GRID MATCHING IN COMPRESSIVE SENSING

HÜSEYİN ŞAR

Master of Science Thesis

Graduate School of Sciences

Electrical and Electronics

Engineering Program

January 2014

JÜRİ VE ENSTİTÜ ONAYI

Hüseyin ŞAR'ın **Grid Matching in Compressive Sensing** başlıklı **Elektrik-Elektronik Mühendisliği** Anabilim Dalı **Telekomünikasyon** Bilim Dalındaki yüksek lisans tezi 27 - 01 - 2014 tarihinde aşağıdaki jüri tarafından Anadolu Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca değerlendirilerek kabul edilmiştir.

	Adı -Soyadı	**İmza**
Üye (Tez Danışmanı)	**: Yrd. Doç. Dr. Nuray AT**	**.....**
Üye	**: Prof. Dr. Ömer Nezih GEREK**	**.....**
Üye	**: Yrd. Doç. Dr. Erol SEKE**	**.....**

Anadolu Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun ..... tarih ve ..... sayılı kararıyla onaylanmıştır.

Enstitü Müdürü

ABSTRACT

Master of Science Thesis

Grid Matching in Compressive Sensing

Hüseyin ŞAR

Anadolu University
Graduate School of Sciences
Electrical and Electronics Engineering Program

Supervisor: Assist. Prof. Dr. Nuray AT

2014, 42 pages

Sparse signal recovery from compressive measurements assumes a grid of possible support points from which to estimate the signal support set. However, reconstruction of high measurement resolution waveforms is very sensitive to small grid offsets and assuming a fixed grid may result to information loss. On the other hand, identifying sparse elements over a very fine grid to minimize information loss is computationally prohibitive. In this work grid matching is performed via a computationally efficient multi-stage Monte Carlo sampling approach. The multi-stage sampling method identifies sparse signal elements and chooses the appropriate grid using information from compressively acquired measurements and any prior information on the signal structure. The effectiveness of the method in reconstructing high resolution waveforms, after compressive acquisition, is demonstrated via simulation study.

Keywords: Bayesian Compressive Sensing, Sparse Signal Reconstruction, Monte Carlo Methods, Grid Matching.

ÖZET

Yüksek Lisans Tezi

Sıkıştırmalı Algılamada Izgara Eşleştirme

Hüseyin ŞAR

Anadolu Üniversitesi
Fen Bilimleri Enstitüsü
Elektrik Elektronik Mühendisliği

Danışman: Yard. Doç. Dr. Nuray AT
2014, 42 Sayfa

Seyrek sinyallerin sıkıştırılmış ölçümlerden rekonstrüksiyonunda, sinyalin destek noktalarının tahmin edildiği muhtemel destek noktalarından oluşan bir ızgara kullanılır. Ancak yüksek ölçüm çözünürlüklü dalga şekillerinin rekonstrüksiyonu küçük ızgara kaymalarına karşı çok hassastır ve sabit bir ızgara varsayımı bilgi kaybına yol açabilir. Öte yandan, seyrek atomların belirlenmesinde bilgi kaybını en aza indirmek için, daha küçük gözenekli ızgara seçmek seyrek atomların bulmak için hesapsal anlamda daha büyük bir iş yükü gerektirmektedir. Bu çalışmada ızgara eşleştirme işlemi, çok aşamalı Monte Carlo örnekleme yaklaşımı ile yapılmıştır. Çok aşamalı örnekleme yöntemi, seyrek sinyal elemanlarını belirler ve sıkıştırılmış ölçümlerdeki bilgiyi ve sinyal yapısı ile ilgili herhangi bir ön bilgiyi kullanarak uygun ızgara seçimini yapar. Yöntemin sıkıştırma işleminden sonraki etkinliği, yüksek çözünürlüklü dalga şekillerinin rekonstrüksiyonunda simülasyon çalışması ile gösterilmiştir.

Anahtar Kelimeler: Bayesçi Sıkıştırma Algılama, Seyrek Sinyal Rekonstrüksiyon, Monte-Carlo Yöntemleri, Izgara Eşleştirme.

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my advisor Assist. Prof. Dr. Nuray AT for the continuous support of my master study and research, for her patience, motivation, enthusiasm, and immense knowledge.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Dr. Ömer Nezih Gerek, and Assist. Prof. Dr. Erol Seke, for their encouragement, insightful comments, and hard questions.

Last but not the least, I would like to thank my wife 'Dilan Güneş Şar' for her unconditional emotionally support -throughout my degree. I could not have imagined overcoming this study without her support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The conventional method for sampling analog signals is based on Nyquist's and Shannon's theorem which guarantees perfect reconstruction of sampled analog signal, if sampling frequency is at least twice as the signal's highest frequency. Most of the signals of interest contain high frequency components hence the conventional sampling theory sets some boundaries on reconstructing or processing these type of signals. For example, an image signal, having sharp edges contains quite high frequency components. Due to the Nyquist theorem, one has to take a lot of samples (pixels) to be able reconstruct the image. Since there will be a redundant information, as a second step, compression is usually performed. In Figure 1.1, the conventional signal compression scheme is seen, the signal x is sampled and then compressed bringing extra computational load, here N denotes signal's dimension and S is a dimension of the compressed signal. Note that if only necessary data could be taken during sampling step there will be no need for an extra compression and decompression steps. Compressive sensing also known as compressive sampling or sparse sampling is due to the work of Candes, Romberg, and Tao [4] and Donoho [5], where they show that a signal compressed by using classical methods such as transform coding can also be efficiently acquired via a small set of nonadaptive, linear and usually randomized measurements during sampling. So that compressive sensing approach can be used instead of Nyquist theorem with reduced sampling rate and processing loads.

The main difference between compressive sensing and classical sampling theory is the way they deal with signal recovery. Classical sampling theory recovers the sampled signal by using sinc interpolation which is a linear, simple and lightweight process. On the other hand,in compressive sensing signal recovery is achieved by using some nonlinear, iterative and costly algorithms (Greedy pursuit, Convex relaxation, Bayesian framework, Nonconvex optimization, etc.) [6].

Compressive acquisition of signals has been shown to preserve sparse

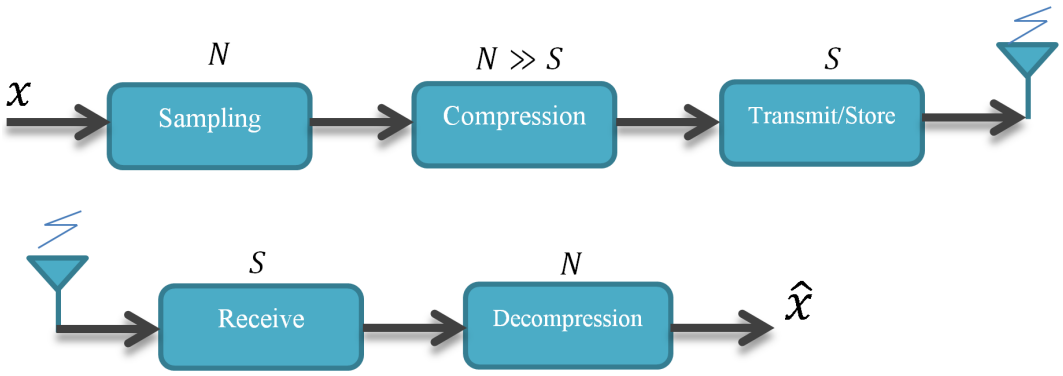


Figure 1.1: The conventional signal compression scheme

signal structure in noise. A reconstruction process following compressive acquisition is then able to reveal the sparse elements in the signal. The reconstruction process assumes a grid on which the sparse signal to be recovered lies. Recovery is then based on a sparsity basis or dictionary [6], [7] which is supported on the assumed grid and defines signal sparsity or compressibility [8–10].

Signals encountered in real world applications are usually determined by continuous parameters. Therefore, the sparse elements in the signal may not lie on the assumed grid and not perfectly match the sparsifying basis or dictionary. If the true signal is not exactly supported on the assumed grid, the basis (grid) mismatch problem occurs [11–13]. The problem created by the mismatch between the true signal grid and the recovery grid, is that signal energy spills onto off-grid components during recovery. This signal energy spread renders the signal incompressible in the grid assumed for recovery and leads to information loss. The grid offset may even cause total loss of information in the case of using high resolution signals [14], [15]. The sensitivity of mismatch has been studied in [11].

One method to solve the grid matching problem is to use a very fine grid. This, however, results in high computational complexity and numerical instability, over shadowing any advantage it might have in sparse recovery [13], [16]. In [16], the issues arising from discretization based on a grid are overcome by working directly on the continuous parameter space. This yields

an infinite dictionary of continuous atoms and arbitrarily high coherence. The reconstruction algorithm is formed as the solution to an atomic norm minimization problem shown to be equivalent to a semi definite program. In [13] the continuous basis pursuit method is developed which uses a dictionary with an auxiliary interpolation function to overcome grid mismatch. In [12] the grid mismatch is modelled using a mismatch parameter which is estimated via an alternating descent algorithm.

In this work the grid matching problem is investigated for the case of reconstructing highly peaked ambiguity function (AF) radar waveforms used in high-resolution target tracking [14]. Due to the fast decay of the AF of such signals a fine grid is required to identify sparse signal elements since for minor grid offsets the signal's energy vanishes. Reconstruction on a very fine grid would, however, be computationally expensive. The grid-matching problem is efficiently solved in this work via a multi-stage Monte Carlo approach. The methods uses a divide and conquer strategy to identify possible signal support point subsets and then samples over the smaller subsets to estimate signal support points and magnitudes. The method combines information from compressive measurements and prior information on the signal structure to achieve the reconstruction goal. Specifically, the proposed methodology allows the use of any prior information on the signal without the assumptions used in Bayesian reconstruction methods which seek a closed form reconstruction solution [10]. Moreover, this work extends prior work on Monte Carlo Bayesian reconstruction using the Monte Carlo Bayesian Compressive Sensing (MC-BCS) [17] by avoiding the assumption of a known grid.

The proposed grid matching reconstruction method is applied to a radar scenario where radar return waveforms composed of atoms with different delay-Doppler shifts are compressively sensed and reconstructed. The effectiveness of the method to accurately reconstruct radar signals is then assessed in terms of mean squared error performance (MSE) in the delay-Doppler plane. Moreover, prior information on the signal structure is shown to improve reconstruction performance.

1.1. Organization of The Thesis

This thesis is organized as follow:

In Chapter 2, basic concepts in compressive sensing, l_p-norms, orthogonal matching pursuit reconstruction approach for sparse signal recovery, some basic probability theory, ambiguity function and complex Björck sequences are reviewed.

In Chapter 3, first grid mismatch problem is defined in general and as an application it is illustrated on ambiguity function. Then grid matching in Monte Carlo Bayesian Compressive Sensing is given in details. The present version of this chapter is mainly drawn from [18].

In Chapter 4, numerical set up and results of the proposed method is given.

In Chapter 5, concluding remarks and possible future work directions are given.

2. PRELIMINARIES

In this chapter mathematical concepts which are used in this thesis are reviewed. First the norm concept and its properties are given. More detailed information about vector norms can be found in [1]. Then, compressive sensing approach and its problem formulation are introduced. Sparse signal recovery algorithms such as l_0-norm, l_1-norm, l_2-norm and orthogonal matching pursuit reconstruction approaches are defined. As an application ambiguity function and its properties are introduced. Finally Monte Carlo methods and Population Monte Carlo with Gibbs sampling method are reviewed.

2.1. Mathematical Background

2.1.1. Norms

In compressive sensing l_1-norm, l_2-norm and l_0-norm are used to find the sparsest solution among in many possible solutions by minimizing the error in a constrained optimization setup.

A vector norm is a metric function that measures the distance or length of a vector in a vector space V. It is denoted by $\|.\|_p$. The norm of a vector $x = [x_1, x_2, ...x_n]^T \in V^n$ is a linear function and must satisfy following three properties.

- Positive Definiteness: $\|x\|_p \geq 0$.
- Linearity: $\|\alpha x\|_p = \alpha \|x\|_p$ where α is any scalar.
- Triangle Inequality: $\|x + y\|_p \leq \|x\|_p + \|y\|_p$ where $y = [y_1, y_2, y_3, ..., y_n]^T \in V^n$ for all vectors in V.

In this thesis the interested norm class is l_p- norm class and specifically the l_0- norm. l_p- norms are defined by

$$\|x\|_p = \begin{cases} (\sum_{i=1}^{n} |x_i|^p)^{1/p} & , \ p \in [1, \infty); \\ \max\limits_{i=1,2,..,n} |x_i| & , \ p = \infty. \end{cases} \tag{2.1}$$

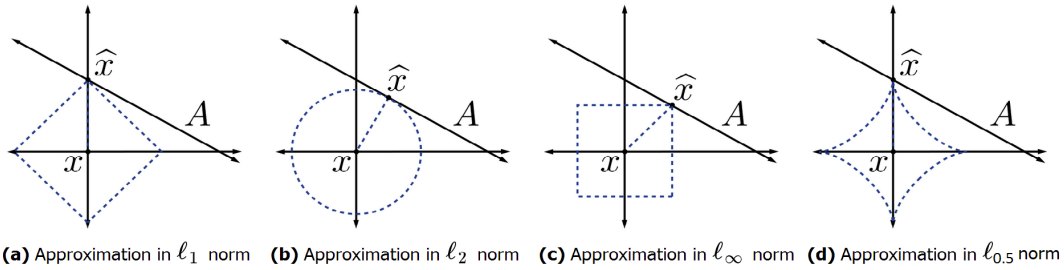
This definition is also valid for $p < 1$ but since these norms do not satisfy the triangle inequality they are known as *pseudo-norms*. As a special case for $p = 0$ the l_0- norm is not a pseudo norm and it is defined as;

$$\|x\|_0 = |x| \quad \text{where} \quad |x| = \text{the number non-zero elements of } x. \tag{2.2}$$

Specifically, l_1- norm and l_2- norm are used as reconstruction tools in this thesis and they are given by

$$\begin{aligned} \|x\|_1 &= \sum_{i=1}^{n} |x_i| \qquad l_1 - \text{norm} \\ \|x\|_2 &= \sqrt{\sum_{i=1}^{n} |x_i|^2} \qquad l_2 - \text{norm} \end{aligned} \tag{2.3}$$

Generally norms are used for measuring strength of a signal or to quantify the error in approximation problems. For example let $x \in \mathbb{R}^2$ is a given vector and A is a one dimensional affine space. To compute the closest point in A to x, the approximation error between x and a point in A is found by $\|x - \hat{x}\|_p$. The choice of p will directly effect the approximation error. The Figure 2.1 illustrates this concept.



(a) Approximation in ℓ_1 norm **(b)** Approximation in ℓ_2 norm **(c)** Approximation in ℓ_∞ norm **(d)** Approximation in $\ell_{0.5}$ norm

Figure 2.1: Approximation errors in $\mathbb{R}^2$ for some l_p norms [1]

2.1.2. Probability Concepts

Conditional Probability

Let A and B be two events, such that $P(B) > 0$. The probability of event A given that event Bhas occurred is called conditional probability and defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{2.4}$$

where $P(A \cap B)$ is probability of intersection event A and B. If the event A and event B are independent events then conditional probability of event A given that event B has occurred is

$$P(A|B) = P(A) \tag{2.5}$$

Bayes' Rule

Let $A_1, A_2, ..., A_n$ be n mutually exclusive events. The union of these events creates the sample space S. Then for every event A, Bayes rule says that

$$P(A_k|A) = \frac{P(A_k \cap A)}{P(A)} = \frac{P(A|A_k)P(A_k)}{\sum_k P(A|A_k)P(A_k)} \tag{2.6}$$

Expectation

Let X be a discrete random variable with probability mass function $p(x)$. Then the expected value of X is defined as

$$E[X] = \sum_x xp(x) \tag{2.7}$$

Similarly for continuous random variable X with probability distribution function $f_X(x)$ the expectation of X is given by

$$E[X] = \int_{-\infty}^{\infty} xf_X(x)dx \tag{2.8}$$

Uniform Distribution

Let X be a uniformly distributed random variable on the interval a to b, $a < b$ then the probability density function of X is given by

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0 \quad , & otherwise \end{cases} \tag{2.9}$$

Normal Distribution

The normal distribution is also known as Gaussian distribution and it is one of the most important distributions in literature. Generally it is denoted as $N(\mu, \sigma^2)$ where μ is mean and σ^2 is variance of the distribution. Probability density function of a Gaussian random variable is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{2.10}$$

Laplace Distribution

Let X be a random variable with Laplace distribution with parameter λ; then the probability density function of X

$$f_X(x) = \frac{\lambda}{2} e^{-\lambda|x|} \tag{2.11}$$

Note that $1/\lambda$ is equal to the mean.

2.1.3. Ambiguity Function

Ambiguity function (AF) of radar waveforms are very important time-frequency signal processing tools for high resolution target tracking in radar systems [19]. Due to their high resolution and fast decayed peak, AF is a suitable candidate to show mismatch effect in compressive sensing. Detailed information about radar signals and ambiguity function can be found in [20]. The AF is the time response of a filter matched to a given finite energy signal when the signal is received with a delay τ and a Doppler shift ν relative to the nominal values. Specifically,

$$|A(\tau, \nu)| = \left| \int_{-\infty}^{\infty} d(t) d^*(t-\tau) e^{j2\pi\nu t} dt \right| \tag{2.12}$$

where $d(t)$ is the complex envelope of the transmitted signal. The ideal AF is a 2 dimensional Dirac delta function, that is, $A(\tau, \nu) = \delta(\tau)\delta(\nu)$.

The AF has four important properties which make it a major tool for analysing radar signals. These properties are explained in the following. First two properties assume the condition of $d(t)$ having unit energy.

Property-1

AF takes the highest value at the origin, i.e. it is maximum at $(0, 0)$. If $d(t)$ is normalized, AF's maximum value will be equal to 1.

$$|A(\tau, \nu)| \leq |A(0, 0)| = 1 \tag{2.13}$$

Property-2

The total volume under the normalized AF surface equals to 1. The waveform type doesn't affect this volume. If the volume of peak squeezed out by some target position changes, it will reappear somewhere else.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |A(\tau, \nu)|^2 d\tau d\nu = 1 \tag{2.14}$$

Property-3

Every points in AF surface are symmetric with respect to the origin. Because AF plots generally contain two adjacent quadrants of the AF.

$$|A(-\tau, -\nu)| = |A(\tau, \nu)| \tag{2.15}$$

Property-4

If the transmitted signal is linear frequency modulated, the resulting AF will be cropped. Let the AF of $d(t)$ be $|A(\tau, \nu)|$ then LFM effect is expressed as follows

$$d(t) \Leftrightarrow |A(\tau, \nu)| \tag{2.16}$$

$$d(t)\mathrm{e}^{(j\pi kt^2)} \Leftrightarrow |A(\tau, \nu - k\tau)| \tag{2.17}$$

2.1.4. Björck Sequences

Björck sequences are known as constant amplitude zero autocorrelation (CAZAC) sequences [21]. CAZAC sequences have perfect periodic autocorrelation function with zero side lobes which makes these sequences excellent candidates for radar tracking. Björck sequences can be defined for prime length P as follows

$$b_P(n) = \begin{cases} e^{\theta(n|p)} & \text{for } p \equiv 1 \, mod \, 4 \\ e^{\varphi[n|p]} & \text{for } p \equiv 3 \, mod \, 4 \end{cases} \tag{2.18}$$

and

$$(n|p) = \begin{cases} 0 & if \;\; n \equiv 0 \, mod \, P \\ +1 & if \;\; n \, is \, a \, quadratic \, residue \\ -1 & if \;\; n \, is \, a \, quadratic \;\; non - residue \end{cases} \tag{2.19}$$

where $(n|P)$ is called as Legendre symbol [22], $\theta = \arccos(\frac{1}{1+\sqrt{P}})$ and $\varphi = \arccos(\frac{1}{\sqrt{1+P}})$. $[n|P]$ is a function which has a similar definition as Legendre symbol; the difference is $[n|P]$ is not equal 0 instead it is equal to 1 if $n \equiv 0 \bmod P$.

2.2. Compressive Sensing Theory

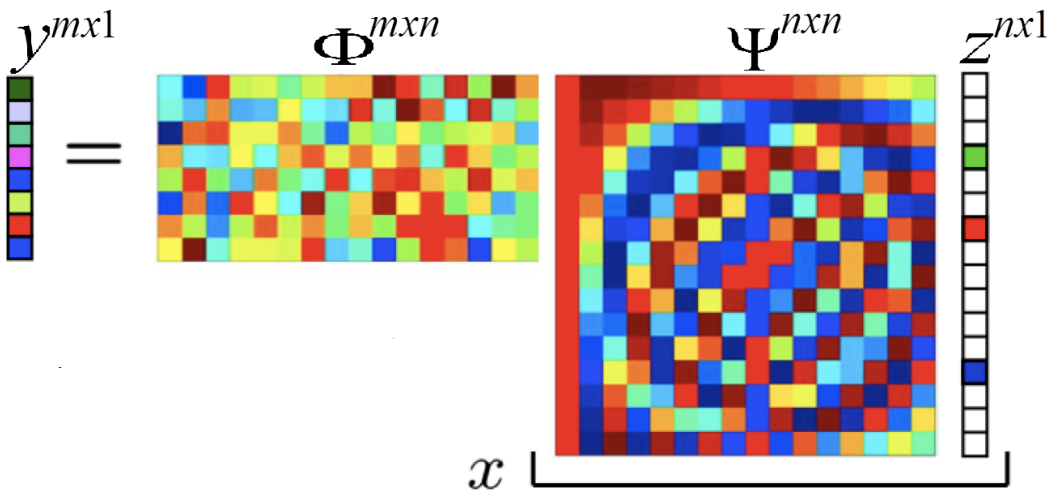


Figure 2.2: Compressive sensing measurement process [2]

Compressed sensing concept consists of two steps; first one is sampling the signal by using a given or chosen transformation matrix and second one is reconstructing the signal back from its transformed version (measurements). Reconstruction step is usually done via solving a constrained optimization problem.

Now let x be a real valued, finite-length, one-dimensional discrete time signal $x = [x_1, x_2, ...x_n]^T$. Here the aim of the compressive sensing is to reconstruct the original signal x from randomly or adaptively taken m measurements where $m \ll n$.

In CS, the reconstruction of the signal depends on three important properties which are Sparsity, Restricted Isometry Property and Mutual Coherence.

Let $\Psi \in \mathbb{C}^{n \times n}$ be an orthonormal matrix where its each column ($\Psi = [\psi_1 \, \psi_2 ... \, \psi_i ... \psi_n]$) is a basis vector. Now we can represent $x \in \mathbb{C}^n$ as a linear combination of these basis vectors by

$$x = \sum_{i=1}^{n} z_i \psi_i \qquad \text{or} \qquad x = \Psi z \tag{2.20}$$

where $z \in \mathbb{C}^n$ is equivalent representation of x .

Measurement of the signal is performed by sampling the signal x or equivalently z using a measurement matrix $\Phi \in \mathbb{C}^{m \times n}$ and $y = \Phi x = \Phi \Psi z = \Theta z$ where $y \in \mathbb{C}^m$ are observations and rows of the Φ span $\mathbb{C}^n$ that we can reconstruct the original signal x from observations y .

$$y = \Phi x = \Phi \Psi z = \Theta z \tag{2.21}$$

where $\Theta = \Phi \Psi$. This is illustrated in Figure 2.2.

2.2.1. Sparsity

Consider a length N signal x. x is called as a $K-$ sparse signal if number of nonzero elements (K) is smaller than the number of zero elements ($N - K$);

$$K << N \tag{2.22}$$

The natural signals could be sparse in the time domain but if not to obtain its sparse representation, they can be represented as sparse signals

in different domains by transformation such as Gabor, Wavelet, Fourier, etc. These are widely used transformations for representing natural signals.

$$x = \Psi z \tag{2.23}$$

where Ψ is a $N \times N$ sparsity basis or transformation matrix and z is natural signal.

2.2.2. Measurement Matrix

The basic CS scheme generally depends on randomly taken measurements which should contain enough information to reconstruct the real signal. The measurement procedure of CS is done via $M < N$ incoherent projections on to a second basis function;

$$y = \Phi x \tag{2.24}$$

where y is a $m \times 1$ column vector (also called measurements) and Φ is an $M \times N$ measurement matrix.

Note that there are infinitely many solutions for x in (2.24), CS theory suggests only one unique solution if measurement matrix satisfies two main conditions. These are Restricted Isometry Property (RIP) [4] and Mutual Coherence [23].

Restricted Isometry Property

The measurement matrix Φ satisfies RIP with parameters (K, ρ) if;

$$(1-\rho)\left\|x\right\|_2^2 < \left\|\phi x\right\|_2^2 < (1+\rho)\left\|x\right\|_2^2 \tag{2.25}$$

holds for all $K-$ sparse vectors. The purpose of the RIP condition is to keep the Euclidean distance of all $K-$ sparse signals after their projection onto a lower dimensional space, measurement space. It means that

$$(1-\rho) < \frac{\left\|\phi x_1 - \phi x_2\right\|_2^2}{\left\|x_1 - x_2\right\|_2^2} < (1+\rho) \tag{2.26}$$

where $\rho \approx (0, 1)$, x_1 and x_2 are $K-$ sparse variables.

Mutual Coherence

The other key feature of the measurement matrix is required to provide is mutual coherence. Mutual coherence is introduced by Donoho et al [23] which is given by

$$\mu(\Phi) \triangleq \sup\{|\langle \phi_i, \phi_j \rangle| : \forall i, j, \, where \, i \neq j\} \tag{2.27}$$

where Φ is the measurement matrix and $\mu(\Phi)$ is coherence between measurement matrix's columns. This measure is easier to calculate than RIP parameter since the complexity of the calculation changes dramatically by the number of measurement matrix's columns.

Suppose that,

$$\|\hat{x}\|_0 < \frac{1}{2}(1 + \frac{1}{\mu(\Phi)}) \tag{2.28}$$

This equation will guarantee $\hat{x}$ which is the sparsest and exact solution for $\Phi\hat{x} = y$ [24].

2.2.3. Signal Reconstruction Algorithms

Given measurements $y = \Phi x = \Phi\Psi z$ where $z, x \in \mathbb{C}^n$ and $y \in \mathbb{C}^m$ there are infinitely many solutions of $y = \Phi x$ for $m << n$. Generally this type of inverse problem is solved by using the least square approach, that is,

$$\tilde{x} = \arg\min_{\hat{x}} \|\hat{x}\|_2 \quad s.t. \quad \Phi\hat{x} = y \tag{2.29}$$

The closed form solution of this problem is $\tilde{x} = (\Phi\Phi^T)^{-1}\Phi^T y$ but this solution is generally not a sparse solution and measures energy not the sparsity as illustrated in Figure 2.3. Hence some other methods are proposed for this pseudo-inverse problem. One of these methods uses $l_0 - norm$ which also enforces sparsity as a constraint. This optimization problem is stated as follows:

$$\tilde{x} = \arg\min_{\hat{x}} \|\hat{x}\|_0 \quad s.t. \quad \Phi\hat{x} = y \tag{2.30}$$

$l_0 - norm$ counts the number of non-zero elements in x so that this method is a proper candidate for finding the solution for the reconstruction problem. However, in this case combinatorial search must be performed. That is

$\binom{n}{K}$ time different combinations must be evaluated for $a\ K-$sparse vector. Therefore $l_0 - norm$ solution has very high complexity which is also known as NP-Hard [25].

The idea of changing l_0- norm by the closest convex norm such as l_1- norm proposed by Chen, Donoho and Saunders in [26]. $l_1 - norm$ approach is known as Basis Pursuit and can be stated as follows:

$$\tilde{x} = \arg\min_{\hat{x}} \|\hat{x}\|_1 \qquad s.t. \qquad \Phi\hat{x} = y \tag{2.31}$$

Reconstruction based on l_1- norm is shown exactly to recover $K-$ sparse signal under the conditions of sparsity of x and incoherence of the measurement matrix Φ. Moreover, it is shown that the solution will rebuild real signal $\tilde{x}$ with high probability by using $m \geq cK\log(N/K)$ iid Gaussian measurements [4] [5].

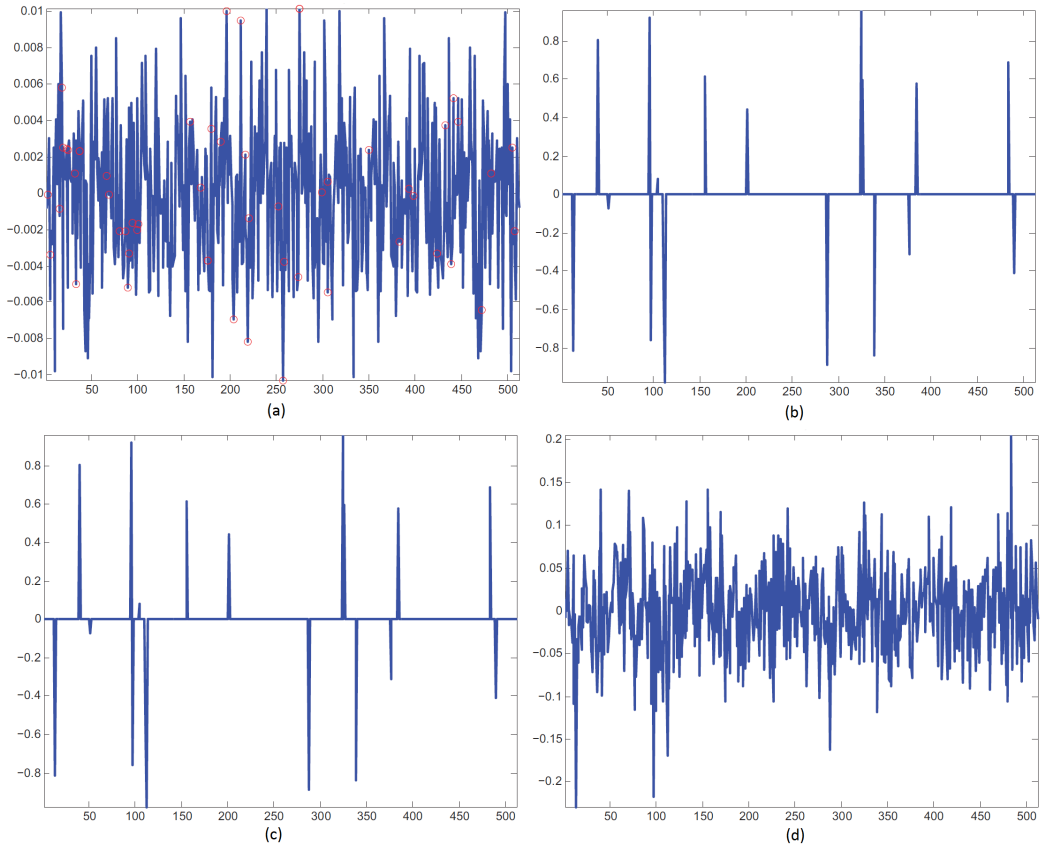


Figure 2.3: (a) Original signal f with random sample points (indicated by red circles); (b) The Fourier transform $\hat{f}$; (c) Perfect recovery of $\hat{f}$ by l_1 minimization; (d) Recovery of f by l_2 minimization [3]

Figure 2.3 illustrates these reconstruction approaches where l_1 and l_2 solutions of original signal's fourier domain representation is shown. It is clearly seen that how well l_1- norm reconstructs the signal and preserve the sparsity versus l_2- norm. As mentioned before l_2- solution does not guarantee the sparsity since it measures the energy.

l_1-norm works properly even under noisy measurements if a carefully chosen error threshold $\varepsilon > 0$ and a particular regularization parameter $\lambda > 0$ which make a balance between Euclidean and sparsity. This problem is equivalent to the unconstrained version which is given by

$$\tilde{x} = \arg\min_{\hat{x}} \|\hat{x}\|_1 \quad s.t. \quad \Phi\hat{x} - y \leq \varepsilon \tag{2.32}$$

$$\tilde{x} = \arg\min_{\hat{x}} \|\Phi\hat{x} - y\|_2^2 + \lambda\|\hat{x}\|_1 \tag{2.33}$$

The l_1-norm recovery is an efficient method in various type of applications but it is not efficient enough for large datasets. So that different methods has been suggested in literature including Greedy algorithms, Bayesian methods, etc. *Orthogonal Matching Pursuit (OMP)* is one of these methods [27]. OMP is a greedy and iterative algorithm which is given in Table 2.1.

Table 2.1: Orthogonal Matching Pursuit

Algorithm 2.2..1: OMP($Orthogonal Matching Pursuit$)

Given:

- $\Phi = (\psi_i)_{i=1}^{n} \in \mathbb{C}^{m \times n}$
- $x \in \mathbb{R}^n$
- Error Threshold ε

Initialization:

- $k = 0,\ x^0 = 0$
- $r^0 = y - \Phi x^0$
- $S^0 = \mathrm{sup} x^0 = \Phi$

while $\left\|r^k\right\|_2 < \varepsilon$ **do**

- $k = k + 1$;
- Choose i_0 such that;

  $\min_c \left\|c\psi_{i_0} - r^{k-1}\right\|_2 \leq \min_c \left\|c\psi_i - r^{k-1}\right\|_2$ for all i
- $S^k = S^{k-1} \cup \{i_0\}$
- $x^k = \arg\min_x \left\|\Phi x - y\right\|_2 \ s.t.\ \mathrm{sup} x = S^k$
- $r^k = y - \Phi x^k$

end

Output:

- x^k

Bayesian compressive sensing is an another method to solve the reconstruction problem which is explained in the next section.

2.3. Monte Carlo Bayesian Compressive Sensing

2.3.1. Bayesian Framework

In Bayesian modeling, solution of the unknown variable relies on the posterior probabilities. In the linear model of CS $y = \Phi x + n$ and according

to Bayesian approach the solution x is being calculated based on posterior probabilities $p(x|y)$. Recall that $p(x|y) = \frac{p(y|x).p(x)}{p(y)}$ where $p(y|x)$ is likelihood, $p(y)$ is the evidence and $p(x)$ is a prior probability. Generally a Bayesian solution is quite hard to find in a closed form since it requires solving complex integrals. Thus Monte Carlo method is used as an approximation tool for the solution of this problem.

If the sparsity of an $n \times 1$ vector x is formalized by a multivariate Laplace prior, the maximum a posterior estimator of x is given by

$$p(x/\lambda) = \frac{\lambda}{2} \exp(-\frac{\lambda}{2}\|x\|_1) \tag{2.34}$$

$$x_{MAP} = \min_x \{\|y - \Phi x\|_2^2 + \alpha\|x\|_1\} \tag{2.35}$$

which is a classic sparse signal recovery by using l_1-norm with a threshold α that balances between sparsity and Euclidean norm [28], [29]. The parameter λ is known as hyper parameter which determines the distribution and obtaining this parameter is vital in the recovery process.

2.3.2. Monte Carlo Methods

Monte Carlo (MC) methods are a set of computational algorithms based on repeated statistical random sampling to obtain some signal or numerical results which have the same statistical properties. MC methods are especially useful for any system with many coupled degrees of freedoms systems such as cellular networks, light transport models etc. There are many types of MC algorithms but general pattern is given in Table 2.2.

For example, let X be a random variable and its expected value or mean is $\mu = E[X]$. If $X_1, ..., X_n$, are n independent random variables generated from same distribution, the approximation for the mean is given by

$$\mu \approx \hat{\mu}_n = \frac{1}{n} \sum_{k=1}^{n} X_k \tag{2.36}$$

This type of approximation algorithm is first introduced by Horvitz and Thompson [30]

Table 2.2: General MC Algorithm Pattern

Algorithm 2.3..1: MC(*General MC Algorithm Pattern*)

Step-1 Determine the possible input domain.

Step-2 Generate random numbers as inputs from a probability distribution.

Step-3 Apply some deterministic computations which could be mean compensation or variance compensation etc.

Step-4 Find the approximation error between produced samples and real samples.

Population Monte Carlo with Gibbs Sampling

As a special method of MC methods, Population Monte Carlo (PMC) with Gibbs sampling is one of the suitable algorithms for a Bayesian framework. PMC is defined as a methodology for approximating joint distributions of unknowns. The unknowns are approximated by randomly measuring samples and weights iteratively where at each iteration samples of unknowns produced from known distribution (A parametric prior distribution) [31]. Gibbs sampling is added to the PMC algorithm to specify the random number generating function. In Gibbs sampling generation of samples is being made from joint probability distribution of two or more unknowns and it is efficient in high dimensions [32]. Djuric and et al combined PCM and Gibbs sampling to generate high dimensional numbers more efficiently in a Bayesian aspect [33]. This algorithm is given in Table 2.3.

Table 2.3: PMC with Gibbs Sampling

Algorithm 2.3..2: MC(*PMC with Gibbs Sampling*)

Variables:

$\pi(x)$: Prior distribution

$y \in \mathbb{R}^{d_y \times 1}$: Observation Vector

$h : \mathbb{R}^{d_\theta \times 1} \times \mathbb{R}^{d_w \times 1} \to \mathbb{R}^{d_y \times 1}$: Observation Generating Function

$\theta \in \mathbb{R}^{d_\theta \times 1}$: Noise Vector with Parametric Distribution

$\theta \in {}^{d_\theta \times 1}$: Unknown vector

$q(\theta)$: Generating Function from Prior

Given:

$\pi(x),\ w,\ h,\ \theta$

Initialization:

Generate sample streams from $\pi(x)$

$\theta_l^m = [\theta_{1,l}^m, \theta_{2,l}^m, \theta_{3,l}^m, ..., \theta_{d_\theta,l}^m]$ where $l = 0, 1, 2, ..., d_\theta$

Compute y from $y = h(\theta, w)$

Assign weights according to $\omega_l^m = p(y|\theta_l^m)$

Procedure:

Step-1 Randomly Choose order of generation of θ

Step-2 Sample a particle from conditioning based on the normalized weights of particles from the previous iteration at jth iteration and $m = 1, 2, ..., M$ where M is number of particles.

$\theta_{l_1,j}^m \sim q_{l_{1,j}}(\theta_{l_1}|\theta_{l_1,j-1}^m, \theta_{l_1,j-1}^m, ..., \theta_{l_1,j-1}^m)$

For $n = 1, 2, ..., d_(\theta - 1)$

$\theta_{l_n,j}^m \sim q_{l_{n,j}}(\theta_{l_n}|\theta_{l_n,j-1}^m, ..., \theta_{l_n,j-1}^m, \theta_{l_{n+1},j-1}^{\lambda m}, ..., \theta_{l_{n+1},j-1}^{\lambda m})$

where λm is selected particle index.

$\theta_{l_{d\theta},j}^m \sim q_{l_{d\theta},j}(\theta_{l_{d\theta}}|\theta_{l_{d\theta},j-1}^m, \theta_{l_{d\theta},j-1}^m, ..., \theta_{l_{d\theta},j-1}^m)$

Step-3 Compute Weights

$\omega_0^m = \frac{p(y|\theta_0^m)p(\theta_0^m)}{q(\theta_0^m)}$

3. GRID MATCHING IN MONTE CARLO BAYESIAN COMPRESSIVE SENSING

In this chapter first the grid mismatch problem is defined. Then the proposed algorithm grid matching in Monte Carlo Bayesian compressive sensing is introduced. The present version of this chapter mainly drawn from [18].

3.1. Grid Matching Problem

The research on compressive sensing is so far centered on signals with sparse representation in finite dictionaries. However, signals we encounter in the real world are usually determined by continuous parameters. In the literature, a discretization approach is adopted to reduce the continuous parameter space to a finite set of grid points. This simple strategy yields superb performance results for many problems provided that the true parameters fall into the grid set. On the other hand, it is not precise that signal fall exactly into on the supported true grid points, the so-called *basis (grid) mismatch* problem occurs. When this is the case, the true signal cannot be sparsely represented by the assumed dictionary specified by the grid points. One might suggest to solve this problem by using a finer discretization. However, using a finer discretization will also increase the coherence of the dictionary which in return would, in general, degrade the performance. Moreover, increasing the discretization level also results in higher computational complexity and numerical instability, overshadowing any advantage it might have in sparse recovery [16].

In this work, target tracking scenario has been assumed with high resolution AF radar waveforms. Note that AFs are highly peaked waveforms and hence, finer grid is needed in their reconstruction. On the other hand, increasing the discretization level would not be a wise choice due to the above-mentioned reasons (ruining incoherence, being computationally expensive, and/or causing numerical instability). In this work, we propose a novel method by using Monte Carlo approach with Bayesian CS framework.

3.2. Signal Model

A discrete complex Björck constant amplitude zero autocorrelation (CAZAC) sequence is denoted as $s_l(m),\ m = 0, 1, \ldots, M-1$ [34]. Here the sequence is expressed as $s_l(m)$ where $l = 1, 2, \ldots, L$ that denotes L different sequences with discrete delay Doppler shifts. s_1 is then symbolize the zero shift sequence. CAZAC sequences are excellent candidates for use in target tracking because of zero-autocorrelation properties which results very low delay and Doppler correlation. That means delay and Doppler correlation of a CAZAC sequence (AF of a CAZAC sequence) is almost an ideal highly peaked Dirac delta function on a fine grid.

A continuous version of complex sequence s_1 is introduced in order to illustrate the effect of grid mismatch. Using a pulse shape $w(t), t \in [0, t_w]$ where t_w is the pulse duration [34], a continuous time CAZAC is given by;

$$d(t) = \sum_{m=0}^{M-1} s_1(m) w(t - m t_w) \tag{3.1}$$

which can be transmitted as a quadrature amplitude modulated (QAM) waveform. After reflection on a target, the transmitted signal in (3.1) returns to the radar receiver having a continuous valued delay τ and Doppler ν shift. The continuous delay-Doppler shifted atom without noise is given by;

$$d_{\tau,\nu}(t) = d(t - \tau) e^{-i2\pi\nu t}. \tag{3.2}$$

In addition to this, a noiseless and a noisy version of a complex signal which composed of T elementary signals or atoms $\underline{\mathrm{s}}_{\tau_\ell,\nu_\ell}(t)$ with shifts $\{\tau_\ell, \nu_\ell\}_{\ell \in |\mathcal{T}|}$ as in (3.2), where $T = |\mathcal{T}|$ is the cardinality of the atom subset $\mathcal{T}$ (The cardinality means the number of components in the set $\mathcal{T}$), are respectively given by;

$$r_{\mathcal{T}}(t) = \sum_{\ell \in \mathcal{T}} \gamma(\ell) d_{\tau_\ell,\nu_\ell}(t) \tag{3.3}$$

$$r_{\mathcal{T}}^{\upsilon}(t) = \sum_{\ell \in \mathcal{T}} \gamma(\ell) d_{\tau_\ell, \nu_\ell}(t) + \upsilon(t) \tag{3.4}$$

where $\gamma(\ell)$ is the random atom strength which assumed to be effect of target and transmission while $\upsilon(t)$ is a zero mean additive random noise. The goal of reconstruction is to estimate (3.3) as accurately as possible. In a compressive receiver, however, only the compressively received version of the radar return signal in (3.4) is available [35, 36].

3.3. Grid Mismatch Effect on Ambiguity Function

Let the vector $d_{\tau,\nu}$ of length M represents an oversampled version of the continuous atom in (3.2) with delay-Doppler shift τ, ν. Then;

$$g_{\tau,\nu} = \mathbf{\Phi} \mathbf{d}_{\tau,\nu} \tag{3.5}$$

Where $g_{\tau,\nu}$ is modelled as the compressed C-dimensional versions of the continuous atoms obtained via a $C \times M$ partial Fourier compressive acquisition matrix (Measurement matrix Φ) [37].

In order to identify an unknown delay-Doppler signal which exists in its compressed form as $g_{\tau,\nu}$ in the compressively received signal, templates are generated using atoms with shifts $\bar{\tau}, \bar{\nu}$ as in (3.2).

$$g_{\bar{\tau},\bar{\nu}} = \mathbf{\Phi} \mathbf{d}_{\bar{\tau},\bar{\nu}} \tag{3.6}$$

The compressed AF surface will be the correlations between compressed templates $g_{\tau,\nu}$ and compressively received atoms with unknown delay-Doppler shifts $g_{\bar{\tau},\bar{\nu}}$ as defined in [38]. The compressed AF surface is;

$$\mathcal{A}_{\bar{\tau},\bar{\nu},\tau,\nu} = \frac{g^*_{\bar{\tau},\bar{\nu}} g_{\tau,\nu}}{||g_{\bar{\tau},\bar{\nu}}||_2 ||g_{\tau,\nu}||_2} \tag{3.7}$$

with $\mathcal{A}_{\bar{\tau},\bar{\nu},\tau,\nu} = 1$. When $\bar{\tau} \approx \tau$ and $\bar{\nu} \approx \nu$, the AF will be peaked and on the other hand the increase in $|\bar{\tau} - \tau|$ and $|\bar{\tau} - \tau|$ will cause a rapidly decrease in compressed AF as shown in Figure 3.1. The highly peaked AF of the CAZAC

sequences [19,38,39] and the RIP property of the acquisition matrix [8] causes that effect.

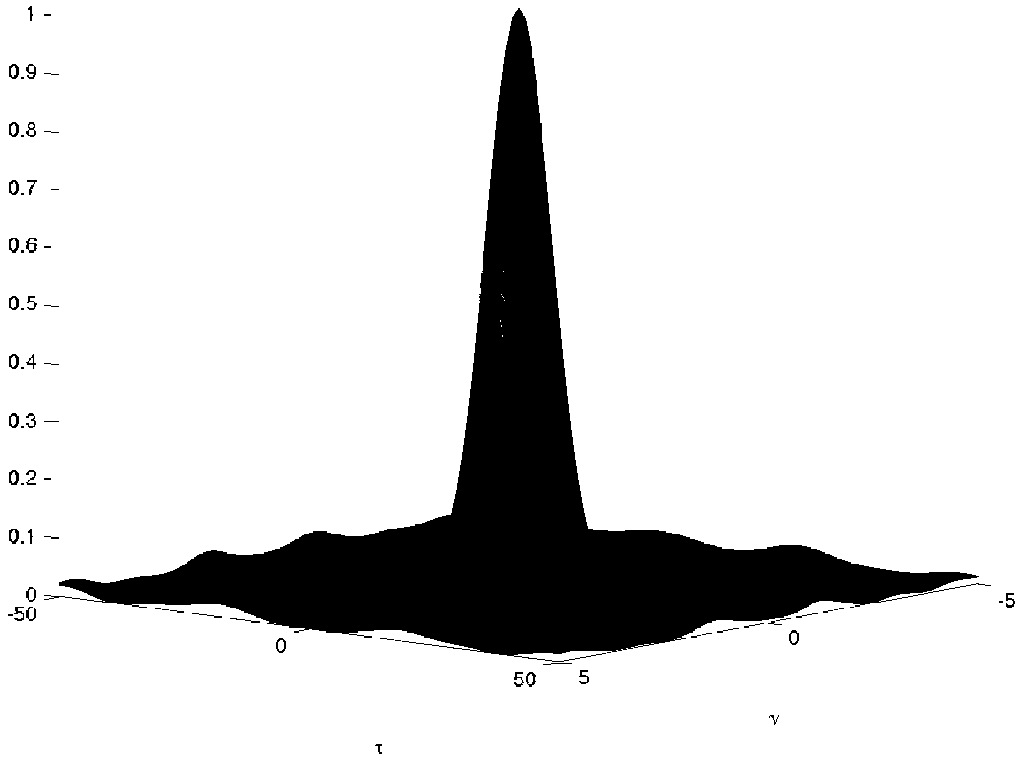


Figure 3.1: AF with no mismatch ($\epsilon_\tau = 0$ and $\epsilon_\nu = 0$).

Due to sharp peak of $\mathcal{A}_{\bar{\tau},\bar{\nu},\tau,\nu}$ as in Figure 3.1 the grid which supposed to AF being on needs to be very fine in order to represent the AF without any distortion and identify true delay-Doppler pair $\{\tau, \nu\}$. As mentioned before the sharpness of AF waveforms causes a challenging grid matching problem during reconstruction and solving this problem by defining a fine grid poses other problems that are expensive computational load and numerical instability. Solving the grid matching problem in a computationally efficient manner is the goal of this thesis. The proposed reconstruction method described in Section 3.4..

In order to solve grid mismatch problem a new grid definition approach is suggested that there will be not only one fixed grid however the new approach defines $\tilde{\mathcal{J}}$ individual grids. Now, let templates which lie on $\jmath$ th grid and defined by $\{l_\tau \delta_\tau,\ l_\nu \delta_\nu\}$ delay-Doppler points are considered. Where $\{\delta_\tau, \delta_\nu\}$ is the spacing between grids.

The grid mismatch errors in delay ϵ_τ and Doppler ϵ_ν can be defined as follow;

$$\min_{l_\tau} |\bar{\tau}_\ell - l_\tau \delta_\tau| = \epsilon_\tau \tag{3.8}$$

$$\min_{l_\nu} |\bar{\nu}_\ell - l_\nu \delta_\nu| = \epsilon_\nu \tag{3.9}$$

Where τ_ℓ and ν_ℓ is delay-Doppler pair of the ℓth atom in the radar return signal in (3.4).

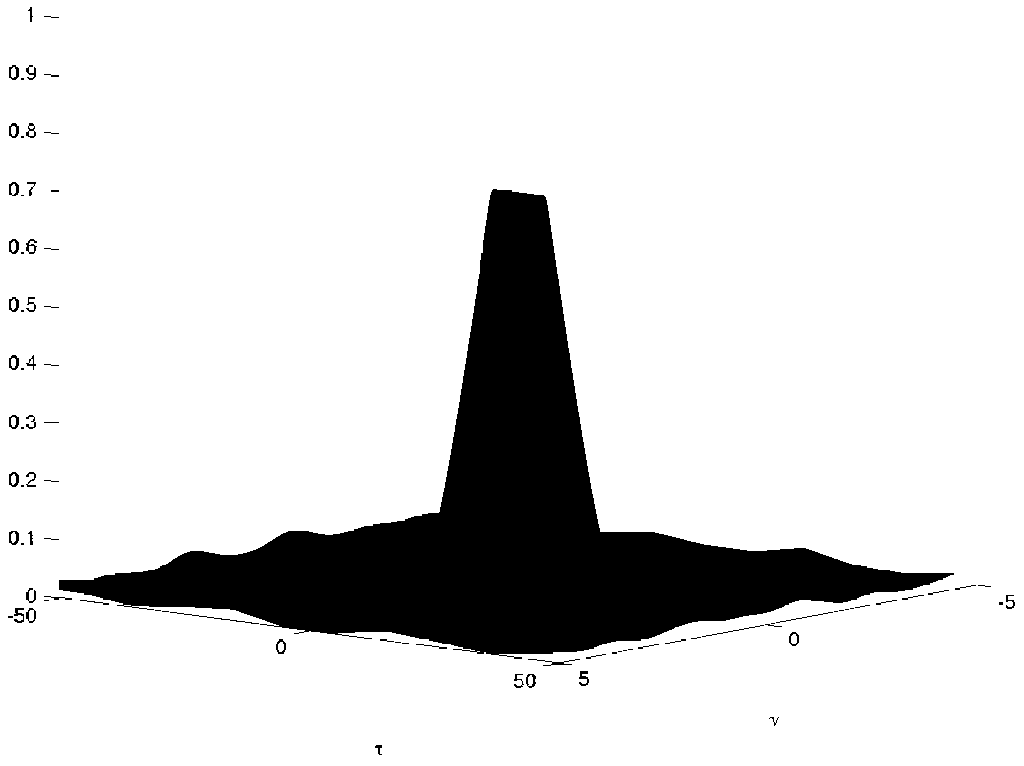


Figure 3.2: AF with minor grid mismatch ($\epsilon_\tau < 1$ and $\epsilon_\nu < 1$).

In case of a grid mismatch, the maximum value of the compressed AF will be $\mathcal{A}_{\tau,\nu,\tau+\epsilon_\tau,\nu+\epsilon_\nu} < 1$ that makes a reduction in received signal strength. The maximum grid mismatch occurs when $\epsilon_\tau = \delta_\tau/2$, $\epsilon_\nu = \delta_\nu/2$.

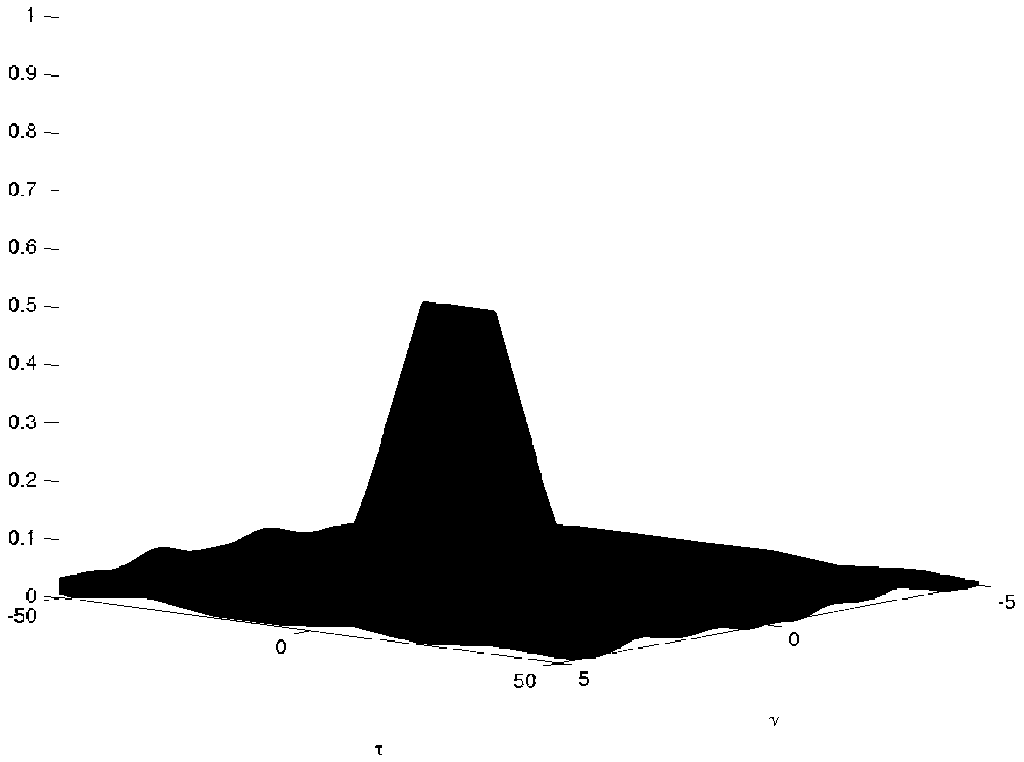


Figure 3.3: AF with major grid mismatch ($\epsilon_\tau \approx \delta_\tau/2$ and $\epsilon_\nu \approx \delta_\nu/2$).

The effect of grid mismatch on the compressed AF is illustrated in Figures 3.1, 3.2, 3.3. In Figure 3.1 perfect matching is assumed and a highly peaked AF results that $\epsilon_\tau = 0$ and $\epsilon_\nu = 0$. In Figure 3.2 the template grid is slightly offseted because the peak value is smaller then 1 that $\epsilon_\tau < 1$ and $\epsilon_\nu < 1$. In Figure 3.3 a severe effect of grid matching occurs resulting to deterioration of the AF and loss of received signal energy that $\epsilon_\tau \approx \delta_\tau/2$ and $\epsilon_\nu \approx \delta_\nu/2$.

3.4. Grid Matching Based Reconstruction

In this section the MC-BCS [17] method is extended to solve the grid matching problem. The method gives the posterior of the solution as a set of N reconstructed signals or solutions $\grave{s}_n$ and probabilities $\omega_n \quad \text{for} \quad n = 1, \ldots, N$, reflecting the belief that the signal $\grave{s}_n$ represents as accurately as possible the original noiseless signal $r_{\mathcal{T}}(t)$ in (3.3).

In this thesis, as a solution of the grid mismatch problem, a new

method is defined where instead of using a fixed fine grid, a set of J individual coarser grids are used with a multi-step sampling technique. With this approach computational expense is lightened. The $\jmath$th grid is defined by the points

$$\begin{aligned} \{\tau_{l,\jmath}, \nu_{l,\jmath}\} =& \{l_\tau \delta_\tau + \jmath f_\tau \ l_\nu \delta_\nu + \jmath f_\nu\} \\ l_\tau = 1, 2, \ldots, L_\tau, \ l_\nu =& 1, 2, \ldots, L_\nu, \ \jmath = 0, 1, \ldots, J-1 \end{aligned} \tag{3.10}$$

where δ_τ, δ_ν represent the spacing between grid points so that the incoherence between atoms in each grid is obtained [11]. $f_\tau = \delta_\tau / J, \ f_\nu = \delta_\nu / J$ symbolize the amount of spacing adjacent grids.

The vector $d_{l,\jmath}^{n,k}$ of length M is an oversampled version of the continuous atom in (3.2) shifted by a delay-Doppler $\tau_{l,\jmath}, \nu_{l,\jmath}$ supported on the $\jmath$th grid point in (3.10).

Similarly, noiseless radar return signal in (3.3) is represented as an oversampled vector with length M and defined as follows

$$r_{\mathcal{T}} = \sum_{l=1}^{|\mathcal{T}|} \gamma(l) d_{l,\jmath}^{n,k} \tag{3.11}$$

where the shifts of (3.3) in subset $\mathcal{T}$ are matched to the nearest grid points in (3.10).

The performance evaluation is performed by using $r_{\mathcal{T}}$ as a target signal for estimating. Furthermore, the compressed C-dimensional measurements are generated with projection of the signal vector onto the compressive measurement domain by $C \times M$ partial Fourier compressive measurement matrix (acquisition matrix) Φ as in [37].

$$g_{l,\jmath} = \Phi d_{l,\jmath} \tag{3.12}$$

Sampling Step-1

Atoms from all grids with the same discrete shifts are first summed as in 3.13. Then for sampling one of the $L_\tau L_\nu$ discrete delay-Doppler pairs $\{l_\tau, l_\nu\}$ in (3.10) which, for all grids $\jmath = 0, \ldots, J-1$, correspond to delay-Doppler location $\{\tau_{l,\jmath}, \nu_{l,\jmath}\}$, templates are formed as;

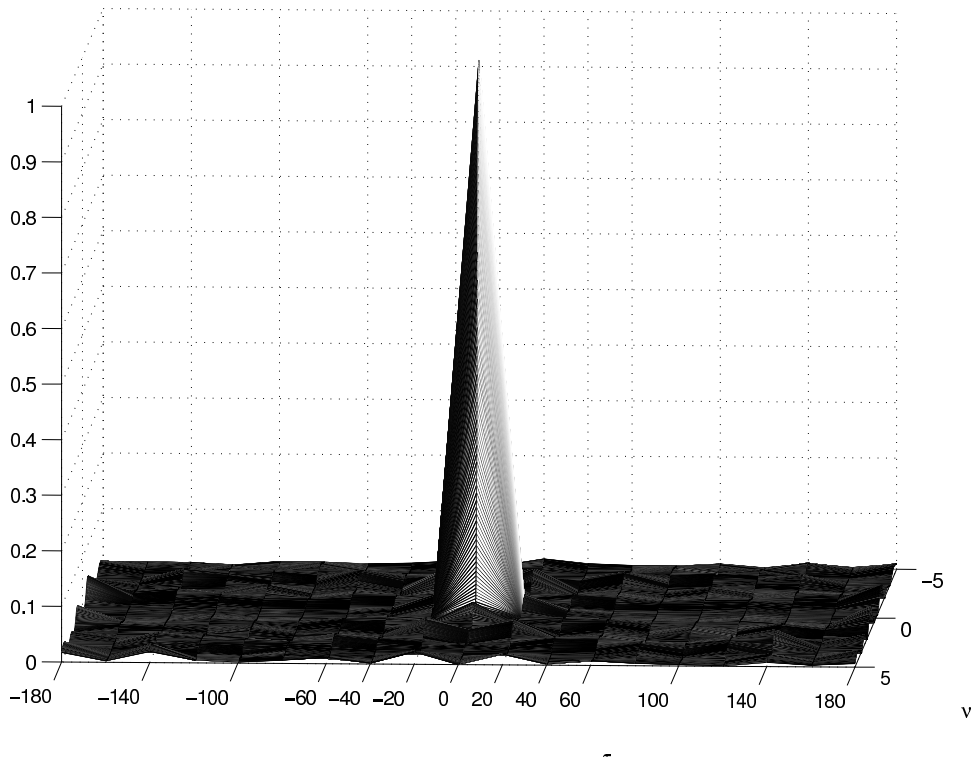


Figure 3.4: AF with grid spacing 20, no-mismatch, true atom lying on the grid at $(0, 0)$ without grid matching

$$\check{g}_l = \sum_{\jmath \in \tilde{\mathcal{J}}} g_{l,\jmath} \tag{3.13}$$

where $g_{l,\jmath}, l = 1, \ldots, L$ represents compressed atom templates in (3.12) that supported on grids $\jmath \in \tilde{\mathcal{J}}$. Set $\tilde{\mathcal{J}}$ is empirically chosen to minimize correlations of $\check{g}_l$ with the residue that yield large values for a true atom ℓ in (3.3) which lies in any grid in (3.10). Residue correlations are found as;

$$\kappa(l) = \frac{\check{g}_l^* h_{k,n}^{res}}{\check{g}_l^* \check{g}_l}, \ l = 1, \ldots, L. \tag{3.14}$$

where for kth iteration $k = 1, \ldots, T_n$ and $h_{k,n}^{res}$ represents residue here and is defined in (3.24). The likelihood ratio can be found by using (3.14) as follows

$$\Lambda(\kappa(l)) = \frac{p_1(\kappa(l)|\mathcal{T})}{p_0(\kappa(l)|\emptyset)}, \ l = 1, \ldots, L. \tag{3.15}$$

where $p_0(\kappa(l)|\emptyset)$ is the noise-only likelihood. Then, the kth atom is sampled

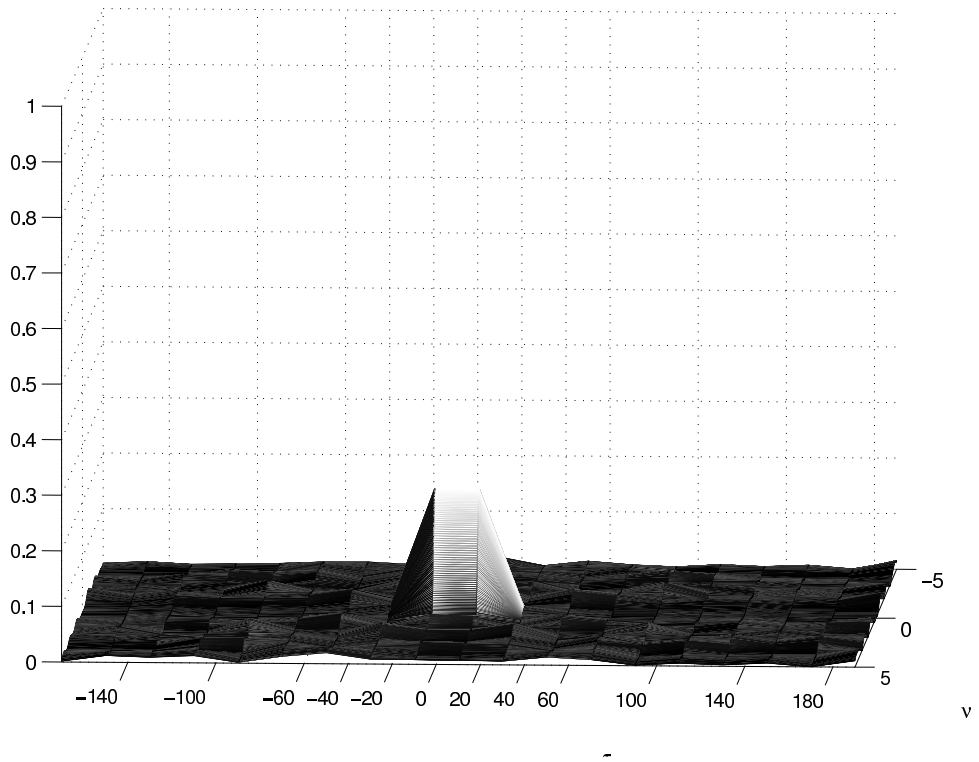


Figure 3.5: AF with grid spacing 20, grid mismatch exists, true atom lying on the grid at $(0, 0)$ without grid matching

according to the likelihood as

$$l \sim \frac{\{\Lambda(\kappa(l))\}_{l=1}^{L}}{\sum_{l'=1}^{L} \Lambda(\kappa(l'))} \tag{3.16}$$

Then the sampling bias is

$$b_{k,n}^{\kappa} = \frac{\Lambda(\kappa(l))}{\sum_{l'=1}^{L} \Lambda(\kappa(l'))}. \tag{3.17}$$

Sampling Step-2

Using compressed atom templates as in (3.13) is part of a divide and conquer strategy. This strategy lightens the algorithm by avoiding to evaluate equation (3.14) and (3.15) to not make $L_\tau L_\nu J$ times process for sampling an equal number of atoms and instead carries out the process with only $L_\tau L_\nu$ times. In addition to that, another sampling step needs to be introduced to improve the probability of correct reconstruction and decrease the MMSE, as shown next, which is carried out $3J$ times. Sampling step 2 consider that

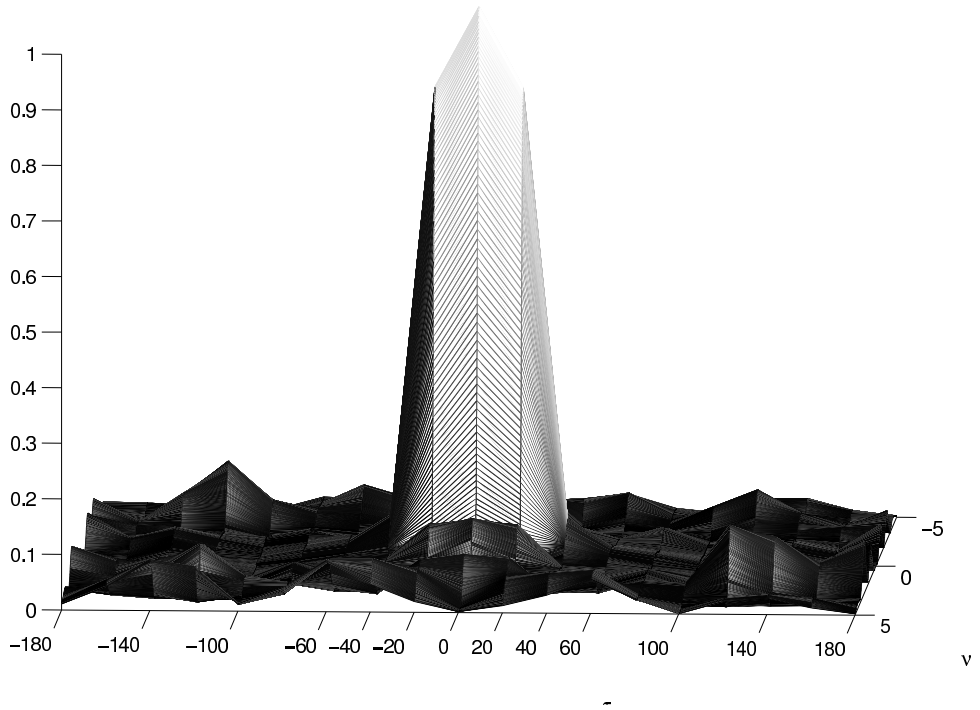


Figure 3.6: AF with grid spacing 20, no-mismatch, true atom lying on the grid at $(10, 0)$ with grid matching

$\check{g}_{l-1}, \check{g}_l$ and $\check{g}_{l+1}$ $\forall l$ are highly correlated compressed atom templates because of summing process in (3.13). That means during sampling one atom the neighbour atoms should be considered. Therefore, sampling atom l, implies that atoms $l-1$ and $l+1$ could also be the true discrete delay-Doppler shift pair.So that, atoms $l-1, l$, and $l+1$ are considered for the second sampling step and then $3J$, atom templates included into likelihood for sampling.

$$\{\{g_{l+\imath,\jmath}\}_{\imath=-1}^{1}\}_{\jmath=1}^{J} \tag{3.18}$$

The numerical based reconstruction method will simply assign a posterior distribution to each support set without seeking a single solution even though the set of compressed atom templates is not an incoherent set [11]. The information in the estimated posterior will also project the spread of the AF. In addition the SNR will also affect the amount of information contained in the compressively received signal that effect of SNR is shown in Figure 4.2.

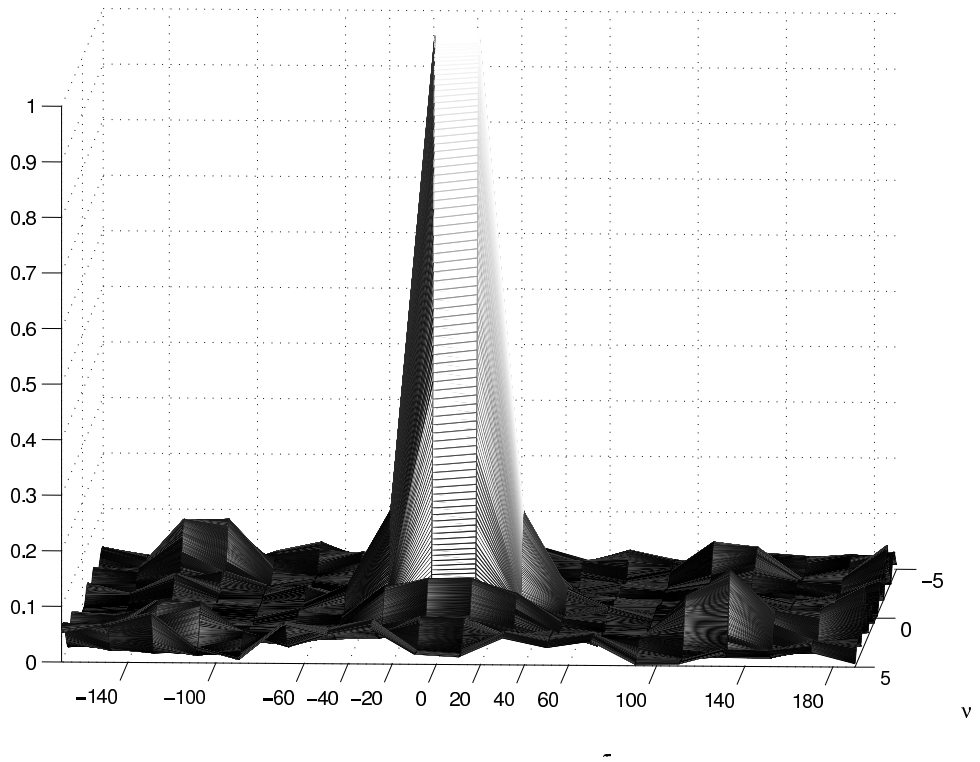


Figure 3.7: AF with grid spacing 20, grid mismatch exists, true atom lied on the grid at 10 and grid matching is applied.

The proposed method's performance for preserving signal information in off-grid components is illustrated in Figures 3.4, 3.5, 3.6, and 3.7.The set up chosen for performance analyses is as follow: for all cases grid spacing chosen as 20, in Figure 3.4, it is assumed that no mismatch occurs that means true atom is lied on the grid at origin 0 and grid matching procedure is not applied, in Figure 3.5; it is assumed that mismatch exists that means true atom is not lied on the grid at origin 0 instead lied at 10 and again grid matching procedure is not applied, in Figure 3.6; it is assumed that no mismatch occurs that means true atom is lied on the grid at origin 0 and grid matching procedure is applied, in Figure 3.7; it is assumed that mismatch exists that means true atom is not lied on the grid at origin 0 instead lied at 10 and grid matching procedure is applied.

The effect of mismatch can be seen clearly from Figures 3.4 and 3.5 that because of the mismatch in Figure 3.5 the highly peak of AF is missed

due to losing the true atom energy. However by applying MCBCS with grid matching procedure atom energy is preserved and distortion in AF is prevented as in Figure 3.6. Even if grid mismatch exists, the atom energy is preserved and a highly peak AF is observed as Figure 3.7.

After choosing neighbour templates next step is finding correlations with the residue for each template as

$$a(\imath, \jmath) = \frac{g^*_{l+\imath,\jmath} h^{res}_{k,n}}{g^*_{l+\imath,\jmath} g_{l+\imath,\jmath}} \tag{3.19}$$

where $\imath = -1, 0, 1$, $\jmath = 1, 2, \ldots, J$ and $h^{res}_{k,n}$ represents residue here and is defined in (3.24). The likelihood ratio is;

$$\Lambda(a(\imath, \jmath)) = \frac{p_1(a(\imath, \jmath)|\mathcal{T})}{p_0(a(\imath, \jmath))|\emptyset)} \tag{3.20}$$

where $\imath = -1, 0, 1$, $\jmath = 1, 2, \ldots, J$ and $p_0(a(\imath, \jmath))|\emptyset)$ is the noise-only likelihood. Then, the kth index pair is sampled as

$$\{\imath, \jmath\}_{k,n} \sim \frac{\{\{\Lambda(a(\imath, \jmath))p(\imath)p(\jmath)\}^1_{\imath=-1}\}^J_{\jmath=1}}{\sum_{\jmath=1}^{J} \sum_{\imath=-1}^{1} \Lambda(a(\imath, \jmath))p(\imath)p(\jmath)} \tag{3.21}$$

where the prior distribution on the grid $p(\jmath)$ and prior distribution of atom $p(l)$ was considered.

The magnitude of k th atom $\hat{\alpha}_n(k)$ is calculated by associating sampled index in (3.21) as;

$$\hat{\alpha}_n(k) = a(\imath, \jmath) \tag{3.22}$$

The normalized sampling bias is computed as;

$$b^a_{k,n} = \frac{\Lambda(\hat{a}_n(k))p(\imath)p(\jmath)}{\sum_{\imath'=-1}^{1} \sum_{\jmath'=1}^{J} \Lambda(a(\imath', \jmath'))p(\imath')p(\jmath')} \tag{3.23}$$

The bias terms (3.17) and (3.23) are used in the weighting the reconstructed signals in (3.27). In order to sample next index pair and atom index k of solution n is incremented by 1 and the residue is updated by removing sampled atom from residue as

$$h_{k,n}^{res} = h_{k-1,n}^{res} - \hat{\alpha}_n(k-1)g_{l,\jmath}^{k,n}, \quad h_{1,n}^{res} = h_\jmath \tag{3.24}$$

Two stop criteria is defined for the above process. The process is halted and $T_n = k$ if the change in residue between two iteration is smaller than a threshold as $\Delta_{res} = |||h_{k,n}^{res}||_2^2 - ||h_{k-1,n}^{res}||_2^2| < \Theta_{res}$ or $k = T_{max}$ where $T_{max} << M$ guarantees sparsity. The appropriate stopping criterion selection may be a future work research problem.

Table 3.1: The Grid Matching MC-BCS Algorithm

Algorithm 3.4..1: (*The Grid Matching MC − BCS Algorithm*)

∘ For each solution $n = 1, \ldots, N$

♢ Let $h_{1,n}^{res} = h_j$ in (3.24)

♢ For $k = 1, \ldots, T_{max}$ until $\Delta_{res} < \Theta_{res}$

Sampling step 1:

- Form templates $\check{g}_l\,,\ l = 1, \ldots, L_\tau L_\nu$ as in (3.13)
- Calculate correlations $\kappa(l)$ in (3.14)
- Calculate likelihood ratios $\Lambda(\kappa(l))$ in (3.15)
- Sample index l in (3.16)
- Calculate bias $b_{k,n}^{\kappa}$ (3.17)

Sampling step 2:

- Identify templates $\{\{g_{l+\imath,\jmath}\}_{\imath=-1}^{1}\}_{\jmath=1}^{J}$ in (3.18)
- Calculate correlations $a(\imath, \jmath)$ (3.19)
- Calculate likelihood ratios $\Lambda(a(\imath, \jmath))$ (3.20)
- Sample index pair $\{\imath, \jmath\}_{k,n}$ (3.21)
- Set $\hat{\alpha}_n(k) = a(\imath, \jmath)$ with $\{\imath, \jmath\}_{k,n}$
- Calculate bias $b_{k,n}^{a}$ (3.23)
- Update residue $h_{k,n}^{res}$ (3.24)
- Halt if $\Delta_{res} < \Theta_{res}$ or $k = T_{max}$ and set $T_n = k$

♢ Form solutions $\grave{s}_n$ (3.25)

♢ Calculate weights ω_n using (3.27) and normalization

♢ Calculate $\hat{s}_{MMSE}$ (3.29)

3.4.1. Reconstructed signal posterior, point estimate, and reconstruction error

The individual sampled atoms, multi-element reconstructed signals are given as

$$\grave{s}_n = \sum_{k=1}^{T_n} \hat{\alpha}_n(k) d_{l,\jmath}^{n,k} \tag{3.25}$$

for each n with $\hat{\alpha}_n(k)$ and vectors $d_{l,\jmath}^{n,k}$ with an index pair $\{\imath, \jmath\}_{k,n}$ from (3.21).

$$\grave{g}_n = \Phi \grave{s}_n \tag{3.26}$$

Compressed sparse solutions are also formed for each n as in 3.26. The unnormalized weights of the reconstructed signals are given by;

$$\tilde{\omega}_n = \frac{\Lambda(\{\hat{\alpha}_n(k)\}_{k=1}^{T_n})}{\prod_{k=1}^{T_n} b_{k,n}^{\alpha} b_{k,n}^{\kappa}},\ \Lambda(\{\hat{\alpha}_n(k)\}_{k=1}^{T_n}) = \frac{p_1(\beta_n|\mathcal{T})}{p_0(\beta_n|\emptyset)} \tag{3.27}$$

where $\beta_n = \grave{g}_n^* h_\jmath$. The weights take into account the likelihood ratio $\Lambda(\{\hat{\alpha}_n(k)\}_{k=1}^{T_n})$ and the biases $b_{k,n}^{\kappa}$ in (3.17) and $b_{k,n}^{a}$ in (3.23) from the two sampling steps are normalized as;

$$\omega_n = \frac{\tilde{\omega}_n}{\sum_{n=1}^{N} \tilde{\omega}_n} \tag{3.28}$$

After that, by using a minimum mean squared error (MMSE) estimation a point estimate can be obtained as;

$$\hat{s}_{MMSE} = \sum_{n=1}^{N} \omega_n \grave{s}_n. \tag{3.29}$$

The MSE in reconstruction w.r.t. the noiseless signal in (3.3) is given by;

$$\mathcal{E} = \sum_{n=1}^{N} \omega_n \frac{||\grave{s}_n - r||_2^2}{||\grave{s}_n||_2 ||r||_2} \tag{3.30}$$

The algorithm scheme is given in Table 3.1.

4. RESULTS

4.1. Numerical Setup

In this thesis the compressive measurement or acquisition matrix Φ with $C \times M$ size is constructed which satisfies the properties mentioned in Section 2.2.2.. RIP and mutual coherence properties provide to preserve the structure of any sparse M-dimensional vector after a dimensionality reduction due to $C < M$. Because of this measurement matrix is chosen as a known partial Fourier [37] matrix with orthogonal rows such that;

$$\Phi\Phi^* = \frac{M}{C} I_{C\times C} \tag{4.1}$$

where $\frac{M}{C}$ compensates for the dimensionality reduction of any signal that transformed from signal domain to compressive measurement domain by projecting the atoms on Φ from M to $C$and C represents number of compressive measurements that is chosen as $500 : 500 : 4000$. The scaling factor $\frac{M}{C}$, however, amplifies the noise variance [9, 40].

A length 10 rectangular pulse was generated to simulate the pulse shape $w(t), t \in [0, t_w]$. The transmitted signal in (3.1) is obtained by combining rectangular pulse $w(t)$ with a CAZAC sequence of prime length 499. Then 2500 delay and Doppler shifted atoms were generated from (3.1) on a very fine grid. The total length of the oversampled transmitted signal which emulates the continuous signal was, considering a maximum delay set to $M = 5090$. The received signal contained 10 delay-Doppler shifted waveforms selected randomly from the pool of 2500 atoms. Atoms had a zero mean complex Gaussian random strength and a zero mean complex Gaussian noise was added to atoms. Both the received signal and templates were projected on the acquisition matrix and the first step of the algorithm was run with second step that is grid selection as described in Section 3.4.. The MC-BCS algorithm was also run with no grid selection in order to investigate the effectiveness of the grid matching method proposed.

4.2. Simulation Results

Figure 4.1 shows the effect of grid selection on the MC-BCS, both grid selection added algorithm and MC-BCS with no grid selection for 20dB SNR and a different number of compressive measurements. It is observed that the MC-BCS with grid selection reconstruction performance improved with an increase in compressive measurements C. The grid selection improved the performance of the reconstruction with significant amount.

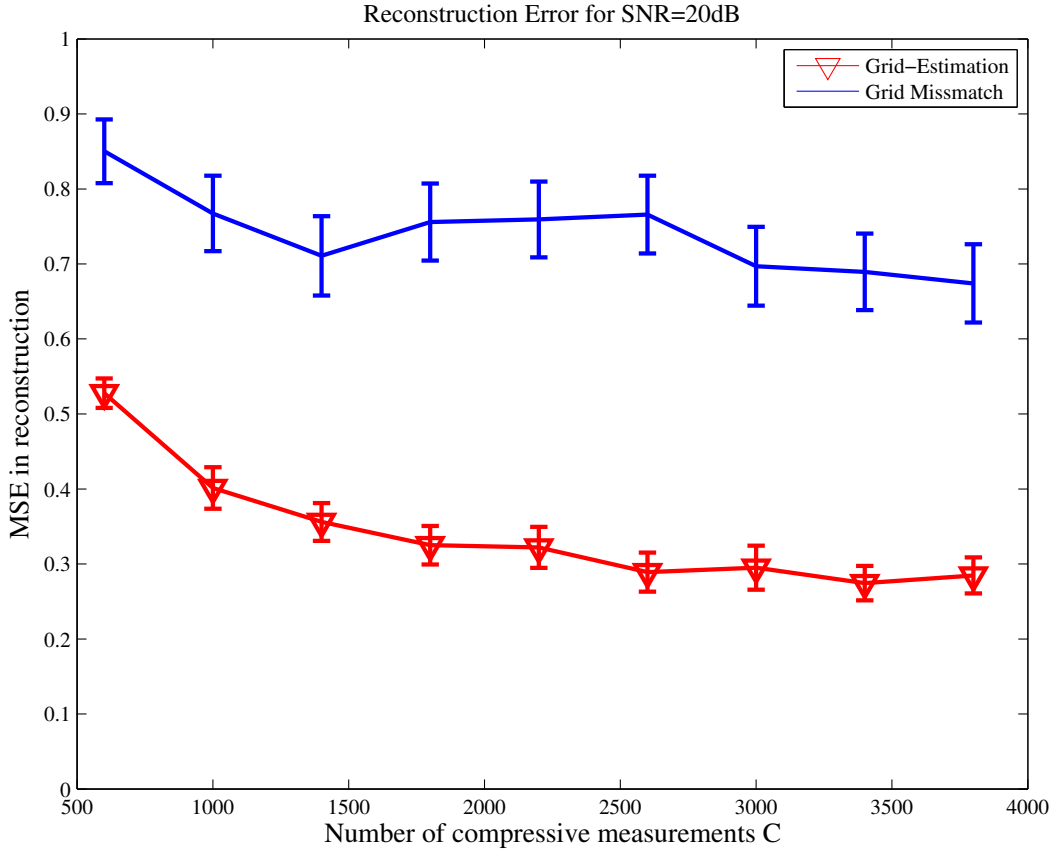


Figure 4.1: MSE reconstruction performance versus the number of compressive measurements for MC-BCS with grid selection and MC-BCS with no grid selection for 20dB SNR.

In Figure 4.2, the performance of the algorithm is observed under $16, 18$ and 20dB SNR values and plotted. As expected the algorithm performance is improved when the SNR increases, even for low SNRs satisfactory performance is observed.

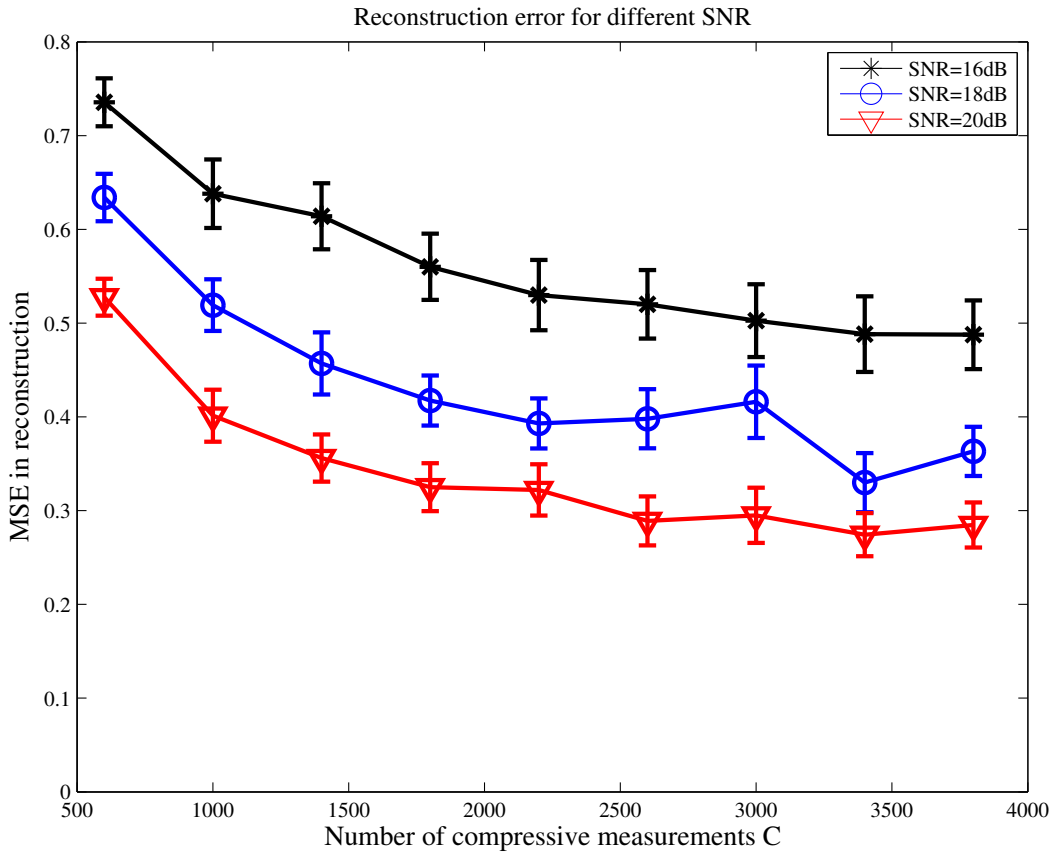


Figure 4.2: MSE reconstruction performance versus the number of compressive measurements for MC-BCS with grid selection for 16, 18, and 20dB SNR.

5. CONCLUSION

In this thesis the grid mismatch problem in compressively sensed radar waveforms with high resolution properties was studied.

In chapter 3 the Monte Carlo Bayesian compressive sensing method (MC-BCS) was extended as multistage MC-BCS to overcome grid mismatch problem. The first step of the multistage MC-BCS is to identify the possible grid configuration in a pre-defined J different grids by using likelihood ratio between grids and residue. Secondly, the multistage MC-BCS recovers sparse signal atoms which lies on a chosen grid and accurately represent the received radar signal. The method avoids computational load of defining one dense grid which might be another solution method for handling the mismatch problem. The performance analysis of the proposed method in sparse reconstruction of high resolution AF that is compressively received in noise was done via a numerical based study in two cases. First, MC-BCS and the multistage MC-BCS were compared to see the grid estimation effect on MC-BCS in 20dB noise. Then the proposed method was tested for different noise levels including 16dB, 18dB and 20dB. The proposed method's performance is improved against MC-BCS as expected, grid estimation made a major contribution to MC-BCS. The proposed method works efficiently even for low SNRs.

In this thesis the MC-BCS method is extended by adding grid estimation stage, as a future work a new approach for grid estimation can be developed which will seek to minimize the reconstruction error while minimizing the computational expense.

REFERENCES

[1] M. F. Duarte and M. A. Davenport. Introduction to vector spaces. In *Connexions*, URL: http://cnx.org/content/m37167/latest/, 2012.

[2] R.G. Baraniuk. Compressive sensing [lecture notes]. *Signal Processing Magazine, IEEE*, 24(4):118–121, 2007.

[3] G. Kutyniok. Theory and applications of compressed sensing. *GAMM-Mitteilungen*, 36:79–101, 2013.

[4] J. Romberg E. Candes and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory*, 52:489–509, 2006.

[5] D.L. Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, 52(4):1289–1306, 2006.

[6] J.A. Tropp and S.J. Wright. Computational methods for sparse solution of linear inverse problems. *Proceedings of the IEEE*, 98(6):948–958, 2010.

[7] J.A. Tropp and A.C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *Information Theory, IEEE Transactions on*, 53(12):4655–4666, 2007.

[8] E.J. Candes and M.B. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, 2008.

[9] M.A. Davenport, P.T. Boufounos, M.B. Wakin, and R.G. Baraniuk. Signal processing with compressive measurements. *Selected Topics in Signal Processing, IEEE Journal of*, 4(2):445–460, 2010.

[10] Shihao Ji, Ya Xue, and L. Carin. Bayesian compressive sensing. *Signal Processing, IEEE Transactions on*, 56(6):2346–2356, 2008.

[11] Yuejie Chi, L.L. Scharf, A. Pezeshki, and A.R. Calderbank. Sensitivity to basis mismatch in compressed sensing. *Signal Processing, IEEE Transactions on*, 59(5):2182–2195, 2011.

[12] R. Jagannath, G. Leus, and R. Pribic. Grid matching for sparse signal recovery in compressive sensing. In *Radar Conference (EuRAD), 2012 9th European*, pages 111–114, 2012.

[13] C. Ekanadham, D. Tranchina, and E.P. Simoncelli. Recovery of sparse

translation-invariant signals with continuous basis pursuit. *Signal Processing, IEEE Transactions on*, 59(10):4735–4744, 2011.

[14] I. Kyriakides, I. Konstantinidis, D. Morrell, J.J. Benedetto, and A. Papandreou-Suppappola. Target tracking using particle filtering and cazac sequences. In *Waveform Diversity and Design Conference, 2007. International*, pages 367–371, 2007.

[15] A. Kebo, I. Konstantinidis, J.J. Benedetto, M.R. Dellomo, and J.M. Sieracki. Ambiguity and sidelobe behavior of cazac coded waveforms. In *Radar Conference, 2007 IEEE*, pages 99–103, 2007.

[16] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht. Compressed sensing off the grid. *CoRR*, abs/1207.6053, 2012.

[17] Pribic R. Kyriakides, I. Bayesian compressive sensing using monte carlo methods. *(Submitted for review)*, 2013.

[18] I. Kyriakides, R. Pribic, H. Sar, and N. At. Grid matching in monte carlo bayesian compressive sensing. In *Information Fusion (FUSION), 2013 16th International Conference on*, pages 2103–2109, 2013.

[19] I. Kyriakides, I. Konstantinidis, D. Morrell, J.J. Benedetto, and A. Papandreou-Suppappola. Target tracking using particle filtering and cazac sequences. In *Waveform Diversity and Design Conference, 2007. International*, pages 367–371, 2007.

[20] Eli Mozeson Nadav Levanon. *Radar Signals*. Wiley-IEEE Press, Boston, MA, 2004.

[21] S.Z. Budisin and P. Spasojevic. Bjorck sequence sets. *Electronics Letters*, 47(8):491–493, 2011.

[22] T. M. Apostol. *Introduction to Analytic Number Theory 'Undergraduate Texts in Mathematics'*. Springer, 1976.

[23] D. L. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory*, 2001.

[24] M. Elad. Optimized projections for compressed sensing. *Signal Processing, IEEE Transactions on*, 55(12):5695–5702, 2007.

[25] S. Muthukrishnan. *Data Streams: Algorithms and Applications*. Now

Publisher, Boston, MA, 2005.

[26] D. L. Chen, S. S.and Donoho, M. Saunders, and A. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20:33–61, 1998.

[27] Y.C. Pati, R. Rezaiifar, and P. S. Krishnaprasad. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*, pages 40–44 vol.1, 1993.

[28] S.D. Babacan, R. Molina, and A.K. Katsaggelos. Bayesian compressive sensing using laplace priors. *Image Processing, IEEE Transactions on*, 19(1):53–63, 2010.

[29] Pribic R. Kyriakides, I. Bayesian compressive sensing in radar systems. In *Proceedings of the International Workshop on Compressed Sensing applied to Radar (CoSeRa 2013).*, Bonn, Germany, 2013.

[30] D. G. Horvitz and D. J. Thompson. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(260):pp. 663–685, 1952.

[31] O. Capp, A. Guillin, J. M. Marin, and C. P. Robert. Population monte carlo. *Journal of Computational and Graphical Statistics*, 13(4):pp. 907–929, 2004.

[32] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-6(6):721–741, 1984.

[33] Shen B. Djuric, P.M. and M. F. Bugallo. Population monte carlo methodology a la gibbs sampling. In *19th European Signal Processing Conference (EUSIPCO 2011)*, Barcelona, Spain, 2011.

[34] A. Kebo, I. Konstantinidis, J.J. Benedetto, M.R. Dellomo, and J.M. Sieracki. Ambiguity and sidelobe behavior of cazac coded waveforms. In *Radar Conference, 2007 IEEE*, pages 99–103, 2007.

[35] M.A. Davenport, S.R. Schnelle, J. P. Slavinsky, R.G. Baraniuk, M.B. Wakin, and P.T. Boufounos. A wideband compressive radio receiver.

In *MILITARY COMMUNICATIONS CONFERENCE, 2010 - MILCOM 2010*, pages 1193–1198, 2010.

[36] M. Mishali, Y. C. Eldar, and J. A. Tropp. Efficient Sampling of Sparse Wideband Analog Signals. *ArXiv e-prints*, 2009.

[37] E.J. Candes and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *Information Theory, IEEE Transactions on*, 52(12):5406–5425, 2006.

[38] I. Kyriakides. Ambiguity functions of compressively sensed and processed radar waveforms. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 4256–4259, 2011.

[39] Donatelli J.and Konstantinidis I. Benedetto, J. J. and C. Shaw. A doppler statistic for zero autocorrelation waveforms.

[40] I. Kyriakides. Adaptive compressive sensing and processing for radar tracking. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 3888–3891, 2011.