shown in Figure 1.1. Sectorscan sonar usually uses a multibeam array for imaging, which allows it to directly produce a 2-dimensional image for each transmitted pulse. Sidescan sonar only uses a single beam or a few beams for each pulse. With the motion of the vessel to obtain coverage of different areas in the imaging domain, the full image can be derived from combining single pulse images at different areas. SAS can be seen as an advanced imaging technique that combines sectorscan sonar and sidescan sonar, which uses multiple pulses from a moving vehicle to create a synthetic aperture array for imaging. In a SAS system, the intensity of voxels in the imaging domain can be obtained by synthetically integrating over multiple pulses that are reflected by the voxels along a known trajectory. Consequently, SAS normally has a higher imaging resolution compared with other conventional sonar imaging system.

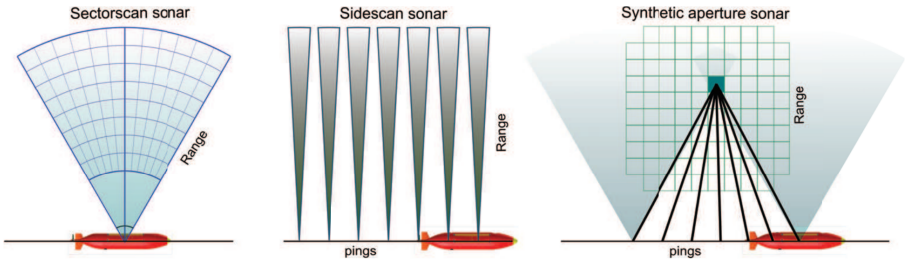


Figure 1.1: Categories of sonar imaging systems: sectorscan sonar, sidescan sonar, and synthetic aperture sonar [2].

A conventional SAS system usually uses high frequency pulses with a carrier frequency around 60-300 kHz [2, 3], which allows the imaging system to achieve cm-level imaging resolution. However, since the high-frequency acoustic wave has higher at-

[1]Continuous signal/wave refers to a signal/wave with 100% duty-cycle in the thesis

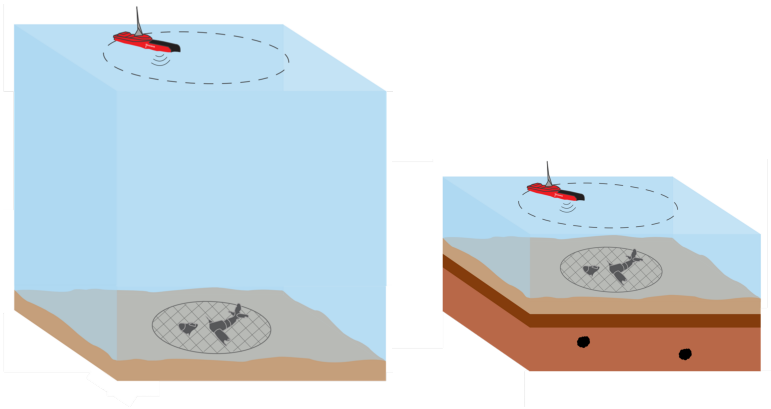
tenuation and higher absorption in an underwater environment compared to the low frequency wave, the high-frequency SAS system needs either a high transmit power or to be deployed near the target of interest to overcome the path loss over a long distance. Moreover, to ensure that the SAS system can integrate over a known trajectory, the navigation accuracy should be higher than at least of a fraction of the wavelength [2, 3]. For example, for a high-frequency SAS system with a carrier frequency of 100 kHz, the navigation accuracy should be at least higher than $\lambda/4 \approx 3.5$ mm. This problem becomes more critical when the SAS system is fully deployed underwater, as the Global Navigation Satellite System (GNSS) is not available in such environment. In the absence of GNSS, the system depends on Inertial Navigation System (INS) and micro-navigation [2, 3] for positioning, which can lead to accumulated positioning errors over a long trajectory. Due to these limitations, high-frequency SAS systems usually cannot integrate over a very long trajectory for imaging.

Recently, Fugro has developed a SAS system, called SASV, a contraction of the Synthetic Aperture Sonar (SAS) and Autonomous Surface Vehicle (ASV), which uses low-frequency[2] continuous or high-duty cycle waves from a moving surface vessel for sonar imaging and bathymetry [1]. The schematic diagram of SASV is shown in Figure 1.2a, where the transmitters and receivers are mounted on the hull of the vessel. The utilization of low-frequency band signals enables the SASV system to image the sea floor a few kilometers beneath the water surface from a surface vehicle. Because of the use of the SAS technique and a high resolution (cm-level accuracy) GNSS system, SASV can achieve a high imaging resolution by integrating over a large synthetic aperture. On the other hand, with the same imaging resolution, the traditional multibeam systems can only image a water depth of around one kilometer and thus have to be deployed using underwater vehicles or towed systems to achieve the same imaging depth, which is more expensive and complicated to deploy. Moreover, for surface-deployed multibeam systems, the imaging resolution depends on the angular resolution of the array. Even a small angular beamwidth can translate into a very large footprint at kilometers of depth. As a result, these systems usually need a very large array to maintain a good angular resolution, which also potentially introduces complexity for the system design.

Additionally, the idea of SASV has been used in another sub-bottom SAS imaging system, called Compact Low-frequency Acoustic Mapping System (CLAMS), which is capable of imaging the boulders 5-10 m below the seafloor using low-frequency continuous waves from a surface vehicle. The schematic of CLAMS is shown in Figure 1.2b. Compared with the conventional sub-bottom profilers which do not use the SAS technique, CLAMS has the potential to provide higher imaging resolution for sub-bottom targets. Another solution for the sub-bottom imaging is the parametric sonar array, which uses directive pulses with high Sound Pressure Level (SPL). The parametric sonar array requires a high instantaneous power for transmitting the pulses and might potentially be harmful to marine life. CLAMS, on the other hand, uses continuous waves for imaging, which spreads the transmit power through the time domain and therefore has lower impact on marine life and imposes lower requirements on transducers and other electronic components.

To improve power efficiency, sonar systems typically use a unimodular signal, which

[2] carrier frequency ranges from 500 Hz to 20 kHz



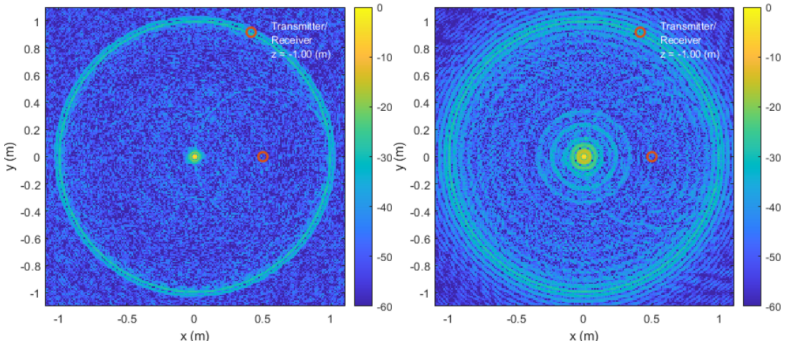
(a) SASV (b) CLAMS

Figure 1.2: Schematic of SAS imaging systems made by Fugro, (a) SASV (b) CLAMS.

has a nearly constant amplitude over the entire transmit duration. In the current SAS system of Fugro, conventional unimodular waveforms, such as random Binary Phase-Shift Keying (random BPSK) signal and Hyperbolic Frequency Modulated (HFM) signal, have been implemented, but their imaging performance is suboptimal [1]. Examples of SAS imaging of a point reflector using random BPSK and HFM are shown in Figure 1.3, where a point reflector at (0 m, 0 m, -1 m) is imaged using a circular trajectory at the $z = 0$ m plane with a radius of 0.5 m. The Point Spread Function (PSF) shows that both these two signals identify the point reflector and circle-shaped ghost artifact around the point reflector caused by the circular trajectory. The sidelobe of random BPSK is constantly high over the entire imaging domain, whereas the sidelobe of HFM is much higher within the circle-shaped ghost artifacts but seems to be lower outside the circle artifacts. If a strong reflector is present, the sidelobe caused by the strong reflector will hide the mainlobe of other small reflectors, which may even result in failed detection. Therefore, waveform design is an important problem for improving the image quality of the SAS system.

The goal of the thesis is to design an optimized waveform that can improve the imaging performance of the SAS system. Due to the use of matched filter techniques in Fugro's SAS system, we have found a high correlation between the ambiguity function property of the waveform and the imaging PSF. That means, the ambiguity function shaping problem plays an important role for the waveform design problem of a SAS system. The ambiguity function shaping is a well-studied problem in radar, but the problem still remains underexplored in the context of sonar because the ambiguity function can no longer be approximated as a narrowband ambiguity function in a wideband sonar scenario. In the thesis, we will design the optimized sequence from the perspective of wideband ambiguity function shaping and then verify the performance of the designed sequence through SAS imaging. The main research question of the thesis can be summarized as:

- *To what extent, can ambiguity function shaping improve the imaging performance of a SAS system?*



(a) PSF for random BPSK (dB) (b) PSF for HFM (dB)

Figure 1.3: The Point Spread Function (PSF) for continuous (a) random BPSK and (b) HFM signals with an Integration time of 5 s [1]. The Pulse Repetition Interval (PRI) of the HFM signal is 50 ms. The bandwidth and the carrier frequency of these two signals are $B = 8000$ Hz and $f_c = 8000$ Hz, respectively. The imaging trajectory is a circle at $z = 0$ m plane with a radius of 0.5 m. The experiments were conducted in air with a speed of sound $c \approx 340$ m/s.

The main research question can be broken down into the following two sub-problems:

- *What is the relationship between the ambiguity properties of a waveform and its corresponding SAS imaging performance? Is ambiguity function shaping equivalent to waveform design for a SAS system?*
- *What is the optimal way to shape the wideband ambiguity function of a sequence for SAS imaging?*

The structure of the thesis is as follows. Firstly, in Chapter 2, the background of the thesis is introduced, including Fugro's SAS imaging systems, the ambiguity function, the conventional sonar waveforms, and a literature review of ambiguity function shaping methods. Here, we also theoretically show the connection between the wideband ambiguity function shaping and the waveform design for a SAS system. Then, four methods for the wideband ambiguity function shaping are proposed in Chapter 3, aiming to determine the optimal approach to shape the ambiguity function of a sequence for SAS imaging. The performance of these methods are evaluated through a series of simulations. Following that, the SAS imaging performance of different waveforms is evaluated through simulations and field trials in Chapter 4. Here, we aim to validate our theoretical conclusion of the relationship between the wideband ambiguity function shaping and waveform design for SAS systems. Finally, in Chapter 5, the main contributions of the thesis are summarized, and suggestions for future work on the waveform design for a SAS system are also provided.

Background

2

In this chapter, we aim to address the first sub-problem of the main research question:

- *What is the relationship between the ambiguity properties of a waveform and its corresponding SAS imaging performance? Is ambiguity function shaping equivalent to the waveform design for a SAS system?*

To partially answer this question, the working principle of the SAS system of Fugro is reviewed in Section 2.1. Here, we show that when the processing interval of a SAS imaging is short enough (e.g., a PRI of the transmit signal), the waveform design problem for the short processing interval can be approximately seen as a wideband ambiguity problem. The relationship between the ambiguity properties and the SAS imaging performance will be further analyzed through simulations and field trials in later chapters. Next, the ambiguity function and its properties are introduced in Section 2.2. Then, the conventional sonar waveforms and their ambiguity function properties are discussed in Section 2.3. Here, the desired ambiguity function properties of a SAS imaging waveform are also mentioned. Finally, a literature review of the current ambiguity function shaping methods is given in Section 2.4.

2.1 SAS imaging system

The SAS imaging systems of Fugro are similar to other conventional SAS systems, but have some differences on the signal processing techniques due to the use of continuous waveforms. To figure out the criteria of a good imaging waveform, the detailed working principle of the SAS system is illustrated in this section.

As shown in Figure 2.1, unlike other SAS systems that process over every transmitted pulse, Fugro's SAS systems implement a technique called Short-Time Matched Filtering (STMF) [1], which segments the continuous transmit signal into short-time windows with a length of T_w and then uses every short-time window as the matched filter to correlate with short segments of the received signal for imaging. To make sure that the time-scaling effect caused by the Doppler effect within the short-time window is negligible, the length of the short-time window T_w is selected such that it satisfies $T_w \ll |\frac{c}{2Bv}|$ [1], where c represents the speed of sound in water, and v represents the speed of the vessel.

Another important aspect of the SAS system is the delay estimation. Let us consider a SAS imaging scenario as shown in Figure 2.2, where a pair of transmitter and receiver are surveying along a line trajectory with a velocity of $\mathbf{v}$ for imaging an area beneath the trajectory. The imaging domain is discretized into a grid of voxels, where each voxel can be seen as a perfect point reflector with a reflection coefficient if the grid resolution is high enough. Assume a short-time window is transmitted at $t = 0$, the transmitted

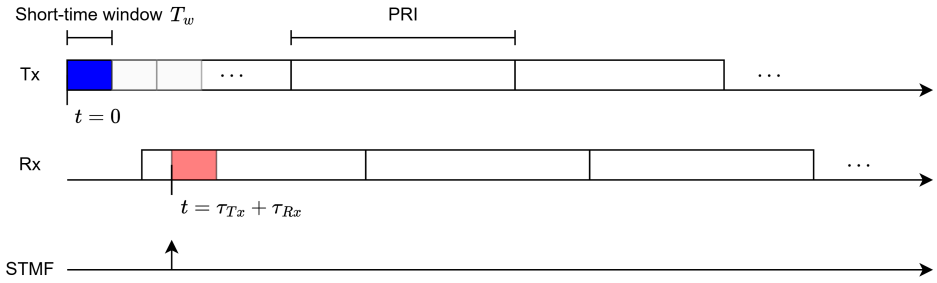


Figure 2.1: Schematic of the short-time matched filtering (STMF) process of the SAS system. The transmit signal is segmented into short-time windows with a window length of T_w. For each short-time window, once the delay corresponding to a certain pixel is found (e.g., $t = \tau_{Tx} + \tau_{Rx}$), the reflection coefficient of the pixel can be estimated by computing the cross-correlation between the short-time window of the transmit signal and the short-time window of the received signal at the given delay.

short-time window is received at $t = \tau_{Tx} + \tau_{Rx}$ after it is reflected by a voxel. Clearly, the delay of the received short-time window consists of two parts: the travel time from the transmitter to the voxel τ_{Tx} and the travel time from the voxel to the receiver τ_{Rx}. Suppose the position of the transmitter at $t = 0$ is $\boldsymbol{\rho}_{Tx} = (x_{Tx}, y_{Tx}, z_{Tx})$, the position of the receiver at $t = 0$ is $\boldsymbol{\rho}_{Rx} = (x_{Rx}, y_{Rx}, z_{Rx})$, and the position of the voxel is $\boldsymbol{\rho}_v = (x_v, y_v, z_v)$, the first part of the delay can be easily calculated as

$$\tau_{Tx} = \frac{\|\boldsymbol{\rho}_v - \boldsymbol{\rho}_{Tx}\|_2}{c} \tag{2.1}$$

Since the vessel is moving with a constant velocity $\mathbf{v}$, the position of the receiver will not stay at $\boldsymbol{\rho}_{Rx}$ when the short-time window is received. Using the law of cosines in trigonometry, the second part of the delay can be expressed in a quadratic equation as:

$$(c\tau_{Rx})^2 = \|\boldsymbol{\rho}_{Rx} - \boldsymbol{\rho}_v\|_2^2 + \|\mathbf{v}(\tau_{Tx} + \tau_{Rx})\|_2^2 - 2(\boldsymbol{\rho}_{Rx} - \boldsymbol{\rho}_v) \cdot (\mathbf{v}(\tau_{Tx} + \tau_{Rx})) \tag{2.2}$$

Then, the second part of the delay can be calculated by solving the above quadratic equation. When the speed of the vessel is much smaller than the speed of sound in water $\|\mathbf{v}\|_2 \ll c$, the displacement of the receiver within the delay $(\tau_{Tx} + \tau_{Rx})$ will become negligible. Then, in this case, the second part of the delay can be approximately calculated as

$$\tau_{Rx} \approx \frac{\|\boldsymbol{\rho}_{Rx} - \boldsymbol{\rho}_v\|_2}{c} \tag{2.3}$$

Another approach to approximate the delay is called the Phase Centre Approximation (PCA), the details of the approach are mentioned in [1, 4].

Once the delay corresponding to a voxel of a short-time window is derived, the reflection coefficient of the voxel can be estimated by using the short-time window to correlate with the short-time segment of the received signal at the delay, as shown in Figure 2.1. By integrating the correlation result for every voxel in the imaging domain for every short-time window, the full SAS image is obtained. The SAS imaging algorithm can be summarized in pseudocode, shown in Algorithm 1. It is important to

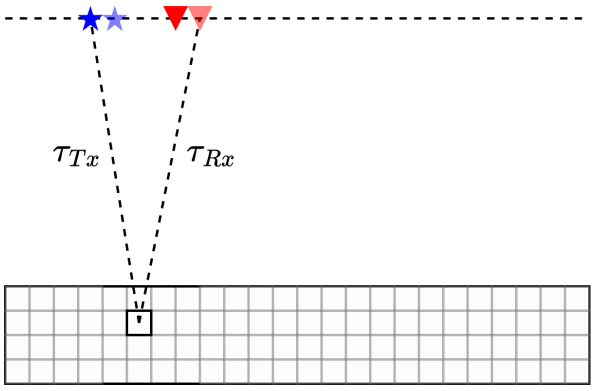


Figure 2.2: Schematic of the delay estimation of a pixel in the imaging domain for a short-time window. Assuming a pair of transmitter and receiver are surveying along a line trajectory with a velocity of $\mathbf{v}$ to image a discretized area that is beneath the trajectory. The initial positions of the transmitter and the receiver are marked with blue star and red triangle, respectively. At $t = 0$, the transmitter sends a short-time window of the continuous signal. After the short-time window is reflected by a certain voxel, it is then received by the transmitter at the position marked with triangle in light red at $t = \tau_{Tx} + \tau_{Rx}$.

note that due to the computation complexity, the current version of the SAS imaging algorithm only uses the matched filter to recover the reflection coefficients. The SAS imaging performance can be further improved by the least squares method (e.g, LSQR [5]).

Algorithm 1 SAS imaging algorithm

1: **for** each short-time window of the transmit signal **do**
2: **for** each voxel in the imaging domain **do**
3: Estimate the delay τ of the voxel for the short-time window
4: Estimate the reflection coefficient of the voxel by calculating the correlation between the short-time window and the short-time segment of the received signal at the delay τ
5: **end for**
6: Coherently integrate the reflection coefficient in the imaging domain for each short-time window of the transmit signal
7: **end for**

Suppose the imaging domain is discretized into a total of N_v voxels, the reflection coefficients of voxels can be represented using a vector $\mathbf{v} = [v_1, \ldots, v_{N_v}]^T \in \mathbb{R}^{N_v \times 1}$. Then, the received signal vector $\mathbf{y} \in \mathbb{C}^{N_{Tx} \times 1}$ can be described as a linear system [1]:

$$\mathbf{y} = \mathbf{H}\mathbf{v} \tag{2.4}$$

with

$$\mathbf{H} = \begin{bmatrix} | & | & & | \\ \mathbf{h}_1 & \mathbf{h}_2 & \ldots & \mathbf{h}_{N_v} \\ | & | & & | \end{bmatrix} \in \mathbb{C}^{N_{Tx} \times N_v} \tag{2.5}$$

where each column of $\mathbf{H}$ can be seen as the signal that is reflected by the corresponding voxel. Here, N_{Tx}, denotes the total number of time-domain samples measured by the receiver, which is determined by the product of the sampling frequency and the duration of the received signal. Due to the relative motion and the change of distance between the vessel and voxels in the imaging domain, each column of $\mathbf{H}$ also represents a copy of the transmit signal that is affected by multiple delays and Doppler effects over each short segment of the signal. Because of the use of STMF for SAS imaging as shown in Algorithm 1, the reflection coefficient vector in the linear system model can be approximately recovered as follows

$$\hat{\mathbf{v}} = \mathbf{H}^H \mathbf{y} = \mathbf{H}^H \mathbf{H} \mathbf{v} \tag{2.6}$$

where $\mathbf{H}^H\mathbf{H}$ determines the imaging performance of the SAS system. The element at mth row and nth column of $\mathbf{H}^H\mathbf{H}$ illustrates the correlation (inner product) of $\mathbf{h}_m$ and $\mathbf{h}_n$. Therefore, to guarantee the optimal imaging performance, the waveform for the SAS system needs to be designed in a way such that any two signals reflected by two different voxels will have a low correlation. In other words, the designed transmit signal needs to satisfy $\mathbf{H}^H\mathbf{H} \approx \mathbf{I}_{N_v}$.

However, since the survey trajectory of the SAS system is usually unknown before the survey is complete, the delays and the Doppler effects on the entire transmit signal are also unknown during the waveform design phase. Thus, designing the waveform for the entire transmit signal might be a difficult problem. One way to alleviate the problem complexity is to only focus on the waveform design within a short duration (PRI), and then repeat the designed waveform in time to generate the transmit signal. In this case, the received signal vector for a PRI of the transmit signal $\mathbf{y}_{PRI} \in \mathbb{C}^{N_{PRI} \times 1}$ ($N_{PRI} \ll N_{Tx}$) can be modeled as

$$\mathbf{y}_{PRI} = \mathbf{H}_{PRI} \mathbf{v} \tag{2.7}$$

with

$$\mathbf{H}_{PRI} = \begin{bmatrix} | & | & & | \\ \mathbf{h}_{PRI,1} & \mathbf{h}_{PRI,2} & \dots & \mathbf{h}_{PRI,N_v} \\ | & | & & | \end{bmatrix} \in \mathbb{C}^{N_{PRI} \times N_v} \tag{2.8}$$

Since $N_{PRI} \ll N_{Tx}$, each column of the matrix $\mathbf{H}_{PRI}$ can be assumed to be a copy of the designed waveform that is affected by a single delay and a single Doppler shift. Similar to Equation 2.6, the reflection coefficient for a single PRI can be restored as

$$\hat{\mathbf{v}}_{PRI} = \mathbf{H}_{PRI}^H \mathbf{y}_{PRI} = \mathbf{H}_{PRI}^H \mathbf{H}_{PRI} \mathbf{v}_{PRI} \tag{2.9}$$

Therefore, the waveform design problem for one PRI of the transmit signal can be seen as designing the repeated waveform such that it satisfies $\mathbf{H}_{PRI}^H \mathbf{H}_{PRI} \approx \mathbf{I}_{N_v}$. Since each element of the matrix $\mathbf{H}_{PRI}^H \mathbf{H}_{PRI}$ illustrates the cross-correlation of two designed waveforms with two sets of delay and Doppler indices, the waveform design problem for the SAS imaging system can be approximately considered as a wideband ambiguity function shaping problem. As the entire SAS imaging process can be approximately regarded as an integration of multiple repeated waveforms from the moving vehicle, we believe that shaping the ambiguity function of the repeated waveform can improve the SAS imaging performance.

2.2 Ambiguity function

Ambiguity Function (AF) $\chi(\tau, \alpha)$ is a two-dimensional function that describes the cross-correlation between a signal and a copy of the signal that is affected by a delay τ and a Doppler α. It has been widely used in radar and sonar systems to analyze the matched filtering performance of a waveform. In this section, the expression of the (wideband) ambiguity function is obtained from a simple moving target detection problem as discussed in Section 2.2.1. Then, under the narrowband condition, the expression of the narrowband ambiguity function is derived in Section 2.2.2. Finally, the properties of the wideband ambiguity function and the narrowband ambiguity function are introduced and evaluated in Section 2.2.3. It is important to note that the narrowband ambiguity function introduced in this section is meant to support the narrowband approximation method for the wideband ambiguity function shaping in Section 3.3.4.

2.2.1 Derivation of the (wideband) ambiguity function

Let us consider the moving target detection task shown in Figure 2.3, where a transducer is placed at the origin and an object is moving towards the transducer with a constant radial velocity of v. The distance between the transducer and the target is R_0 at $t = 0$. Then, the distance between the transducer and the target is given by

$$R(t) = R_0 - vt \tag{2.10}$$

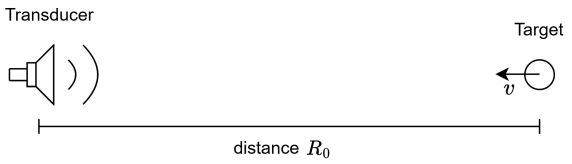


Figure 2.3: Schematic of a moving target detection task with a transducer

Suppose the transducer uses a passband signal $x(t)$ for range detection, and the propagation speed of the signal in the medium is c. After the transmit signal is reflected by the target, the transducer will receive a copy of the transmit signal with an instantaneous delay $\tau(t)$:

$$y_0(t) = x(t - \tau(t)) \tag{2.11}$$

Since the signal received at t is reflected by the target at $t_0 = t - \frac{\tau(t)}{2}$, we obtain [6]

$$\tau(t) = \frac{2R(t_0)}{c} = \frac{2(R_0 - vt_0)}{c} = \frac{2(R_0 - vt + v\tau(t)/2)}{c} \tag{2.12}$$

Solving Equation 2.12 gives us the function for the instantaneous delay with respect to time:

$$\tau(t) = \frac{2R_0 - 2vt}{c - v} \tag{2.13}$$

By substituting Equation 2.13 into the received signal (Equation 2.11), we get

$$y(t) = x((\frac{c+v}{c-v})(t - \frac{2R_0}{c+v})) \tag{2.14}$$

where $\frac{c+v}{c-v}$ can be seen as a time-scaling factor on the received signal due to the Doppler effect of the velocity of the moving target, and $\frac{2R_0}{c+v}$ can be seen as a global delay corresponding to the initial distance between the transducer and the target. Let $\alpha = \frac{c+v}{c-v}$ and $\tau = \frac{2R_0}{c+v}$. The received signal can then be modeled as the transmit signal that is affected by a delay $\tau \in (-\infty, \infty)$ and a Doppler $\alpha \in (0, \infty)$:

$$y(t) = \sqrt{\alpha} x(\alpha(t - \tau)) \tag{2.15}$$

where $\sqrt{\alpha}$ is a normalization factor [6] that ensures the received signal has the same energy as the transmit signal. Then the expression of the AF can be obtained by calculating the cross-correlation between the transmit signal and the received signal:

$$\chi(\tau, \alpha) = \sqrt{\alpha} \int_{-\infty}^{\infty} x(t) x^*(\alpha(t - \tau)) dt \tag{2.16}$$

Since Equation 2.16 can be applied to the waveform analysis in wideband scenarios, such as wideband radar or sonar, it is also referred to as the **Wideband Ambiguity Function (WAF)**.

Suppose the transmit signal $\tilde{x}(t)$ is periodic in time with a PRI of T_{PRI}, its periodicity can be expressed as

$$\tilde{x}(t) = \tilde{x}(t + nT_{PRI}), \ n = 0, \pm 1, \pm 2, \ldots \tag{2.17}$$

Then, the **periodic Wideband Ambiguity Function (periodic WAF)** can be defined as

$$\chi_P(\tau, \alpha) = \sqrt{\alpha} \int_{0}^{T_{PRI}} \tilde{x}(t) \tilde{x}^*(\alpha(t - \tau)) dt \tag{2.18}$$

2.2.2 Derivation of the narrowband ambiguity function

Suppose the passband signal $x(t)$ is composed of its complex baseband signal $x_{bb}(t)$ and a complex carrier with a carrier frequency of f_c:

$$x(t) = x_{bb}(t) e^{j2\pi f_c t} \tag{2.19}$$

The WAF can then be expanded as

$$\chi(\tau, \alpha) = \sqrt{\alpha} \int_{-\infty}^{\infty} x_{bb}(t) e^{j2\pi f_c t} x_{bb}^*(\alpha(t - \tau)) e^{-j2\pi f_c \alpha(t-\tau)} dt \tag{2.20}$$

$$= \sqrt{\alpha} e^{j2\pi f_c \alpha \tau} \int_{-\infty}^{\infty} x_{bb}(t) x_{bb}^*(\alpha(t - \tau)) e^{j2\pi f_c (1-\alpha) t} dt \tag{2.21}$$

When the speed of the target is much lower than the speed of the wave and the bandwidth to the carrier frequency ratio is much smaller than 1 (the narrowband condition [7, 8])

$$|v| \ll c, \ \frac{B}{f_c} \ll 1 \tag{2.22}$$

we have

$$\alpha = \frac{c+v}{c-v} \approx 1 \tag{2.23}$$

$$x_{bb}^*(\alpha(t-\tau)) \approx x_{bb}^*(t-\tau) \tag{2.24}$$

In this case, the time scaling caused by the Doppler effect can be ignored, and then the WAF can be approximated as

$$\chi(\tau,\alpha) \approx e^{j2\pi f_c \tau} \int_{-\infty}^{\infty} x_{bb}(t) x_{bb}^*(t-\tau) e^{j2\pi f_c(1-\alpha)t} dt \tag{2.25}$$

The integral part of Equation 2.25 is known as the **narrowband ambiguity function (NAF)**:

$$\chi_N(\tau, f_D) = \int_{-\infty}^{\infty} x_{bb}(t) x_{bb}^*(t-\tau) e^{j2\pi f_D t} dt \tag{2.26}$$

where $f_D = f_c(1-\alpha) = -f_c\frac{2v}{c-v} \approx -f_c\frac{2v}{c}$ represents the Doppler shift. The NAF can be seen as a simplified version of the WAF under the narrowband condition, and it is widely used in narrowband radar, but generally not applicable for sonar systems because the propagation speed of a sound wave in water is around 1,500 m/s, which is much lower compared to the speed of an electromagnetic wave in air of approximately 3×10^8 m/s.

2.2.3 Properties of the ambiguity function

The NAF has the following three prominent properties [9]:

- **Maximum value property**: the maximum value of a NAF locates at the origin $(\tau, f_D) = (0, 0)$ and is equal to the energy of the transmit signal:

$$|\chi_N(\tau, f_D)| \leq |\chi_N(0,0)| = \|x(t)\|^2, \ \forall(\tau, f_D) \tag{2.27}$$

- **Symmetric property**: the NAF is symmetric about the origin:

$$|\chi_N(\tau, f_D)| = |\chi_N(-\tau, -f_D)| \tag{2.28}$$

- **Volume invariance property**: the volume under the surface of a NAF is a constant equal to the square of the energy of the transmit signal:

$$V_{NAF} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\chi_N(\tau, f_D)|^2 d\tau df_D = |\chi_N(0,0)|^2 = (\|x(t)\|^2)^2 \tag{2.29}$$

The maximum value property and the symmetric property of the NAF can be easily observed from the formula of the NAF (Equation 2.26). Since the derivation of the volume invariance property is not straightforward, a comprehensive proof is given in Appendix A.1.

For a WAF, the maximum value property still holds:

$$|\chi(\tau,\alpha)| \leq |\chi(0,1)| = \|x(t)\|^2, \forall(\tau,\alpha) \tag{2.30}$$

where the maximum value locates at $(\tau,\alpha) = (0,1)$ and is equal to the energy of the transmit signal. However, because the Doppler effect is modeled as a time stretch or compression in the WAF, the symmetric property and the volume invariance property of the WAF do not hold anymore. As shown in [10], the volume of a WAF can be even infinite in some specific cases.

Luckily, for a band-limited signal with its Fourier transform decreasing to 0 as the frequency approaches 0, the volume of its WAF is still asymptomatically conserved and has a lower bound equal to the volume of the NAF of the signal [11]. The relationship between the volume of a WAF and the fractional bandwidth is shown in Figure 2.4, if the spectrum is assumed to be a rectangular function. For example, for a band-limited sonar signal with a bandwidth of 2000 Hz and a carrier frequency of 2000 Hz, the volume of its WAF is approximately 1.1 times the volume of its NAF. Therefore, the WAF of a band-limited signal has the asymptotic volume invariance property.

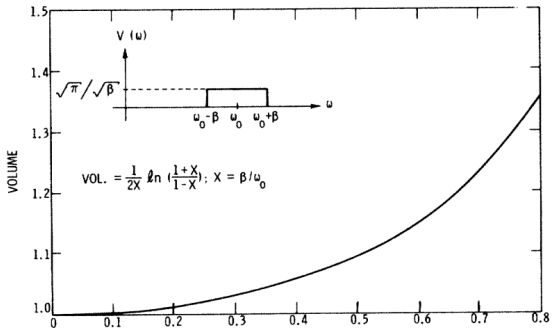


Figure 2.4: Volume of a WAF versus fractional bandwidth (β/ω_0) for a rectangular spectrum, where β represents half of the total bandwidth and ω_0 represents the center frequency [11]. The volume of a NAF is assumed to be 1 in this figure. The volume of a WAF approaches 1, as the fractional bandwidth approaches 0.

The maximum value property and the (asymptotic) volume invariance property are two essential properties for the idea of AF shaping. The maximum value property limits the height of the mainlobe of AF, which cannot be higher than the transmit energy. The volume invariance property implies the trade-off between the size of the sidelobe minimization region and the performance of sidelobe minimization in that region. As the volume under the sidelobe of an AF is also a constant, it is impossible to reduce the total sidelobe energy of the entire region of an AF. Generally, a larger sidelobe minimization region will lead to worse sidelobe minimization performance, and vice versa.

2.3 Review of conventional waveforms

The general formula for the complex envelope of a unimodular waveform is given by

$$x(t) = e^{j(2\pi f_c t + \varphi_i(t))},\ 0 \leq t < T \tag{2.31}$$

where $\varphi_i(t)$ is the instantaneous phase of the waveform, and T represents the duration of the waveform. The instantaneous frequency of the unimodular waveform is then obtained by calculating the derivative of the phase with respect to time:

$$f_i(t) = f_c + \frac{1}{2\pi}\frac{\partial \varphi_i(t)}{\partial t} \tag{2.32}$$

The unimodular property means the absolute value of the complex envelope of the waveform equals to 1 over the waveform duration:

$$|x(t)| = 1, 0 \leq t < T \tag{2.33}$$

In sonar and radar systems, unimodular waveforms, such as the frequency modulated waveform or the phase modulated waveform, are implemented to improve the energy efficiency and reduce the requirement for the peak transmit power.

In this section, some conventional unimodular waveforms and their AF properties are reviewed. Following that, the ideal AF properties for the SAS imaging waveform are discussed.

2.3.1 Frequency modulated waveform

Frequency modulation modulates the instantaneous frequency of a waveform. One of the most common frequency modulated waveforms is the linear frequency modulated (LFM, also known as chirp) waveform, whose instantaneous frequency is as follows

$$f_{i,LFM}(t) = (f_c - \frac{B}{2}) + \frac{B}{T}t,\ 0 \leq t < T \tag{2.34}$$

Another popular frequency modulated waveform is the HFM waveform, whose instantaneous frequency is given by

$$f_{i,HFM}(t) = \frac{(f_c^2 - \frac{B^2}{4})T}{f_c + \frac{B}{2} - Bt},\ 0 \leq t < T \tag{2.35}$$

The HFM waveform is very similar to the LFM waveform. As shown in Figure 2.5, the instantaneous frequency of the HFM has the same trend as that of the LFM, but the shape of $f_{i,HFM}(t)$ is a hyperbolic curve instead of a straight line.

Both the LFM and HFM waveforms are known for their Doppler invariance property [12], which means LFM and HFM waveforms will have strong correlation output for the matched filter even if the Doppler is present. Examples of the WAF for the LFM and HFM waveforms are shown in Figure 2.6a and 2.6b, where we see that the LFM and HFM signals have a ridge-shaped mainlobe and low sidelobes on two sides of the mainlobe. The ridge-shaped mainlobe provides good range detection for targets with

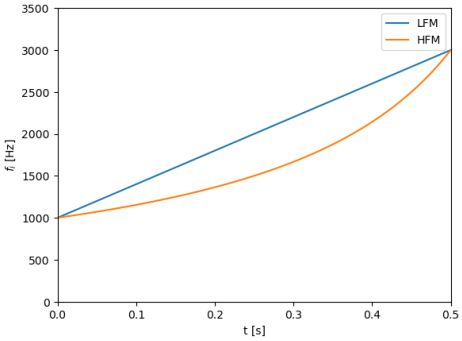
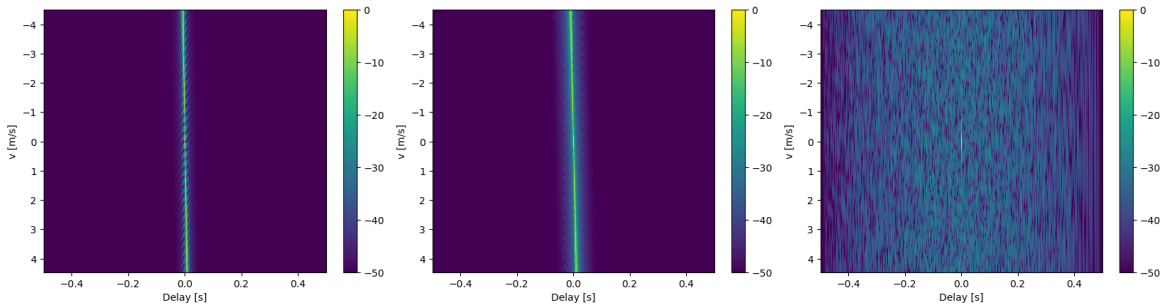


Figure 2.5: Instantaneous frequency versus time plot of LFM and HFM signals with $B = 2000$ Hz, $f_c = 2000$ Hz, and $T = 0.5$ s.

any velocity but introduces ambiguity for the velocity estimation. Generally, the LFM and HFM waveforms are not ideal for SAS imaging, because their ridge shape mainlobe in the WAF domain may cause mainlobe interference (extra artifacts) in the imaging domain, as shown in Figure 1.3b.



(a) LFM (b) HFM (c) random BPSK

Figure 2.6: WAF for (a) LFM, (b) HFM, and (c) random BPSK signal with $B = 2000$ Hz, $f_c = 2000$ Hz, and a signal duration of $T = 0.5$ s. The y-axis of the WAF (α, Doppler-axis) is converted into the velocity-axis [m/s]. The value of the WAF is normalized by the maximum value in the WAF, and the color bar is set to a dB scale ($20\log_{10}(\cdot)$).

2.3.2 Phase modulated waveform

Phase modulation modulates the instantaneous phase of a waveform. Compared with the frequency modulated waveforms whose instantaneous frequency is a continuous function, the instantaneous phase of the phase modulated waveforms is usually discrete. The general formula of a phase modulated waveform with a discrete phase can be expressed as

$$x_{PM}(t) = \sum_{n=1}^{N} s_n g(t - (n-1)\Delta T) e^{j2\pi f_c t}, \ 0 \le t < N\Delta T \tag{2.36}$$

with

$$g(t-(n-1)\Delta T) = \begin{cases} 1, & (n-1)\Delta T \leq t < n\Delta T \\ 0, & \text{otherwise} \end{cases} \tag{2.37}$$

where $\Delta T = \frac{1}{B}$ represents the duration of each phase code, $g(t-(n-1)\Delta T)$ is a rectangular pulse shaping function, and $s_n = e^{j\phi_n},\ n \in \{1,\ldots,N\}$ represents the nth phase code of the phase modulated waveform. One of the most common phase modulated waveforms is the random BPSK waveform, whose phase code is randomly selected from $\{-1,1\}$:

$$\mathbf{s}_{BPSK}[n] \in \{-1,1\},\ n = 1,\ldots,N \tag{2.38}$$

An example of the WAF of a random-BPSK signal is shown in Figure 2.6c. The sidelobe of the random-BPSK signal is higher than that of chirp signals, but the mainlobe of the random-BPSK signal has a thumbtack shape, which provides unambiguous detection of both the range and velocity of the target of interest.

Some phase modulated waveforms, such as Golomb sequence [13], and m-sequence or Maximal Length Sequence (MLS) [14], have a perfect periodic autocorrelation property, which means the periodic autocorrelation of these sequences is 0. The m-sequence is a periodic code sequence generated by maximal linear-feedback shift registers. To ensure its perfect autocorrelation property, the phase code of the m-sequence needs to satisfy

$$\mathbf{s}_{MLS}[n] \in \{e^{j\tan^{-1}(\sqrt{N})}, e^{-j\tan^{-1}(\sqrt{N})}\},\ n = 1,\ldots,N \tag{2.39}$$

The length of the m-sequence is $N = 2^l - 1$, where $l \in \mathbb{Z}$ and $l \geq 2$. When N approaches infinity, $\tan^{-1}(\sqrt{N})$ approaches $\pi/2$ and the phase code of the m-sequence approaches

$$\mathbf{s}_{MLS}[n] \in \{j,-j\},\ n = 1,\ldots,N \tag{2.40}$$

where the phase difference of two codes is π, which is the same as that of the random BPSK signal. Generally, the m-sequence can be approximately regarded as a BPSK signal with a fixed sequence order.

The Golomb sequence can be generated by

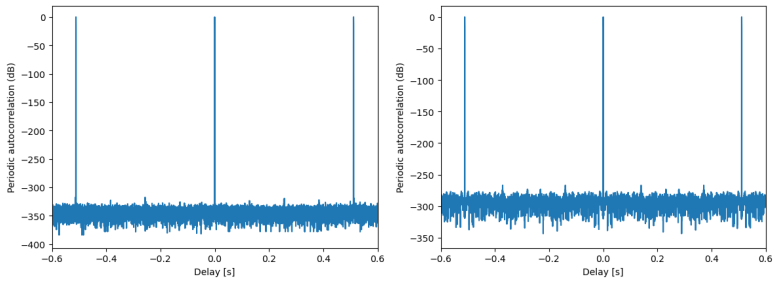
$$\mathbf{s}_{Golomb}[n] = e^{j\pi\frac{n(n-1)}{N}},\ n = 1,\ldots,N \tag{2.41}$$

where the length of the Golomb sequence N has to be an odd number. The derivative of the phase of $\mathbf{s}_{Golomb}[n]$ with respect to n gives the instantaneous frequency of the Golomb sequence:

$$\mathbf{f}_{Golomb}[n] = \frac{\partial \boldsymbol{\varphi}_{Golomb}[n]}{2\pi\partial n} = \frac{\partial(j\pi\frac{n(n-1)}{N})}{2\pi\partial n} = \frac{j}{N}(n-\frac{1}{2}) \tag{2.42}$$

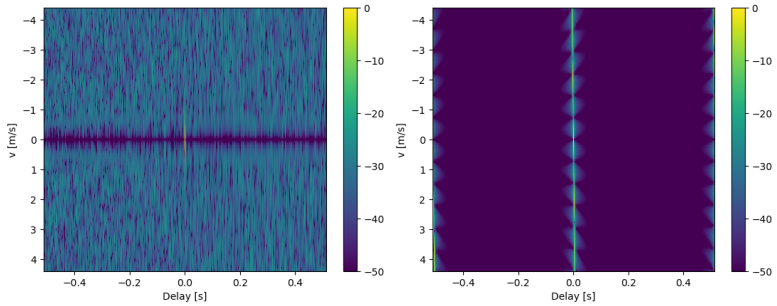
where the frequency sweep is linear with n. Therefore, the Golomb sequence can be considered as the discrete version of an LFM signal.

The periodic autocorrelation of the m-sequence and the Golomb sequence is shown in Figure 2.7, which shows that both of these two waveforms have a periodic "zero" sidelobe. Due to the floating point error, the sidelobe of the periodic autocorrelation



(a) m-sequence (b) Golomb sequence

Figure 2.7: Periodic autocorrelation of (a) m-sequence and (b) Golomb sequence with $N = 1023$, $B = 2000$ Hz, and $Fc = 2000$ Hz. The mainlobe of the periodic autocorrelations is periodic in time, with a period of $N/B = 0.5115$ s.



(a) periodic WAF of m-sequence (b) periodic WAF of Golomb sequence

Figure 2.8: Periodic WAF for (a) m-sequence and (b) Golomb sequence with $N = 1023$, $B = 2000$ Hz, and $Fc = 2000$ Hz. The mainlobe of the periodic WAF is periodic in time, with a period of $N/B = 0.5115$ s.

is not exactly equal to 0. An example of periodic WAF of the m-sequence and the Golomb sequence is shown in Figure 2.8. The periodic WAF can be considered as a combination of multiple periodically delayed WAFs. Because of the perfect periodic autocorrelation property, a zero sidelobe region can be found at $v = 0$ m/s of the periodic WAF of the m-sequence and the Golomb sequence. The periodic WAF of the m-sequence has the same thumbtack-shaped mainlobe and uniform high sidelobe as the WAF of the random BPSK. Similarly, the periodic WAF of the Golomb sequence has the same ridge-shaped mainlobe and low sidelobe on two sides of the mainlobe as the WAF of the LFM and HFM.

2.3.3 Other waveforms and ideal properties for the SAS imaging waveform

Apart from the random BPSK signal and m-sequence, another popular thumbtack-shaped waveform is the Costas code [15], which can be generated by segmenting and permuting an LFM waveform. Examples of the application of the Costas code in continuous active sonar are shown in [16, 17], where the idea of the Costas code is used to

create a non-recurrent continuous waveform for joint range and velocity detection for underwater targets. Recently, the family of Generalized Sinusoidal Frequency Modulated (GSFM) waveforms is gradually gaining traction because it offers high spectral efficiency and an optimal peak-to-average power ratio [18, 19]. The GSFM waveform can be seen as a frequency-modulated waveform with a sinusoidal-shaped instantaneous frequency, and the WAF of a GSFM waveform can have a thumbtack shape under some parameter settings. Recent Research on the GSFM waveform focuses on the design of a pulse train of GSFM pulses with a low cross-correlation between any two pulses using the genetic algorithm [20, 21] such that the WAF of the pulse train can have a thumbtack shape.

Ideally, we want the waveform for the SAS imaging to have a thumbtack-shaped WAF, so that the waveform is unambiguous in both range and velocity in the WAF domain. The aforementioned waveforms with a thumbtack shape WAF are unambiguous in the WAF domain, but have a uniform high sidelobe across the entire WAF domain. Consequently, the uniform high sidelobe in the WAF domain will result in the high sidelobe of the PSF in the imaging domain, as shown in Figure 1.3a. Therefore, it is also essential to reduce the sidelobe of the WAF.

As shown in Section 2.2.3, since the WAF has the maximum value property and the asymptotic volume invariance property, it is impossible to reduce the entire region of the sidelobe of a WAF. An alternative way to bypass the constraints is to only reduce the sidelobe in a region that is close to the mainlobe in the WAF domain, as illustrated in Figure 2.9. In this way, the sidelobe is reduced with a cost of high sidelobe in other regions. If the region of interest (RoI) in the WAF domain covers all the possible delays and Dopplers in the imaging domain of a SAS survey, the high sidelobe outside the RoI will not lead to a degraded imaging performance. Usually, the RoI is set to a rectangular area around the mainlobe but does not include the mainlobe. In a SAS scenario, the maximum distance of the RoI can be set to the maximum distance between the clutter and the voxels in the imaging domain. Because the voxels in the imaging domain are assumed to be static with respect to the background, the radial velocity of voxels with respect to the vessel cannot be higher than the speed of the vessel. Thus, the maximum velocity of the RoI can be set to the surveying speed of the vessel. In general, to design a waveform with minimized sidelobe level in a customized region in the AF domain, generalized ambiguity function shaping methods are required.

2.4 Literature review of ambiguity function shaping methods

The ambiguity function shaping methods generally aim to formulate the waveform design problem with AF properties into an optimization problem, which can then be solved with iterative algorithms. The maximum sidelobe level or the average sidelobe level in the RoI of the AF domain can be used as the objective function of the optimization problem. These ambiguity function shaping methods usually focus on the design of a phase code sequence with a unimodular constraint and then convert the sequence into a signal using Equations 2.36 and 2.37 with the given bandwidth and carrier frequency.

Since direct ambiguity function shaping is usually hard to solve, early research, such

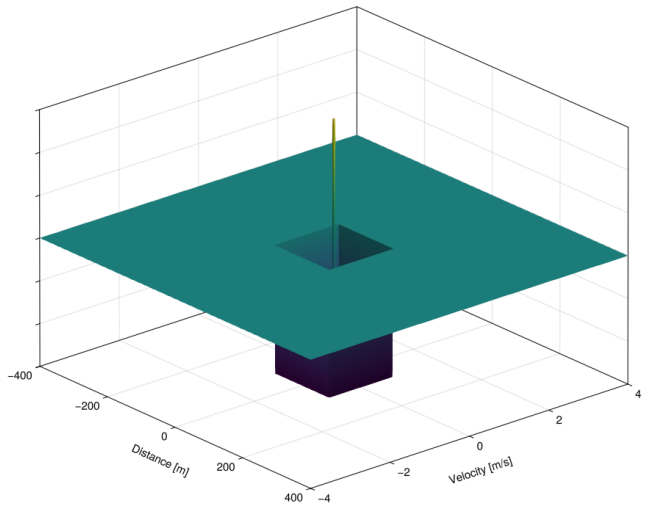


Figure 2.9: Ideal WAF for SAS imaging. The delay-axis and the Doppler-axis of the WAF are converted into the Distance-axis [m] and the Velocity-axis [m/s], respectively. The z-axis represents the value of the WAF at the given distance and velocity on a dB scale.

as cyclic algorithms (CAs) (e.g., [22]), has primarily focused on shaping the autocorrelation function rather than the entire ambiguity function. Sequences generated by CAs exhibit much better autocorrelation properties compared with the traditional sequences such as Golomb sequence and Frank sequence. However, since only the sidelobe at the zero-Doppler ($v = 0$ m/s) region is optimized, these sequences usually do not have desired ambiguity properties. As a follow-up, [23, 24] introduced how a variation of CAs can be applied to shape the NAF. Since the CA-based NAF shaping algorithm relies on iterative computing of the singular value decomposition (SVD) of a matrix with its dimension linearly scaled with the length of the sequence, it is not efficient to use the algorithm to generate a long sequence ($N \sim 10^3$). Following that, [25] proposed a gradient descent based method to update the phases of the designed sequence, which greatly reduces the computational complexity for NAF shaping. It has been shown in [26, 27] that the NAF shaping problem can be formulated into a unimodular quadratic program (UQP), which can be solved by the randomization method illustrated in [28] or the Power method-like algorithm shown in [26]. Following [26], [29] introduced the Accelerated Iterative Sequential Optimization (AISO) algorithm that can efficiently solve the NAF shaping problem and has better sidelobe minimization performance compared to the gradient descent method. Further, [30] proposed a NAF synthesizing method, which greatly suppresses the sidelobe of ambiguity regions by designing a train of unimodular sequences. Recently, [31] presented a majorization-minimization (MM) based algorithm to jointly design the unimodular sequence and the receiving filter for NAF shaping, which achieves a lower sidelobe level in the RoI with a cost of lower SNR.

It is important to note that the aforementioned NAF shaping methods predominately rely on the symmetric property of the NAF. This symmetric property significantly simplifies the computational complexity associated with the NAF shaping problem. However, due to the lack of a symmetry property in the WAF, the WAF shaping is usually hard to solve. The paper [32] introduced the Unimodular Signal Synthesis via

Minimization (USSM) algorithm that minimizes the maximum sidelobe in the RoI of the WAF of the designed sequence through iterative optimization. For each iteration of the USSM algorithm, a Semidefinite Programming (SDP) problem with numerous constraints is solved by a convex optimization solver to obtain the update of the sequence. Even though [32] also introduced methods to reduce the computational complexity by solving the SDP problem with only a small portion of the total constraints for each iteration, the USSM algorithm is still computationally expensive and does not scale well with the length of the sequence.

In general, current literature on the AF shaping methods mostly focuses on the NAF shaping. Due to the complex structure of the WAF, only few papers cover this topic and the WAF shaping methods are all computationally expensive. Future research on reducing the computational complexity and improving the optimality of the WAF shaping methods may help to fill this gap.

2.5 Conclusion

In this chapter, we partially answered the first sub-problem of the main research question: *"What is the relationship between the ambiguity properties of a waveform and its corresponding SAS imaging performance? Is ambiguity function shaping equivalent to waveform design for a SAS system?"*. In Section 2.1, we introduced the working principle of Fugro's SAS system and theoretically showed that the waveform design problem of a SAS system for a short integration time (e.g., a PRI) can be considered as a WAF shaping problem. Within a short time window, the waveform reflected by a voxel in the imaging domain can be seen as the transmit waveform affected by a single delay and a single Doppler shift. If we can reduce the cross-correlation between waveforms reflected by any two different voxels, the SAS imaging performance with a short integration time can be significantly improved. However, with a full integration trajectory, the SAS imaging performance of the waveform obtained from the WAF shaping is still unknown and requires further investigations.

The subsequence sections provided the necessary background for the WAF and the WAF shaping. In Section 2.2, we discussed the AF and its properties, including the derivation of the wideband and narrowband ambiguity functions. Here, we also highlighted the important properties of the WAF, such as the maximum value property and the asymptotic volume invariance property, which are essential for the idea of WAF shaping. In Section 2.3, we reviewed conventional waveforms used in sonar systems and their WAF properties. Here, we also briefly mentioned the importance of a thumbtack-shaped mainlobe in the WAF of a waveform and discussed the ideal WAF properties for SAS imaging. In Section 2.4, we provided a literature review of current ambiguity function shaping methods. Here, we found that most of the research targets the NAF shaping and only a few papers mentioned the WAF shaping.

In conclusion, this chapter established a theoretical connection between the waveform design and the WAF shaping. The next step is to explore how to effectively shape the WAF to improve the SAS imaging performance. In Chapter 3, we will formulate the WAF shaping problem as a non-convex problem and propose several algorithms to solve it. In Chapter 4, we will validate the SAS imaging performance of sequences

obtained from the WAF shaping through simulations and field trials, aiming to fully answer the first sub-problem of the main research question.

Wideband Ambiguity Function Shaping

3

In this chapter, we focus on the second sub-problem of the main research question:

- *What is the optimal way to shape the wideband ambiguity function of a sequence for SAS imaging?*

To address this question, the WAF shaping problem is formulated into a non-convex optimization problem in Section 3.1. The periodic case of the WAF shaping problem is considered in Section 3.2. Then, in Section 3.3, four algorithms are proposed to handle the non-convex problems introduced in Sections 3.1 and 3.2. To evaluate the performance of the proposed algorithms, we use the proposed algorithms to generate sequences with minimized sidelobe in the RoI under different settings.

3.1 Non-convex formulation of the wideband ambiguity function problem

Suppose we want to design a unimodular sequence with a length of N, the vector of the sequence is represented as follows:

$$\mathbf{s} = [s_1, s_2, \ldots, s_N] \tag{3.1}$$

with

$$s_n = e^{j\theta_n}, \quad n = 1, \ldots, N \tag{3.2}$$

where the phase code of the sequence is a complex number with a unit amplitude and a phase of $\theta_n \in [0, 2\pi)$. Once the unimodular sequence is obtained, we can convert the sequence into a waveform through phase modulation (Equations 2.36 and 2.37). The schematic of the process is shown in Figure 3.1. Then, the waveform of the given

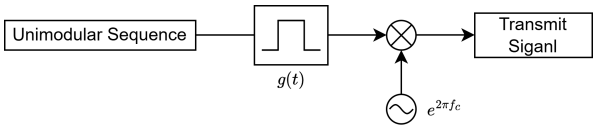


Figure 3.1: Schematic demonstrating how to convert a unimodular sequence into the waveform. $g(t)$ is a rectangular pulse shaping function in baseband.

sequence is written as

$$x(t) = \sum_{n=1}^{N} s_n g(t - (n-1)\Delta T) e^{j2\pi f_c t}, \ 0 \leq t < N\Delta T \tag{3.3}$$

where $\Delta T = \frac{1}{B}$ represents the duration of each phase code and $g(t)$ is assumed to be a rectangular pulse shaping function as shown in Equation 2.37. By substituting Equation 3.3 into the WAF (Equation 2.16), we obtain

$$\begin{aligned}\chi(\tau,\alpha) =&\sqrt{\alpha}\int_{-\infty}^{\infty}\sum_{n=1}^{N} s_n g(t-(n-1)\Delta T)e^{j2\pi f_c t}\\ &\sum_{m=1}^{N} s_m^* g^*(\alpha(t-\tau)-(m-1)\Delta T)e^{-j2\pi f_c\alpha(t-\tau)}dt\\ =&\sqrt{\alpha}e^{j2\pi f_c\alpha\tau}\sum_{n=1}^{N}\sum_{m=1}^{N} s_n s_m^*\int_0^{N\Delta T} g(t-(n-1)\Delta T)\\ &g^*(\alpha(t-\tau)-(m-1)\Delta T)e^{j2\pi f_c(1-\alpha)t}dt\end{aligned} \tag{3.4}$$

Since the energy of the waveform is $\|x(t)\|_2^2 = N\Delta T$, the above WAF can be normalized as follows

$$\begin{aligned}\bar{\chi}(\tau,\alpha) =&\frac{\sqrt{\alpha}e^{j2\pi f_c\alpha\tau}}{N\Delta T}\sum_{n=1}^{N}\sum_{m=1}^{N} s_n s_m^*\int_0^{N\Delta T} g(t-(n-1)\Delta T)\\ &g^*(\alpha(t-\tau)-(m-1)\Delta T)e^{j2\pi f_c(1-\alpha)t}dt\end{aligned} \tag{3.5}$$

To simplify the above expression, we introduce a complex matrix $\mathbf{A}_{(\tau,\alpha)} \in \mathbb{C}^{N\times N}$. The entry at the nth row and mth column of the matrix is defined as

$$\begin{aligned}\mathbf{A}_{(\tau,\alpha)}[m,n] =&\frac{\sqrt{\alpha}e^{j2\pi f_c\alpha\tau}}{N\Delta T}\int_0^{N\Delta T} g(t-(n-1)\Delta T)\\ &g^*(\alpha(t-\tau)-(m-1)\Delta T)e^{j2\pi f_c(1-\alpha)t}dt\end{aligned} \tag{3.6}$$

The above integral can be expanded as

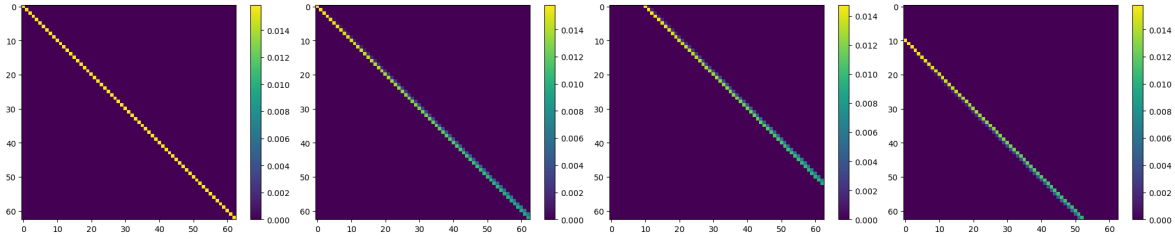
$$\mathbf{A}_{(\tau,\alpha)}[m,n] = \begin{cases}\frac{j\sqrt{\alpha}e^{j2\pi f_c\alpha\tau}}{2\pi f_c(\alpha-1)N\Delta T}(e^{j2\pi f_c(1-\alpha)t_2} - e^{j2\pi f_c(1-\alpha)t_1}) & t_1 < t_2, \alpha \neq 1\\ \frac{e^{j2\pi f_c\tau}}{N\Delta T}(t_2 - t_1) & t_1 < t_2, \alpha = 1\\ 0 & \text{otherwise}\end{cases} \tag{3.7}$$

where

$$t_1 = \max(0, (n-1)\Delta T, (m-1)\Delta T/\alpha + \tau) \tag{3.8}$$

$$t_2 = \min(N\Delta T, n\Delta T, m\Delta T/\alpha + \tau) \tag{3.9}$$

Examples of the matrix $\mathbf{A}_{(\tau,\alpha)}$ with different sets of the delay and Doppler are shown in Figure 3.2, which shows that $\mathbf{A}_{(\tau,\alpha)}$ is a sparse matrix and only has values on a "diagonal" line. When the Doppler is nonzero, the matrix $\mathbf{A}_{(\tau,\alpha)}$ is not Toeplitz, because the "diagonal" line is not perfectly diagonal. Under the narrowband condition, it is easy to show that the matrix $\mathbf{A}_{(\tau,\alpha)}$ is Toeplitz even when the Doppler is present.



(a) (τ, v)=(0,0) (b) (τ, v)=(0,-5) (c) (τ, v)=($\frac{10}{B}$,-5) (d) (τ, v)=($-\frac{10}{B}$,5)

Figure 3.2: 2D image of the matrix $|\mathbf{A}_{(\tau,\alpha)}|$ with different sets of delays and Dopplers. The length of the designed sequence is $N = 63$. The bandwidth and the carrier frequency of the waveform are $B = 2000$ Hz and $f_c = 2000$ Hz, respectively.

With the help of Equation 3.7, the normalized WAF can be written in a quadratic form as

$$\bar{\chi}(\tau, \alpha) = \sum_{n=1}^{N} \sum_{m=1}^{N} s_n s_m^* \mathbf{A}_{(\tau,\alpha)}[m, n] = \mathbf{s}^H \mathbf{A}_{(\tau,\alpha)} \mathbf{s} \tag{3.10}$$

Suppose the RoI is represented by a set of delay and Doppler pairs $\mathcal{Q}$. The objective function for the WAF shaping problem can then be defined as the average of the square of the sidelobes in the RoI of the WAF:

$$F_{avg}(\mathbf{s}) = \frac{1}{|\mathcal{Q}|} \sum_{(\tau,\alpha)\in\mathcal{Q}} w_{(\tau,\alpha)} |\bar{\chi}(\tau, \alpha)|^2 = \frac{1}{|\mathcal{Q}|} \sum_{(\tau,\alpha)\in\mathcal{Q}} w_{(\tau,\alpha)} |\mathbf{s}^H \mathbf{A}_{(\tau,\alpha)} \mathbf{s}|^2 \tag{3.11}$$

where $w_{(\tau,\alpha)}$ is a weight for the sidelobe at (τ, α). With the unimodular constraint on the sequence we aim to design, the corresponding optimization problem is given as follows:

$$\begin{aligned} &\min_{\mathbf{s}\in\mathbb{C}^N} \frac{1}{|\mathcal{Q}|} \sum_{(\tau,\alpha)\in\mathcal{Q}} w_{(\tau,\alpha)} |\mathbf{s}^H \mathbf{A}_{(\tau,\alpha)} \mathbf{s}|^2 \\ &\text{s.t. } |s_n| = 1 \ \forall n \in \{1, \ldots, N\} \end{aligned} \tag{3.12}$$

Since $\mathbf{A}_{(\tau,\alpha)}$ is not positive semidefinite and the objective function is not in a standard quadratic form, the objective function is not convex. Since the unimodular constraint is equivalent to a set of points on a unit circle in the complex plane, its constraint is also not convex.

In the USSM algorithm [32], the maximum sidelobe in the RoI is used as the objective function:

$$F_{max}(\mathbf{s}) = \max_{(\tau,\alpha)\in\mathcal{Q}} |\bar{\chi}(\tau, \alpha)| = \max_{(\tau,\alpha)\in\mathcal{Q}} |\mathbf{s}^H \mathbf{A}_{(\tau,\alpha)} \mathbf{s}| \tag{3.13}$$

The non-convex optimization problem of the USSM algorithm can be expressed as

$$\begin{aligned} &\min_{\mathbf{s}\in\mathbb{C}^N} \max_{(\tau,\alpha)\in\mathcal{Q}} |\mathbf{s}^H \mathbf{A}_{(\tau,\alpha)} \mathbf{s}| \\ &\text{s.t. } |s_n| = 1 \ \forall n \in \{1, \ldots, N\} \end{aligned} \tag{3.14}$$

Both the average sidelobe energy and the maximum sidelobe objective functions facilitate lower sidelobe level in the RoI, but the optimization outcomes of these two objective functions are quite different, as shown in Section 3.4.2.1. Using the maximum sidelobe objective function prevents a high peak sidelobe in the RoI, but does not guarantee an optimal total sidelobe in the RoI. On the other hand, using the average sidelobe objective function minimizes the total sidelobe level in the RoI, but may lead to an undesirable high peak sidelobe in the RoI. From the perspective of the WAF shaping for a SAS waveform, since we wish to minimize the sidelobe energy produced by the strong clutter, the average sidelobe energy might be a better choice for the optimization formulation.

3.2 Non-convex formulation of the periodic wideband ambiguity function problem

Since the designed waveform is transmitted repeatedly with a 100% duty-cycle during a SAS imaging survey, it would be more realistic to consider the shaping of the periodic WAF. Suppose the periodic version of the designed sequence is represented by a vector $\tilde{\mathbf{s}}$ with an infinite length. The entries of the periodic sequence $\tilde{\mathbf{s}}$ are obtained as

$$\tilde{s}_n = s_{n \bmod N},\ n = -\infty, \ldots, -1, 0, 1, \ldots, \infty \tag{3.15}$$

Then, the periodic waveform of the periodic sequence is derived as

$$\tilde{x}(t) = \sum_{n=-\infty}^{\infty} \tilde{s}_n g(t - (n-1)\Delta T) e^{j2\pi f_c t},\ -\infty \leq t < \infty \tag{3.16}$$

The normalized periodic ambiguity function is obtained by calculating the inner product of the periodic transmit signal and the periodic received signal that is affected by the delay and Doppler within a duration of $N\Delta T$:

$$\begin{aligned}
\bar{\chi}_P(\tau, \alpha) =& \frac{\sqrt{\alpha}}{N\Delta T} \int_0^{N\Delta T} \sum_{n=-\infty}^{\infty} \tilde{s}_n g(t - (n-1)\Delta T) e^{j2\pi f_c t} \\
& \sum_{m=-\infty}^{\infty} \tilde{s}_m^* g^*(\alpha(t-\tau) - (m-1)\Delta T) e^{-j2\pi f_c \alpha(t-\tau)} dt \\
=& \frac{\sqrt{\alpha} e^{j2\pi f_c \alpha \tau}}{N\Delta T} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \tilde{s}_n \tilde{s}_m^* \int_0^{N\Delta T} g(t - (n-1)\Delta T) \\
& g^*(\alpha(t-\tau) - (m-1)\Delta T) e^{j2\pi f_c (1-\alpha) t} dt
\end{aligned} \tag{3.17}$$

Using Equation 3.7, the normalized periodic WAF can be simplified as [32]

$$\begin{aligned}\bar{\chi}_P(\tau,\alpha) &= \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} \tilde{s}_n \tilde{s}_m^* \mathbf{A}_{(\tau,\alpha)}[m,n] \\ &= \sum_{p=-\infty}^{\infty}\sum_{q=-\infty}^{\infty}\sum_{n=1}^{N}\sum_{m=1}^{N} s_n s_m^* \mathbf{A}_{(\tau,\alpha)}[m+qN, n+pN] \\ &= \sum_{n=1}^{N}\sum_{m=1}^{N} s_n s_m^* \left(\sum_{p=-\infty}^{\infty}\sum_{q=-\infty}^{\infty} \mathbf{A}_{(\tau,\alpha)}[m+qN, n+pN] \right)\end{aligned} \tag{3.18}$$

It is important to note that here we treat $\mathbf{A}_{(\tau,\alpha)}[m+qN, n+pN]$ as a function of two inputs rather than as an entry of an $N \times N$ matrix, since the indices $m+qN$ and $n+pN$ may fall outside the matrix's index range. Let

$$\tilde{\mathbf{A}}_{(\tau,\alpha)}[m,n] = \sum_{p=-\infty}^{\infty}\sum_{q=-\infty}^{\infty} \mathbf{A}_{(\tau,\alpha)}[m+qN, n+pN] \tag{3.19}$$

then the normalized periodic WAF can be rewritten as a quadratic form:

$$\bar{\chi}_{PWAF}(\tau,\alpha) = \sum_{n=1}^{N}\sum_{m=1}^{N} s_n s_m^* \tilde{\mathbf{A}}_{(\tau,\alpha)}[m,n] = \mathbf{s}^H \tilde{\mathbf{A}}_{(\tau,\alpha)} \mathbf{s} \tag{3.20}$$

Since the Doppler is relatively small in practical scenario, it is safe to assume that [32]

$$\tilde{\mathbf{A}}_{(\tau,\alpha)}[m,n] = \sum_{p=-1}^{1}\sum_{q=-1}^{1} \mathbf{A}_{(\tau,\alpha)}[m+qN, n+pN] \tag{3.21}$$

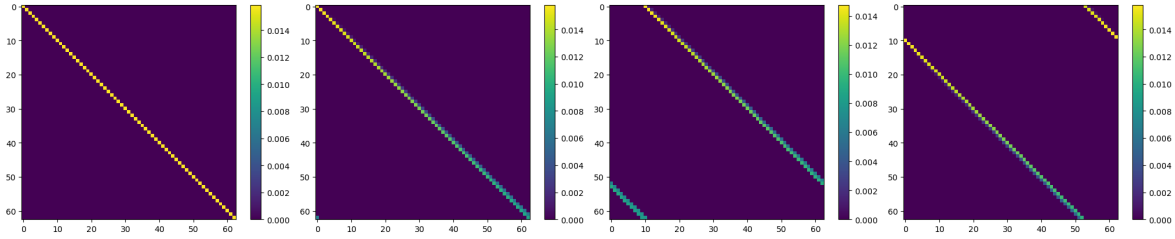
Therefore, the non-convex formulation for the periodic WAF shaping problem with an average sidelobe objective function is represented by

$$\begin{aligned}&\min_{\mathbf{s}\in\mathbb{C}^N} \frac{1}{|\mathcal{Q}|} \sum_{(\tau,\alpha)\in\mathcal{Q}} w_{(\tau,\alpha)} |\mathbf{s}^H \tilde{\mathbf{A}}_{(\tau,\alpha)} \mathbf{s}|^2 \\ &\text{s.t. } |s_n| = 1 \ \forall n \in \{1,\dots,N\}\end{aligned} \tag{3.22}$$

which is exactly the same as the non-convex formulation of Equation 3.12 in Section 3.1 except for the matrix $\tilde{\mathbf{A}}_{(\tau,\alpha)}$. Similarly, the non-convex formulation of the USSM algorithm for the periodic WAF shaping problem can be easily obtained by replacing $\mathbf{A}_{(\tau,\alpha)}$ in the WAF shaping formulation of Equation 3.14 with $\tilde{\mathbf{A}}_{(\tau,\alpha)}$:

$$\begin{aligned}&\min_{\mathbf{s}\in\mathbb{C}^N} \max_{(\tau,\alpha)\in\mathcal{Q}} |\mathbf{s}^H \tilde{\mathbf{A}}_{(\tau,\alpha)} \mathbf{s}| \\ &\text{s.t. } |s_n| = 1 \ \forall n \in \{1,\dots,N\}\end{aligned} \tag{3.23}$$

Examples of the matrix $\tilde{\mathbf{A}}_{(\tau,\alpha)}$ with different sets of the delay and Doppler are shown in Figure 3.3. These matrices are still sparse and similar to the matrices shown in Figure 3.2. Due to the sum operation shown in Equation 3.21, small diagonal segments can be found at the corners of Figures 3.3c and 3.3d. It is also easy to show that the matrix $\tilde{\mathbf{A}}_{(\tau,\alpha)}$ is always a Toeplitz matrix under the narrowband condition.



(a) (τ, v)=(0,0) (b) (τ, v)=(0,-5) (c) (τ, v)=($\frac{10}{B}$,-5) (d) (τ, v)=($-\frac{10}{B}$,5)

Figure 3.3: 2D image of the matrix $|\tilde{\mathbf{A}}_{(\tau,\alpha)}|$ with different sets of delays and Dopplers. The length of the designed sequence is $N = 63$. The bandwidth and the carrier frequency of the waveform are $B = 2000$ Hz and $f_c = 2000$ Hz, respectively.

3.3 Algorithms for the wideband ambiguity function shaping

In Sections 3.1 and 3.2, we formulated the WAF and periodic WAF shaping problems as non-convex optimization problems. In this section, four methods for solving these non-convex optimization problems are proposed.

3.3.1 USSM-avg algorithm

The USSM algorithm [32] uses the maximum sidelobe inside the RoI as the objective function for WAF shaping. Instead of using the maximum sidelobe as the objective function for the optimization, another possible choice would be the average sidelobe in the RoI, as the average sidelobe provides a better representation of the sidelobe level in the RoI. The first algorithm we introduced is called the USSM-avg, because the algorithm mostly follows the relaxation methods mentioned in the original USSM algorithm [32] and its objective function is changed to the average sidelobe in the RoI.

The initial non-convex optimization formulation of the USSM-avg algorithm is shown in Equation 3.12. To convert the complex variables $\mathbf{s} \in \mathbb{C}^N$ into real variables, we introduce a real vector $\mathbf{x} \in \mathbb{R}^{2N}$ that contains the real and imaginary components of $\mathbf{s}$:

$$\begin{aligned} \mathbf{x} &= [x_1 \dots x_{2N}]^T \\ &= [\Re(s_1)\Im(s_1) \dots \Re(s_N)\Im(s_N)]^T \end{aligned} \tag{3.24}$$

The quadratic term in the objective function of Equation 3.12 can then be expanded as

$$\mathbf{s}^H \mathbf{A}_{(\tau,\alpha)} \mathbf{s} = \mathbf{x}^T \hat{\mathbf{A}}_{(\tau,\alpha)} \mathbf{x} + j\mathbf{x}^T \check{\mathbf{A}}_{(\tau,\alpha)} \mathbf{x} \tag{3.25}$$

where $\hat{\mathbf{A}}_{(\tau,\alpha)} \in \mathbb{R}^{2N\times 2N}$ and $\check{\mathbf{A}}_{(\tau,\alpha)} \in \mathbb{R}^{2N\times 2N}$ are obtained by expanding the real and

imaginary components of $\mathbf{A}_{(\tau,\alpha)}$. The expressions of $\hat{\mathbf{A}}_{(\tau,\alpha)}$ and $\check{\mathbf{A}}_{(\tau,\alpha)}$ are given by

$$\hat{\mathbf{A}}_{(\tau,\alpha)} = \begin{bmatrix} \Re(\mathbf{A}_{(\tau,\alpha)}[1,1]) & -\Im(\mathbf{A}_{(\tau,\alpha)}[1,1]) & \Re(\mathbf{A}_{(\tau,\alpha)}[1,2]) & -\Im(\mathbf{A}_{(\tau,\alpha)}[1,2]) & \dots \\ \Im(\mathbf{A}_{(\tau,\alpha)}[1,1]) & \Re(\mathbf{A}_{(\tau,\alpha)}[1,1]) & \Im(\mathbf{A}_{(\tau,\alpha)}[1,2]) & \Re(\mathbf{A}_{(\tau,\alpha)}[1,2]) & \dots \\ \Re(\mathbf{A}_{(\tau,\alpha)}[2,1]) & -\Im(\mathbf{A}_{(\tau,\alpha)}[2,1]) & \Re(\mathbf{A}_{(\tau,\alpha)}[2,2]) & -\Im(\mathbf{A}_{(\tau,\alpha)}[2,2]) & \dots \\ \Im(\mathbf{A}_{(\tau,\alpha)}[2,1]) & \Re(\mathbf{A}_{(\tau,\alpha)}[2,1]) & \Im(\mathbf{A}_{(\tau,\alpha)}[2,2]) & \Re(\mathbf{A}_{(\tau,\alpha)}[2,2]) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{3.26}$$

$$\check{\mathbf{A}}_{(\tau,\alpha)} = \begin{bmatrix} \Im(\mathbf{A}_{(\tau,\alpha)}[1,1]) & \Re(\mathbf{A}_{(\tau,\alpha)}[1,1]) & \Im(\mathbf{A}_{(\tau,\alpha)}[1,2]) & \Re(\mathbf{A}_{(\tau,\alpha)}[1,2]) & \dots \\ -\Re(\mathbf{A}_{(\tau,\alpha)}[1,1]) & \Im(\mathbf{A}_{(\tau,\alpha)}[1,1]) & -\Re(\mathbf{A}_{(\tau,\alpha)}[1,2]) & \Im(\mathbf{A}_{(\tau,\alpha)}[1,2]) & \dots \\ \Im(\mathbf{A}_{(\tau,\alpha)}[2,1]) & \Re(\mathbf{A}_{(\tau,\alpha)}[2,1]) & \Im(\mathbf{A}_{(\tau,\alpha)}[2,2]) & \Re(\mathbf{A}_{(\tau,\alpha)}[2,2]) & \dots \\ -\Re(\mathbf{A}_{(\tau,\alpha)}[2,1]) & \Im(\mathbf{A}_{(\tau,\alpha)}[2,1]) & -\Re(\mathbf{A}_{(\tau,\alpha)}[2,2]) & \Im(\mathbf{A}_{(\tau,\alpha)}[2,2]) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{3.27}$$

By substituting Equation 3.25 into the average sidelobe formulation of Equation 3.12, we obtain the following optimization problem with only real variables:

$$\begin{aligned} &\min_{\mathbf{x}\in\mathbb{R}^{2N}} \frac{1}{|\mathcal{Q}|} \sum_{(\tau,\alpha)\in\mathcal{Q}} w_{(\tau,\alpha)} \left((\mathbf{x}^T \hat{\mathbf{A}}_{(\tau,\alpha)} \mathbf{x})^2 + (\mathbf{x}^T \check{\mathbf{A}}_{(\tau,\alpha)} \mathbf{x})^2 \right) \\ &\text{s.t. } x_{2n-1}^2 + x_{2n}^2 = 1 \ \forall n \in \{1,\dots,N\} \end{aligned} \tag{3.28}$$

We can further simplify the optimization problem by introducing a positive semidefinite matrix $\mathbf{X} \in \mathbb{S}_+^{2N}$, where

$$\mathbf{X} = \mathbf{x}\mathbf{x}^T \tag{3.29}$$

Since we have

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathrm{Tr}(\mathbf{x}^T \mathbf{A} \mathbf{x}) = \mathrm{Tr}(\mathbf{x}\mathbf{x}^T \mathbf{A}^T) = \mathrm{Tr}(\mathbf{X}\mathbf{A}^T) = \langle \mathbf{A}, \mathbf{X} \rangle \tag{3.30}$$

the optimization problem can then be rewritten in an equivalent form as

$$\begin{aligned} &\min_{\mathbf{X}\in\mathbb{S}_+^{2N}} \frac{1}{|\mathcal{Q}|} \sum_{(\tau,\alpha)\in\mathcal{Q}} w_{(\tau,\alpha)} \Big(\langle \hat{\mathbf{A}}_{(\tau,\alpha)}, \mathbf{X} \rangle^2 + \langle \check{\mathbf{A}}_{(\tau,\alpha)}, \mathbf{X} \rangle^2 \Big) \\ &\text{s.t. } \mathbf{X}[2n-1, 2n-1] + \mathbf{X}[2n, 2n] = 1 \ \forall n \in \{1,\dots,N\}, \\ &\quad \mathrm{rank}(\mathbf{X}) = 1 \end{aligned} \tag{3.31}$$

where the rank-1 constraint is introduced to ensure that $\mathbf{X} = \mathbf{x}\mathbf{x}^T$. It is important to note that the above optimization problem is an SDP problem with a rank-1 constraint.

To resolve the rank-1 constraint on $\mathbf{X}$, we use a rank reduction method mentioned in [33, 32], which relaxes the SDP problem with a rank constraint as an alternating minimization of two sub-problems. The first sub-problem of the relaxation is given by

$$\begin{aligned} &\min_{\mathbf{X}\in\mathbb{S}_+^{2N}} \zeta \left(\frac{1}{|\mathcal{Q}|} \sum_{(\tau,\alpha)\in\mathcal{Q}} w_{(\tau,\alpha)} \Big(\langle \hat{\mathbf{A}}_{(\tau,\alpha)}, \mathbf{X} \rangle^2 + \langle \check{\mathbf{A}}_{(\tau,\alpha)}, \mathbf{X} \rangle^2 \Big) \right) + \langle \mathbf{W}_{opt}, \mathbf{X} \rangle \\ &\text{s.t. } \mathbf{X}[2n-1, 2n-1] + \mathbf{X}[2n, 2n] = 1 \ \forall n \in \{1,\dots,N\} \end{aligned} \tag{3.32}$$

where $\mathbf{W}_{opt} \in \mathbb{S}_+^{2N}$ is a search direction matrix that can be obtained from the optimal solution of the second sub-problem, and ζ is a regularization parameter that regularizes the rank-1 constraint on $\mathbf{X}$. The second sub-problem minimizes the inner product of $\mathbf{W}$ and $\mathbf{X}_{opt}$ with respect to $\mathbf{W}$:

$$\begin{aligned} &\min_{\mathbf{W} \in \mathbb{S}_+^{2N}} \langle \mathbf{W}, \mathbf{X}_{opt} \rangle \\ \text{s.t. } &0 \preceq \mathbf{W} \preceq \mathbf{I} \\ &\mathrm{Tr}(\mathbf{W}) = 2N - 1 \end{aligned} \tag{3.33}$$

where $\mathbf{X}_{opt}$ represents the optimal solution of the first sub-problem. The second sub-problem ensures that the direction matrix $\mathbf{W}$ aligns with the eigenspace of $\mathbf{X}_{opt}$ corresponding to all eigenvalues except the largest one. Suppose $\mathbf{X}_{opt}$ has the diagonalization:

$$\mathbf{X}_{opt} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T \tag{3.34}$$

It has been proven in [33, 32] that the second sub-problem has an analytical solution,

$$\mathbf{W}_{opt} = \mathbf{Q}[:, 2:2N]\mathbf{Q}[:, 2:2N]^T \tag{3.35}$$

where $\mathbf{Q}[:, 2:2N]$ represents the second to the last columns of $\mathbf{Q}$. The first sub-problem of the alternating minimization is an SDP problem and can be solved by SDP solvers.

The objective function of the first sub-problem is a weighted sum of the average sidelobe level in the RoI and the inner product of the matrices $\mathbf{X}$ and $\mathbf{W}_{opt}$. The inner product of the matrices $\mathbf{X}$ and $\mathbf{W}_{opt}$ represents the sum of all eigenvalues of $\mathbf{X}$ except for the largest eigenvalue. By minimizing the inner product, the rank-1 solution of $\mathbf{X}$ is promoted. The choice of the regularization parameter, ζ, affects whether the optimization focuses more on the sidelobe minimization or keeping the rank-1 constraint of $\mathbf{X}$. A higher value of ζ will lead to a lower sidelobe level, but it may break the rank-1 constraint. On the other hand, a lower value of ζ will lead to a rank-1 $\mathbf{X}$ but it may not guarantee a minimized sidelobe level in the RoI. Thus, the choice of ζ is crucial while using the USSM-avg algorithm for the sequence design. Usually, we need to initialize the algorithm with a set of ζ to find the unimodular sequence with the optimal sidelobe level.

As shown in [32], the alternating minimization in the USSM-avg algorithm can be initialized by letting $\mathbf{W}_{opt} = \mathbf{0}_{2N}$. In this way, the first iteration of the USSM-avg can be seen as solving the optimization problem of Equation 3.31 without the rank-1 constraint, and its optimized solution $\mathbf{X}_{opt}$ ensures the average sidelobe energy is minimized. Even if $\mathbf{X}_{opt}$ is not likely to be rank-1 after the first iteration, the rank of $\mathbf{X}_{opt}$ can be reduced through the alternating minimization of $\mathbf{X}$ and $\mathbf{W}$. Once $\mathbf{X}_{opt}$ has a rank close to 1, the vector $\mathbf{x}_{opt}$ can be approximately derived from the principal eigenvector of $\mathbf{X}_{opt}$. Then, the designed sequence $\mathbf{s}_{opt}$ can be easily obtained from $\mathbf{x}_{opt}$ by reversing the process in Equation 3.24. The USSM-avg algorithm is summarized in Algorithm 2. Because an eigenvalue decomposition problem and an SDP problem are required in each iteration of the USSM-avg algorithm, the computational complexity for each iteration of the USSM-avg algorithm is approximately $\mathcal{O}(N^3)$.

During the optimization process, the average sidelobe in the RoI,

$$\hat{F}_{avg}(\mathbf{X}_{opt}) = \frac{1}{|\mathcal{Q}|} \sum_{(\tau,\alpha)\in\mathcal{Q}} w_{(\tau,\alpha)} \Big(\langle \hat{\mathbf{A}}_{(\tau,\alpha)}, \mathbf{X}_{opt} \rangle^2 + \langle \check{\mathbf{A}}_{(\tau,\alpha)}, \mathbf{X}_{opt} \rangle^2 \Big) \tag{3.36}$$

and the eigenvalue residual of $\mathbf{X}_{opt}$,

$$r_\lambda = \sum_{i=2}^{2N} \lambda_i(\mathbf{X}_{opt}) \tag{3.37}$$

can be used as the evaluation metrics for the USSM-avg algorithm. It is worth noting that $\hat{F}_{avg}(\mathbf{X}_{opt})$ in Equation 3.36 is not equal to $F_{avg}(\mathbf{s})$ in Equation 3.11, unless $\mathbf{X}_{opt}$ has a rank of 1. When $r_\lambda < 10^{-3}$, $\mathbf{X}_{opt}$ can be approximately seen as a rank-1 matrix.

Algorithm 2 USSM-avg algorithm for the wideband ambiguity function shaping

1: Define the bandwidth B, the carrier frequency f_c, and the sequence length N
2: Initialize the regularization parameter ζ, and define the RoI $\mathcal{Q}$
3: Generate $\hat{\mathbf{A}}_{(\tau,\alpha)}$ and $\check{\mathbf{A}}_{(\tau,\alpha)}$ $\forall(\tau,\alpha) \in \mathcal{Q}$ using Equations 3.7, 3.26, and 3.27
4: **while** convergence conditions are not satisfied **do**
5: **if** this is the first iteration **then**
6: Let $\mathbf{W}_{opt} = \mathbf{0}_{2N}$
7: **else**
8: Solve $\mathbf{W}_{opt}$ from the second sub-problem (Equation 3.33) using Equations 3.34 and 3.35
9: **end if**
10: Solve $\mathbf{X}_{opt}$ from the first sub-problem (Equation 3.32)
11: **end while**
12: Derive $\mathbf{x}_{opt}$ from the principle eigenvector of $\mathbf{X}_{opt}$
13: Derive the optimal sequence $\mathbf{s}_{opt}$ from $\mathbf{x}_{opt}$ by reversing the process in Equation 3.24

3.3.2 Randomization-based method

In Section 3.3.1, we have shown that the WAF shaping problem can be rewritten as an SDP problem with a rank-1 constraint (Equation 3.31). The relaxation method mentioned in [32, 33] requires the USSM-avg algorithm to solve an SDP problem for each iteration until convergence in order to obtain the rank-1 solution of $\mathbf{X}$, which can be computationally expensive when the convergence is slow. Another way to handle the rank-1 constraint is the randomization method depicted in [28]. The randomization method [28] is originally designed to solve the non-convex quadratically constrained quadratic program (QCQP), and it has been shown in some literature (e.g., [34, 35]) that the method can obtain the solution with a high approximation accuracy for both the minimization and maximization versions of the QCQP problems. However, as the initial non-convex formulation for the WAF shaping (Equation 3.12) is not a QCQP problem, the optimality of the randomization method is not guaranteed.

The first step of the randomization method is to directly solve the SDP problem with a rank-1 constraint (Equation 3.31) by dropping the rank-1 constraint:

$$\min_{\mathbf{X}\in\mathbb{S}_+^{2N}} \frac{1}{|\mathcal{Q}|} \sum_{(\tau,\alpha)\in\mathcal{Q}} w_{(\tau,\alpha)} \Big(\langle \hat{\mathbf{A}}_{(\tau,\alpha)}, \mathbf{X} \rangle^2 + \langle \check{\mathbf{A}}_{(\tau,\alpha)}, \mathbf{X} \rangle^2 \Big) \tag{3.38}$$
$$\text{s.t. } \mathbf{X}[2n-1, 2n-1] + \mathbf{X}[2n, 2n] = 1 \ \forall n \in \{1, \dots, N\}$$

which is the same as the optimization problem solved in the first iteration of the USSM-avg algorithm. Suppose the optimal solution of the above optimization problem is $\mathbf{X}_{opt,R}$. A multivariate normal distribution with zero mean and covariance matrix $\mathbf{X}_{opt,R}$ is then considered:

$$\mathcal{N}(\mathbf{0}, \mathbf{X}_{opt,R}) \tag{3.39}$$

Finally, the optimal sequence $\mathbf{s}_{opt,R}$ can be approximated by taking random samples $\boldsymbol{\xi}_k$ from the above distribution and selecting the sample that has the smallest objective. In this way, we only need to solve an SDP problem once at the start of the algorithm. The randomization method for WAF shaping is summarized in Algorithm 3.

Algorithm 3 Randomization method for the wideband ambiguity function shaping

1: Define the bandwidth B, the carrier frequency f_c, and the sequence length N
2: Define the RoI $\mathcal{Q}$
3: Generate $\hat{\mathbf{A}}_{(\tau,\alpha)}$ and $\check{\mathbf{A}}_{(\tau,\alpha)}$ $\forall(\tau,\alpha) \in \mathcal{Q}$ using Equations 3.7, 3.26, and 3.27
4: Solve $\mathbf{X}_{opt,R}$ from the SDP problem of Equation 3.38
5: Define the number of samples K
6: **for** $k = 1, \dots, K$ **do**
7: Generate $\boldsymbol{\xi}_k$ from the distribution $\mathcal{N}(\mathbf{0}, \mathbf{X}_{opt,R})$
8: Derive $\mathbf{s}_k$ from $\boldsymbol{\xi}_k$ by reversing the process in Equation 3.24
9: Project $\mathbf{s}_k$ to the unimodular space, $\mathbf{s}_k \leftarrow e^{j\mathbf{s}_k}$
10: **end for**
11: Derive the optimal sequence $\mathbf{s}_{opt,R}$ by finding the $\mathbf{s}_k$, $\forall k \in \{1, \dots, K\}$ that minimizes the average sidelobe energy in the RoI $\mathcal{Q}$ (Equation 3.11)

3.3.3 Wideband gradient descent method

The idea of the wideband gradient descent method originates from the narrowband gradient descent method [25], which uses the gradient descent to update the phase of a sequence for the NAF shaping. The initial optimization problem for the WAF shaping is shown in Equation 3.12, where the objective function is the average sidelobe energy in the RoI ($F_{avg}(\mathbf{s})$, Equation 3.11), subject to a unimodular constraint on the sequence $\mathbf{s}$. Since the direct gradient descent update on phases of the sequence bypasses the unimodular constraint, the main challenge of the wideband gradient descent method is to derive the expression for the partial derivative of the average sidelobe objective function with respect to phases.

Suppose the phase vector of the sequence $\mathbf{s}$ is represented by

$$\boldsymbol{\theta} = [\theta_1, \dots, \theta_N]^H \tag{3.40}$$

By using the chain rule [36], the partial derivative of $F_{avg}(\mathbf{s})$ with respect to θ_n can be written as [25]

$$\begin{aligned}\frac{\partial F_{avg}(\mathbf{s})}{\partial \theta_n} &= \frac{\partial F_{avg}(\mathbf{s})}{\partial \Re(s_n)}\frac{\partial \Re(s_n)}{\partial \theta_n} + \frac{\partial F_{avg}(\mathbf{s})}{\partial \Im(s_n)}\frac{\partial \Im(s_n)}{\partial \theta_n} \\ &= -\Im(s_n)\frac{\partial F_{avg}(\mathbf{s})}{\partial \Re(s_n)} + \Re(s_n)\frac{\partial F_{avg}(\mathbf{s})}{\partial \Im(s_n)}\end{aligned} \tag{3.41}$$

Then, the vector form of Equation 3.41 can be expressed as

$$\frac{\partial F_{avg}(\mathbf{s})}{\partial \boldsymbol{\theta}} = -\Im(\mathbf{s}) \odot \frac{\partial F_{avg}(\mathbf{s})}{\partial \Re(\mathbf{s})} + \Re(\mathbf{s}) \odot \frac{\partial F_{avg}(\mathbf{s})}{\partial \Im(\mathbf{s})} \tag{3.42}$$

Since $|z|^2 = \Re(z)^2 + \Im(z)^2$, the partial derivatives of $F_{avg}(\mathbf{s})$ with respect to $\Re(\mathbf{s})$ and $\Im(\mathbf{s})$ can also be obtained by applying the chain rule [36],

$$\begin{aligned}\frac{\partial F(\mathbf{s})}{\partial(\cdot)} =& \frac{1}{|\mathcal{Q}|}\sum_{(\tau,\alpha)\in\mathcal{Q}} w_{(\tau,\alpha)}\frac{\partial|\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s}|^2}{\partial(\cdot)} \\ =& \frac{2}{|\mathcal{Q}|}\sum_{(\tau,\alpha)\in\mathcal{Q}} w_{(\tau,\alpha)}\Bigg(\Re(\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s})\frac{\partial\Re(\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s})}{\partial(\cdot)} \\ &+ \Im(\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s})\frac{\partial\Im(\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s})}{\partial(\cdot)}\Bigg)\end{aligned} \tag{3.43}$$

where

$$\partial(\cdot) = \begin{cases}\partial\Re(\mathbf{s}) \\ \partial\Im(\mathbf{s})\end{cases} \tag{3.44}$$

The complex vector $\mathbf{s}$ can be separated into its real part and imaginary part:

$$\mathbf{s} = \Re(\mathbf{s}) + j\Im(\mathbf{s}) \tag{3.45}$$

Similarly, the matrix $\mathbf{A}_{(\tau,\alpha)}$ can also be rewritten as

$$\mathbf{A}_{(\tau,\alpha)} = \Re(\mathbf{A}_{(\tau,\alpha)}) + j\Im(\mathbf{A}_{(\tau,\alpha)}) \tag{3.46}$$

By substituting Equations 3.45 and 3.46 into $\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s}$, we obtain

$$\begin{aligned}\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s} =& (\Re(\mathbf{s}) - j\Im(\mathbf{s}))^T(\Re(\mathbf{A}_{(\tau,\alpha)}) + j\Im(\mathbf{A}_{(\tau,\alpha)}))(\Re(\mathbf{s}) + j\Im(\mathbf{s})) \\ =& \Re(\mathbf{s})^T(\Re(\mathbf{A}_{(\tau,\alpha)}) + j\Im(\mathbf{A}_{(\tau,\alpha)}))\Re(\mathbf{s}) \\ &+ j\Re(\mathbf{s})^T(\Re(\mathbf{A}_{(\tau,\alpha)}) + j\Im(\mathbf{A}_{(\tau,\alpha)}))\Im(\mathbf{s}) \\ &- j\Im(\mathbf{s})^T(\Re(\mathbf{A}_{(\tau,\alpha)}) + j\Im(\mathbf{A}_{(\tau,\alpha)}))\Re(\mathbf{s}) \\ &+ \Im(\mathbf{s})^T(\Re(\mathbf{A}_{(\tau,\alpha)}) + j\Im(\mathbf{A}_{(\tau,\alpha)}))\Im(\mathbf{s}) \\ =& \Re(\mathbf{s})^T\Re(\mathbf{A}_{(\tau,\alpha)})\Re(\mathbf{s}) + j\Re(\mathbf{s})^T\Im(\mathbf{A}_{(\tau,\alpha)})\Re(\mathbf{s}) \\ &+ j\Re(\mathbf{s})^T\Re(\mathbf{A}_{(\tau,\alpha)})\Im(\mathbf{s}) - \Re(\mathbf{s})^T\Im(\mathbf{A}_{(\tau,\alpha)})\Im(\mathbf{s}) \\ &- j\Im(\mathbf{s})^T\Re(\mathbf{A}_{(\tau,\alpha)})\Re(\mathbf{s}) + \Im(\mathbf{s})^T\Im(\mathbf{A}_{(\tau,\alpha)})\Re(\mathbf{s}) \\ &+ \Im(\mathbf{s})^T\Re(\mathbf{A}_{(\tau,\alpha)})\Im(\mathbf{s}) + j\Im(\mathbf{s})^T\Im(\mathbf{A}_{(\tau,\alpha)})\Im(\mathbf{s})\end{aligned} \tag{3.47}$$

Then, the real part of $\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s}$ is given by

$$\begin{aligned}\Re(\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s}) =&\Re(\mathbf{s})^T\Re(\mathbf{A}_{(\tau,\alpha)})\Re(\mathbf{s}) - \Re(\mathbf{s})^T\Im(\mathbf{A}_{(\tau,\alpha)})\Im(\mathbf{s})\\ &+ \Im(\mathbf{s})^T\Im(\mathbf{A}_{(\tau,\alpha)})\Re(\mathbf{s}) + \Im(\mathbf{s})^T\Re(\mathbf{A}_{(\tau,\alpha)})\Im(\mathbf{s})\end{aligned} \tag{3.48}$$

and the imaginary part of $\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s}$ is given by

$$\begin{aligned}\Im(\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s}) =&\Re(\mathbf{s})^T\Im(\mathbf{A}_{(\tau,\alpha)})\Re(\mathbf{s}) + \Re(\mathbf{s})^T\Re(\mathbf{A}_{(\tau,\alpha)})\Im(\mathbf{s})\\ &- \Im(\mathbf{s})^T\Re(\mathbf{A}_{(\tau,\alpha)})\Re(\mathbf{s}) + \Im(\mathbf{s})^T\Im(\mathbf{A}_{(\tau,\alpha)})\Im(\mathbf{s})\end{aligned} \tag{3.49}$$

Since we have [36]

$$\frac{\partial\mathbf{x}^T\mathbf{a}}{\partial\mathbf{x}} = \frac{\partial\mathbf{a}^T\mathbf{x}}{\partial\mathbf{x}} = \mathbf{a} \tag{3.50}$$

$$\frac{\partial\mathbf{x}^T\mathbf{A}\mathbf{x}}{\partial\mathbf{x}} = (\mathbf{A}+\mathbf{A}^T)\mathbf{x} \tag{3.51}$$

the corresponding partial derivatives of $\Re(\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s})$ and $\Im(\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s})$ can then be calculated as follows

$$\frac{\partial\Re(\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s})}{\partial\Re(\mathbf{s})} = (\Re(\mathbf{A}_{(\tau,\alpha)})+\Re(\mathbf{A}_{(\tau,\alpha)})^T)\Re(\mathbf{s})-\Im(\mathbf{A}_{(\tau,\alpha)})\Im(\mathbf{s})+\Im(\mathbf{A}_{(\tau,\alpha)})^T\Im(\mathbf{s}) \tag{3.52}$$

$$\frac{\partial\Re(\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s})}{\partial\Im(\mathbf{s})} = -\Im(\mathbf{A}_{(\tau,\alpha)})^T\Re(\mathbf{s}) + \Im(\mathbf{A}_{(\tau,\alpha)})\Re(\mathbf{s}) + (\Re(\mathbf{A}_{(\tau,\alpha)}) + \Re(\mathbf{A}_{(\tau,\alpha)})^T)\Im(\mathbf{s}) \tag{3.53}$$

$$\frac{\partial\Im(\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s})}{\partial\Re(\mathbf{s})} = (\Im(\mathbf{A}_{(\tau,\alpha)})+\Im(\mathbf{A}_{(\tau,\alpha)})^T)\Re(\mathbf{s})+\Re(\mathbf{A}_{(\tau,\alpha)})\Im(\mathbf{s})-\Re(\mathbf{A}_{(\tau,\alpha)})^T\Im(\mathbf{s}) \tag{3.54}$$

$$\frac{\partial\Im(\mathbf{s}^H\mathbf{A}_{(\tau,\alpha)}\mathbf{s})}{\partial\Im(\mathbf{s})} = \Re(\mathbf{A}_{(\tau,\alpha)})^T\Re(\mathbf{s})-\Re(\mathbf{A}_{(\tau,\alpha)})\Re(\mathbf{s})+(\Im(\mathbf{A}_{(\tau,\alpha)})+\Im(\mathbf{A}_{(\tau,\alpha)})^T)\Im(\mathbf{s}) \tag{3.55}$$

Because $\mathbf{A}_{(\tau,\alpha)}$ is sparse as shown in Figures 3.2 and 3.3, the product operation in the above partial derivatives can be computed with a complexity of $\mathcal{O}(N)$ instead of $\mathcal{O}(N^2)$. By substituting Equations 3.52, 3.53, 3.54, and 3.55 into Equation 3.43, the partial derivatives of $F_{avg}(\mathbf{s})$ with respect to $\Re(\mathbf{s})$ and $\Im(\mathbf{s})$ are obtained. Finally, the partial derivative of the average sidelobe with respect to the phases is derived by substituting the obtained partial derivatives into Equation 3.42. Because of the sum operation in Equation 3.43, the computational complexity for each iteration of the wideband gradient descent method is approximately $\mathcal{O}(|\mathcal{Q}|N)$. Since the set size of the RoI scales linearly with the sequence length, the computational complexity of the algorithm can be approximated as $\mathcal{O}(N^2)$.

The gradient descent method for the WAF shaping is summarized in Algorithm 4. The Adam optimizer [37] is used to determine the step size of the gradient descent for each iteration. It is important to note that, since the original optimization problem is non-convex, the gradient descent method does not guarantee convergence to a global minimum. The optimality of the gradient descent method may depend on the initialization of the sequence $\mathbf{s}$. For simplicity, a randomly generated BPSK sequence is used to initialize the algorithm. Alternative initialization strategies will be discussed in Section 3.4.3.3.

Algorithm 4 Gradient descent method for wideband ambiguity function shaping

1: Define the bandwidth B, the carrier frequency f_c, and the sequence length N
2: Initialize the unimodular sequence $\mathbf{s}$, and define the RoI $\mathcal{Q}$
3: Generate $\mathbf{A}_{(\tau,\alpha)}$ $\forall(\tau,\alpha) \in \mathcal{Q}$ using Equations 3.7
4: **while** convergence conditions are not satisfied **do**
5: Derive the gradient $\boldsymbol{\delta} := \frac{\partial F_{avg}(\mathbf{s})}{\partial \boldsymbol{\theta}}$ using Equations 3.42, 3.43, 3.52, 3.53, 3.54, and 3.55
6: Use Adam optimizer [37] to update the phases $\boldsymbol{\theta}$ of the unimodular sequence
7: Derive the update of the sequence $\mathbf{s}$ by projecting the phase vector to the unimodular space: $\mathbf{s} \leftarrow e^{j\boldsymbol{\theta}}$
8: **end while**

3.3.4 Narrowband approximation method

Since the actual moving speed in the current implementation of Fugro's SAS systems is usually less than 1-2 m/s, the velocity range for the RoI of the WAF shaping can be selected to be a small range (e.g., $v \in [-1, 1]$ m/s). In this case, the difference between the WAF and the NAF may become negligible. Therefore, we may use the NAF shaping algorithms for the WAF shaping with an acceptable error. In this section, we aim to figure out when the NAF algorithms can be used to generate wideband sequences (waveforms).

The WAF of the baseband signal $x_{bb}(t)$ with a carrier frequency of f_c is given by

$$\chi(\tau,\alpha) = \sqrt{\alpha}\int_{-\infty}^{\infty} x_{bb}(t)e^{j2\pi f_c t}x_{bb}^*(\alpha(t-\tau))e^{-j2\pi f_c\alpha(t-\tau)}dt \tag{3.56}$$

The above expression can be expressed using convolution as follows

$$\chi(\tau,\alpha) = \left(x_{bb}(\tau)e^{j2\pi f_c\tau}\right) * \left(\sqrt{\alpha}x_{bb}^*(-\alpha\tau)e^{j2\pi f_c\alpha\tau}\right) \tag{3.57}$$

For a fixed Doppler time-scaling factor α, the Fourier transform of the WAF on the delay axis is given by

$$\mathcal{F}(\chi(\tau,\alpha)) = \mathcal{F}(x_{bb}(\tau)e^{j2\pi f_c\tau})\mathcal{F}(\sqrt{\alpha}x_{bb}^*(-\alpha\tau)e^{j2\pi f_c\alpha\tau}) \tag{3.58}$$

Under the narrowband condition (Equation 2.22), the WAF can be approximated as the NAF. The NAF of the baseband signal $x_{bb}(t)$ with a carrier frequency of f_c is given by,

$$\chi_N(\tau,v) = \int_{-\infty}^{\infty} x_{bb}(t)e^{j2\pi f_c t}x_{bb}^*(t-\tau)e^{-j2\pi f_c\frac{c+v}{c-v}(t-\tau)}dt \tag{3.59}$$

which can be expressed using convolution as

$$\chi_N(\tau,v) = \left(x_{bb}(\tau)e^{j2\pi f_c\tau}\right) * \left(x_{bb}^*(-\tau)e^{j2\pi f_c\frac{c+v}{c-v}\tau}\right) \tag{3.60}$$

For a fixed target velocity v, the Fourier transform of the NAF on the delay axis is given by

$$\mathcal{F}(\chi_N(\tau,v)) = \mathcal{F}(x_{bb}(\tau)e^{j2\pi f_c\tau})\mathcal{F}(x_{bb}^*(-\tau)e^{j2\pi f_c\frac{c+v}{c-v}\tau}) \tag{3.61}$$

In the frequency domain, the error because of using the narrowband condition can be written as

$$\begin{aligned}\epsilon &= \int_{-\infty}^{\infty} |\mathcal{F}(\chi(\tau,\alpha)) - \mathcal{F}(\chi_N(\tau,v))|df \\ &= \int_{-\infty}^{\infty} |\mathcal{F}(x_{bb}(\tau)e^{j2\pi f_c\tau})(\mathcal{F}(\sqrt{\alpha}x_{bb}^*(-\alpha\tau)e^{j2\pi f_c\alpha\tau}) - \mathcal{F}(x_{bb}^*(-\tau)e^{j2\pi f_c\alpha\tau}))|df\end{aligned} \tag{3.62}$$

Assuming the Fourier transform of $x_{bb}^*(\tau)$ is given by

$$\mathcal{F}(x_{bb}(\tau)) = \int_{-\infty}^{\infty} x_{bb}(\tau)e^{-j2\pi f\tau}d\tau = X_{bb}(f) \tag{3.63}$$

the Fourier transform of $x_{bb}^*(-\alpha\tau)e^{j2\pi f_c\alpha\tau}$ can be expressed as

$$\begin{aligned}\mathcal{F}(x_{bb}^*(-\alpha\tau)e^{j2\pi f_c\alpha\tau}) &= \int_{-\infty}^{\infty} x_{bb}^*(-\alpha\tau)e^{j2\pi f_c\alpha\tau}e^{-j2\pi f\tau}d\tau \\ &= \int_{-\infty}^{\infty} x_{bb}^*(-\alpha\tau)e^{-j2\pi(f-f_c\alpha)\tau}d\tau\end{aligned} \tag{3.64}$$

Defining $\tau' = -\alpha\tau$, we obtain

$$\tau = -\frac{\tau'}{\alpha} \tag{3.65}$$

$$d\tau = -\frac{1}{\alpha}d\tau' \tag{3.66}$$

By substituting Equations 3.65 and 3.66 into Equation 3.64, the Fourier transform of $x_{bb}^*(-\alpha\tau)e^{j2\pi f_c\alpha\tau}$ is derived as

$$\begin{aligned}\mathcal{F}(x_{bb}^*(-\alpha\tau)e^{j2\pi f_c\alpha\tau}) &= \frac{1}{\alpha}\int_{-\infty}^{\infty} x_{bb}^*(\tau')e^{-j2\pi(f_c-f/\alpha)\tau'}d\tau' \\ &= \left(\frac{1}{\alpha}\int_{-\infty}^{\infty} x_{bb}(\tau')e^{-j2\pi(f/\alpha-f_c)\tau'}d\tau'\right)^* \\ &= \frac{1}{\alpha}X_{bb}^*(f/\alpha - f_c)\end{aligned} \tag{3.67}$$

Similarly, the Fourier transforms of $x_{bb}(\tau)e^{j2\pi f_c\tau}$ and $x_{bb}^*(-\tau)e^{j2\pi f_c\alpha\tau}$ are given by

$$\mathcal{F}(x_{bb}(\tau)e^{j2\pi f_c\tau}) = X_{bb}(f - f_c) \tag{3.68}$$

$$\mathcal{F}(x_{bb}^*(-\tau)e^{j2\pi f_c\alpha\tau}) = X_{bb}^*(f - f_c\alpha) \tag{3.69}$$

Therefore, the absolute error between the WAF and the NAF in the frequency domain is written as

$$\epsilon = \int_{-\infty}^{\infty} |X_{bb}(f - f_c)\left(\frac{1}{\sqrt{\alpha}}X_{bb}^*(\frac{f}{\alpha} - f_c) - X_{bb}^*(f - f_c\alpha)\right)|df \tag{3.70}$$

Clearly, when $\alpha = \frac{c+v}{c-v} = 1$, the absolute error between the WAF and the NAF is 0.

To easily numerically evaluate the absolute error, we assume that the spectrum $X_{bb}(f)$ is uniform in the frequency domain and can be represented by

$$X_{bb}(f) = \begin{cases} 1, & -\frac{B}{2} \leq f < \frac{B}{2} \\ 0, & otherwise \end{cases} \tag{3.71}$$

Suppose f_c, B. and α are given, the error between the WAF and the NAF can be calculated through the integral shown in Equation 3.70. When $f_c = 2000$ Hz and $c = 1500$ m/s, the error between the WAF and the NAF for different sets of B and v (α is converted into v) is shown in Figure 3.4. The error map result generally matches with the narrowband condition (Equation 2.22). Lower $|v|$ and B/f_c will reduce the absolute error between the WAF and the NAF. It is also worth noting that when the speed of the target is small enough (e.g., less than 1 m/s), the absolute error will be low even for $B/f_c \geq 1$. Therefore, the NAF shaping algorithms described in [23, 24, 26, 27, 25, 29, 31] might be used as a good approximation for the WAF shaping algorithms when the velocity range of the RoI is small enough. In this thesis, the narrowband gradient descent method [25] and the AISO algorithm [29] are used as the approximation for the shaping of the WAF, because of their good sidelobe minimization performance.

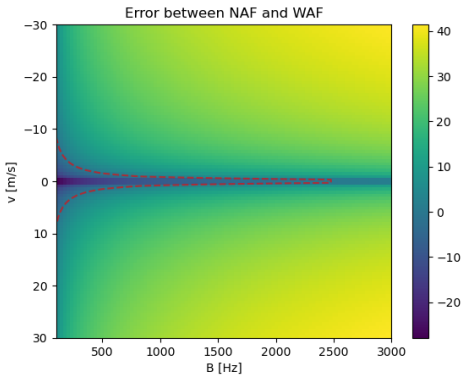


Figure 3.4: Error map illustrating the difference between the WAF and the NAF with different bandwidths and velocities. The carrier frequency was set to 2000 Hz, and the propagation speed of the wave was assumed to be 1500 m/s. The bandwidth ranges from 100 to 3000 Hz. The velocity ranges from -30 to 30 m/s. The absolute error is presented on a dB scale. The red dashed line represents the 0 dB contour line.

3.4 Evaluation and comparison of the proposed algorithms

In this section, we aim to evaluate the sidelobe minimization performance of the four proposed algorithms by using them to generate sequences with desired WAF properties. The simulation settings and evaluation metrics are described in Section 3.4.1. Then, the simulation results are presented in Section 3.4.2. Finally, the results are discussed and evaluated in Section 3.4.3.

3.4.1 Simulation settings and evaluation metrics

Since WAF shaping algorithms are generally not as efficient as NAF shaping algorithms due to the complexity of the WAF shaping problem, the length of the sequences we aimed to design was set to $N = 63$ in the first simulation. The bandwidth and the carrier frequency were set to $B = 2000$ Hz and $f_c = 2000$ Hz, respectively. The propagation speed of the sound wave was assumed to be $c = 1500$ m/s in all simulations. We considered a normal WAF shaping problem with a RoI

$$\mathcal{Q}_1 = \mathcal{Q}_{A1} \setminus \mathcal{Q}_{B1} \tag{3.72}$$

with

$$\mathcal{Q}_{A1} = \{(\tau, v) | -8/B \leq \tau \leq 8/B \text{ and } -3v_n \leq v \leq 3v_n\} \tag{3.73}$$
$$\mathcal{Q}_{B1} = \{(\tau, v) | -1/B < \tau < 1/B \text{ and } -v_n \leq v \leq v_n\} \tag{3.74}$$

where

$$v_n = \frac{cB}{2f_cN} \tag{3.75}$$

v_n represents the normalized velocity and is the first zero crossing point on the velocity-axis of the NAF of a rectangular pulse, as proven in [9]. Thus, it can also be considered as the velocity resolution of a waveform whose AF has a thumbtack shape. The RoI was selected to be a rectangular region around the mainlobe, but did not include the region corresponding to the mainlobe, $\mathcal{Q}_{B1}$. To ensure that the sidelobe in the RoI is minimized without missing any significant sidelobe, the sampling rate for the velocity v and the delay τ in the RoI was set to $v_n/4$ and $1/B$, respectively. The weight $w_{(\tau,\alpha)}$ in Algorithms 2, 3, and 4 is defined as

$$w(\tau, \alpha) = \begin{cases} 1, & (\tau, \alpha) \in \mathcal{Q}_1 \\ 0, & \text{otherwise} \end{cases} \tag{3.76}$$

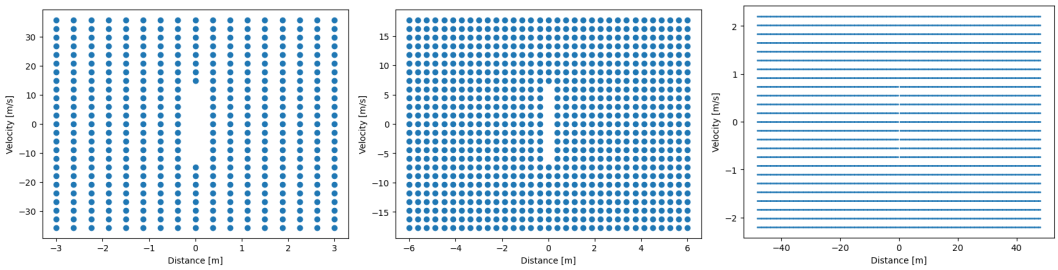
The RoI $\mathcal{Q}_1$ is shown in Figure 3.5a, where the delay-axis is converted into the distance-axis. In the first simulation, all algorithms proposed in Section 3.3 were evaluated. At the same time, the original USSM algorithm was simulated in the same setting for comparison. Following the selection of the parameter ζ in [32], we set $\zeta = \{100, 150, 200\}$ in the USSM algorithm. Because of the extra normalization term $1/|\mathcal{Q}|$ in the objective function of the USSM-avg formulation (Equation 3.31), we chose $\zeta = \{2500, 5000, 7500, 10000\}$ for the USSM-avg algorithm. Additionally, random BPSK sequences were used for the initialization of the wideband gradient descent method, unless otherwise specified.

In the second simulation, we considered a WAF shaping problem for a sequence with a length of $N = 127$. The settings for B, f_c, and $w_{(\tau,\alpha)}$ were the same as that of the first simulation. The RoI of the second simulation was defined as

$$\mathcal{Q}_2 = \mathcal{Q}_{A2} \setminus \mathcal{Q}_{B2} \tag{3.77}$$

with

$$\mathcal{Q}_{A2} = \{(\tau, v) | -16/B \leq \tau \leq 16/B \text{ and } -3v_n \leq v \leq 3v_n\} \tag{3.78}$$
$$\mathcal{Q}_{B2} = \{(\tau, v) | -1/B < \tau < 1/B \text{ and } -v_n \leq v \leq v_n\} \tag{3.79}$$



(a) RoI $\mathcal{Q}_1$ ($N = 63$) (b) RoI $\mathcal{Q}_2$ ($N = 127$) (c) RoI $\mathcal{Q}_3$ ($N = 1023$)

Figure 3.5: RoI of the simulations for evaluating the sidelobe minimization performance of the proposed algorithms in Section 3.3. The delay axis is converted into the distance axis using $R = c\tau/2$.

The RoI $\mathcal{Q}_2$ is shown in Figure 3.5b. In the second simulation, all algorithms proposed in Section 3.3 were simulated. The ζ in the USSM-avg algorithm was set to $\zeta = \{10000, 15000, 20000, 25000\}$ for the second simulation.

In the third simulation, a more realistic waveform design scenario is considered, where we considered the design of the periodic WAF of a periodic sequence with a length of $N = 1023$. The RoI of the third simulation was defined as

$$\mathcal{Q}_3 = \mathcal{Q}_{A3} \setminus \mathcal{Q}_{B3} \tag{3.80}$$

with

$$\mathcal{Q}_{A3} = \{(\tau, v) | -128/B \leq \tau \leq 128/B \text{ and } -3v_n \leq v \leq 3v_n\} \tag{3.81}$$
$$\mathcal{Q}_{B3} = \{(\tau, v) | -1/B < \tau < 1/B \text{ and } -v_n \leq v \leq v_n\} \tag{3.82}$$

The RoI of the third simulation is shown in Figure 3.5c. It is worth noting that since the periodic sequence was considered in the third simulation, the actual RoI in the periodic WAF can be approximately seen as a periodic region on the delay-axis with a PRI of $1023/B$. The settings for B, f_c, and $w_{(\tau,\alpha)}$ were the same as that of the first two simulations. In the third simulation, only the wideband gradient descent method and the narrowband approximation method (with the narrowband gradient descent method) were evaluated, because these two algorithms are relatively more efficient compared with the rest of the algorithms for the WAF shaping. For a sequence length of $N = 1023$, it became computationally prohibitive to run the other algorithms within a reasonable amount of time.

The maximum sidelobe (Equation 3.13), the average sidelobe (Equation 3.11), and the algorithm running time were used as the evaluation metrics for different algorithms. The SCS solver [38, 39] was used to solve the SDP problems in the USSM-avg algorithm, randomization-based method, and original USSM algorithm [32].

3.4.2 Simulation results

3.4.2.1 Results of WAF shaping of sequences with $N = 63$

Table 3.1 illustrates the running time, the maximum sidelobe, and the average sidelobe of different algorithms in the first simulation. Among all algorithms in the first simulation, the wideband gradient descent method has the best average sidelobe minimization performance, achieving an average sidelobe of -60.15 dB. Since the original USSM algorithm aims to minimize the maximum sidelobe in the RoI, sequences generated by USSM algorithms generally have lower maximum sidelobe in the RoI compared with sequences generated by other algorithms. Among the three values for the parameter ζ of the USSM, the USSM with $\zeta = 200$ generated the sequence with the lowest maximum sidelobe in the RoI. For a proper selection of ζ, the average sidelobe minimization performance of the USSM-avg algorithm is close to that of the wideband gradient descent method, but the USSM-avg is not computationally efficient and requires much longer running time. When $\zeta = \{7500, 10000\}$, we failed to obtain unimodular sequences from the USSM-avg, since the eigenvalue residuals of $\mathbf{X}_{opt}$ are larger than 10^{-3}. Narrowband approximation methods using the narrowband gradient descent method [25] and the AISO algorithm [29] require the least running time of all methods, but introduce an extra error because of using the narrowband assumption in the wideband scenario, which makes the sidelobe minimization suboptimal. The randomization-based method has a similar running time as that of the narrowband approximation method, but its sidelobe minimization performance is the worst of the four proposed methods. Additionally, in algorithms where the average sidelobe is used as the objective function, there is a general correlation between the maximum sidelobe level and the average sidelobe level. Minimizing the average sidelobe in the RoI during optimization simultaneously reduces the maximum sidelobe in the RoI.

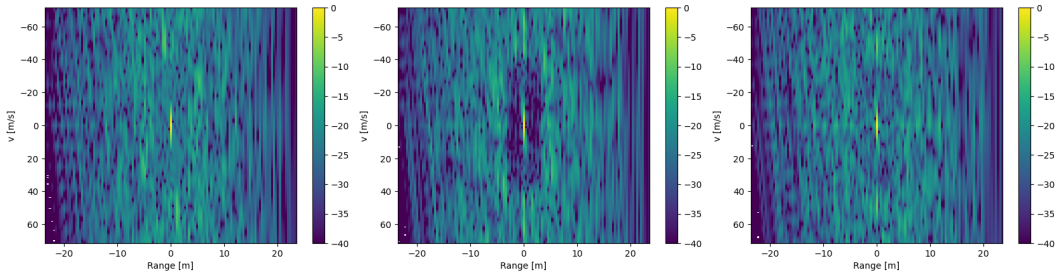
The WAF of sequences generated by different algorithms in the first simulation are shown in Figure 3.6. Among all generated sequences, sequences generated by the USSM-avg ($\zeta = 5000$) and wideband gradient descent have the best sidelobe minimization performance in the designed RoI of their WAF. The WAF of sequences generated by the USSM ($\zeta = 200$) and the randomization-based method is similar to that of a random BPSK signal with a thumbtack shape mainlobe as shown in Figure 2.6c. It is even difficult to determine the designed RoI from their WAFs. Narrowband approximation methods using the narrowband gradient descent method and the AISO algorithm managed to reduce the average sidelobe, but their RoI is distorted because of the approximation error.

3.4.2.2 Results of WAF shaping of sequences with $N = 127$

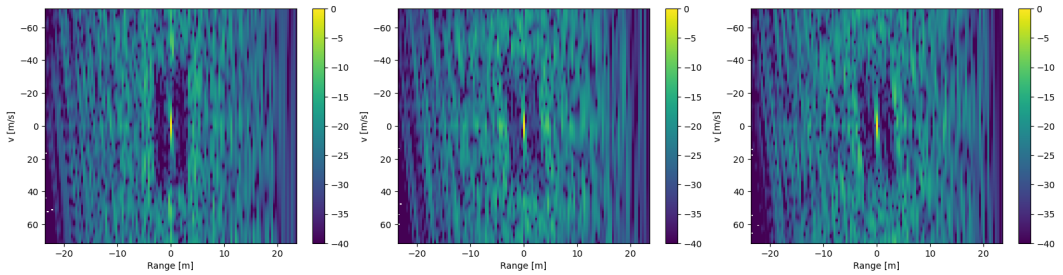
The running time, the maximum sidelobe, and the average sidelobe of different algorithms in the second simulation are shown in Table 3.2. The wideband gradient descent method still has the best average sidelobe minimization performance, and the sequence generated by the algorithm achieves an average sidelobe of -65.04 dB in the RoI of its WAF. The USSM-avg algorithm can generate sequences with a similar average sidelobe level, but requires much longer computation time. Additionally, the USSM-avg

Table 3.1: Running time, maximum sidelobe, and average sidelobe in the RoI of WAF shaping algorithms in the first simulation ($N = 63$, $f_c = 2000$ Hz, and $B = 2000$ Hz)

WAF shaping algorithm	Running time (minutes:seconds)	Maximum sidelobe (dB)	Average sidelobe (dB)
USSM [32] ($\zeta = 100$)	13:32.3	-21.11	-48.31
USSM ($\zeta = 150$)	13:51.4	-21.23	-47.71
USSM ($\zeta = 200$)	17:47.7	**-21.26**	-47.66
USSM-avg ($\zeta = 2500$)	02:03.2	-16.09	-57.93
USSM-avg ($\zeta = 5000$)	01:31.7	-16.10	-58.07
USSM-avg ($\zeta = 7500$)	01:40.1	-	-
USSM-avg ($\zeta = 10000$)	01:47.9	-	-
Randomization-based	00:13.3	-15.40	-48.07
Wideband gradient descent	00:25.4	-17.10	**-60.15**
Narrowband gradient descent [25]	00:08.5	-15.88	-53.23
Narrowband AISO [29]	**00:05.5**	-12.64	-50.32



(a) USSM ($\zeta = 200$) (b) USSM-avg ($\zeta = 5000$) (c) Randomization-based



(d) Wideband gradient descent (e) Narrowband gradient descent (f) Narrowband AISO

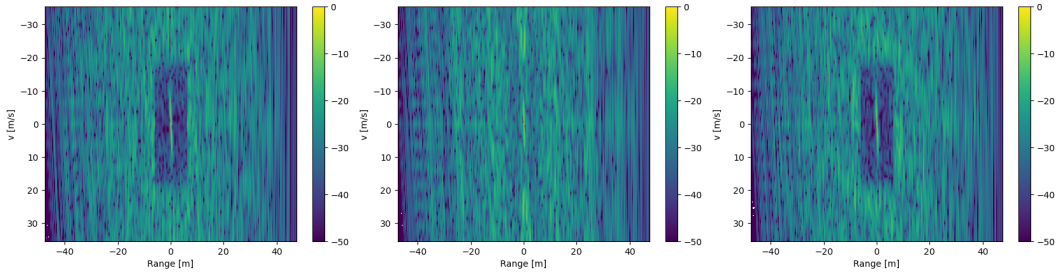
Figure 3.6: WAF for sequences ($N = 63$, $f_c = 2000$ Hz, and $B = 2000$ Hz) generated by different algorithms in the first simulation. The Doppler-axis and delay-axis of the WAF were converted into the velocity-axis and the range-axis, respectively. The value of the WAF is presented on a dB scale.

algorithm failed to generate the sequence, when $\zeta = 25000$. Among the algorithms simulated in the second simulation, narrowband approximation methods require the least running time, but still have the suboptimal sidelobe minimization performance because of the approximation error. The sequence generated by the randomization-based method has the worst performance on the minimization of the maximum sidelobe and the average sidelobe. The reason for the poor performance of the randomization-based method will be discussed in Section 3.4.3.2. The WAF of sequences generated by different algorithms in the second simulation is shown in Figure 3.7. These WAF results

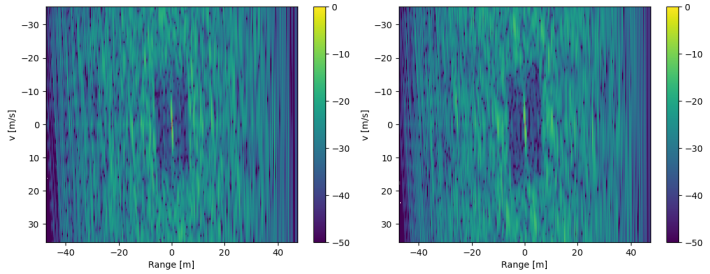
are generally similar to the results shown in Figure 3.6.

Table 3.2: Running time, maximum sidelobe, and average sidelobe in the RoI of WAF shaping algorithms in the second simulation ($N = 127$, $f_c = 2000$ Hz, and $B = 2000$ Hz)

WAF shaping algorithm	Running time (minutes:seconds)	Maximum sidelobe (dB)	Average sidelobe (dB)
USSM-avg ($\zeta = 10000$)	11:56.2	-16.67	-64.38
USSM-avg ($\zeta = 15000$)	11:51.3	-16.46	-63.90
USSM-avg ($\zeta = 20000$)	12:10.6	-16.47	-64.40
USSM-avg ($\zeta = 25000$)	13:40.6	-	-
Randomization-based	01:31.6	-16.03	-53.56
Wideband gradient descent	01:31.4	**-17.05**	**-65.04**
Narrowband gradient descent	00:11.0	-13.88	-58.49
Narrowband AISO	**00:10.3**	-13.22	-59.01



(a) USSM-avg ($\zeta = 20000$) (b) Randomization-based (c) Wideband gradient descent



(d) Narrowband gradient descent (e) Narrowband AISO

Figure 3.7: WAF for sequences ($N = 127$, $f_c = 2000$ Hz, and $B = 2000$ Hz) generated by different algorithms in the second simulation.

3.4.2.3 Results of periodic WAF shaping of periodic sequences with $N = 1023$

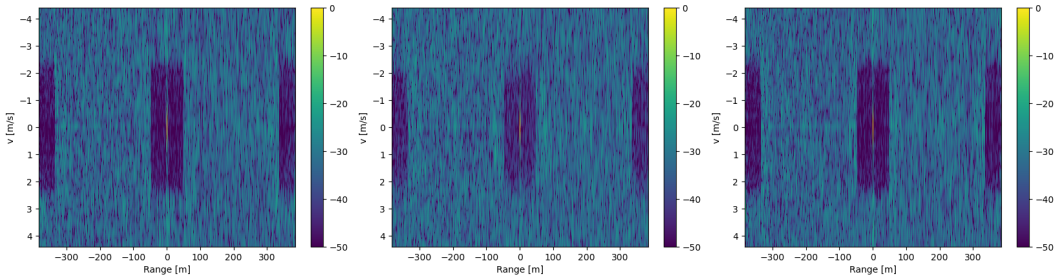
The running time, the maximum sidelobe, and the average sidelobe of different algorithms in the third simulation are shown in Table 3.3. It is important to note that the narrowband gradient descent method [25] is the only approach that has considered the shaping of the periodic AF in its original paper. Therefore, only the narrowband gradient descent method was used for the narrowband approximation. The sequence generated by the wideband gradient descent method still has the minimum average sidelobe in the RoI. The sidelobe minimization performance of the wideband gradient descent method might be slightly improved by a warm start (initializing the algorithm

with the sequence generated by the narrowband gradient descent method). The warm start technique for the wideband gradient descent method will be discussed in detail in Section 3.4.3.3. The narrowband gradient descent method only used 40.5 seconds to generate the sequence with a length of $N = 1023$, because it has a computational complexity of $\mathcal{O}(N \log N)$. On the other hand, the wideband gradient descent method is less efficient and used approximately 60 minutes to generate sequences with the same length.

The periodic WAF of sequences generated by different algorithms in the third simulation is shown in Figure 3.8. Since the periodic WAF shaping is considered in the third simulation, the low sidelobe region is approximately periodic on the delay-axis. Due to the choice of a smaller velocity range in the RoI in the third simulation, the approximation error of the narrowband gradient descent method is reduced. However, as shown in Figure 3.8b, the distortion caused by using the narrowband assumption in the wideband scenario can still be found on the upper and the lower edges of the RoI.

Table 3.3: Running time, maximum sidelobe, and average sidelobe in the RoI of WAF shaping algorithms in the third simulation ($N = 1023$, $f_c = 2000$ Hz, and $B = 2000$ Hz)

WAF shaping algorithm	Running time (minutes:seconds)	Maximum sidelobe (dB)	Average sidelobe (dB)
Wideband gradient descent	61:31.4	-16.79	-82.68
Narrowband gradient descent	**00:40.5**	-13.80	-76.12
Wideband gradient descent with a warm start[1]	64:15.3	-16.98	**-83.04**



(a) Wideband gradient descent (b) Narrowband gradient descent (c) Wideband gradient descent with a warm start

Figure 3.8: Periodic WAF for sequences ($N = 1023$, $f_c = 2000$ Hz, and $B = 2000$ Hz) generated by different algorithms in the third simulation.

3.4.3 Discussions and conclusion

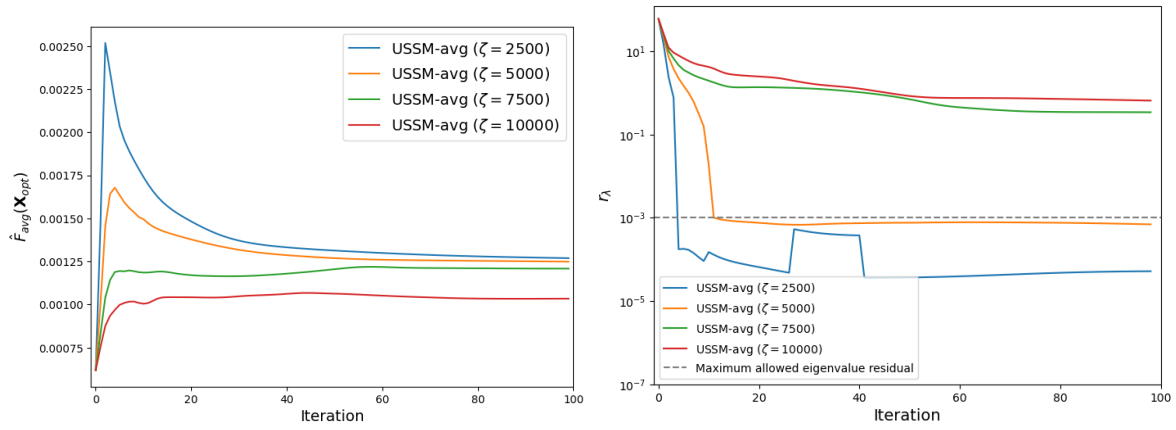
3.4.3.1 Effects of the parameter ζ in the USSM-avg algorithm

The convergence of the average sidelobe $\hat{F}_{avg}$ and the eigenvalue residual r_λ of the USSM-avg algorithm in the first and second simulation are shown in Figures 3.9 and 3.10. Due to the initialization with $\mathbf{W}_{opt} = \mathbf{0}_{2N}$, the first iteration of the USSM-avg

[1]The wideband gradient descent method was initialized with the sequence generated by the narrowband gradient descent method

algorithm is the same for all ζ. The difference of the convergence with different ζ can be observed starting from the second iteration. Since the first iteration only focuses on the minimization of the averages sidelobe and ignores the rank minimization, an increase on $\hat{F}_{avg}$ is observed for the first few iterations of the USSM-avg, as shown in Figures 3.9a and 3.10a. Following that, the average sidelobe $\hat{F}_{avg}$ decreases as the iteration number increases until the convergence. As shown in Figures 3.9b and 3.10b, the eigenvalue residual generally decreases as the iteration number increases, but may have some fluctuations.

In the first and second simulation, the effects of ζ in the USSM-avg algorithm generally match with our analysis in Section 3.3.1. That is, a higher value of ζ usually results in a lower average sidelobe $\hat{F}_{avg}$ but may lead to a non-rank-1 solution of $\mathbf{X}_{opt}$. On the other hand, a lower value of ζ ensures the rank-1 constraint on $\mathbf{X}_{opt}$, but does not necessarily guarantee the average sidelobe minimization performance. However, one exception to these properties can also be found in the second simulation. As shown in Figure 3.10, after approximately 80 iterations, the USSM-avg with $\zeta = 10000$ has the same average sidelobe minimization performance as the USSM-avg with $\zeta = 20000$. Therefore, we suggest using a grid search to find the optimal parameter ζ when applying the USSM-avg algorithm for the WAF shaping.



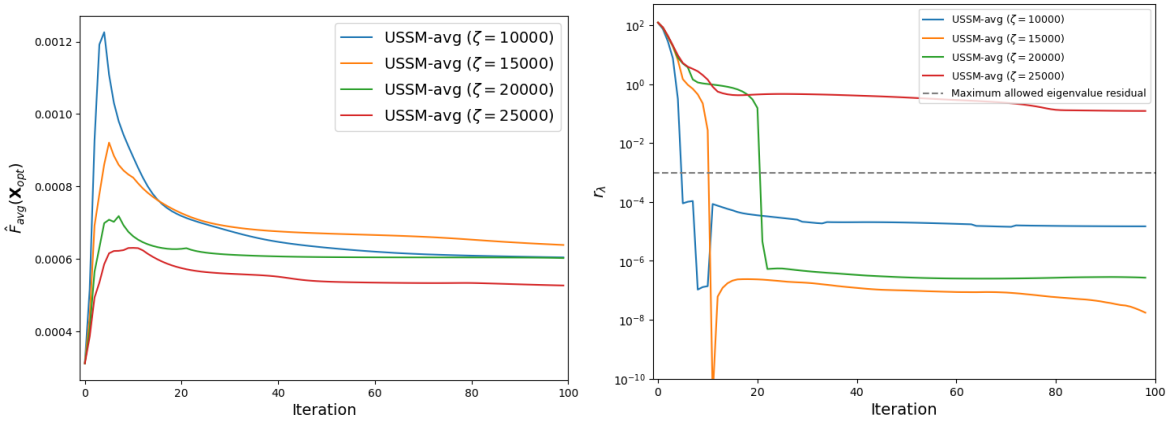
(a) $\hat{F}_{avg}$ vs. iteration

(b) r_λ vs. iteration

Figure 3.9: Convergence of the average sidelobe $\hat{F}_{avg}$ and the eigenvalue residual r_λ of the USSM-avg for $\zeta = \{2500, 5000, 7500, 10000\}$ in the first simulation ($N = 63$, $f_c = 2000$ Hz, and $B = 2000$ Hz). The maximum allowed eigenvalue residual was set to 10^{-3}. After projecting $\mathbf{X}_{opt}$ to a unimodular sequence, the USSM-avg with $\zeta = 5000$ leads to the waveform with the lowest average sidelobe.

3.4.3.2 Discussion on the randomization-based method

In this section, we aim to determine the reason for the poor sidelobe minimization performance of the randomization-based method, and then we want to improve the performance of this method. The first step of the randomization-based method is to solve $\mathbf{X}_{opt,R}$ from the convex optimization problem 3.38. The eigenvalue distributions



(a) $\hat{F}_{avg}$ vs. iteration

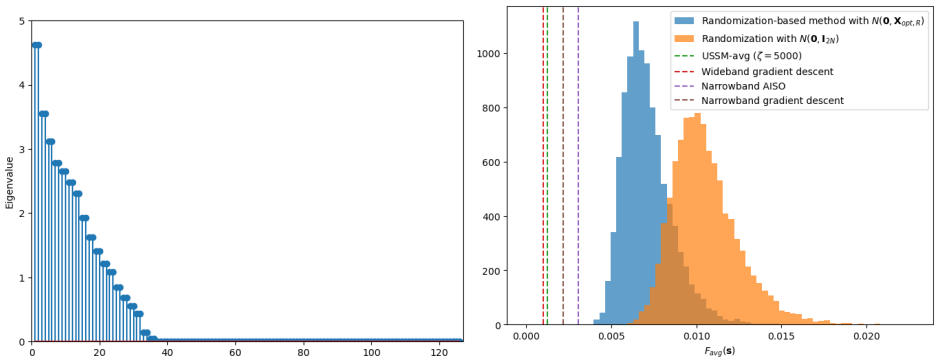
(b) r_λ vs. iteration

Figure 3.10: Convergence of the average sidelobe $\hat{F}_{avg}$ and the eigenvalue residual r_λ of the USSM-avg for $\zeta = \{10000, 15000, 20000, 25000\}$ in the second simulation ($N = 127$, $f_c = 2000$ Hz, and $B = 2000$ Hz). After projecting $\mathbf{X}_{opt}$ to a unimodular sequence, the USSM-avg with $\zeta = 20000$ leads to the waveform with the lowest average sidelobe.

of $\mathbf{X}_{opt,R}$ in the first and second simulation are shown in Figures 3.11a and 3.12a. Then, $\mathbf{X}_{opt,R}$ is used to generate a multivariate normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{X}_{opt,R})$ for generating random sequences. The distribution of the average sidelobe of random sequences generated from the distribution $\mathcal{N}(\mathbf{0}, \mathbf{X}_{opt,R})$ are shown in the blue histogram of Figures 3.11b and 3.12b. Compared with random sequences generated from $\mathcal{N}(\mathbf{0}, \mathbf{I}_{2N})$, random sequences generated from $\mathcal{N}(\mathbf{0}, \mathbf{X}_{opt,R})$ are likely to have a lower average sidelobe level in the RoI, but their average sidelobe level is not as good as of sequences generated by other methods. That means that the randomization-based method can generate sequences with a lower average sidelobe level to some extent, but its average sidelobe minimization performance is limited. As we discussed in Section 3.3.2, one reason for underperformance in sidelobe minimization of the randomization-based method is that its initial optimization problem is not a QPCP problem.

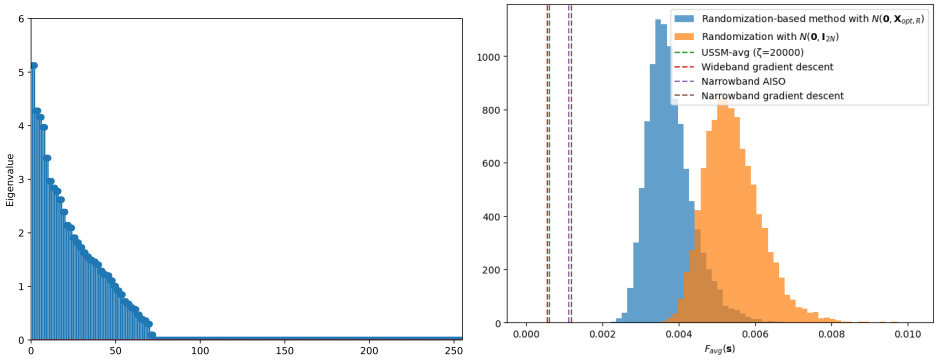
The performance of the randomization-based method is directly determined by the property of the matrix $\mathbf{X}_{opt,R}$. In the original randomization-based method, the derivation of $\mathbf{X}_{opt,R}$ can be seen as solving $\mathbf{X}_{opt}$ from the first iteration of the USSM-avg algorithm. Since the rank and the average sidelobe of the matrix $\mathbf{X}_{opt}$ is gradually reduced through iterative alternating minimization, the optimality of the randomization-based method might be improved by using the matrix $\mathbf{X}_{opt}$ obtained from a few iterations of the USSM-avg algorithm as the covariance matrix for the randomization. To verify the feasibility of this idea, the same simulation scenario described in the first simulation was used. Suppose the matrix $\mathbf{X}_{opt}$ obtained from n iterations of the USSM-avg ($\zeta = 10000$) is denoted as $\mathbf{X}_{opt,Rn}$. We used the matrices $\mathbf{X}_{opt,R2}$, $\mathbf{X}_{opt,R5}$, $\mathbf{X}_{opt,R10}$, $\mathbf{X}_{opt,R20}$ as the covariance matrix to construct multivariate distributions for the randomization.

The eigenvalue distribution of these matrices is shown in Figure 3.13. The rank of the matrix $\mathbf{X}_{opt,Rn}$ decreases as the number of iterations increases, which matches with the eigenvalue residual vs. iteration plot for the USSM-avg ($\zeta = 10000$) shown



(a) Eigenvalues of $\mathbf{X}_{opt,R}$ (b) Distribution of the average sidelobe $F_{avg}(\mathbf{s})$

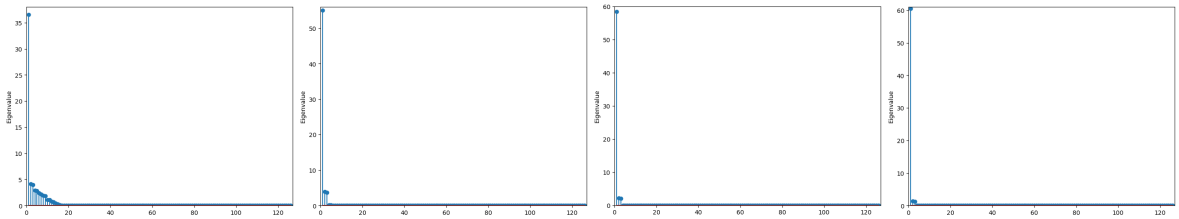
Figure 3.11: (a) Eigenvalues of $\mathbf{X}_{opt,R}$, and (b) distribution of the average sidelobe of sequences generated by the randomization-based method in the first simulation ($N = 63$, $f_c = 2000$ Hz, and $B = 2000$ Hz). A total of 10000 sequences were generated for the given distribution using the randomization-based method.



(a) Eigenvalues of $\mathbf{X}_{opt,R}$ (b) Distribution of the average sidelobe $F_{avg}(\mathbf{s})$

Figure 3.12: (a) Eigenvalues of $\mathbf{X}_{opt,R}$, and (b) distribution of the average sidelobe of sequences generated by the randomization-based method in the second simulation ($N = 127$, $f_c = 2000$ Hz, and $B = 2000$ Hz). A total of 10000 sequences were generated for the given distribution using the randomization-based method.

in Figure 3.9b. The distribution of the average sidelobe of sequences generated by the different multivariate distributions is shown in Figure 3.14. As we increase the number of iterations, the distribution of the random samples becomes more concentrated and deterministic, which might be due to the decrease on the rank of the matrix $\mathbf{X}_{opt,R}$. At the same time, the mean value of the distribution is gradually decreasing as the number of iterations increases. It is also easy to see that the average sidelobe of sequences obtained from the modified randomization-based method has a lower bound equal to the minimum average sidelobe of the USSM-avg sequence. The WAF of the sequence with the minimum average sidelobe obtained from 10000 random samples of



(a) $\mathbf{X}_{opt,R2}$ (b) $\mathbf{X}_{opt,R5}$ (c) $\mathbf{X}_{opt,R10}$ (d) $\mathbf{X}_{opt,R20}$

Figure 3.13: Eigenvalue distribution of $\mathbf{X}_{opt,R2}$, $\mathbf{X}_{opt,R5}$, $\mathbf{X}_{opt,R10}$, and $\mathbf{X}_{opt,R20}$.

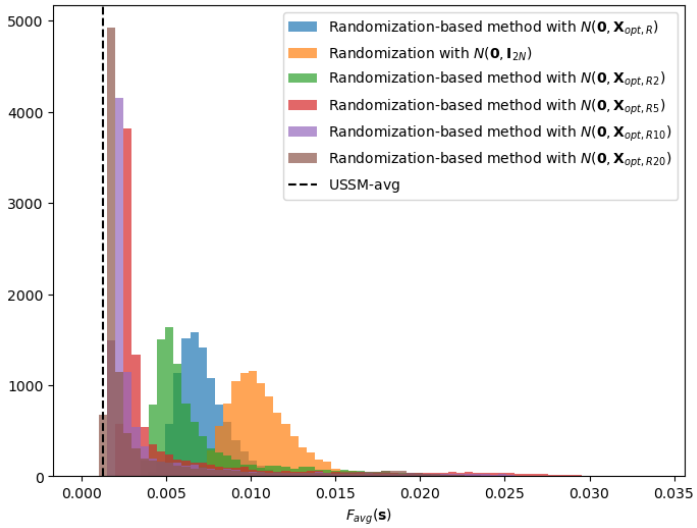


Figure 3.14: Histogram showing the distribution of the average sidelobe $F_{avg}(\mathbf{s})$ of sequences generated by multivariate distributions with different covariance matrices. A total of 10000 sequences ($N = 63$, $f_c = 2000$ Hz, and $B = 2000$ Hz) were generated for each multivariate distribution using the randomization-based method.

$\mathcal{N}(\mathbf{0}, \mathbf{X}_{opt,R20})$ is shown in Figure 3.15. The maximum sidelobe and the average sidelobe in the RoI of the sequence are -15.53 dB and -56.82 dB respectively, which are close to the sidelobe level of USSM-avg sequences. Therefore, instead of directly solving the optimization problem of Equation 3.38 to obtain the covariance matrix, we can improve the sidelobe minimization performance of the randomization-based method by using the covariance matrix $\mathbf{X}_{opt}$ derived from the first few iterations of the USSM-avg algorithm.

3.4.3.3 Warm start for the wideband gradient descent method

Since the optimization problem for the WAF shaping is non-convex, the gradient descent method does not guarantee finding the global minimum of the optimization problem. One way to improve the optimality of the gradient descent for a non-convex optimization is by using the warm start technique. Warm start refers to initializing the optimization problem with a solution that is close to the optimal point, which can help

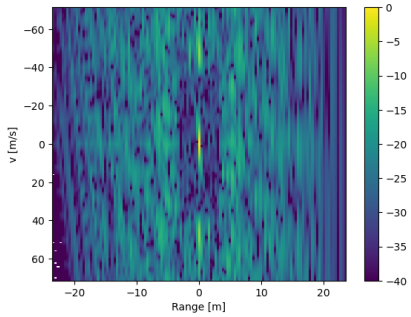


Figure 3.15: WAF for the sequence ($N = 63$, $f_c = 2000$ Hz, and $B = 2000$ Hz) with the minimum average sidelobe obtained from 10000 random samples of $\mathcal{N}(\mathbf{0}, \mathbf{X}_{opt,R20})$.

the convergence of the gradient descent method and potentially improve the solution's optimality. In this section, we aimed to determine whether the warm start can improve the optimality of the wideband gradient descent method.

In terms of the warm start for the wideband gradient descent method, we can initialize the gradient descent with sequences with a low average sidelobe level in the RoI. For example, the sequence generated by the narrowband gradient descent method was used as a warm start, as shown in Table 3.3. Since the narrowband gradient descent method is much more efficient compared to the wideband gradient descent method, the warm start process only used approximately 1% of the total running time. To evaluate the performance of the warm start technique, we used the same simulation setup as in the second simulation. In the simulation setup, a sequence with $B = 2000$ Hz, $f_c = 2000$ Hz, and $N = 127$ was designed to minimize its average sidelobe level in a RoI, as shown in Figure 3.5b. The wideband gradient descent method with different initializations was used to generate sequences. The sequences generated by the narrowband gradient descent method [25] and the narrowband AISO algorithm [29] were used for the warm start. For comparison, we also used random BPSK sequences for the initialization of the wideband gradient descent method.

Table 3.4: Maximum sidelobe, and average sidelobe in the RoI of sequences generated by the wideband gradient descent method with different initializations

Initialization	Maximum sidelobe (dB)	Average sidelobe (dB)
Random BPSK (average of 30 trials)	-16.85	-65.11
Narrowband gradient descent sequence	-16.78	**-65.50**
Narrowband AISO sequence	**-16.90**	-64.76

The maximum sidelobe, and the average sidelobe in the RoI of sequences generated by the wideband gradient descent method with different initializations are shown in Table 3.4. Compared to initializing with random BPSK sequences, warm starts using narrowband gradient descent sequences and AISO sequences do not significantly improve the maximum sidelobe level or the average sidelobe level of sequences generated by the wideband gradient descent method. In fact, using the narrowband AISO sequence as a warm start even results in a higher average sidelobe level than the mean

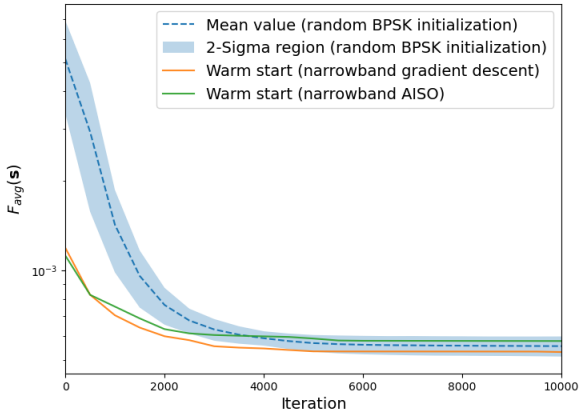


Figure 3.16: Convergence of the average sidelobe of the wideband gradient descent method with different initializations.

obtained from random BPSK initializations. However, as shown in Figure 3.16, the warm start technique still improves the convergence of the wideband gradient descent method. When warm-started, the wideband gradient descent method generally converges faster than when initialized with random BPSK sequences. In general, even if the warm start technique may not improve the optimality of the average sidelobe minimization performance, it still effectively reduces the number of iterations required for convergence. Therefore, we suggest using a warm start when applying the wideband gradient descent method for the WAF shaping.

3.4.3.4 Conclusion

In this chapter, four methods are proposed to handle the WAF and the periodic WAF shaping problem. In contrast to the USSM algorithm [32] that aims to minimize the maximum sidelobe, these four methods focus on the minimization of the average sidelobe in the RoI of the WAF. In the first simulation, we have shown that the proposed methods that focus on minimizing the average sidelobe generally have more significant sidelobe minimization performance in the RoI, compared to the USSM algorithm.

Among the four proposed methods, the wideband gradient descent method has the optimal average sidelobe minimization performance. For each iteration, the computational complexity of this method is $\mathcal{O}(N^2)$. Even if the narrowband approximation method is more efficient and generally has a computational complexity of $\mathcal{O}(N \log N)$ for the entire algorithm, it will underperform by approximately 6 dB to 10 dB on the average sidelobe in the RoI, compared to the wideband gradient descent method. As we discussed in Section 3.4.3.3, the convergence of the gradient descent method can be improved using a warm start (initializing the algorithm with sequences generated by the narrowband approximation method).

The USSM-avg algorithm has the second-best performance on the average sidelobe minimization. For a proper selection of the parameter ζ, the USSM-avg algorithm can have a sidelobe minimization performance close to that of the wideband gradient descent method. For each iteration, the computation complexity of this method is

$\mathcal{O}(N^3)$, which is also worse than that of the wideband gradient descent method. When using the USSM-avg algorithm, a grid search for the parameter ζ is suggested for improving the optimality of the algorithm, which is another fundamental drawback of the USSM algorithms.

The randomization-based method was proposed to alleviate the heavy computational complexity of the USSM-avg algorithm. However, the average sidelobe minimization performance of this method is not good, as shown in the results of the first and second simulation. In Section 3.4.3.2, we proposed a method that can improve the optimality of the randomization-based method by using the solution $\mathbf{X}_{opt}$ obtained from the first few iterations of the USSM-avg algorithm as the covariance matrix for the randomization. The optimality of the randomization-based method is improved as the number of iterations for the USSM-avg algorithm increases. In this way, a trade-off between the optimality and the running time is introduced. Additionally, it is also worth noting that the performance of the modified randomization-based method is upper bounded by that of the USSM-avg algorithm.

The narrowband approximation method refers to directly using the narrowband algorithms, such as the narrowband gradient descent algorithm [25] and the AISO algorithm [29], for the shaping of the WAF. The results in the first, second, and third simulations show that when the velocity range of the RoI is small, the narrowband approximation method can obtain good sidelobe minimization performance with a relative low error. Among the four proposed methods, the narrowband approximation method is the most efficient. However, the sidelobe minimization performance of this method is not as good as that of the USSM-avg algorithm and the wideband gradient descent method.

In general, among the algorithms for the WAF shaping, the wideband gradient descent method with a warm start has the best performance, considering both the optimality of the average sidelobe minimization and the computational complexity. Therefore, we decided to use the wideband gradient descent method to shape the periodic WAF of sequences for SAS imaging in the subsequent chapter.

Waveform Design for SAS imaging

4

In Chapter 2, we theoretically demonstrated that the waveform design problem for SAS systems can be approximately assumed to be a WAF shaping problem. However, it remains unclear how effective this approximation is in practice. In this chapter, we aim to validate the effectiveness of this approach through both simulations and real-world experiments. Specifically, we revisit the first sub-question of the main research question:

- *What is the relationship between the ambiguity properties of a waveform and its corresponding SAS imaging performance? Is ambiguity function shaping equivalent to waveform design for a SAS system?*

To answer this, we conduct a series of SAS imaging simulations using different waveforms, as described in Section 4.1. In addition, we also carry out field trials to further evaluate the real-world performance of waveforms obtained through WAF shaping. Details of the trials are provided in Section 4.2.

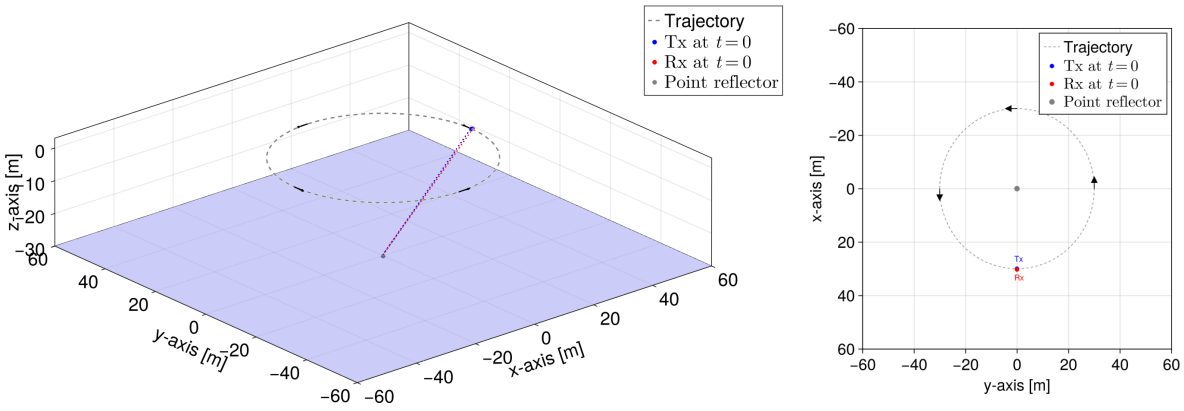
4.1 SAS imaging simulation

4.1.1 Simulation settings

To verify the relationship between the ambiguity properties and the SAS imaging performance, we considered the SAS imaging scenario shown in Figure 4.1. We aimed to image a point reflector placed at (0 m, 0 m, -30 m) with a transmitter and a receiver. The transmitter and the receiver are co-located with a spacing of 0.3 m. The transmitter-receiver pair was moving along a circular trajectory with a radius of 30 m at $z = 0$ m for SAS imaging. Unless otherwise stated, the vessel speed was set to approximately 1.05 m/s, such that the total integration time of the full circle trajectory is 180 s. While they were moving, a continuous waveform was transmitted and received simultaneously for imaging. The imaging domain is shown in the purple area in Figure 4.1a, which is a 120×120 m^2 square area centered at $z = -30$ m plane. The imaging resolution was set to 0.2 m. The bandwidth and the carrier frequency of waveforms were $B = 2000$ Hz and $f_c = 2000$ Hz, respectively. The algorithm mentioned in Algorithm 1 was used for SAS imaging.

For the first simulation, the travel trajectory of the first 1 s of the transmitter-receiver pair was integrated to image the point reflector with different waveforms (100% duty cycle), including the wideband gradient descent sequence ($N = 1023$) generated in the last chapter and other conventional waveforms. Here, we aim to show whether the waveform design for a short synetic aperture of the SAS imaging is equivalent to the WAF shaping. For the second simulation, the full circle trajectory was integrated to image the point reflector, using the same waveforms used in the first simulation.

Here, we aim to test the imaging performance of different waveforms in a more realistic scenario.



(a) 3D view (b) Top view

Figure 4.1: Simulation setup for the SAS imaging of a point reflector. The transmitter and the receiver were co-located with a spacing of 0.3 m. They were moving along a circle trajectory with a radius of 30 m at $z = 0$ m plane to image a point reflect placed at (0 m, 0 m, -30 m). The moving speed was set to approximately 1.05 m/s, such that the total integration time of the full circle trajectory is 180 s. At $t = 0$ s, the positions of the transmitter and the receiver are (30 m, 0 m,0 m) and (30.3, 0 m, 0 m), respectively.

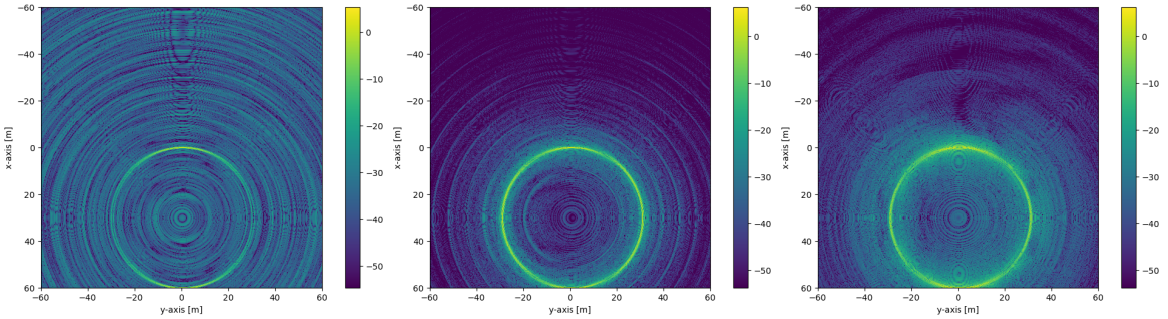
4.1.2 Simulation results

4.1.2.1 SAS imaging of a point reflector over a short trajectory

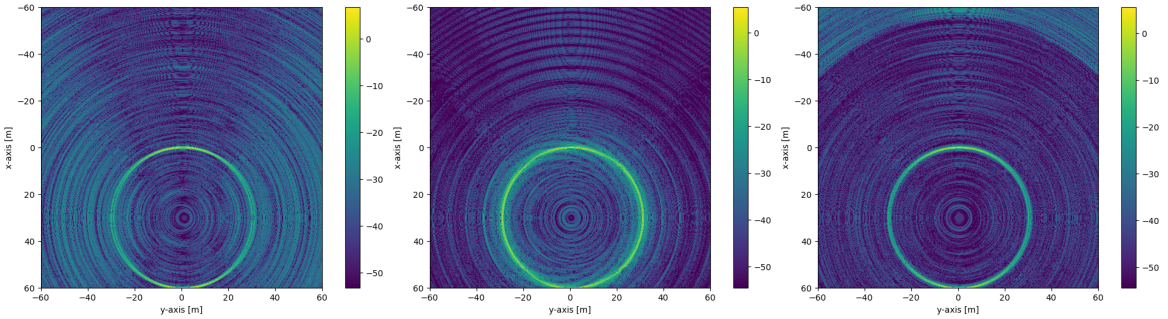
The SAS imaging results of a point reflector using different waveforms with an integration time of 1 s are shown in Figure 4.2. Here, the PRI of all waveforms was set to $N/B = 0.5115$ s, except for that of the non-repeating Random BPSK waveform. To analyze the relationship between WAF properties of waveforms and SAS imaging results, the WAF computed from approximately 1 s of corresponding waveforms is presented in Figure 4.3.

Because we consider a SAS imaging system with a single receiver, it can only estimate the range of the target in one dimension for a short integration time, and thus we observed a circular ambiguity of the point reflector in all SAS imaging results. This circular ambiguity is unavoidable for SAS systems with one receiver, and can only be resolved with an array of receivers or advanced imaging algorithms (e.g., the least squares methods). It is important to note that the actual ambiguity is a sphere with a radius corresponding to the distance between the transmitter-receiver pair and the point reflector. The circular ambiguity in the imaging domain is a cross-section of the sphere ambiguity.

Since the chirp-like waveforms, LFM, HFM, and Golomb sequences, have similar ambiguity function properties, their SAS imaging results are similar, as shown in Figures 4.2b, 4.2c, and 4.2e. Because of their ridge-shaped mainlobe in the WAF as shown in Figures 4.3b, 4.3c, and 4.3e, these waveforms are insensitive to Doppler shifts, only



(a) random BPSK (non-repeating) (b) LFM (PRI = 0.5115 s) (c) HFM (PRI = 0.5115 s)

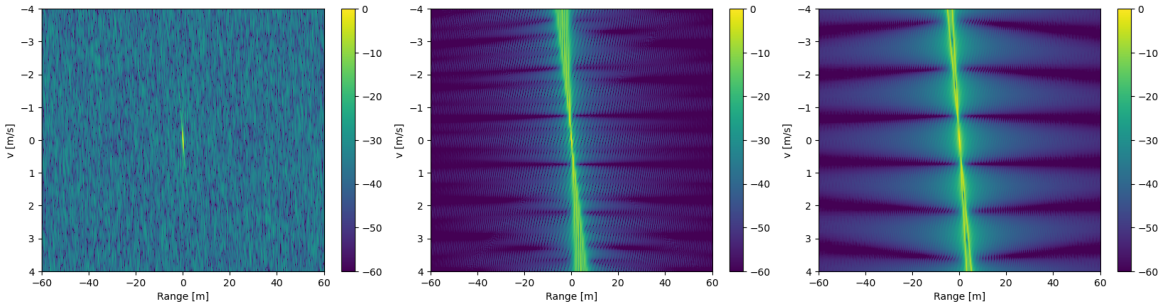


(d) m-sequence ($N = 1023$) (e) Golomb sequence ($N = 1023$) (f) Wideband gradient descent sequence ($N = 1023$)

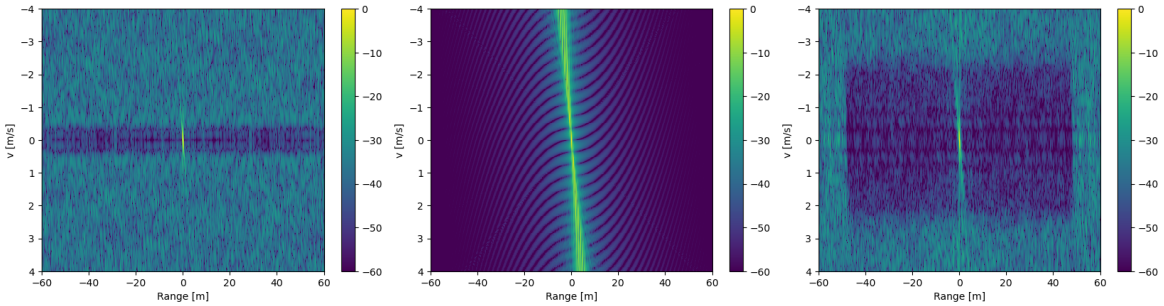
Figure 4.2: SAS imaging reconstructions of a point reflector using different waveforms, with an integration time of 1 s. At $t = 0$, the positions of the transmitter and the receiver are (30 m, 0 m, 0 m) and (30.3 m, 0 m, 0 m) respectively. Apart from the non-repeating random BPSK sequence, the other waveforms are periodic in time with a PRI of $N/B = 0.5115$ s. The image intensity is presented in a dB scale.

causing a slight displacement of the main lobe location. Consequently, the circular ambiguity of the point reflector of these waveforms has a higher intensity than that of other waveforms with a thumbtack-shaped mainlobe in the WAF. On the other hand, since the chirp-like waveforms have much lower sidelobes on two sides of the ridge-shaped mainlobe of their WAFs, these waveforms provide a relatively lower imaging sidelobe elsewhere as shown in their SAS imaging results, compared to m-sequence and random BPSK.

For waveforms with a thumbtack-shaped mainlobe in the WAF, their SAS imaging results have lower intensity on parts of the circular ambiguity of the point reflector, because of their Doppler sensitivity. Among these waveforms, the non-repeating random BPSK sequence has the worst imaging sidelobe performance, because of the uniform high sidelobes in its WAF as shown in Figure 4.3a. In Figure 4.2d, the SAS imaging result of the m-sequence has a lower sidelobe in a fan-shaped region at the center of the imaging domain, because of the perfect autocorrelation property of m-sequence. As shown in Figure 4.3d, the m-sequence has a low sidelobe in the ambiguity domain when $|v| < 0.4$ m/s. At $t = 0$ s, the absolute radial velocity map of voxels in the imaging domain is shown in Figure 4.4, where we see that the region corresponding to $|v| < 0.4$



(a) random BPSK (non-repeating) (b) LFM (PRI = 0.5115 s) (c) HFM (PRI = 0.5115 s)



(d) m-sequence ($N = 1023$) (e) Golomb sequence ($N = 1023$) (f) Wideband gradient descent sequence ($N = 1023$)

Figure 4.3: WAF for approximately 1 s of different waveforms. Apart from the non-repeating random BPSK sequence, the other waveforms are periodic in time with a PRI of $N/B = 0.5115$ s. Within a duration of 1 s, waveforms with a PRI of 0.5115 s have approximately two periods.

m/s happens to have the same shape as the low imaging sidelobe region we observed in Figure 4.3d.

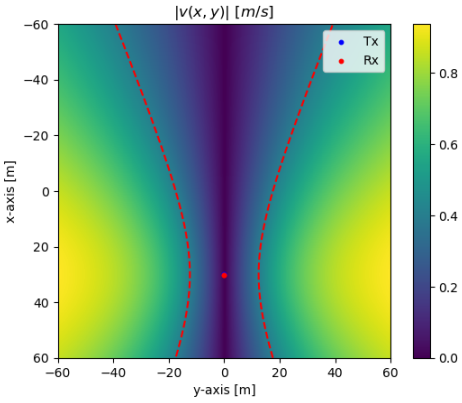


Figure 4.4: Absolute radial velocity of voxels in the imaging domain at $t = 0$ s. The red dashed line represents the 0.4 m/s contour line of the absolute radial velocity map. The value of the radial velocity is presented on a linear scale.

Among the waveforms with a PRI of 0.5115 s, the wideband gradient descent se-

quence ($N = 1023$) has the best SAS imaging performance. Not only does it have a narrower and low-intensity circular ambiguity of the point reflector because of its thumbtack mainlobe in the WAF, the imaging sidelobe of the wideband gradient descent sequence is even better than that of the chirp-like waveforms. However, a high imaging sidelobe region corresponding to the high sidelobe region in its WAF can also be observed at the upper corners of its SAS image. When the imaging domain increases from 120×120 m^2 to 240×240 m^2, we find that the low sidelobe region is within a large circle with a radius of approximately 85 m, as shown in Figure 4.5. Because the radial velocity in the imaging domain is less than the maximum velocity of the RoI of the wideband gradient descent sequence, the size of the low sidelobe region in the imaging domain should be determined by the maximum range of the RoI, which is around 48 m. For a short integration time, the low average sidelobe in the RoI will translate into a low imaging sidelobe region around the point reflector with a radius of 48 m. Because of the ambiguity of the point reflector in the imaging domain, the low imaging sidelobe region can be approximately seen as a region within a sphere with a radius of $48 + 30\sqrt{2}$ m, as shown in Figure 4.6. Thus, the cross-section of the sphere region with the imaging domain at $z = -30$ m is a circle region, and the radius of the low sidelobe region in the imaging domain is 85.3 m, which matches with our observations in Figure 4.5.

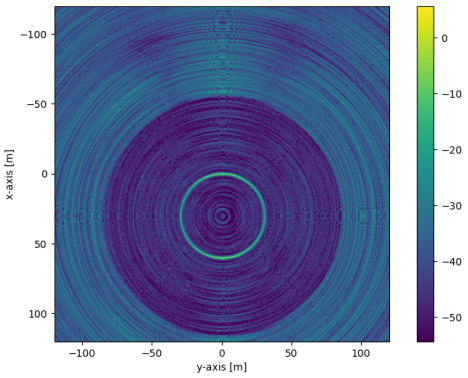


Figure 4.5: SAS imaging reconstruction of a point reflector using wideband gradient descent sequence (N=1023), with an integration time of 1 s. In this figure, the imaging domain was increased to 240×240 m^2. The low sidelobe region is within a circle with a radius of approximately 85 m, due to the maximum range within the low sidelobe region of the waveform's WAF.

The maximum unambiguous detection range of a periodic waveform is determined by its PRI or sequence length:

$$R_{max} = \frac{c\text{PRI}}{2} = \frac{cN}{2B} \tag{4.1}$$

For a PRI of 0.5115 s, the waveform can provide a maximum unambiguous detection range of approximately 383.6 m. As the detection range is much higher than needed, we also investigated the performance of unoptimized waveforms with a shorter PRI or sequence length. The SAS imaging results of the LFM waveform, Golomb sequence,

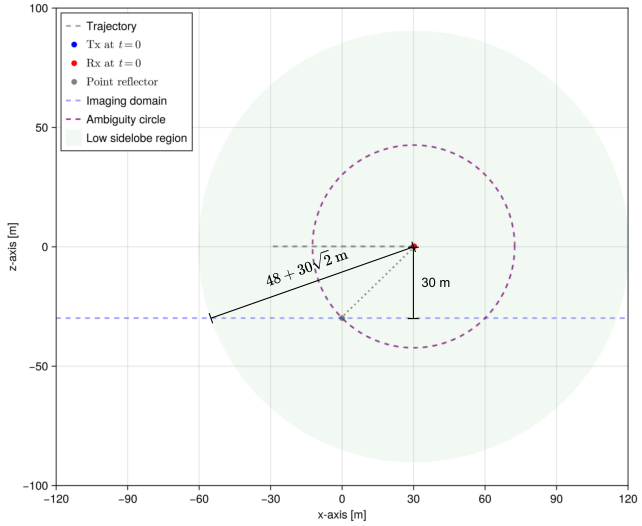


Figure 4.6: Schematic explaining the actual low sidelobe region of SAS imaging of the wide-band gradient descent sequence ($N = 1023$) with a short integration time of 1 s. At $t = 0$ s, the transmitter is placed at (30 m, 0 m, 0 m) on the water surface. According to the ambiguity function properties of the sequence, the low sidelobe region is roughly a sphere region with a radius of 48 m around the point reflector. After integrating the low sidelobe region over the ambiguity sphere (circle), the actual low sidelobe region can be approximately seen as a sphere region with a radius of $48 + 30\sqrt{2}$ m. The cross-section of the region with the imaging domain (a 240×240 m^2 plane at $z = -30$ m) is a circle with a radius of approximately 85.3 m.

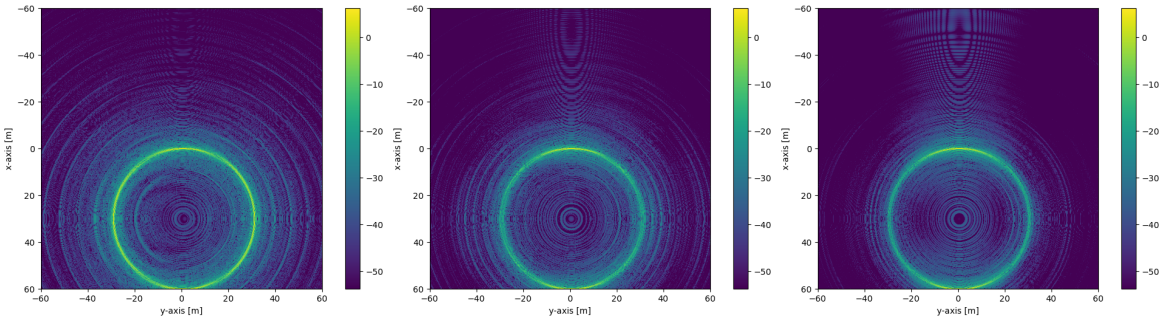
and m-sequence with a different PRI are shown in Figures 4.7, 4.9, and 4.11. The corresponding WAF of approximately 1 s of these waveforms is presented in Figures 4.8, 4.10, and 4.12.

For chirp-like waveforms (the LFM and the Golomb sequence), a lower PRI results in more waveform periods being integrated over a 1 s interval. This increased integration decreases the intensity of the ridge-shaped mainlobe in the WAFs (as shown in Figures 4.8 and 4.10), which in turn reduces the circular ambiguity of the point reflector in the SAS images.

For the m-sequence, a short sequence length does not seem to significantly affect the mainlobe of the WAF. However, it expands the low sidelobe region along the velocity axis and reduces the sidelobe level within that region, as shown in Figure 4.12. The width of the low sidelobe region along the velocity axis increases approximately in proportion to $1/N$. When the velocity range of the low sidelobe region in the periodic WAF is larger than $[-1, 1]$ m/s, the low imaging sidelobe region will cover the entire imaging domain. As a result, the SAS imaging sidelobe is reduced with decreasing sequence length, as presented in Figure 4.11.

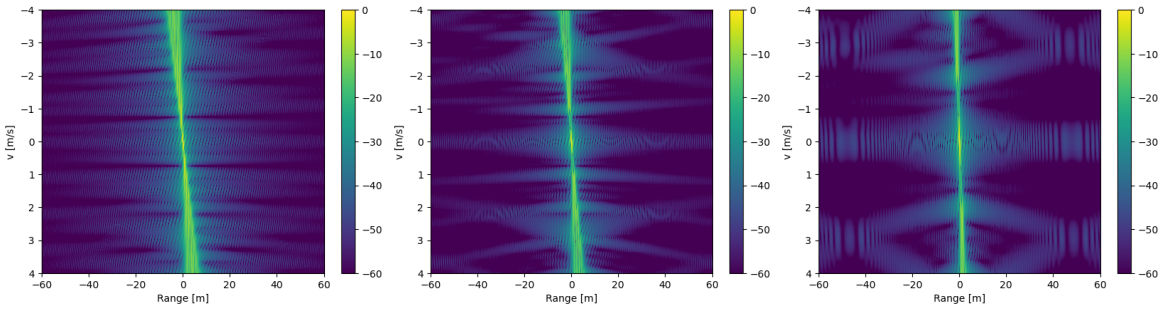
It is also important to note that as we decrease the PRI from 0.5115 s to 0.1275 s, the maximum detection range also decreases from 383.62 m to 95.625 m for the given simulation setting. For a practical SAS imaging scenario, the PRI cannot be set

to a very low value. Otherwise, extra imaging artifacts due to the range ambiguity will appear in SAS imaging results. According to our experimental data, the maximum detection range should be 3 times higher than the distance between the target of interest and the transmitter-receiver pair, to avoid the effects of reverberations in an underwater environment.



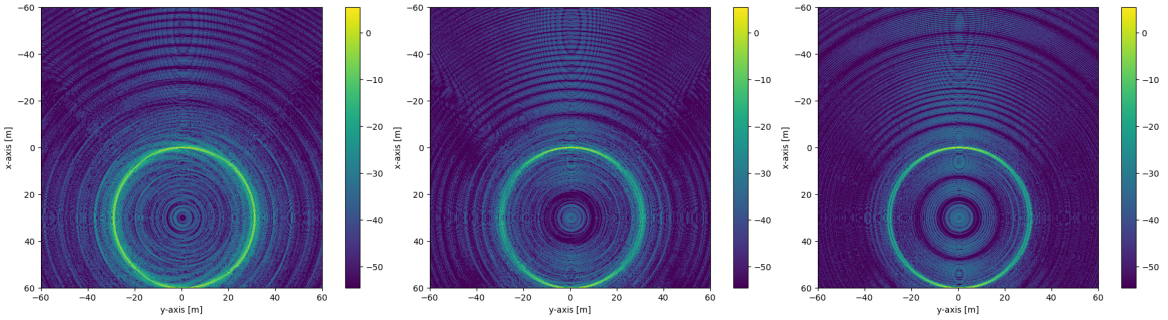
(a) LFM (PRI = 0.5115 s) (b) LFM (PRI = 0.2555 s) (c) LFM (PRI = 0.1275 s)

Figure 4.7: SAS imaging reconstructions of a point reflector using LFM waveforms with a different PRI, with an integration time of 1 s. As the PRI of the LFM decreases, the circular ambiguity of the point reflector decreases in intensity.



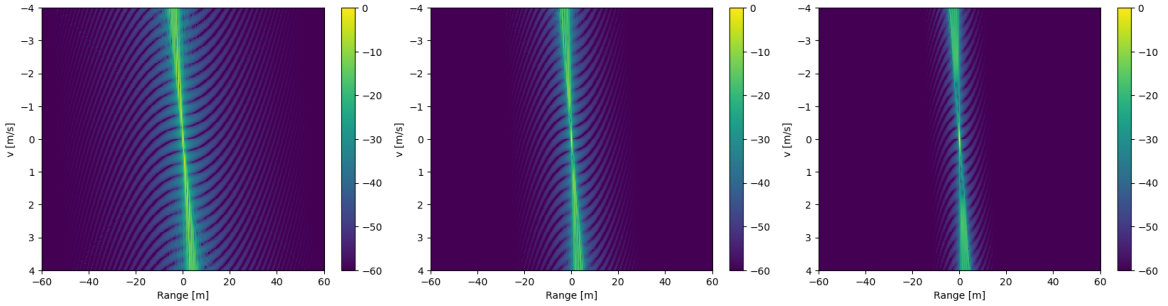
(a) LFM (PRI = 0.5115 s) (b) LFM (PRI = 0.2555 s) (c) LFM (PRI = 0.1275 s)

Figure 4.8: WAF for approximately 1 s of LFM waveforms with a different PRI. As the PRI of the LFM decreases, the waveform periods integrated within a duration of 1 s increases.



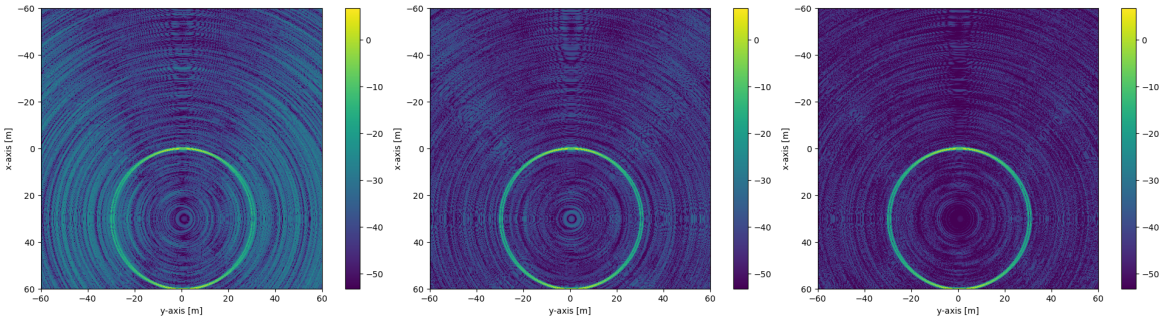
(a) Golomb sequence ($N = 1023$) (b) Golomb sequence ($N = 511$) (c) Golomb sequence ($N = 255$)

Figure 4.9: SAS imaging reconstructions of a point reflector using Golomb sequences with a different sequence length, with an integration time of 1 s. As the length of the Golomb sequence decreases, the circular ambiguity of the point reflector decreases in intensity.



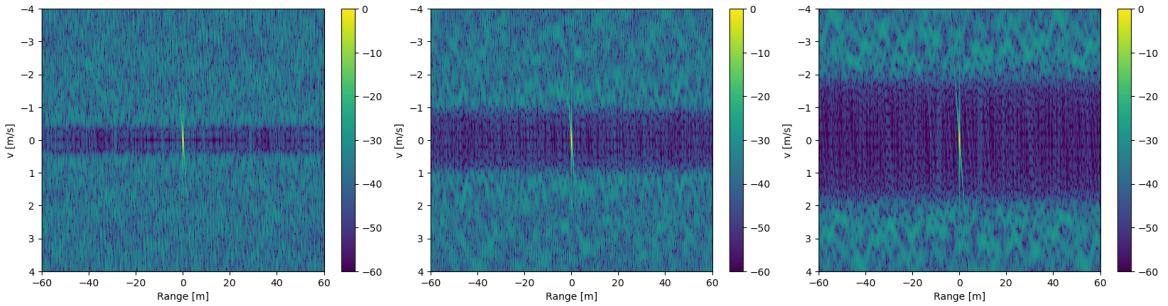
(a) Golomb sequence ($N = 1023$) (b) Golomb sequence ($N = 511$) (c) Golomb sequence ($N = 255$)

Figure 4.10: WAF for approximately 1 s of Golomb sequences with a different length. As the sequence length of the Golomb sequence decreases, the number of waveform periods integrated within a duration of 1 s increases, reducing sidelobes and Doppler ambiguities.



(a) m-sequence ($N = 1023$) (b) m-sequence ($N = 511$) (c) m-sequence ($N = 255$)

Figure 4.11: SAS imaging reconstructions of a point reflector using m-sequence with a different sequence length, with an integration time of 1 s. As the length of the m-sequence decreases, the imaging sidelobe of the point reflector is reduced.



(a) m-sequence ($N = 1023$) (b) m-sequence ($N = 511$) (c) m-sequence ($N = 255$)

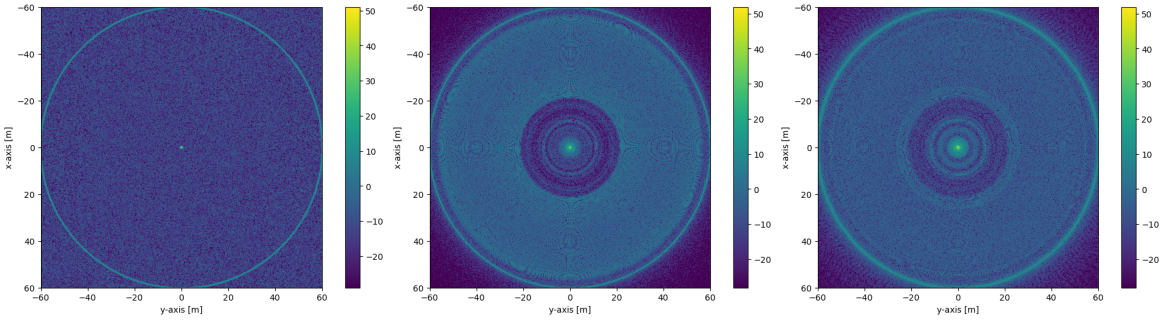
Figure 4.12: WAF for approximately 1 s of m-sequences with a different length. As the length of the m-sequence decreases, the waveform periods integrated within a duration of 1 s increases.

4.1.2.2 SAS imaging of a point reflector over the full circle trajectory

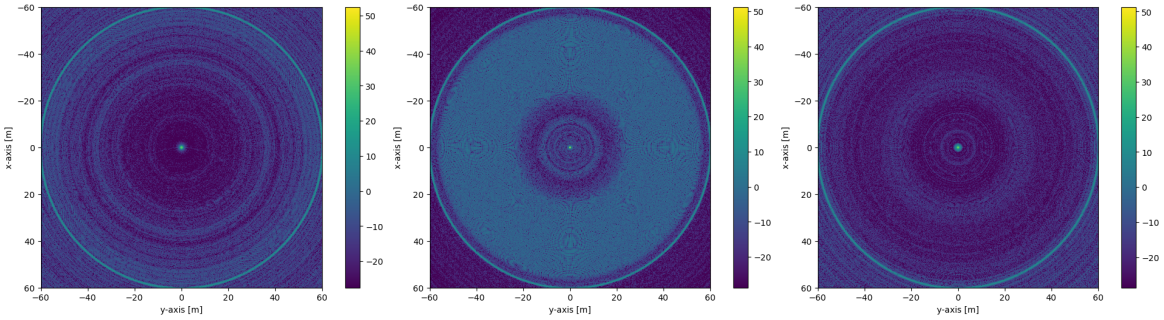
The SAS imaging results of a point reflector over the full circle trajectory are shown in Figures 4.13, 4.14, 4.15, and 4.16. The imaging results show that all sequences managed to image the point reflector at (0 m, 0 m, -30 m), after the SAS integration over the full circle trajectory. However, a circular imaging artifact with a radius of 60 m is also observed in all SAS imaging results. Since the SAS imaging over the full circle trajectory can be seen as an integration of multiple SAS images with a short integration trajectory, the circular ambiguity of the point reflector observed in the previous section is also integrated during the SAS imaging. As shown in Figure 4.17, the integration of multiple ambiguity circles not only identifies the point reflector but also results in a circular imaging artifact with a radius of 60 m.

As shown in Figures 4.13a and 4.13c, the SAS imaging results of the random BPSK waveform and the HFM waveform (PRI = 0.5115 s) are similar to the imaging PSFs in Figure 1.3 [1]. The imaging sidelobe of the random BPSK waveform is still constantly high after the full circle integration. Because chirp-like waveforms have much worse circular ambiguity effects, extra imaging artifacts are observed near the center of the imaging domain, after integrating over the full circle trajectory. Among the waveforms with a PRI of 0.5115 s, the wideband gradient descent sequence has the best SAS imaging performance, whose sidelobe level is around 8 dB lower than that of the random BPSK sequence. However, some extra imaging artifacts can also be observed near the center in Figure 4.13f, which might be caused by the integration of multiple ambiguity circles.

As shown in Figures 4.14, 4.15, and 4.16, waveforms with a lower PRI generally provide better SAS imaging performance, at a cost of a reduced maximum unambiguous detection range. For the LFM waveform and the Golomb sequence, a lower PRI reduces the circular ambiguity of the point reflector in the SAS imaging over a short integration trajectory due to the reduced Doppler ambiguities, which consequently improves the SAS imaging performance after the full circle integration. For the m-sequence, a lower PRI reduces the imaging sidelobe of the SAS imaging over a short trajectory, which

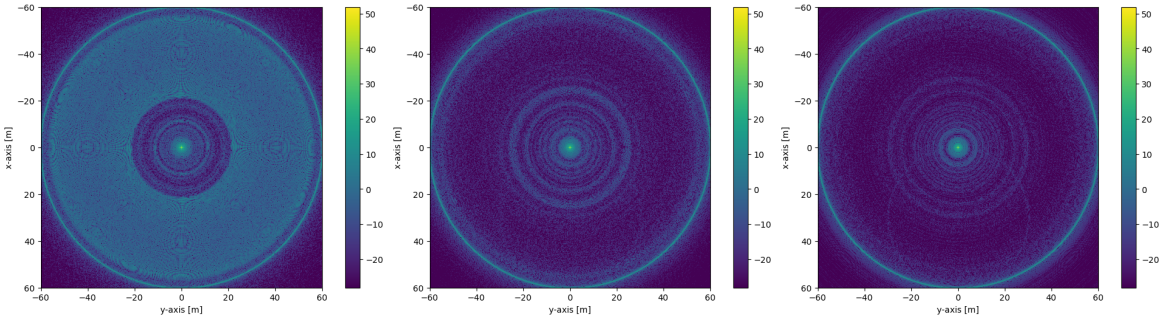


(a) random BPSK (non-repeating) (b) LFM (PRI = 0.5115 s) (c) HFM (PRI = 0.5115 s)



(d) m-sequence ($N = 1023$) (e) Golomb sequence ($N = 1023$) (f) Wideband gradient descent sequence ($N = 1023$)

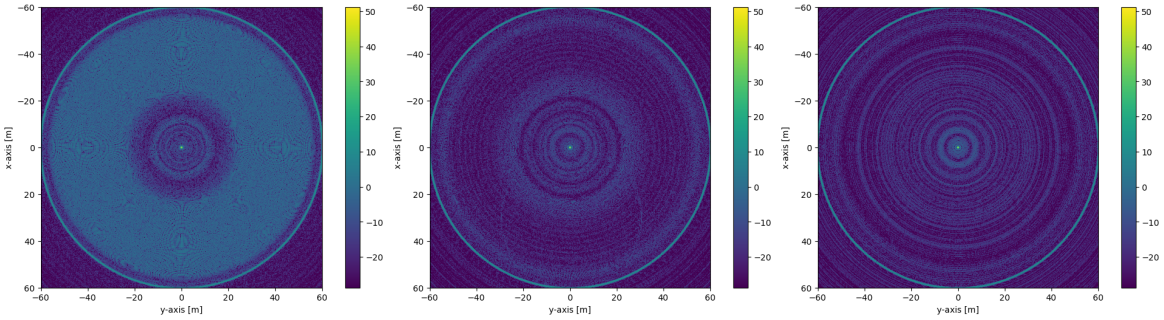
Figure 4.13: SAS imaging reconstructions of a point reflector using different waveforms, with an integration time of 180 s. Apart from the non-repeating random BPSK sequence, the other waveforms are periodic in time with a PRI of $N/B = 0.5115$ s. The circular imaging artifact with a radius of approximately 60 m is observed in all SAS imaging results.



(a) LFM (PRI = 0.5115 s) (b) LFM (PRI = 0.2555 s) (c) LFM (PRI = 0.1275 s)

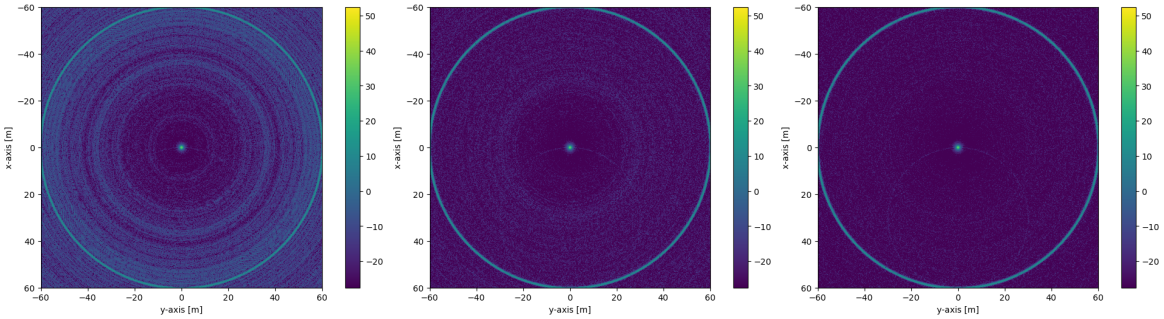
Figure 4.14: SAS imaging reconstructions of a point reflector using LFM waveforms with different PRI, with an integration time of 180 s.

also enhances the imaging performance after integration over the full circle trajectory. Among all sequences tested in the second simulation, the m-sequence with $N = 255$ has the best SAS imaging performance.



(a) Golomb sequence ($N = 1023$) (b) Golomb sequence ($N = 511$) (c) Golomb sequence ($N = 255$)

Figure 4.15: SAS imaging reconstructions of a point reflector using Golomb sequences with different sequence length, with an integration time of 180 s



(a) m-sequence ($N = 1023$) (b) m-sequence ($N = 511$) (c) m-sequence ($N = 255$)

Figure 4.16: SAS imaging reconstructions of a point reflector using m-sequences with different sequence length, with an integration time of 180 s

4.1.3 Discussion

The simulation results of SAS imaging of a point reflector suggest a strong connection between the ambiguity function properties of a waveform and its corresponding imaging performance. For a short synthetic aperture equivalent to 1 s of the integration time, the SAS imaging result of a waveform can be approximately viewed as a projection of the waveform's WAF evaluated over that 1 s duration. This projection relationship is detailedly illustrated in Figure 4.18. For a short integration time (e.g, a PRI of the transmit waveform), the waveform design problem of a SAS system can be approximately seen as a WAF shaping problem. Here, "a short integration time" refers to a duration during which the position and the velocity of the SAS system can be approximately considered constant, such that the waveforms reflected by voxels within that duration can be seen as the transmit waveform affected by a single delay and a single Doppler shift. Even if ambiguity function shaping is not strictly equivalent to the waveform design problem for a SAS system, the waveform obtained through this approach still delivers good imaging performance, compared to conventional waveforms.

Additionally, we found that the imaging performance also depends on the PRI, or

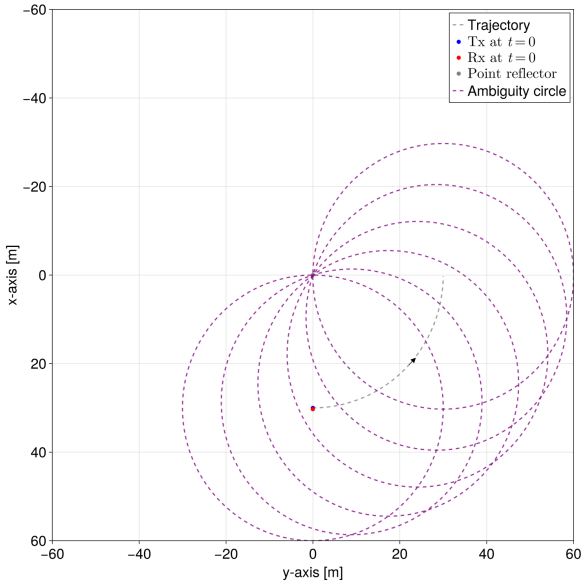


Figure 4.17: Schematic explaining the reason for the circular artifacts with a radius of 60 m shown in Figures 4.13, 4.14, 4.15, and 4.16. The SAS system is moving along a circle trajectory to image the point reflector. While the SAS system moving along the trajectory, multiple ambiguity circles are generated. The integration of multiple ambiguity circles along the trajectory not only identifies the point reflector at the center of the imaging domain but also introduces an imaging artifact with a radius of 60 m.

equivalently, the sequence length of the waveform. A lower PRI allows more waveform periods to be transmitted and integrated within a fixed time window, which enhances the coherent gain and improves the WAF properties of waveforms. For chirp-like waveforms, a lower PRI decreases the ridge-shaped ambiguity in their WAF, thereby reducing the intensity of the circular ambiguity. For the m-sequence, a shorter sequence length expands the low sidelobe region in its WAF and reduces the sidelobe level in that region, thereby reducing the SAS imaging sidelobe. However, reducing the PRI or the sequence length comes at the cost of a shorter maximum unambiguous detection range, which can cause imaging artifacts due to range ambiguity in practical sonar imaging scenarios. Therefore, the PRI or the sequence length of waveforms should not be set too short.

When comparing equal length sequences (using a PRI of 0.5115 s), the wideband gradient descent sequence (N=1023) generated in Chapter 3 demonstrates the best SAS imaging. Its imaging sidelobe is around 8 dB lower than that of the non-repeating random BPSK sequence. However, when we decreased the sequence length to $N = 255$, the SAS imaging performance of the m-sequence ($N = 255$) outperforms that of the wideband gradient descent sequence ($N = 1023$), because the sidelobe level of the m-sequence in its WAF is slightly lower than that of the wideband gradient descent sequence ($N = 1023$) as shown in Figures 4.12c and 4.3f. Moreover, the mainlobe of the m-sequence in its WAF is narrower than that of the wideband gradient descent sequence, which provides a lower intensity on the circular ambiguity of the point reflector as shown

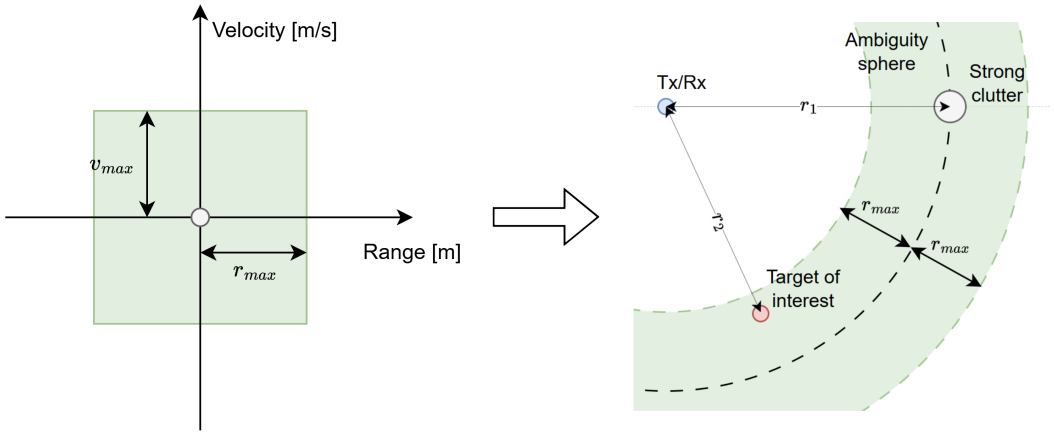


Figure 4.18: Schematic illustrating the projection of the WAF of a waveform to the SAS imaging domain when considering a short synthetic aperture. The low sidelobe regions in the ambiguity domain and the SAS imaging domain are marked with green color. The ambiguity sphere of the strong clutter is represented by the black dashed line. The maximum range and the maximum velocity of the low sidelobe region in the ambiguity domain are represented by v_{max} and r_{max}, respectively. It is assumed that the distances of the strong clutter (r_1) and the target of interest (r_2), and the radial velocities of the strong clutter (v_1) and the target of interest (v_2), are given. To ensure that the target of interest is not covered by the sidelobe of the clutter, we need to satisfy $|r_1 - r_2| \leq r_{max}$ and $|v_1 - v_2| \leq v_{max}$. In the practical sonar imaging scenario, we typically have a large field of voxels, which are all potential targets. For all of these voxels, the above two conditions need to be satisfied.

in Figures 4.11c and 4.2f. Therefore, among the all sequences tested, the m-sequence with $N = 255$ has the best imaging performance for the given simulation scenario.

The m-sequence only provides a low sidelobe region near the zero-Doppler axis in its WAF. When the speed of the vessel is high enough, the m-sequence may fail to provide a good SAS imaging performance because the radial velocity of voxels in the imaging domain may be higher than the maximum velocity of the low sidelobe region of the m-sequence's WAF. As we discussed in Section 2.3.2, the length of the m-sequence has to be $N = 2^l - 1$, where $l \in \mathbb{Z}$ and $l \geq 2$, which further limits the design choice of the m-sequence. On the other hand, the WAF shaping algorithms (e.g., the wideband gradient descent methods) allow designing sequences with an arbitrary sequence length and an arbitrary shape for the RoI. An example demonstrating that the m-sequence may not provide an ideal SAS imaging performance is described in Section 4.2.2. Theoretically, with a proper choice of the sequence length N and the RoI $\mathcal{Q}$, sequences generated by the wideband gradient descent method should give better SAS imaging performance.

4.1.3.1 Sequence design for SAS imaging

Based on our observations, we recommend that a good SAS imaging sequence for a short synthetic integration time should possess the following properties

- The mainlobe of the WAF should have a thumbtack shape.

- The RoI $\mathcal{Q}$ in the WAF should cover all possible delay-Doppler combinations of voxels in the imaging domain.

- The sidelobe level in the RoI should be as low as possible.

The thumbtack-shaped mainlobe in the WAF provides an unambiguous joint estimation of the range and the velocity of the target, which can help to alleviate the circular ambiguity of the target when using a short synthetic aperture for SAS imaging. Since the SAS imaging result of a waveform for a short integration time can be approximately seen as a projection of the waveform's WAF, the choice of RoI $\mathcal{Q}$ and the sidelobe level within the RoI $\mathcal{Q}$ will directly affect the SAS imaging performance of a waveform. As shown in Figure 4.18, to ensure that the target of interest is still detectable in the presence of a strong clutter, the target of interest need to stay within the low sidelobe region of the strong clutter. Based on the relative distance and the relative radial velocity between the strong clutter and the target of interest, we can obtain a suitable RoI for the WAF shaping. It is also important to note that, even though we observed that a low PRI (or equivalently, a short sequence length) can lead to better SAS imaging results, we did not suggest an ideal sequence length in the above properties. This is because the main reason for this phenomenon is that changes in the PRI (or the sequence length) only affect the WAF properties of a waveform within a fixed time window, which has already been mentioned in the other three properties.

Additionally, we have extra constraints on the sequence length N and the RoI $\mathcal{Q}$. To ensure that the sequence has a maximum unambiguous detection range of R_{max}, the sequence length should be at least higher than $\frac{2BR_{max}}{c}$:

$$N \geq \frac{2BR_{max}}{c} \tag{4.2}$$

Because of the (asymptotic) volume invariance property of the WAF, the RoI $\mathcal{Q}$ cannot be too large, otherwise the optimality of sidelobe minimization in the RoI is not guaranteed. Ideally, the RoI in the (periodic) WAF can be defined as

$$\mathcal{Q}_{opt}{}^{1} = \{(\tau, v) | -\rho_T N/B \leq \tau \leq \rho_T N/B \text{ and } -\rho_v v_n \leq v \leq \rho_v v_n\} \tag{4.3}$$

where the expression of v_n is given in Equation 3.75, and we introduced the RoI parameters $\rho_T \in (0, 1/2)$ and $\rho_v \in (0, +\infty)$ which depend on the imaging scenario. We recommend that the product of ρ_T and ρ_v has to be small enough:

$$\rho_{opt} = \rho_T \rho_v \ll 1. \tag{4.4}$$

For example, the ratio ρ_{opt} for the wideband gradient descent sequence used in this section is approximately 0.375, which ensures that the RoI is relatively small compared to the total region in the WAF.

[1] This RoI also includes the mainlobe of the WAF. It should be excluded in any waveform optimization techniques.

4.2 Field Trial

In this section, the results of SAS imaging in field trials are provided. The goal of these trials is to evaluate the SAS imaging performance of sequences obtained from the WAF shaping in a practical scenario.

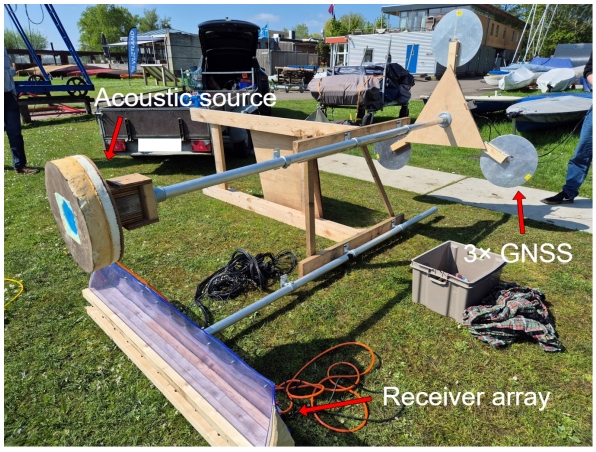
4.2.1 SAS imaging setup

In the field trials, we aim to image two corner reflectors placed on the lake bed with the CLAMS SAS imaging system. The picture of the two reflectors is shown in Figure 4.19. The distance between the water surface and the lake bed is approximately 20-30 m. Images of the CLAMS system are shown in Figure 4.20, where the CLAMS system was mounted on a small boat for SAS imaging. Instead of using only one receiver as we discussed earlier, the SAS imaging system used a receiver array consisting of 10 hydrophones with a spacing of 15 cm. The delay-and-sum beamformer and the null steering beamformer are used to process the data obtained from the array. To ensure a cm-level positioning accuracy, three GNSS antennas are attached to the frame of the system. The frequency band used for the trials is 8000-12000 Hz. To avoid causing extra turbulence noises, the speed of the boat was less than 1 m/s during the trial. The SAS imaging performance of the sequences obtained from the WAF shaping is compared against that of the random BPSK sequence.

Figure 4.19: Image of two corner reflectors to be imaged. The radii of the corner reflectors are 20 cm and 40 cm, respectively, with estimated target strengths of -24 dB and -12 dB.

4.2.2 Sequence design for the field trial

As shown in Figure 4.21, the two main sources of clutter for the field trials are the direct wave and the specular reflection. The direct wave refers to the line-of-sight transmission from the transmitter to the receiver, and typically has a signal strength a few orders of magnitude higher than that of the signal reflected by the target of interest. The specular reflection refers to the mirror-like reflection of waves from a smooth surface. The intensity of the specular reflection is also a few orders of magnitude higher than that of the scatter reflection. In the field trial, the lake bed can be approximately considered a flat and smooth surface. When the transmit waveform hits the lake bed,



(a) Entire frame of the source and sensors

(b) Receiver array mounted over the side of a boat

Figure 4.20: Images of the CLAMS system used for the SAS imaging trial. The receiver array consists of 10 hydrophones with a spacing of 15 cm. Three GNSS antennas are used to provide a cm-level positioning accuracy. The frequency band used for the trial is 8000-12000 Hz.

it will be reflected directionally and preserve most of the energy. From a modeling perspective, the specular reflection can be approximated as a strong, persistent point clutter located on the lake bed directly beneath the transmitter and the receiver array.

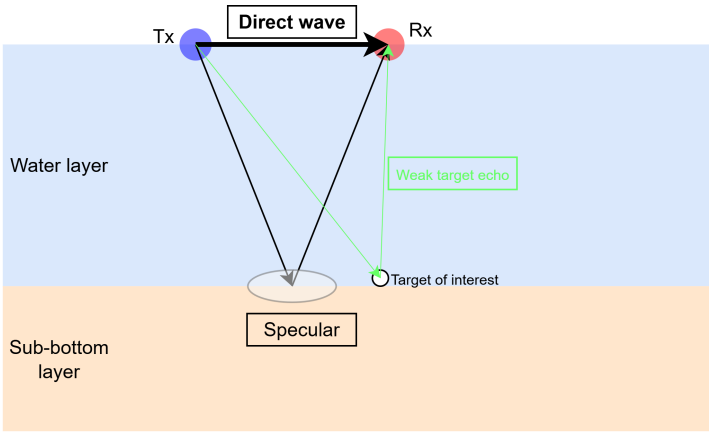


Figure 4.21: Schematic illustrating the two main sources of clutter in the field trial: the direct wave and the specular reflection.

To ensure that the corner reflectors are visible in the presence of the direct wave and the specular reflection, we designed a $N = 1500$ periodic sequence using the wideband gradient descent method based on the suggestion in Section 4.1.3.1. The periodic WAF of the sequence is shown in Figure 4.22. Since the speed of the boat was less than 1 m/s, the maximum velocity of the RoI was set to $v_{max} = 1$ m/s. To ensure that the

corner reflectors stay inside the low sidelobe region of the strong clutters, the maximum range of the RoI was set to $r_{max} = 35$ m.

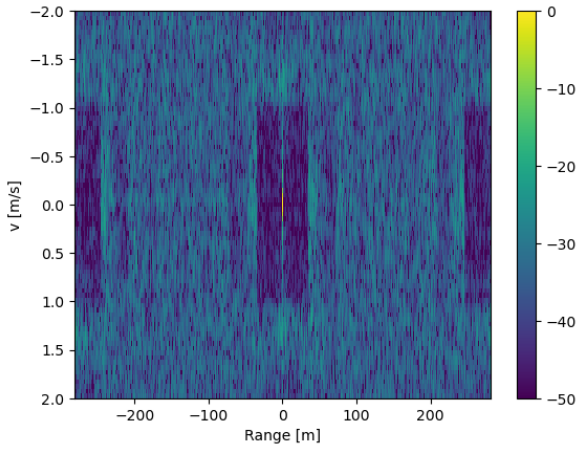


Figure 4.22: Periodic WAF for the wideband gradient descent sequence with $N = 1500$, $B = 4000$ Hz, and $f_c = 10000$ Hz.

Apart from the sequence designed from the WAF shaping, the m-sequence was also considered. With $B = 4000$ Hz and $f_c = 10000$ Hz, the length of the m-sequence has to be at least $N = 511$ to guarantee a maximum unambiguous detection range of $R_{max} = \frac{cN}{2B} = 96$ m. Since the m-sequence has better ambiguity function properties with a lower sequence length, the ideal length for the m-sequence is $N = 511$. The WAF for approximately 1s of the m-sequence ($N = 511$) is shown in Figure 4.23a. Compared to the wideband gradient descent sequence ($N = 1500$), the WAF of the m-sequence only has a low sidelobe region within $|v| < 0.5$ m/s, which is much less than the speed of the boat. Therefore, the m-sequence ($N = 511$) was not used in the field trial.

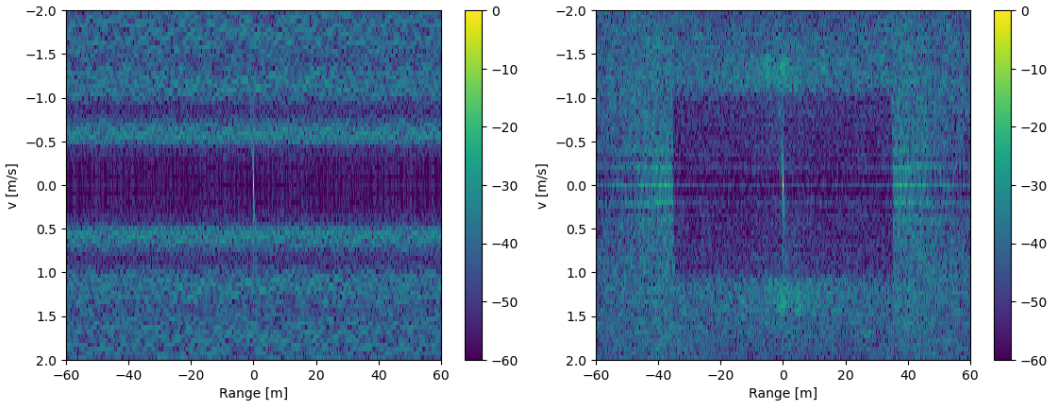
4.2.3 Field trial results

4.2.3.1 Preprocessing results

The SAS imaging trajectory of the field trials with the random BPSK sequence and the wideband gradient descent sequence ($N = 1500$) is shown in Figure 4.24. The integration time of the two trials is 450 s.

To evaluate the SAS imaging performance of the two waveforms before the SAS imaging, we applied the STMF techniques introduced in Section 2.1 to a single hydrophone recording of both waveforms. The STMF results are shown in Figure 4.25, in which the direct wave at $z = 1$ m and the specular reflection at approximately $z = 22$ m are clearly identified. In the STMF result of the wideband gradient descent sequence ($N = 1500$), the sidelobes of the direct wave within a 35 m range are significantly lower than those observed in the STMF of the random BPSK sequence, because of the ambiguity function properties of the wideband gradient descent sequence.

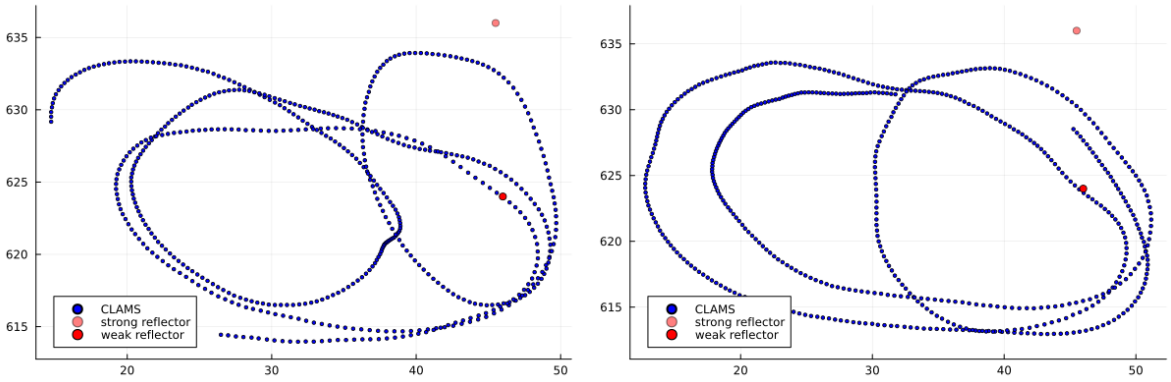
Since the relative position between the transmitter and the receiver array is always a constant during the SAS imaging, the direct wave can be easily removed using a decon-



(a) m-sequence ($N = 511$)

(b) Wideband gradient descent sequence ($N = 1500$)

Figure 4.23: WAF for approximately 1 s of the m-sequence ($N = 511$) and the wideband gradient descent sequence ($N = 1500$) with $B = 4000$ Hz and $f_c = 10000$ Hz.

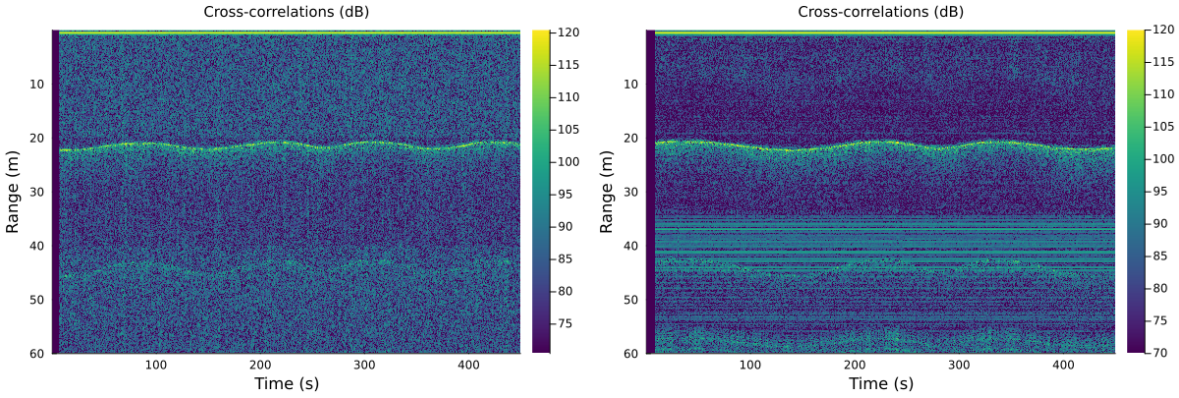


(a) Random BPSK

(b) Wideband gradient descent sequence ($N = 1500$)

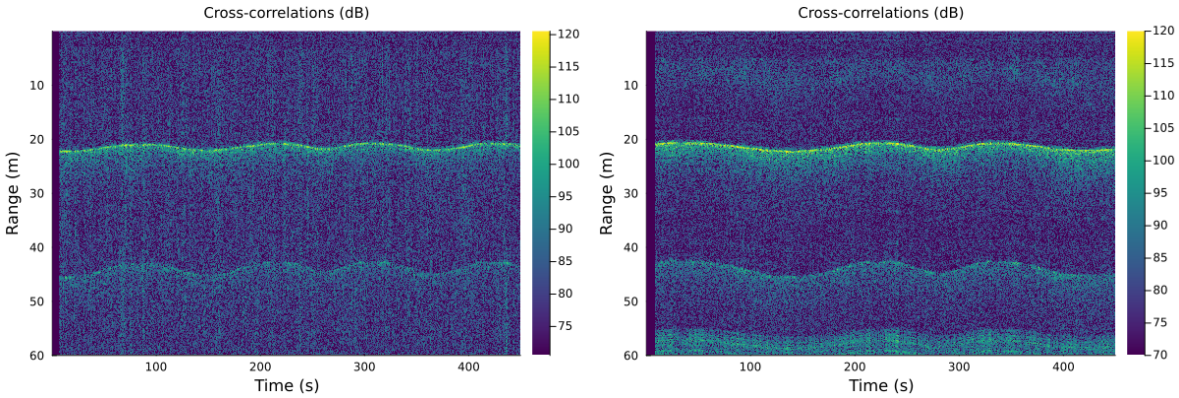
Figure 4.24: Imaging trajectories of the trial with (a) the random BPSK sequence and (b) the wideband gradient descent sequence ($N = 1500$). The positions of the strong corner reflector and the weak corner reflector are approximately (46 m, 624 m) and (45 m, 636 m), respectively. The integration time of the two trials is 450 s.

volution algorithm (e.g., the CLEAN algorithm mentioned in [1]) given the knowledge of the transmit waveform. After we applied the CLEAN algorithm to remove the direct wave, the STMF results are shown in Figure 4.26, in which only the specular reflection at approximately $z = 22$ m is identified. In the STMF result of the wideband gradient descent sequence ($N = 1500$), the sidelobes of the specular reflection, within a range of 35 m, are still lower than those observed in the STMF of the random BPSK sequence, because of the ambiguity function properties of the wideband gradient descent sequence.



(a) Random BPSK (b) Wideband gradient descent sequence ($N = 1500$)

Figure 4.25: STMF results of the recording from a single hydrophone before applying the CLEAN algorithm.



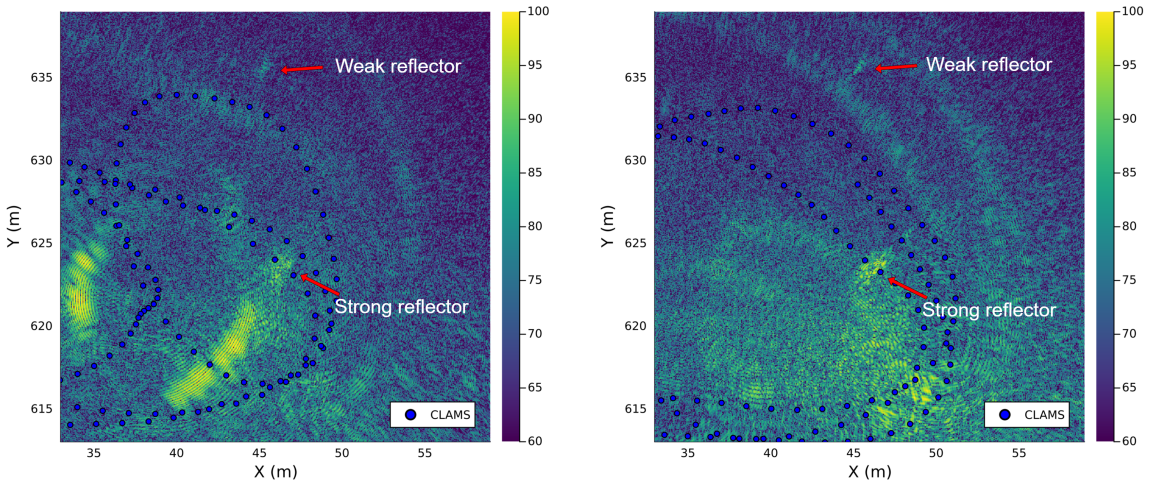
(a) Random BPSK (b) Wideband gradient descent sequence ($N = 1500$)

Figure 4.26: STMF results of the recording from a single hydrophone after applying the CLEAN algorithm [1] to remove the direct wave.

4.2.3.2 SAS imaging results with the delay-and-sum beamformer

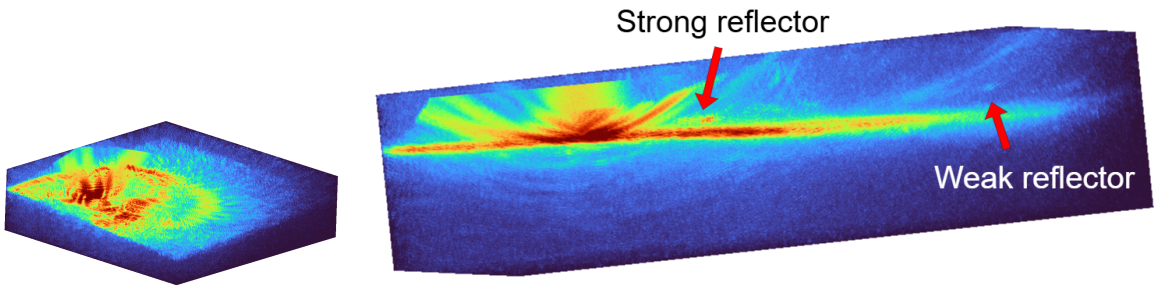
The SAS imaging results obtained using the two waveforms with the delay-and-sum beamformer are illustrated in Figures 4.27, 4.28, and 4.29. The strong corner reflector at (46 m, 624 m) is successfully imaged by both the random BPSK sequence and the wideband gradient descent sequence ($N = 1500$). In contrast, the weak reflector is poorly imaged, which can be attributed to its 12 dB lower target and the insufficient coverage of the SAS trajectories in its vicinity. Notably, the SAS imaging performance of the wideband gradient descent sequence is slightly better than that of the random BPSK sequence, as it provides marginally improved imaging of the corner reflectors. However, since the imaging trajectories are different for the two trials, this performance difference cannot be conclusively attributed to the waveforms themselves. Apart from the corner reflectors, the bathymetry of the lake bed between $z = -26$ m and $z = -25$

m is also imaged in the two trials. The voxels with high intensity on the lake bed refers to the specular reflection, which is generally aligned with the SAS imaging trajectory of our system.



(a) Random BPSK (b) Wideband gradient descent sequence ($N = 1500$)

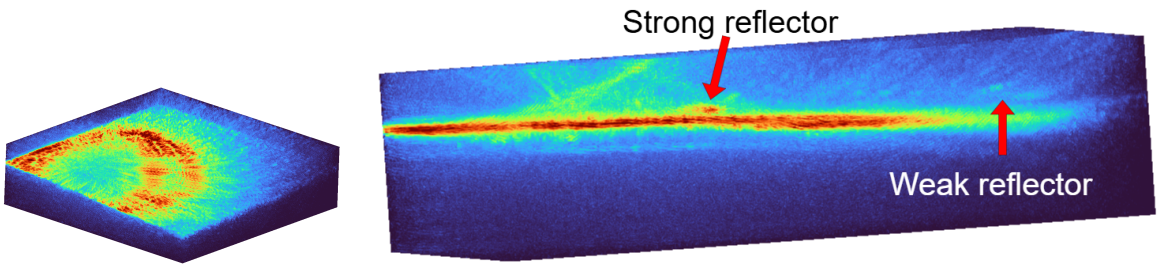
Figure 4.27: SAS imaging results obtained using two waveforms at $z = -24.85$ m, with the delay-and-sum beamformer. The imaging trajectories are marked with blue dots.



(a) (b)

Figure 4.28: SAS imaging results obtained using the random BPSK sequence, with the delay-and-sum beamformer. The imaging depth ranges from -23 m to -30 m. The dynamic range is 40 dB.

The mean energy distribution along the z-axis for the random BPSK sequence and the wideband gradient descent sequence ($N = 1500$) is shown in Figure 4.30. Since the targets of interest are sparsely located in the imaging domain, the mean energy can be seen as the SAS imaging sidelobe at a given water depth. Since the lake bed is located at approximately $z = -25.5$ m, a peak in mean energy is observed at $z = -25.5$ m for both sequences. Because the SAS image obtained using the wideband gradient descent sequence has stronger specular reflections on the lake bed, the peak mean energy for the wideband gradient descent sequence is higher than that for the random BPSK sequence.



(a) (b)

Figure 4.29: SAS imaging results obtained using the wideband gradient descent sequence ($N = 1500$), with the delay-and-sum beamformer.

When $z \in [-26.1, -23)$, the mean energy for the wideband gradient descent sequence is higher than that for the random BPSK sequence. However, when $z \in [-30, -26.1)$, the mean energy for the wideband gradient descent sequence is lower than that for the random BPSK sequence.

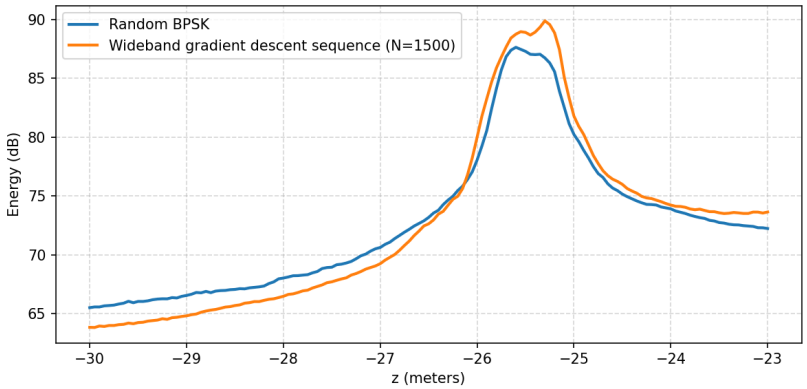


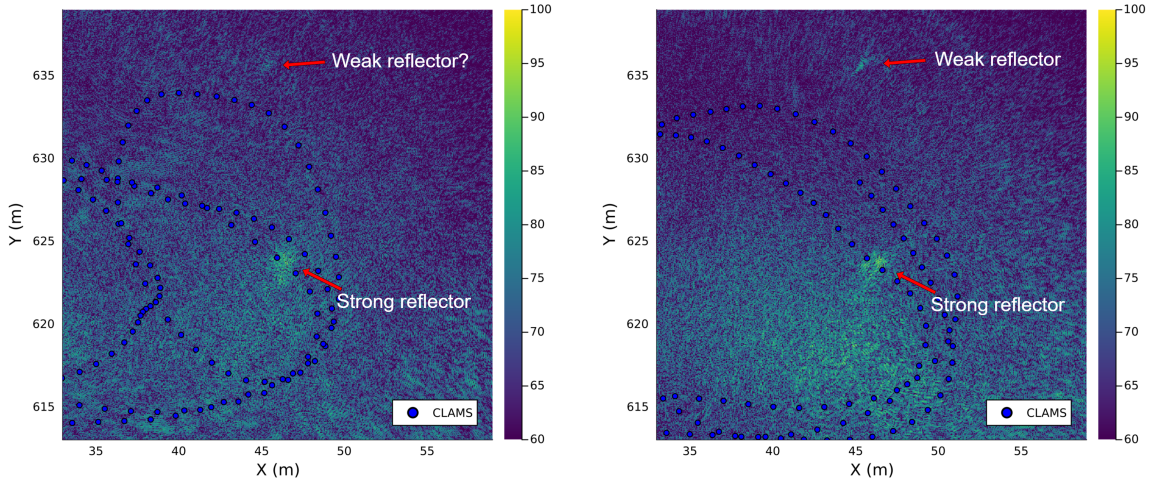
Figure 4.30: Mean energy distribution along the z-axis in the SAS imaging domain for the random BPSK sequence and the wideband gradient descent sequence ($N = 1500$), with the delay-and-sum beamformer.

4.2.3.3 SAS imaging results with the null steering beamformer

As shown in the previous section, the SAS imaging performance of the corner reflectors on the lake bed is limited by the specular reflection. Since the specular reflection is directly beneath the transmitter and the receiver array, its intensity can be reduced by using the null steering beamformer. This beamforming technique can cancel the reflected signal coming from directly beneath the transmitter and the receiver array, thereby allowing only the scatter reflection to contribute to the SAS imaging.

The SAS imaging results obtained using the two waveforms with the null steering beamformer are illustrated in Figures 4.31, 4.32, and 4.33. Applying the null steering beamformer significantly reduces the specular reflection on the lake bed. As a result,

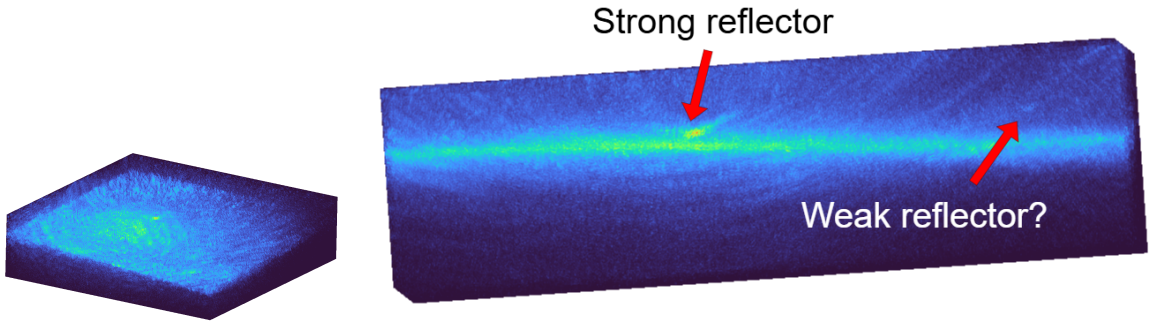
the imaging performance of corner reflectors is improved in both trials using the random BPSK sequence and the wideband gradient descent sequence. It is also important to note that the intensity of the corner reflectors is also reduced after applying the null steering. In the field trial using the random BPSK sequence, the weak corner reflector is barely visible, as shown in Figures 4.31a and 4.32.



(a) Random BPSK

(b) Wideband gradient descent sequence ($N = 1500$)

Figure 4.31: SAS imaging results obtained using two waveforms at $z = -24.85$ m, with the null steering beamformer. The imaging trajectories are marked with blue dots.

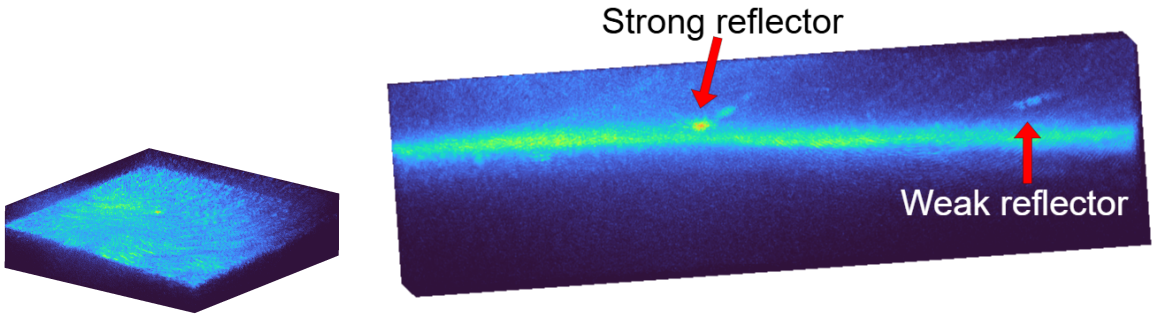


(a)

(b)

Figure 4.32: SAS imaging results obtained using the random BPSK sequence, with the null steering beamformer. The imaging depth ranges from -23 m to -30 m. The dynamic range is 40 dB.

With the null steering beamformer, the mean energy distribution along the z-axis in the SAS imaging domain for the random BPSK sequence and the wideband gradient descent sequence ($N = 1500$) is shown in Figure 4.34. When $z \in [-25.5, -23)$, the mean energy for the random BPSK sequence is nearly the same as that for the wideband gradient descent sequence. When $z \in [-30, -25.5)$, the mean energy for the wideband



(a) (b)

Figure 4.33: SAS imaging results obtained using the wideband gradient descent sequence ($N = 1500$), with the null steering beamformer.

gradient descent sequence is lower than that for the random BPSK sequence, because of its WAF properties.

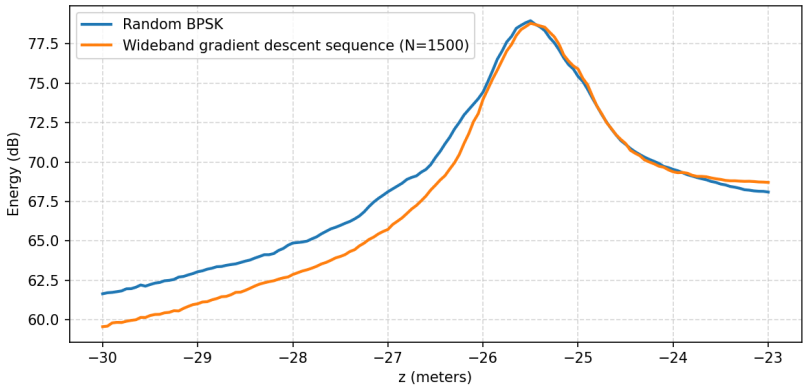


Figure 4.34: Mean energy distribution along the z-axis in the SAS imaging domain for the random BPSK sequence and the wideband gradient descent sequence ($N = 1500$), with the null steering beamformer.

4.2.4 Discussion

The field trial results show that although the wideband gradient descent sequence ($N = 1500$) produces better STMF results compared to the random BPSK sequence, its SAS imaging performance on corner reflectors does not reflect its optimal ambiguity function properties. The wideband gradient descent sequence has much lower sidelobes in the RoI in its WAF, but its SAS imaging performance on the corner reflectors is only slightly better than that of the random BPSK sequence.

One possible explanation for this phenomenon is that the SAS imaging performance of the corner reflectors is dominated by the ambiguities of the lake bed in the field trial. In the SAS imaging results with the delay-and-sum beamformer, while the SAS imaging system are moving on the water surface along a trajectory, an ambiguity of the specular

reflection is generated for every hydrophone in every short-time window, as shown in Figure 4.35. Since the intensity of the ambiguity of the specular reflection is a few orders of magnitude higher than the other sidelobes, the SAS imaging sidelobe level in the water layer is dominated by the ambiguity of the specular reflection rather than the sidelobe of the clutter. With the delay-and-sum beamformer, the ambiguity of the specular reflection obtained from every hydrophone is integrated in the imaging domain, which creates a high imaging sidelobe region in the water layer. Therefore, we observed that both the random BPSK sequence and the wideband gradient descent sequence ($N = 1500$) have a similar imaging performance on the corner reflectors placed on the lake bed, as shown in Figures 4.27, 4.28, and 4.29.

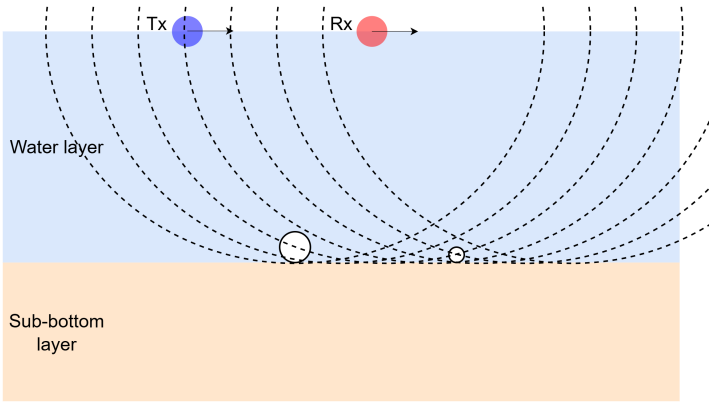


Figure 4.35: Schematic explaining the high imaging sidelobe in the water layer of SAS imaging results of the field trials with the delay-and-sum beamformer. The black dashed lines represent the ambiguities of the specular reflection. While the SAS imaging system are moving on the water surface along a trajectory, an ambiguity of the specular reflection is generated for every hydrophone in every short-time window. The white circles represent the corner reflectors on the lake bed.

In the SAS imaging results with the null steering beamformer, the specular reflection and its corresponding ambiguity are canceled. However, since nulling the specular reflection does not remove the backscatter of the lake bed caused by features and inhomogeneities on the lake bed itself, each voxel on the lake bed can still be considered as a point reflector but has a much lower reflection coefficient compared to the specular reflection. As shown in Figure 4.36, the voxels illuminated by the transmitter on the lake bed still create ambiguities in the water layer. As a result, the imaging sidelobe in the water layer is still dominated by the ambiguities of these voxels on the lake bed. Therefore, we observed that the mean energy for the random BPSK is nearly the same as that for the wideband gradient descent sequence when $z \in [25.5, -23)$, as shown in Figure 4.34.

According to our theoretical analysis, the imaging sidelobes within the first few meters below the lake bed are still dominated by the ambiguity of voxels on the lake bed. As the depth increases, the imaging sidelobe will gradually become dominated by the sidelobe of the clutter. Thus, we observed that the SAS imaging results obtained using the wideband gradient descent sequence ($N = 1500$) have a lower imaging sidelobe for the sub-bottom layer compared to those obtained using the random BPSK sequence,

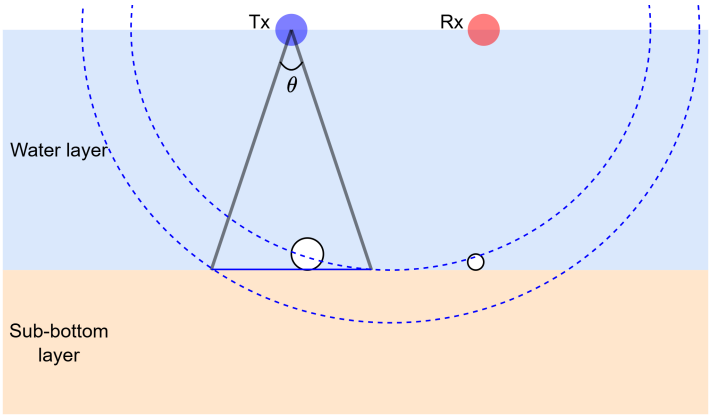
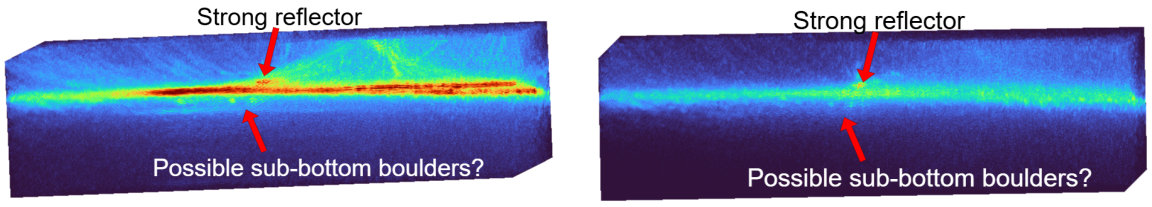


Figure 4.36: Schematic explaining the high imaging sidelobe in the water layer of SAS imaging results of the field trials with the null steering beamformer. The beamwidth of the transmitter is assumed to be θ. The blue dashed lines represent the ambiguities of the voxels illuminated by the transmitter. These ambiguities also appear in the sub-bottom layer, but only up to a range that depends on the transmitter's beamwidth and the position of the receiver array. It is important to note that the intensity of the ambiguity of these voxels is much lower than that of the specular reflection.

as shown in Figures 4.30 and 4.34. These low sidelobes in the sub-bottom layer may potentially help us find sub-bottom targets. As shown in Figure 4.37, several distinct dots were observed below the lake bed in the SAS imaging results obtained using the wideband gradient descent ($N = 1500$), which might correspond to sub-bottom boulders.

In general, because the imaging sidelobes in the water layer are dominated by the ambiguity of the specular reflection or the backscatter of the lake bed, the sequence designed through WAF shaping only provides marginally improvement in SAS imaging of corner reflectors on the lake bed, compared to the random BPSK sequence. However, through our analysis, we found that this sequence can provide lower imaging sidelobes within the sub-bottom layer, which may be beneficial for the sub-bottom target detection.



(a) Delay-and-sum beamformer

(b) Null steering beamformer

Figure 4.37: Possible sub-bottom boulders in the SAS imaging results obtained using the wideband gradient descent sequence ($N = 1500$), with the delay-and-sum beamformer and the null steering beamformer.

4.3 Conclusion

In this chapter, we evaluated the SAS imaging performance of waveforms designed through WAF shaping. Through both simulations and fields, we investigated the relationship between the WAF properties of a waveform and its corresponding SAS imaging performance.

Simulation results demonstrated a strong connection between the WAF of waveform and its SAS imaging performance. For a short synthetic aperture, the SAS imaging results can be approximately interpreted as a projection of the waveform's WAF. Waveforms with a thumbtack-shaped mainlobe and a low average sidelobe energy in a well-defined RoI generally provide better SAS imaging performance. Among the tested waveforms with a sequence length of $N = 1023$ (or equivalently, a PRI of 0.5115 s), the sequence generated by the wideband gradient descent algorithm in Section 3.4 showed the best SAS imaging performance, outperforming conventional waveforms, such as the LFM, and the Random BPSK sequence. We also observed that reducing the sequence length can improve imaging performance by increasing the number of periods integrated over a fixed time window. However, this comes at the cost of a reduced maximum unambiguous detection range, which may introduce range ambiguity artifacts in practical applications. Based on our observations of simulation results, ideal properties for a SAS imaging waveform are provided in Section 4.1.3.1.

Field trial results further validated the practical advantages of waveforms obtained through the WAF shaping. Since the imaging sidelobes in the water layer are dominated by the ambiguity of the specular reflection or the backscatter of the lake bed, the wideband gradient descent sequence ($N = 1500$) only offered slight improvement in SAS imaging of corner reflectors on the lake bed, compared to a random BPSK sequence. However, this sequence did exhibit lower imaging sidelobes in the sub-bottom layer, suggesting the sequences designed through the WAF shaping may offer advantages for sub-bottom target detection, where the sidelobe suppression is crucial for detecting weak targets beneath the strong surface reflector.

In general, through simulations and field trials, we answered the first sub-question of the main research question: *"What is the relationship between the ambiguity properties of a waveform and its corresponding SAS imaging performance? Is ambiguity function shaping equivalent to waveform design for a SAS system?"*. Although the WAF shaping is not strictly equivalent to waveform design for a SAS system, it still provides a useful and practical framework for improving the SAS imaging performance.

5 Conclusion

In this thesis, we addressed the optimal continuous waveform design problem for a SAS system from the perspective of the wideband ambiguity function shaping. To ensure that the designed waveform has an optimal peak-to-average power ratio, an unimodular constraint was considered.

In Chapter 2, we began by establishing a theoretical connection between the wideband ambiguity function shaping of a waveform and its impact on SAS imaging performance. Specifically, we showed that the waveform design problem for a SAS system, within a short integration time (during which the position and velocity of the system can be considered approximately constant), can be seen as a wideband ambiguity function shaping problem. Using this approximation, we can simplify the waveform design problem for the entire synthetic aperture as a wideband ambiguity function shaping problem for a short waveform and then repeat the designed waveform periodically to obtain the waveform for transmission.

In Chapter 3, we formulated the wideband ambiguity shaping problem as a non-convex optimization problem and proposed four relaxation methods to solve it, including the USSM-avg algorithm (based on the USSM [32]), the randomization-based method, the wideband gradient descent method, and the narrowband approximation method. The proposed methods use the average sidelobe energy in the RoI as the objective function for optimization, resulting in better overall sidelobe minimization performance compared to the USSM algorithm [32], which aims to minimize the maximum sidelobe. Among these methods, the wideband gradient descent method with a warm start initialization showed the best performance in terms of average sidelobe minimization in the RoI and computational efficiency.

In Chapter 4, the SAS imaging performance of waveforms designed by the wideband ambiguity function shaping was evaluated through both simulations and field trials. Through SAS imaging simulations of a point reflector, we showed that waveforms with thumb-tack shaped mainlobe in its WAF and a low sidelobe region in a well-designed RoI yield superior SAS imaging performance, particularly in imaging scenarios with a short synthetic aperture. Additionally, we observed that reducing the sequence length (or PRI) can enhance imaging quality by increasing the number of waveform periods integrated within a fixed duration, thereby reducing the velocity ambiguity of the waveform. However, this improvement comes at the cost of a reduced maximum unambiguous detection range, which must be carefully balanced in practical applications. Experiments based on field trials further supported our findings. Even though it was expected that the SAS imaging performance of corner reflectors on the lake bed is limited by ambiguities of the specular reflection or the backscatter of the lake bed, we still observed slightly better imaging performance when using the sequence obtained through the wideband ambiguity function shaping. Since the sidelobe suppression is crucial for detecting weak sub-bottom targets beneath the strong surface

reflector, where we expect that ambiguity artifacts from the lake bed no longer play a role, the waveform obtained through the wideband ambiguity function may have potential advantages for sub-bottom imaging.

In conclusion, this thesis demonstrates that the wideband ambiguity function shaping offers a powerful and practical framework for continuous waveform design for SAS systems. Although ambiguity function shaping is not strictly equivalent to the waveform design problem for SAS systems, it still provides a tractable and effective method for improving the SAS imaging performance.

5.1 Future work

5.1.1 Future work on wideband ambiguity shaping algorithms

Despite the promising performance of the wideband gradient descent method for the wideband ambiguity function shaping, several questions still remain for the wideband ambiguity function shaping algorithms. Firstly, in terms of computational efficiency, our approaches are generally more expensive than algorithms for narrowband ambiguity function shaping, especially when the sequence length is long (e.g., $N > 1000$). Future work could explore more efficient methods for relaxing the non-convex optimization problem for the wideband ambiguity function shaping to reduce the computational complexity without compromising performance.

Secondly, the objective function of our algorithms only minimizes the average sidelobe energy in the RoI, which does not necessarily reduce the maximum sidelobe that the USSM algorithm aims to minimize and may result in several high peak sidelobes in the RoI. An alternative objective function could be a weighted sum of the average sidelobe energy and the maximum sidelobe in the RoI. In this way, both the maximum sidelobe and the average sidelobe energy are taken into account during the optimization, which may lead to improved performance in the wideband ambiguity function shaping.

Thirdly, it is also worth exploring the joint design of transmit waveform and receive filter for the wideband ambiguity function shaping, as investigated in [31] for the narrowband ambiguity function shaping. Such joint design of two separate sequences has the potential to further improve the optimality of the sidelobe minimization in the RoI, with the cost of a slightly lower SNR.

Lastly, the wideband gradient descent method was specifically used for the waveform design for SAS systems in this thesis. It could be easily extended to other wideband applications, such as the waveform design for underwater target tracking, where the range-Doppler ambiguity of the waveform is also crucial for the tracking performance.

5.1.2 Future work on waveform design for SAS imaging

In this thesis, we have focused on periodic waveform design for SAS imaging, where a single waveform is repeatedly transmitted. Although the designed periodic waveform demonstrates better imaging performance compared with the conventional waveforms, there are still several promising directions for further improving the SAS imaging performance.

Firstly, although some ideal properties of a good SAS imaging sequence are suggested in Section 4.1.3.1, it is still unclear what would be the optimal selection for the RoI $\mathcal{Q}$ and the sequence length N for SAS imaging. For example, we only considered a rectangular-shaped RoI with uniform weights for the wideband ambiguity function shaping in this thesis. However, since some delay-Doppler pairs (τ, α) in the RoI will not be mapped to the SAS imaging domain, we may further improve the SAS imaging performance by not including those pairs during the wideband ambiguity function shaping. The design of the sequence length N and the RoI $\mathcal{Q}$ still requires further investigations.

Secondly, instead of designing a single periodic waveform, another possible direction could be the design of a train of waveforms, where each waveform may occupy a different frequency band. This approach introduces greater flexibility in ambiguity function shaping, and may enable the use of advanced coding strategies to further suppress the imaging sidelobe, improve the imaging robustness, and mitigate the range ambiguity caused by using periodic waveforms.

Bibliography

[1] A. Demir, "Coherent integration for imaging and detection using active sonar," 12 2023. [Online]. Available: https://resolver.tudelft.nl/uuid:eaa5818c-6b0f-4380-9055-9fbf794576bf

[2] R. Hansen, "Synthetic aperture sonar technology review," *Marine Technology Society Journal*, vol. 47, pp. 117–127, 09 2013.

[3] M. P. Hayes and P. T. Gough, "Synthetic aperture sonar: A review of current status," *IEEE Journal of Oceanic Engineering*, vol. 34, no. 3, pp. 207–224, 2009.

[4] M. Pinto, "Design of synthetic aperture sonar systems for high-resolution seabed imaging (tutorial slides)," *NATO Undersea Research Center NURC-PR-2006-029*, 2006.

[5] C. C. Paige and M. A. Saunders, "Lsqr: An algorithm for sparse linear equations and sparse least squares," *ACM Transactions on Mathematical Software (TOMS)*, vol. 8, no. 1, pp. 43–71, 1982.

[6] L. Weiss, "Wavelets and wideband correlation processing," *IEEE Signal Processing Magazine*, vol. 11, no. 1, pp. 13–32, 1994.

[7] M. Ruggiano and P. van Genderen, "Wideband ambiguity function and optimized coded radar signals," in *2007 European Radar Conference*, 2007, pp. 142–145.

[8] G. Kaiser, "Physical wavelets and radar: a variational approach to remote sensing," *IEEE Antennas and Propagation Magazine*, vol. 38, no. 1, pp. 15–24, 1996.

[9] N. Levanon, *RADAR - CONCISE COURSE Volume 1*, 2021. [Online]. Available: https://www.nadavlevanon.sites.tau.ac.il/_files/ugd/dfbcdc_0f87a816725f42fdbc43777690c237e0.pdf

[10] R. Gassner and G. Cooper, "Note on a generalized ambiguity function (corresp.)," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 126–126, 1967.

[11] L. Sibul and E. Titlebaum, "Volume properties for the wideband ambiguity function," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-17, no. 1, pp. 83–87, 1981.

[12] J. Yang and T. K. Sarkar, "Doppler-invariant property of hyperbolic frequency modulated waveforms," *Microwave and Optical Technology Letters*, vol. 48, no. 6, pp. 1174–1179, 2006. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/mop.21573

[13] S. W. Golomb, *Shift register sequences: secure and limited-access code generators, efficiency code generators, prescribed property generators, mathematical models*. World Scientific, 2017.

[14] T. Birdsall, K. Metzger, and R. Spindel, "Signal processing for ocean tomography with moving ships," in *Twenty-Second Asilomar Conference on Signals, Systems and Computers*, vol. 1, 1988, pp. 436–439.

[15] J. Costas, "A study of a class of detection waveforms having nearly ideal range—doppler ambiguity properties," *Proceedings of the IEEE*, vol. 72, no. 8, pp. 996–1009, 1984.

[16] G. Hickman and J. L. Krolik, "Non-recurrent wideband continuous active sonar," in *2012 Oceans*, 2012, pp. 1–6.

[17] J. B. Soli, "Non-recurrent wideband continuous active sonar," Duke.edu, 2014. [Online]. Available: https://dukespace.lib.duke.edu/items/a03c1086-16f6-44b8-9df7-6f5ee9dfc355

[18] D. A. Hague and J. R. Buck, "The generalized sinusoidal frequency-modulated waveform for active sonar," *IEEE Journal of Oceanic Engineering*, vol. 42, no. 1, pp. 109–123, 2017.

[19] ——, "An experimental evaluation of the generalized sinusoidal frequency modulated waveform for active sonar systems," *The Journal of the Acoustical Society of America*, vol. 145, no. 6, pp. 3741–3755, 06 2019. [Online]. Available: https://doi.org/10.1121/1.5113581

[20] C. Guan, Z. Zhou, and X. Zeng, "Optimal waveform design using frequency-modulated pulse trains for active sonar," *Sensors*, vol. 19, no. 19, 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/19/4262

[21] G. Kim, K. Lee, K. Yoon, and S. Lee, "A study on pulse train waveforms for high duty cycle sonar systems: Optimization scheme and relationship between orthogonality and bandwidth," *IEEE Access*, vol. 9, pp. 119 800–119 817, 2021.

[22] P. Stoica, H. He, and J. Li, "New algorithms for designing unimodular sequences with good correlation properties," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1415–1425, 2009.

[23] H. He, "Waveform design for active sensing systems – a computational approach," Ph.D. dissertation, University of Florida, 2011. [Online]. Available: https://ufdcimages.uflib.ufl.edu/UF/E0/04/31/19/00001/he_h.pdf

[24] H. He, J. Li, and P. Stoica, *Waveform Design for Active Sensing Systems: A Computational Approach*. Cambridge University Press, 2012, p. 88–122.

[25] F. Arlery, R. Kassab, U. Tan, and F. Lehmann, "Efficient gradient method for locally optimizing the periodic/aperiodic ambiguity function," in *2016 IEEE Radar Conference (RadarConf)*, 2016, pp. 1–6.

[26] M. Soltanalian and P. Stoica, "Designing unimodular codes via quadratic optimization," *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1221–1234, 2014.

[27] H. He, P. Stoica, and J. Li, "On synthesizing cross ambiguity functions," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 3536–3539.

[28] Z.-q. Luo, W.-k. Ma, A. M.-c. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, 2010.

[29] G. Cui, Y. Fu, X. Yu, and J. Li, "Local ambiguity function shaping via unimodular sequence design," *IEEE Signal Processing Letters*, vol. 24, no. 7, pp. 977–981, 2017.

[30] J. Zhang, C. Shi, X. Qiu, and Y. Wu, "Shaping radar ambiguity function by l -phase unimodular sequence," *IEEE Sensors Journal*, vol. 16, no. 14, pp. 5648–5659, 2016.

[31] F. Wang, S. Feng, J. Yin, C. Pang, Y. Li, and X. Wang, "Unimodular sequence and receiving filter design for local ambiguity function shaping," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–12, 2022.

[32] M. Chitre, J. Tian, and H. Vishnu, "On ambiguity function shaping for broadband constant-modulus signals," *Signal Processing*, vol. 166, p. 107224, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165168419302701

[33] J. Dattoro, "Convex optimization & euclidean distance geometry," Stanford.edu, 2018. [Online]. Available: https://ccrma.stanford.edu/~dattorro/mybook.html

[34] Z.-Q. Luo, N. D. Sidiropoulos, P. Tseng, and S. Zhang, "Approximation bounds for quadratic optimization with homogeneous quadratic constraints," *SIAM Journal on optimization*, vol. 18, no. 1, pp. 1–28, 2007.

[35] S. Zhang and Y. Huang, "Complex quadratic optimization and semidefinite programming," *SIAM Journal on Optimization*, vol. 16, no. 3, pp. 871–890, 2006.

[36] K. B. Petersen, M. S. Pedersen *et al.*, "The matrix cookbook," *Technical University of Denmark*, vol. 7, no. 15, p. 510, 2008.

[37] D. P. Kingma, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[38] B. O'Donoghue, "Operator splitting for a homogeneous embedding of the linear complementarity problem," *SIAM Journal on Optimization*, vol. 31, pp. 1999–2023, August 2021.

[39] ——, "Scs — scs 3.2.7 documentation," Cvxgrp.org, 2021. [Online]. Available: https://www.cvxgrp.org/scs/

Appendix A

A.1 Proof: the volume invariance property of narrowband ambiguity function

Recall that the expression of a NAF is

$$\chi_N(\tau, f_D) = \int_{-\infty}^{\infty} x_{bb}(t) x_{bb}^*(t-\tau) e^{j2\pi f_D t} dt \tag{A.1}$$

Then we have

$$\begin{aligned} |\chi_N(\tau, f_D)|^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_{bb}(t_a) x_{bb}^*(t_a-\tau) e^{j2\pi f_D t_a} x_{bb}^*(t_b) x_{bb}(t_b-\tau) e^{-j2\pi f_D t_b} dt_a dt_b \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_{bb}(t_a) x_{bb}^*(t_a-\tau) x_{bb}^*(t_b) x_{bb}(t_b-\tau) e^{j2\pi f_D (t_a-t_b)} dt_a dt_b \end{aligned} \tag{A.2}$$

The volume under the surface of a narrowband ambiguity function can be expressed as

$$V_{NAF} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\chi_N(\tau, f_D)|^2 d\tau df_D \tag{A.3}$$

By substituting Equation A.2 into Equation A.3, we obtain

$$\begin{aligned} V_{NAF} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_{bb}(t_a) x_{bb}^*(t_a-\tau) x_{bb}^*(t_b) x_{bb}(t_b-\tau) e^{j2\pi f_D (t_a-t_b)} dt_a dt_b d\tau df_D \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_{bb}(t_a) x_{bb}^*(t_a-\tau) x_{bb}^*(t_b) x_{bb}(t_b-\tau) [\int_{-\infty}^{\infty} e^{j2\pi f_D (t_a-t_b)} df_D] dt_a dt_b d\tau \end{aligned} \tag{A.4}$$

Since $\int_{-\infty}^{\infty} e^{j2\pi f_D (t_a-t_b)} df_D = \delta(t_a - t_b)$, we have

$$\begin{aligned} V_{NAF} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_{bb}(t_a) x_{bb}^*(t_a-\tau) x_{bb}^*(t_b) x_{bb}(t_b-\tau) \delta(t_a-t_b) dt_a dt_b d\tau \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_{bb}(t_a) x_{bb}^*(t_a-\tau) x_{bb}^*(t_a) x_{bb}(t_a-\tau) dt_a d\tau \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x_{bb}(t_a)|^2 |x_{bb}(t_a-\tau)|^2 d\tau dt_a \end{aligned} \tag{A.5}$$

Let $t_c = t_a - \tau$, and substituting it into Equation A.5, we get

$$\begin{aligned} V_{NAF} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x_{bb}(t_a)|^2 |x_{bb}(t_c)|^2 dt_c dt_a \\ &= \int_{-\infty}^{\infty} |x_{bb}(t_c)|^2 dt_c \int_{-\infty}^{\infty} |x_{bb}(t_a)|^2 dt_a = (\|x_{bb}(t)\|^2)^2 \\ &= |\chi_N(0,0)|^2 \end{aligned} \tag{A.6}$$

Therefore, the volume under a NAF is a fixed constant that only depends on the energy of the transmit signal.