THE CONCEPT OF SINGLE PIXEL FIBER OPTICAL IMAGING VIA COMPRESSED SENSING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

AHMET AZGIN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MICRO AND NANOTECHNOLOGY

SEPTEMBER 2022

Approval of the thesis:

THE CONCEPT OF SINGLE PIXEL FIBER OPTICAL SPECKLE IMAGING VIA COMPRESSED SENSING

submitted by **AHMET AZGIN** in partial fulfillment of the requirements for the degree of **Master of Science in Micro and Nanotechnology, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Deniz Üner
Head of the Department, **Micro and Nanotechnology** _____

Assoc. Prof. Emre YÜCE
Supervisor, **Physics, METU** _____

Examining Committee Members:

Prof. Dr. Raşit Turan
Physics, METU _____

Assoc. Prof. Emre Yüce
Physics, METU _____

Asst. Prof. Ihor Pavlov
Physics, METU _____

Assoc. Prof. Onur Ferhanoğlu
Electronics and Communications Eng., ITU _____

Assoc. Prof. Serdar Kocaman
Electrical - Electronics Eng., METU _____

Date: 01.09.2022

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name Last name : Ahmet Azgın

Signature :

ABSTRACT

CONCEPT OF SINGLE PIXEL FIBER OPTICAL IMAGING VIA COMPRESSED SENSING

Azgın, Ahmet
Master of Science, Micro and Nanotechnology
Supervisor: Assoc. Prof. Emre Yüce

September 2022, 65 pages

The field of Compressed Sensing is nearly two decades old, hence a relatively new signal reconstruction routine, pioneered by Donoho, Candes, Baraniuk, Romberg, and Tao (E. Candes et al., 2004, 2005; E. J. Candes & Wakin, 2008; E. Candès & Romberg, 2005b, 2005a; E. Candes & Tao, 2004; D. L. Donoho, 2004; Romberg, 2007) . It exploits the sparsity property of the real-world signals to allow reconstruction at even sub-Nyquist measurements. However, there are still several key properties for a signal to be fully reconstructed, namely Restricted Isometry Property and Incoherence. The aim of this study is to investigate the modality of fiber optical imaging via a compressed sensing paradigm to check the multiple Lorentzian (Cauchy) distributed nature of the fiber optical speckles are able to give rise to the reconstruction of a 2D scene illuminated and doing it so with multimode optical fibers (MMFs) and single-pixel-imaging (SPI) approach.

Keywords: Fiber Imaging, Compressed Sensing, Compressive Imaging, Single Pixel Imaging

ÖZ

SIKIŞTIRMALI ALGILAMA İLE TEK PİKSEL FİBER OPTİK GÖRÜNTÜLEME

Azgın, Ahmet
Yüksek Lisans, Mikro ve Nanoteknoloji
Tez Yöneticisi: Doç. Dr. Emre Yüce

Eylül 2022, 65 sayfa

Yaklaşık yirmi yaşına giren sıkıştırmalı algılama; Donoho, Candes, Romberg ve Tao'nun öncülüğündeki, görece yeni bir sinyal yeniden oluşturma rutinidir (E. Candes et al., 2004, 2005; E. J. Candes & Wakin, 2008; E. Candès & Romberg, 2005b, 2005a; E. Candes & Tao, 2004; D. L. Donoho, 2004; Romberg, 2007). Gerçek sinyallerin seyreklik özelliğinden yararlanılarak, Nyquist-Shannon örnekleme sınırının altında ölçüm ile tekrar oluşturulabilmeleri mümkün kılınmaktadır. Fakat, bu yeniden oluşturma için sinyalin sahip olması gereken özellikler tanımlanmıştır, ve bunlar sınırlı izometri özelliği ve uyumsuzluk özellikleridir. Bu çalışmanın amacı, çoklu Lorentzian (Cauchy) dağılıma sahip olan çok modlu fiber optik beneklerin ışıklandırdığı iki boyutlu alanın sıkıştırmalı algılama kullanarak objenin görüntüsünün oluşturulmasının sağlanabileceğinin araştırılmasıdır.

Anahtar Kelimeler: Fiber Görüntüleme, Sıkıştırmalı Algılama, Sıkıştırmalı Örnekleme, Tek Piksel Görüntüleme

to my family

ACKNOWLEDGMENTS

It has been an interesting journey.

First and foremost, I would like to express wholehearted gratitude and appreciation to my advisor, Assoc. Prof. Emre Yüce whom I was extremely fortunate to cross paths. Undoubtedly, his vision and creativity shed light on this long, hard road many times. His insightful feedbacks, enthusiasm and encouragement have been invaluable not only on academia, but on various topics in my life.

I would like to thank a fellow group member Şahin Kürekçi for his knowledge, support, help, and humor.

I have been blessed by the company of a number of valuable friends. My special thanks to, Ayşen Elşen Aydın, Merve Demir, Elif Ünsal, Enfal Sartaş for their amazing friendship, fruitful conversations. They never cease to challenge and inspire me for the better.

I believe it would not be fair to mention everyone that have altered my course positively throughout these years, because eventually I would end up forgetting someone. Nevertheless, I would like to also cherish the opportunity to thank Mustafa Mert Yüksekdağ, Mehmet Boğatay Akata, Feriha Sartaş, Caner Tuz, İbrahim Burak Öztürk, Samet Böcü, Şükrü Aydın, Mert Elmas, Avni Burak Yenice, Seçil Güler, Onur Uçanok and many more.

Finally, I would like to thank my parents, Hatice and Yaşar for their endless support, caring and encouragement. I would surely not be who and where I am today without them. My sister, Şerife for aiding me to maintain joy in my life, and given me an awesome niece, Eylül.

We would like to acknowledge the financial support of TÜBA-GEBİP project.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

ABBREVIATIONS

1D	One dimensional
2D	Two Dimensional
3D	Three Dimensional
BP	Basis Pursuit
BS	Basis Scan
BPDN	Basis Pursuit Denoising
CCD	Charge Coupled Device
CMOS	Complementary Metal Oxide Semiconductor
CoSaMP	Compressive Sampling Matching Pursuit
CS	Compressive Sensing
DCT	Discrete Cosine Sampling
DMD	Digital Micromirror Device
DWT	Discrete Wavelet Transform
FPA	Focal Plane Array
IHT	Iterative Hard Thresholding
IR	Infrared
ISTA	Iterative Soft Thresholding Algorithm
LASSO	Least Absolute Shrinkage and Selection Operator
LIDAR	Light Detection and Ranging
MMF	Multi-mode Fiber

MP	Matching Pursuit
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
NP	Nondeterministic Polynomial Time
OMP	Orthogonal Matching Pursuit
PSNR	Peak Signal-to-Noise Ratio
RADAR	Radio Detection and Ranging
RGB	Red Green Blue
RIC	Restricted Isometry Constant
RIP	Restricted Isometry Property
RMSE	Root Mean Squared Error
ROMP	Regularized Orthogonal Matching Pursuit
RS	Raster Scan
SLM	Spatial Light Modulator
SNR	Signal-to-Noise Ratio
SPC	Single-Pixel Camera
SPD	Single-Pixel
SPI	Single-Pixel Imaging
SSIM	Structural Similarity Index
StOMP	Stagewise Orthogonal Matching Pursuit

LIST OF SYMBOLS

SYMBOLS

$\mathbb{R}^N$	The N-dimensional real vector space
c	A scalar
$\boldsymbol{x}$	A vector
x_i	i^{th} entry of vector **x**
A	A matrix
A_i	i^{th} row vector of **A**
$A_{i,j}$	i^{th} row, j^{th} column entry of matrix **A**
A^T	Transpose of matrix **A**
A^{-1}	Inverse of matrix **A**
$\|\| \cdot \|\|_{\mathrm{p}}$	p-norm notation for $1 \leq p < \infty$
$\|\| \cdot \|\|_0$	ℓ_0-norm notation for $1 \leq p < \infty$
$< \boldsymbol{u}, \boldsymbol{w} >$	Inner product of vectors $\boldsymbol{u}$ and $\boldsymbol{w}$
$\boldsymbol{x}^{[i]}$	The k^{th} iteration of the vector $\boldsymbol{x}$
$\boldsymbol{x}_K$	Sparsity of vector $\boldsymbol{x}$
${}^{H,S}T$	Hard and soft thresholding operator

CHAPTER 1

INTRODUCTION

Fiber optical imaging has been a relatively recent area that attracts more interest for the couple of decades. Whereas the bundles of single mode fibers were able to be shown to be used for imaging, multimode fibers imaging possibilities have been investigated by the works of Yariv and Gover as early as 1978 (Gover et al., 1976; Yariv, 1976) . Lately, as the technology matures with the industry driven demand, mainly information sector, fiber-optic costs are hitting an all-time-low. Multimode optical fibers are appealing more than ever since they are able to transport much information in a relatively very small cross section area considerably fast, and moreover, they are scalable. As a result, there is an increasing grow of interest for fiber optical imaging, particularly for biomedical purposes like in vivo imaging, or imaging of such areas that conventional cameras would not reach.

The conventional method for imaging involves an optical system to map the light intensity at designated parts of the imaged scene onto a detector array correspondingly. As opposed to standard Focal Plane Arrays (FPAs), Charge Coupled Devices (CCD) and Complementary Metal Oxide Semiconductor (CMOS) cameras. which are comprised of thousands to millions of pixels, and every pixel is deticated to detect the light intensity of the specific region in the scene, Single Pixel Imaging (SPI) systems, as its name suggests, promote just a single detector and take measurements of the scene in a sequential manner. Although the CCDs, CMOSs and FPAs are practically unmatched by their cost, speed and resolution, due to their mass production availibilities and the fortunate coincidence that the visible range in the electromagnetic spectrum and the second most abundant element in Earth's crust happen to spectrally match, unfortunately this advantage is mostly lost outside of this range or for non-standard purposes. The promises of SPI becomes prominent, as

it is evidently easier to have just one sensor of high quality and manage it, rather than getting each of millions of sensors from CCDs and FPAs.

There are two main methods of practice for SPI systems, namely Raster Scan (RS) and Basis Scan (BS). Compressed Sensing (CS) may be considered an extension of the BS. BS utilizes a single pixel detector to take full measurement of the scene successively by using vectors from a known basis whereas CS does the same, just with lesser measurements. It would not be exaggaration if one is to say Compressed Sensing is a fairly new signal processing paradigm that considerably revoulitionized the field. It proposes that given a signal satisfies several properties, namely Sparsity and Incoherence Property (E. J. Candes & Wakin, 2008), one can reconstruct the initial signal through finding an unique solution to an underdetermined system.

1.1 Applications

Since its publication, CS has revolutionized and attracted variety of fields. The CS paradigm already found applications in areas such as signal processing, statistics, medical imaging, remote sensing, computer science and electrical engineering, to name a few. Real world problems often can be stated approximately as sparse recovery problems in an appropriate domain. This combines well with the CS which have a great potential of applications upon considering its premise of recovery from incomplete information. In short, if a problem involves signals, CS most probably has a potential application in that area. CS is almost two decades old. Although CS has been trialed in many application areas, the bridging between hardware and software is still in its early stages.

1.1.1 RADAR

Conventional detection systems generally compose of two stages. Transmission phase where a pulse is emitted, then the receiving phase that makes use of a filter to correlate the scattered signal with that pulse. Two stages can appear interchangeably

at the receiving phase, pulse compression and digitization. The analog signal is processed by using high-rate analog-to-digital (A/D) converters to digitize the analog signal (Richards et al., 2010).

After the debut of CS, several papers have recognized its potential in RADARs. Baraniuk suggested that the CS framework could benefit the conventional RADAR scheme, by eliminating the need for match filter (Baraniuk & Steeghs, 2007). Moreover, Herman proposed several articles on the topic, and suggested discretizing the time-frequency plane into a grid and showing it as a matrix where small number of targets make it sparse in order to reconstruct with higher resolution (M. A. Herman & Strohmer, 2009; M. Herman & Strohmer, 2008). Furthermore, the CS RADAR scheme will be unaffected by the same uncertainty principle that classical RADAR schemes have, which caps their performances (Yoon & Amin, 2008). Several more interesting papers came out on this topic to this day (Potter et al., 2010) (Harding & Milla, 2013) and the author advises interested readers to the review by Yang (Yang et al., 2019).

1.1.2 LIDAR

Light Detection and Ranging systems (LIDAR) is relatively a hot topic for the last couple of decades, as one of the byproducts of the emerging autonomous technologies. The working scheme is fairly simple, and maybe somewhat similar to the RADAR. A laser source emits pulsed light, where the reflected light is collected at the receiver and time-of-flight (ToF) is calculated to measure distance. Most popular LIDAR systems operate by scanning the region with a defined scan rate to acquire data about the surroundings, up to high distances.

Since 2010s, LIDAR application with CS is investigated on several occasions. Howland provided one of the first prominent results, where not only he shows the applicability of CS to the LIDAR systems, but also 32x32 pixel video at 14 fps, at with 30% of the original size (Howland, Dixon, et al., 2011; Howland, Zerom, et al.,

2011). In 2014, Lau proposed a new resampling method to reconstruct an 3D image from as low as 20% of a point cloud data (Lau & Woodward, 2014). In 2016, Gong et al. published their results for a 3D LIDAR with ghost imaging and CS, showing the reconstruction of a 3D scene about 1 km at 49.6% of the Nyquist limit. Edgar published a real-time LIDAR with CS principles, where it performed 3 fps for 64x64 reconstruction (Edgar et al., 2017).

In signal processing, direction-of-arrival remains as one of the most important issues. As of the 21st century, the technology has enabled the wide reach and use of wireless communications.

1.1.3 MRI

One of the most prominent studies for CS application take place in medical imaging, specifically, magnetic resonance imaging (MRI). MRI utilizes MR images which are taken as points in the 2D or 3D Fourier space. The process for an MRI can took between 15 to 90 minutes, due to its innate slow data taking operation. Ensuring the people hold still for these long durations, especially young people can be challenging. CS based MRI, can reduce this acquisition time immensely by taking a lot fewer samples to reconstruct the same image. Less samples result with less image acquisition time, by also keeping the resulting image diagnostically recognizable.

First works came by Lustig et al., where he showed that medical images are highly compressible, hence compliant with CS scheme (Lustig et al., 2007, 2008). The company Siemens Healthineers is pioneering the medical imaging field with the first CS products (Siemens-Healthineers, n.d.) [26] and got CS MRI approved from FDA in February 2017 (Imaging Technoogy News, 2017). Subsequentially, Philips also introduced their MRI imaging machine that uses CS, called Phillips Compressed SENSE (Philips, n.d.). This thesis will not delve into the individual papers for the MRI with CS, since it is huge and still a hot topic. However, readers are advised to

refer to the informative papers (Jaspan et al., 2015) and (Feng et al., 2017), in pursuit to caught up with the improvements and developments in the topic.

1.1.4 Single Pixel Camera

A traditional imaging scheme involves acquiring all the pixels of the scene, up to tens of megapixels, then compresses it according to the need of transmission and storage. This procedure is highly expensive and seemingly redundant. Compressive imaging challenges this idea, by fusing the acquisition and compression phases. It promotes acquiring random linear measurements directly from an image, without the information need of every pixel in the scene. The benefits outweigh the cost, especially in the range outside of visible spectrum, where every pixel is extremely valuable.

The notorious single-pixel camera that was developed at the Rice University was the first implementation of CS to an imaging system (Takhar, Laska, Baron, et al., 2006; Takhar, Laska, Wakin, et al., 2006). The system was a fairly simple setup that utilized a digital micromirror device (DMD) as a spatial light modulator (SLM), that resulted in a simple, compact, brilliant and low-cost solution for broader ranges. Shortly after, the CS concept was again successfully trialed for a terahertz imager at the Rice University where the modulating patterns are provided with printed circuit boards (Chan et al., 2008).

CHAPTER 2

BACKGROUND

This section will include a brief introduction to compressed sensing history and theory. This will include the explaining of the aforementioned properties a signal needs to have in order to be full reconstructed. The current state of the applications for compressed sensing will also be discussed. Proceedingly, different algorithms will be explained, and their use will be compared.

2.1 Brief History of Compressed Sensing

In the so-called information age, data generation has surpassed the available storage as of 2007, according to John Gantz, IDC (Gantz et al., 2008). In today's technology, individuals are able to create data, then it was ever possible before. This is combined with the quality of the sensors, that are capable of sampling ever faster and in higher resolution than before, it is reasonable to point out, the gap between the data created and the data stored is in a diverging trend. The diverging gap can be observed for the years between 2005 and 2012 in Fig. 1.

Coincidentally, in 2004, the first publication and the appearance of Compressed Sensing, which is a signal processing routine that leverages data acquision scheme by exploiting sparsity (i.e. having most of the elements zero or approximately zero) of signals is introduced by Donoho (D. L. Donoho, 2006a).

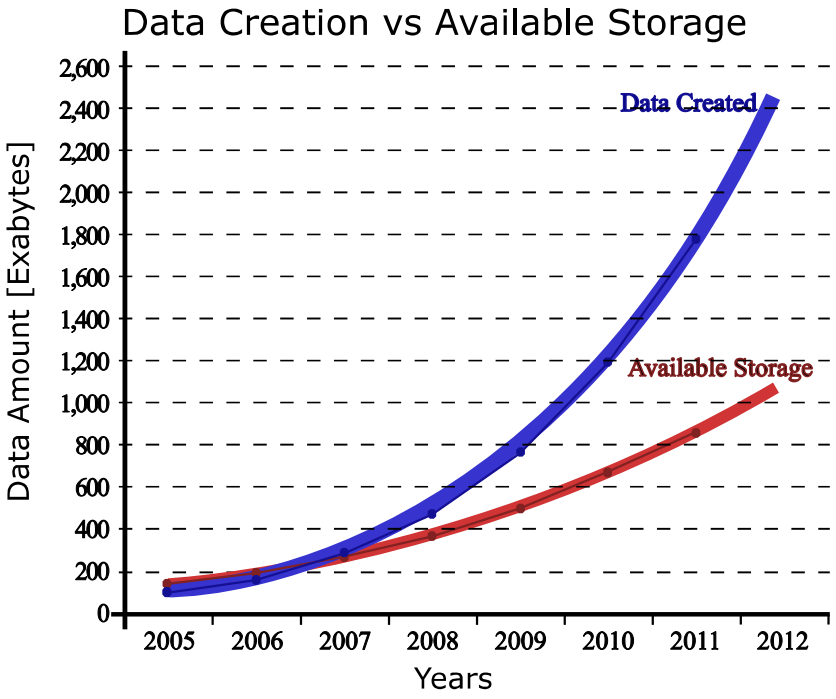


Figure 1. Graph showing the data creation and available storage between 2005 – 2012 (Readapted from [34])

The mathematical foundations behind the compressed sensing was as old as 1980s. Moreover, Least Absolute Deviations (ℓ_1 norm) was introduced centuries earlier by Boscovich in 1757, and discussed by prominent scientists, like Laplace, Gauss and Edgeworth (Farebrother Richard William, 2013). ℓ_1 norm since then reappears in sparse signal reconstruction around 1987 (Santosa et al., 1987). The leveraging idea of exploiting sparsity was already a common practice with some of the 2000s compression formats. As an example, JPEG standard was similary relying on exploiting sparsity of images by rerepresenting them in a favorable, advantageous transformation (i.e. Discrete Cosine Transform [DCT], wavelet) , then keeping only the largest coefficients (Wallace, 1992) (Taubman & Marcellin, 2002). Fig. 2 illustrates the compressibility by exploiting the image sparsity by DCT. Even when, as high as 75% of the DCT coefficients are missing, the reconstructed image can be still regarded as recognizable and indifferent. Keeping only 10% and lower, although the image is recognizable, the defects overwhelm and the quality is noticable upon sight.

Figure 2. A compression by DCT example showing a) Original image b) Reconstruction with 25% and c) 10% of the data.

Since the data transmission and storage is a priority in a lot of applications, the compression is a common practice to apply. The absurdity is, these practices include compression and uncompression of the signal after acquiring the full signal (See Fig. 3a). Compressed sensing, precisely deals with this absurdity, and basically tackles the questions: "Why do bother to take the whole data, just to discard most of it?" and "Is there a way to just directly acquire the compressed form?" (See Fig. 3b).

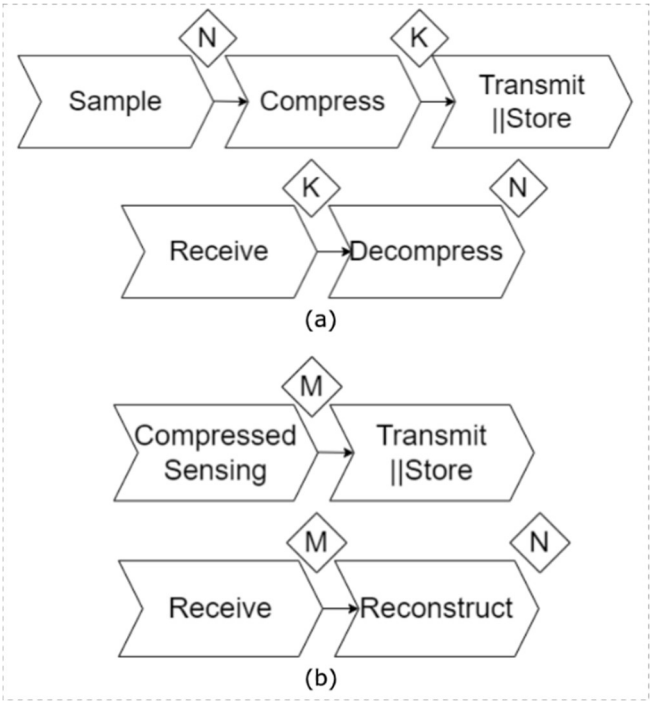


Figure 3. Comparison of acquisition schemes; a) Conventional acquisition scheme b) Compressed sensing acquisition scheme

Upon its commencement in 2004, compressed sensing has been met with a lot of suspicions due to its challenging nature against the Nyquist-Shannon Theorem by claiming it is possible to fully or mostly reconstruct a signal from far less samples that the Nyquist-Shannon Theorem dictates, given that the signal satisfies the conditions, sparsity and incoherence (E. J. Candes & Wakin, 2008). The Nyquist-Shannon Theorem states that, in order to reconstruct a signal perfectly, it has to be sampled at least twice the maximum frequency that the signal comprises (Shannon, 1949). Otherwise, the reconstructed signal risks aliasing as shown in Fig. 4.

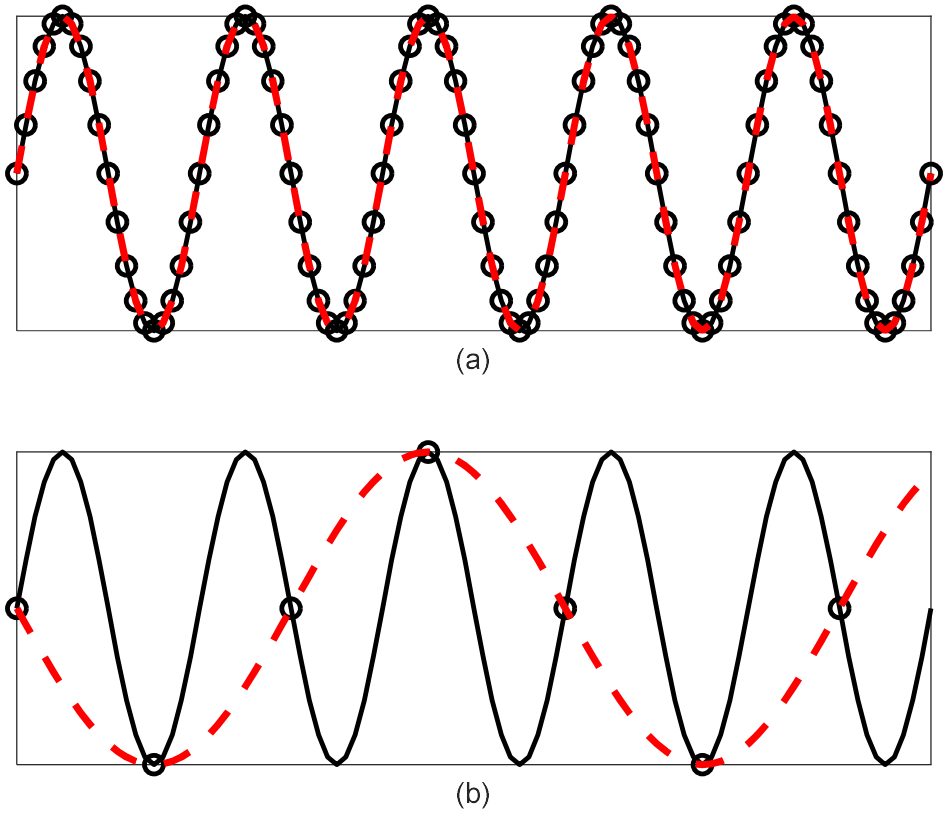


Figure 4. a) Sufficiently sampled signal (black) and reconstruction (red) b) Aliased signal (red) due to undersampling

2.2 The problem of interest

Let us consider the following equation where a real-world continuous signal acquisition scheme is shown as discretized signal vector $\boldsymbol{x} \in \mathbb{R}^N$, $\mathbf{y}$ is the

measurement vector, and A is the relating system that often also referred as the measurement matrix,

$$A\boldsymbol{x} = \boldsymbol{y} \qquad \text{(Eq. 1)}$$

If we know how the imaging system behaves, and the input signal, it is straightforward to foresee what the output must be, which is called the *forward problem*. One is given the inputs of the system and expects to find or predict the outcome. However, in physics, a good deal of interesting problems including imaging require mostly the opposite called the *inverse problem*, in which, by having the system knowledge and the set of rules governing it, one is asked to find the underlying question or cause by looking at the output and effect (Razavy, 2020) (Vauhkonen Marko and Tarvainen, 2016).

CS theory requires a signal to have sparsity and incoherence properties, in order to reconstruct it with using a relatively few measurements (E. J. Candes & Wakin, 2008). It turns out, most of the real-world signals are almost sparse in an appropriate domain (E. J. Candes & Wakin, 2008) (Lee et al., 2016). For example, images are sparse in DCT, wavelet domain, while audio is sparse in Short Time Fourier Transform (STFT), which is a generalization of the Gabor transform. So, mathematically, every signal $\boldsymbol{x} \in \mathbb{R}^N$ can be represented in a given orthonormal basis $\{\psi_i\}_{i=1}^N$ in $\mathbb{R}^N$with N coefficients $\{s_i\}_{i=1}^N$ where most of the coefficients are zero or approximately zero.

$$\boldsymbol{x} = \sum_{i=1}^{N} \psi_i s_i \qquad \text{(Eq. 2)}$$

One can also state it in the matrix notation by arranging the basis vectors into the columns of the sparsifying basis $\Psi \in \mathbb{R}^{N\times N}$, and stacking the N coefficients into a coefficient vector $\boldsymbol{s} \in \mathbb{R}^{N\times 1}$. The sparsifying matrix is also often called as dictionary or sparsifying basis. Some exemplary common orthonormal sparsifying bases Ψ are shown in Fig. 5. One can utilize these bases to transform a signal into another domain to benefit the sparsity. As an example, a sinusoid wave that can be described by a few coefficients in Fourier domain, needs much more coefficients to describe it

in the Kronecker domain. For $\boldsymbol{s}$ to be a sparse signal, one must have most of the constituents of the vector $\boldsymbol{s}$ as zeros. So, by saying a signal is K-sparse in some domain Ψ, it is understood that there are K coefficients ($K \ll N$) in the coefficient vector $\boldsymbol{s}$ that are non-zero entries. It can also be put that, cardinality or ℓ_0 of the signal $\boldsymbol{x}$ is K. The norm notation will be explained in detail later in Section 2.5, for now, it can be defined as the total number of non-zero entries in a vector.

$$\boldsymbol{x} = \Psi \boldsymbol{s} \qquad \text{(Eq. 3)}$$

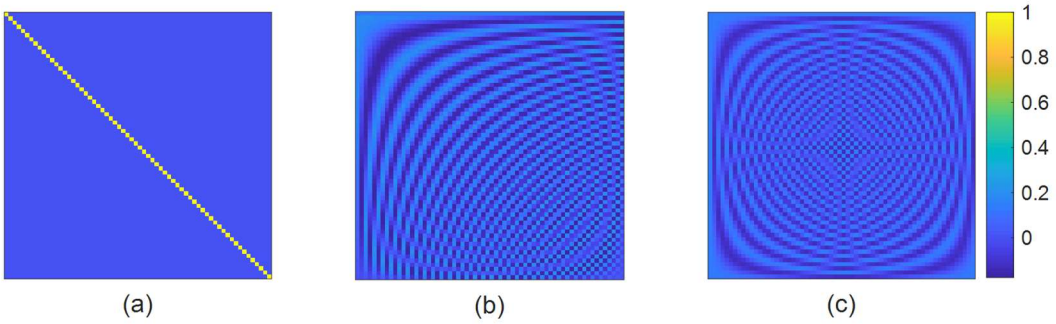


Figure 5. Exemplary bases Ψ of a) Spike b) DCT c) Fourier

In the compressed sensing framework, the problem then can be reformulated by inserting the signal (Eq.3) into the sensing scheme (Eq. 1),

$$\boldsymbol{y} = A\boldsymbol{x} = \Phi\Psi\boldsymbol{s} \qquad \text{(Eq. 4)}$$

where $\boldsymbol{y} \in \mathbb{R}^N$ is the measurement vector, $\boldsymbol{x}$ is the signal vector of interest, Ψ is the transform/sparsifying matrix commonly of size $N \times N$, $\boldsymbol{s}$ is the coefficient vector of size $N \times 1$. The relating system, in other words, the measurement matrix, in some texts referred as *sensing matrix*, A of size $M \times N$ is rewritten as Φ, in order to follow the conventionally preferred notation in the compressed sensing texts, rather than its form of use in the inverse problem texts as formulation of the general systems.
Restricted Isometry Property

In order to be able to quantitavely compare different sensing matrices, and their usability within the scope of CS, a crucial approach proposed by Candés and Tao, called the Restricted Isometry Property (successor of the so-called Uniform

Uncertainty Principle) has proven to be useful (E. Candes & Tao, 2004) (E. J. Candes & Wakin, 2008). It simply provides a numerical measurement scheme to understand whether a measurement/sensing matrix retains the distance of vector **x** under transformation by the system. A matrix A of size $M \times N$, said to have Restricted Isometry Property (RIP) of order K, with Restricted Isometry Constant (RIC) δ, if the equation below holds true for all K-sparse vectors **x**, and $0 < \delta_{2K} < 1$,

$$(1 - \delta_{2K})||\boldsymbol{x}||_2^2 \leq ||A\boldsymbol{x}||_2^2 \leq (1 + \delta_{2K})||\boldsymbol{x}||_2^2 \qquad \text{(Eq. 5)}$$

where δ_{2K} is preferably as close to zero as possible to favor a stable K-sparse signal recovery. As close as it is to zero, the better reconstruction arises even in the presence of noise. For sufficiently small values of δ_{2K}, the solution is expected to be exact and unique. The reason of picking the RIC for $2K$-sparsity value, is due to the superpositional nature of vectors. One can have two k-sparse vectors, namely x_1 and x_2, adding these vectors can result in a sum vector of K-sparse to $2K$-sparse, depending on the overlapping support locations. Furthermore, even if the support locations do not coincide, one cannot deduce the resulting vector is still sparse. The RIP condition ensures that these two K-sparse vectors do not map onto the same measurement vector **y**, so that,

$$\boldsymbol{x}_1 \neq \boldsymbol{x}_2 \qquad \text{(Eq. 6)}$$

$$A\boldsymbol{x}_1 = y \qquad \text{(Eq. 7)}$$

$$A\boldsymbol{x}_2 = y \qquad \text{(Eq. 8)}$$

do not hold. Hence, (Eq. 5) can be rewritten accordingly as,

$$(1 - \delta_{2K})||\boldsymbol{x}_1 - \boldsymbol{x}_2||_2^2 \leq ||A(\boldsymbol{x}_1 - \boldsymbol{x}_2)||_2^2 \leq (1 + \delta_{2K})||\boldsymbol{x}_1 - \boldsymbol{x}_2\,||_2^2 \qquad \text{(Eq. 9)}$$

$$(1 - \delta_{2K})||\boldsymbol{x_1} - \boldsymbol{x}_2||_2^2 \leq ||A\boldsymbol{x}_1 - A\boldsymbol{x}_2||_2^2 \leq (1 + \delta_{2K})||\boldsymbol{x}_1 - \boldsymbol{x}_2\,||_2^2 \qquad \text{(Eq. 10)}$$

results in squeezing the sensing matrix A, to ensure that transformation under matrix A, does preserve the Euclidian distance. Consider an orthogonal matrix A where the

columns will have its inner products as zeros. Hence, it is an isometry that preserves the distance and rotation perfectly,

$$||A\boldsymbol{x}||_2 = ||\boldsymbol{x}||_2 \qquad \text{(Eq. 11)}$$

and results in an ideal RIC $\delta_{2K} = 0$. However, the sensing matrices A in CS are commonly of $M \times N$ size where $M < N$ and $2K < M$, resulting in a RIC of $\delta_{2K} > 0$. So, the RIP is a great tool to ensure any subset of K-columns of A, acts on K-sparse signal $\boldsymbol{x}$, without change of any considerable information. If the sensing matrix A has its RIC $\delta_{2K} < 1$, than it is almost behaving as an isometry. Moreover, it has been shown by Candes (E. J. Candes & Wakin, 2008), that if $\boldsymbol{x}$ is a K-sparse signal where the RIC $\delta_{2K} < \sqrt{2} - 1$, then the solution is exact.

Unfortunately, showing a matrix **A** satisfies the RIP is a NP-Hard problem (Bandeira et al., 2013), if not NP-complete (E. Candès, 2006). Since the non-zero entries in the coefficient vector $\boldsymbol{s}$ in (Eq. 4), is not known beforehand, a sensing matrix **A** is to be constructed such that, it has a good RIP without the prior knowledge of the signal entries. Fortunately, randomness has been proven to be an useful aproach for such construction of measurement matrices (E. Candes & Tao, 2004). Namely; standard probability distributions, such as Gaussian, Bernoulli, Fourier ensemble random matrices with independent and identically distributed entries (D. L. Donoho, 2004)(E. Candès, 2006)(E. J. Candes & Tao, 2006). Furthermore, these matrices tend to share *universality property*, which means they perform regardless of the orthonormal basis choice Ψ, and there is a high chance that the reconstruction matrices ΦΨ would still have the RIP (Marco Duarte, 2009).

2.3 Incoherence

The coherence of a matrix $A \in \mathbb{R}^{M \times N}$, can be calculated to construct a low coherence measurement matrix Φ (e.g. sensing matrix) that has linearly independent vectors, which allows a quantitative metric for sensing matrix quality. It can be formulated as,

$$\mu(A) = \max_{j \neq k} |< a_j, a_k >| \qquad \text{(Eq. 12)}$$

where a_j and a_k are the normalized columns of **A**, and $\mu \in [1, \sqrt{N}]$ is the coherence. Furthermore, this formulation can be reduced to a somewhat simpler form to calculate a bound (Foucart & Rauhut, 2013) for the coherence value μ as,

$$\mu \geq \sqrt{\frac{N-M}{M(N-1)}} \qquad \text{(Eq. 13)}$$

It can be shown that, in case where $M << N$, the equation above form reduces to,

$$\mu \geq \frac{1}{\sqrt{M}} \qquad \text{(Eq. 14)}$$

Incoherence property can be understood as, two opposite domains give the maximum information if they are uncorrelated. It feeds from the duality that if a signal is sparse in domain Ψ, then the domain Φ must be incoherent, so that the signal information is dispersed in the Φ domain (E. J. Candes & Wakin, 2008). Hence, each measurement contains some information about the whole signal, and each are equally important. The easiest option for a domain pair that is incoherent would be Kronecker (i.e. Dirac, spike domain) and frequency (i.e. Fourier domain) where the coherence value $\mu = 1$ [1999 Donoho]. On the other hand, the complete coherence case would result in a $\mu = \sqrt{N}$, where the measurement matrix and the sparsifying basis would be equal Φ = Ψ.

The incoherence property is one of the two pillar conditions that CS relies on along with sparsity. Although the RIP condition quantitatively ensures a robust, unique signal reconstruction occurance, the incoherence property is named as a fundamental premise (E. J. Candes & Wakin, 2008). The reason for this apparent conflict is, because incoherency is already implied inside the RIP, and it is somewhat looser condition to verify as opposed to the RIP condition. If the measurement matrix Φ and sparsifying matrix Ψ have large mutual coherence, then it can be mostly deduced that ΦΨ would be coherent too (E. J. Candès et al., 2011). Also, very large coherence would imply, how much of a just one basis vector is present in the signal, therefore it would take the whole N number of basis vectors to accurately represent the signal,

and defeats the purpose of CS, which is undersampling. In other words, smaller the coherence value μ, means fewer measurement required. Fortunately again, i.i.d. random matrices (e.g. Gaussian, Bernoulli) are incoherent with almost any fixed sparsifying basis in any dimension (E. J. Candes & Tao, 2006) (Duarte et al., 2006).

2.4 Normed Vector Spaces

Normed vector spaces are defined where every vector in a vector space is associated with a norm, which helps to introduce the concept of length beyond just 2D or 3D. It is useful when working with vectors to measure a signals strength, or an error size(Eldar & Kutyniok, 2012). It is formulated as below where p denotes the norm,

$$\left|\left|\boldsymbol{x}\right|\right|_p = \begin{cases} \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, p\epsilon(0,\infty) \\ \max\limits_{i=1,2,\ldots,n} |x_i|\,, \; p = \infty \end{cases} \qquad \text{(Eq. 15)}$$

2.4.1 ℓ_0 Norm

The ℓ_0-norm is often called as ℓ_0-pseudonorm, formally because it does not satisfy all the properties of the norm definition, namely homogeneity. It can be formulized as,

$$\left|\left|x\right|\right|_0 \coloneqq \left|\{j : x_j \neq 0\}\right| \qquad \text{(Eq. 16)}$$

It simply counts the non-zero elements in a vector. It can also be put as sparsity of $\boldsymbol{x}$, $|supp(\boldsymbol{x})|$, or $card(supp(\boldsymbol{x}))$ where card stands for the cardinality function and corresponds to the number of elements in a set.

2.4.2 ℓ_2 Norm

In the order of importance, ℓ_2-norm (Euclidian norm, square norm, second order norm) is given priority before understanding ℓ_1-norm. Since the problem at hand is an imaging problem, that is represented by vectors and matrices, it is intuitive to use

ℓ_2-norm. It has been the common practice for solving and fitting for datasets to approximate signals and minimize the error for decades, which inherently are vectors. Its use has been regarded widely because it serves as an energy norm and often its physical interpretation is obvious (Kutz, 2013). The ℓ_2-norm can be formulated as,

$$\left|\left|\boldsymbol{x}\right|\right|_2 = \sqrt{\sum_{i=1}^{n}|x_i|^2} = \sqrt{<\boldsymbol{x},\boldsymbol{x}>} \qquad \text{(Eq. 17)}$$

where angle bracket notation < , > indicates the inner product.

2.4.3 ℓ_1 Norm

In the case of ℓ_1-norm (Manhattan norm, Taxi-cab norm, first order norm), it sums the vectors magnitudes. Contrary to ℓ_2-norm where the vector length is measured, ℓ_1-makes use of the total of absolute length of the vector's constituents.

$$\left|\left|\boldsymbol{x}\right|\right|_1 = \sum_{i=1}^{n}|x_i| \qquad \text{(Eq. 18)}$$

2.4.4 Quasinorms

The use of norms where $0 < p < 1$ is not recognized fully, but instead referred as *quasinorms*, because they do not satisfy all the properties of a norm, namely triangle inequality. They do not classify as convex relaxation, since in the range $0 < p < 1$, the problem is NP-hard in general. Although not as much as ℓ_1 and ℓ_2 norms, they do have their use in non-convex optimization algorithms against ℓ_0-minimization problem (Yin et al., 2015).

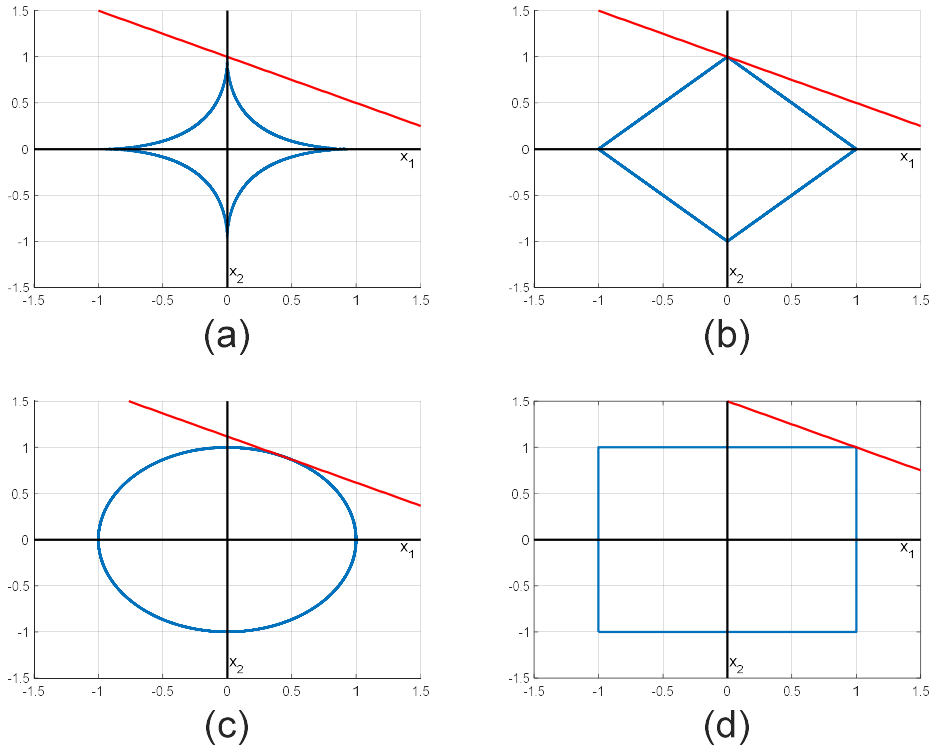


Figure 6. Unit p-balls in R^2 for a) $\ell_{1/2}$-norm b)ℓ_1-norm c) ℓ_2-norm d) ℓ_∞-norm

In Fig. 6, common p-ball drawings are drawn for $\mathbb{R}^2$. $x \in \mathbb{R}^2$ is evaluated for various norms. One can easily deduce from the figures, as $p \to 0$ for ℓ_p, the p-ball drawing becomes just a spike ball for ℓ_0 whose minimization is the actual solution since it counts non-zero entries. The solution space is represented by the red line. It is evident that $\ell_{1/2}$ and ℓ_1-norm minimizations tend to promote sparsity in a solution, as they coincide along on axis with the corresponding unitary p-balls. So the solution found after the related minimization procedure is, indeed the sparsest solution. Unlike, $\ell_{1/2}$-norm which is nonconvex, ℓ_1-norm can be solved by convex optimization. ℓ_2-norm coincides with the p-ball at a tangential point. In other words, the solution that is found by minimizing the ℓ_2-norm gives a solution that has significant value of all the constituents, hence the solution it finds is not sparse. As $p \to \infty$ for ℓ_p, it gives equal weight to the all components, and has no viable solution for the inverse problem $y = Ax$.

2.5 Sparse Signal Recovery

In the compressed sensing framework, the problem then can be expressed by inserting the signal (Eq. 3) into the sensing scheme (Eq. 1),

$$y = A\boldsymbol{x} = A\Psi S = \Phi\Psi\boldsymbol{s} \qquad \text{(Eq. 4)}$$

where $\boldsymbol{y}$ is the $N \times 1$ measurement vector, $\boldsymbol{x}$ is the signal vector of interest, Ψ is the transform/sparsifying matrix commonly of size $N \times N$, $\boldsymbol{s}$ is the $N \times 1$ coefficient vector. The relating system, in other words, the measurement matrix (in some texts referred as sensing matrix) A of size $M \times N$ is rewritten as Φ, in order to follow the conventionally preferred notation in the compressed sensing texts, rather than its form of use in the linear algebra texts as formulation of the general systems. Reforming the equation by using the product of the measurement matrix and the sparsifying matrix, and representing them under one variable, the equation then becomes,

$$y = \Phi\Psi S = \Theta\boldsymbol{s} \qquad \text{(Eq. 19)}$$

where Θ is the reconstruction matrix of size $M \times N$ in the compressed sensing context.

Consider the case where the matrix $A \in \mathbb{R}^{M \times N}$ in (Eq. 4), has $M = N$ where all of its columns are linearly independent (i.e., full-rank matrix). Assuming it has a solution, then the linear system would have a unique solution. In the case where $M > N$, the system is called overdetermined. There are more equations than unknowns and the matrix A is often called as a tall matrix. This results in a system where there are no solutions. Oppositely when $M < N$, the matrix A is referred as fat, then the system is called underdetermined and ill-posed since there are infinitely many solutions. This paves the way for an optimization problem. There are infinitely many solutions, and an optimization procedure is needed to constrain and solve for the most valid answer among these infinitely many solutions. An optimization problem can be formulated as (Foucart & Rauhut, 2013),

$$\min_{x \in R^N} f_O\,(\boldsymbol{x})\; subject\; to\; f_c(\boldsymbol{x}) \leq c_i, 1 \leq i \leq n \qquad \text{(Eq. 20)}$$

where $f_0 \colon \mathbb{R}^N \to (-\infty, \infty]$ is the objective function and $f_c \colon \mathbb{R}^N \to (-\infty, \infty]$ is referred as the *constraint functions* for which c and i are real and integer values respectively. If the locations of the non-zero entries were known beforehand, the equation (Eq. 4) would be fairly straightforward to solve, considering the signal is K-sparse. However, without this information, one possible approach to take advantage of the already present knowledge (that the signal is K-sparse), would be looking for all the K-sparse vectors. This can be posed as,

$$\min_{x \in R^N} \left|\left|\boldsymbol{x}\right|\right|_0 \quad subject\; to\; y = A\boldsymbol{x} \qquad \text{(Eq. 21)}$$

The above result is for the ideal case when there is no error present. In order to account for the noise, presumably a Gaussian noise, the constraint function is changed as,

$$\min_{x \in R^N} \left|\left|\boldsymbol{x}\right|\right|_0 \quad subject\; to\; \left|\left|y - A\boldsymbol{x}\right|\right|_2 \leq \epsilon \qquad \text{(Eq. 22)}$$

where the ideal measurement $\boldsymbol{y}^* = A\boldsymbol{x}$ is subtracted from the actual measurement $\boldsymbol{y}$ to make sure its Euclidian distance is less than the specified error ϵ.

As mentioned before in Section 2.5, ℓ_0-pseudonorm of a vector only counts the number of non-zero entries. This is a non-convex problem and solving for the above equation, requires $C(N, k)$ possibilities of trial, considering each K-sparse vector in solution space where $N >> K$, hence it is referred as a combinatorial problem, or NP-Hard.

Although there are several classes of algorithms developed to address this problem in different ways, these classes can be grouped under two approaches. One approach is to iteratively seek the solution space and build up the signal from bottom-up heuristically by selecting the components that result in the best performance. This class of approach is often called *greedy search*. This approach is simpler to implement, and often under certain conditions, they perform adequately and even guarantee a recovery. Unfortunately, brute search is often oblivious and not expected

to find a global extremum. The other approach is to weaken the constraints of the equation, such that it is easier to deal with. Hence, this class of approach is named *convex-relaxation*. Rish and Grabarnik (Rish & Grabarnik, 2014) encapsulates this statement best possible way as; one can either solve the exact problem approximately, or an approximate problem exactly.

2.6 Convex Relaxation

2.6.1 ℓ_1 Minimization

Chen, Donoho and Saunders introduced the ℓ_1-norm usage in 1998 for atomic decomposition of dictionaries and called it Basis Pursuit (Chen et al., 1998). Since ℓ_0-norm is a pseudonorm that is NP-hard and non-convex, the objective function of (Eq. 21) is dialed down to a weaker form which is easier to solve, namely ℓ_1-norm, thus, it is called convex-relaxation. For the noiseless case of (Eq. 21), it can be restated as,

$$\min_{\boldsymbol{x}\in R^N} \left|\left|\boldsymbol{x}\right|\right|_1 \quad subject\ to\ \boldsymbol{y} = A\boldsymbol{x} \qquad \text{(Eq. 23)}$$

and same reasoning applies for the noisy case of (Eq. 22), called as *Basis Pursuit Denoising (BPDN)*, and restated as,

$$\min_{\boldsymbol{x}\in R^N} \left|\left|\boldsymbol{x}\right|\right|_1 \quad subject\ to\ \left|\left|\boldsymbol{y} - A\boldsymbol{x}\right|\right|_2 \leq \epsilon \qquad \text{(Eq. 24)}$$

Notice where $\epsilon = 0$, the above formulation (Eq. 24) becomes the same with (Eq. 23). An equivalent approach is called LASSO (Least Absolute Shrinkage and Selection Operator) as,

$$\min_{\boldsymbol{x}\in R^N} \left|\left|\boldsymbol{y} - A\boldsymbol{x}\right|\right|_2 \quad subject\ to\ \left|\left|\boldsymbol{x}\right|\right|_1 \leq \delta \qquad \text{(Eq. 25)}$$

where the objective function and the constraint functions are replaced. Whereas BPDN looks for the sparsest vector that fits in the inverse problem within certain error tolerance, LASSO looks for the minimum error vector that have its sparsity limited, in other words puts a constrain on its sparsity. Although these are called

equivalent, they are not the same. BPDN is an ℓ_1-norm minimization problem whereas LASSO is an ℓ_1-regularized problem. They are developed by different communities, end referred as equivalent since for a certain positive value for ϵ and δ, they lead to the same solution. BPDN was developed in signal processing community by Chen, Donoho and Saunders in 1998 (Chen et al., 1998), whereas LASSO introduced in statistics community by Tibshirani in 1996 (Tibshirani, 1996) (Ament, 2017). These two equations can also be shown in Lagrangian form as,

$$\mathcal{L}(\boldsymbol{x}, \lambda) = \min_{\boldsymbol{x} \in R^N} \left|\left|\boldsymbol{x}\right|\right|_1 + \lambda \left|\left|\boldsymbol{y} - A\boldsymbol{x}\right|\right|_2 \qquad \text{(Eq. 26)}$$

$$\mathcal{L}(\boldsymbol{x}, \lambda) = \min_{\boldsymbol{x} \in R^N} \left|\left|\boldsymbol{y} - A\boldsymbol{x}\right|\right|_2 + \lambda \left|\left|\boldsymbol{x}\right|\right|_1 \qquad \text{(Eq. 27)}$$

where, the objective function is often called as loss function, and instead of a constraint function, an added penalty function term (i.e. regularizer) can be tuned for changing the Lagrange multiplier (i.e. regularization parameter) λ, where $\lambda \to \infty$, the vector $\boldsymbol{x} \to 0$. It enables the control over sparsity and approximation, in order to favor a balance to emphasize one over another. Sometimes, the LASSO equation can be also seen where square of the objective function itself gets executed with the constraint as in (Eq. 27), so the constrained minimization problem is,

$$\min_{\boldsymbol{x} \in R^N} ||\boldsymbol{y} - A\boldsymbol{x}||_2^2 \; subject\ to \; \left|\left|\boldsymbol{x}\right|\right|_1 < t \qquad \text{(Eq. 28)}$$

for some parameter t. Furthermore, in case **A** is an orthogonal matrix, then the Euclidian norm of the vector gets preserved, as

$$||A\boldsymbol{x}||_2^2 = (A\boldsymbol{x})^T(A\boldsymbol{x}) = \boldsymbol{x}^T(A^TA)\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{x} = ||\boldsymbol{x}||_2^2 \qquad \text{(Eq. 29)}$$

$$||\boldsymbol{y} - A\boldsymbol{x}||_2^2 = ||A^T(\boldsymbol{y} - A\boldsymbol{x})||_2^2 = ||\hat{\boldsymbol{x}} - \boldsymbol{x}||_2^2 = \sum_{i=1}^{n} (\hat{\boldsymbol{x}}_{\boldsymbol{i}} - \boldsymbol{x}_{\boldsymbol{i}})^2 \qquad \text{(Eq. 30)}$$

where $\hat{\boldsymbol{x}}$ is the measured vector (Rish & Grabarnik, 2014).

2.6.2 ℓ_2 Minimization

ℓ_2-norm minimization is also a convex relaxation method that is commonly used in engineering fields. The convex problem is often referred as *least-squares problem* and given as,

$$\min_{\boldsymbol{x}\in R^N} ||\boldsymbol{y} - A\boldsymbol{x}||_2^2 \qquad \text{(Eq. 31)}$$

where the minimization has a unique solution $\boldsymbol{x} = (A^T A)^{-1} A^T \boldsymbol{y}$ for underdetermined systems, which can be easily checked by the objective function's derivative in (Eq. 28). However, since ℓ_2-norm is an energy preserving norm, and does not promote sparsity as ℓ_1-norm does, in sparse signal recovery, the value resulted from ℓ_2-minimization disperses the energy onto the whole signal. Hence, the solution it gives is almost never sparse. The square terms in the norm definition itself, makes ℓ_2-norm usage in sparse signal recovery more susceptible to the outliers, the absolute distance gets squared in ℓ_2-norm, whereas ℓ_1-norm minimization is more democratic by taking only the absolute distance itself.

Nevertheless, ℓ_2-norm is still used in some CS algorithms with a slight modification. Although, by itself it does not promote a sparse solution like ℓ_1-does, ℓ_2-norm lays in the foundation of a lot of models and recovery algorithms. Besides LASSO, it is used mainly in *greedy algorithms, non-convex minimization* and *nuclear norm minimization algorithms*, often with some penalty function. This method in general, is studied under the name *regularization theory*.

2.7 Greedy Approach

Besides the convex relaxation approach, greedy approach motivates and performs quite differently. Whereas the convex relaxation approach weakens the constraint on a problem to resemble a different, easily solvable problem, greedy approaches perform iteratively to build a solution from ground up to approximate as much as to the original signal until a stopping criterion is met. They iteratively seek and update

the signal according to the best resemblances. Although convex relaxation presents a complete framework to find a global minimum to the problem, its complexity makes it difficult to implement to any hardware. Greedy approaches are more machine friendly, considered by their computational cost and implementability. These attributes result in greedy approaches to perform better in both time wise and complexity wise (i.e. for higher dimensional signals).

The greedy approach algorithms that will be explained in this chapter will be the classes of *Greedy Iterative Algorithms* and *Iterative Thresholding Algorithms*. The main difference between them is due to their residual update treatment.

2.7.1 Greedy Iterative Algorithms

This class of algorithms are tackling with the original problems in (Eq. 21) and (Eq. 22) iteratively, and solve the below formulation of the problem,

$$\min_{x \in R^N} ||y - Ax||_2^2 \quad subject\ to\ \left|\left|x\right|\right|_0 \leq k \qquad \text{(Eq. 32)}$$

First an empty version of the vector $\boldsymbol{x}$ and an empty support vector $\boldsymbol{s}$ are initialized. Then a residual vector $\mathbf{r}$ that is equal to the measurement vector $\mathbf{y}$, is introduced. Iteratively, the vector $\boldsymbol{x}$ and $\boldsymbol{r}$ are updated, according to some selection criteria, until a stopping condition is met. The vector $\boldsymbol{x}$ keeps getting reconstructed from zero, whilst the residual vector $\boldsymbol{r}$ gets smaller. Greedy iterative algorithms mostly follow a similar scheme. The difference inherently arises from their selection criteria, the residual estimation and the satisfaction conditions amongst these algorithms.

2.7.1.1 Matching Pursuit (MP)

Mallat and Zhang first published the *Matching Pursuit* (MP) scheme in 1993 (Mallat & Zhang, 1994). It is the most basic algorithm among the greedy approach family. The algorithm iterates through by looking for the correlation between the columns

of A and the residual vector $\boldsymbol{r}$, and selects the most correlated column vector from A, and reflects it as a coefficient for newly building $\boldsymbol{x}$.

Algorithm 1 Matching Pursuit (MP)

Input: Measurement vector $\boldsymbol{y}$, Measurement matrix A
Output: Sparse vector $\boldsymbol{x}$
Initialize: $\boldsymbol{x}^{[0]} \leftarrow 0, \boldsymbol{r}^{[0]} \leftarrow \boldsymbol{y}, i \leftarrow 0$

1: **while** halting criterion false **do**
2: $\quad i \leftarrow i + 1$
3: $\quad c^{[i]} \leftarrow A^T r^{[i-1]}$ ▷ Compute the correlation
4: $\quad j^{[i]} \leftarrow \underset{j}{\operatorname{argmax}} |c_j^{[i]}|$ ▷ Find the index of the most correlated
5: $\quad x_{j^{[i]}}^{[i]} \leftarrow x^{[i-1]} + c_{j^{[i]}}^{[i]}$ ▷ Update the approximation
6: $\quad r^{[i]} \leftarrow r^{[i-1]} - A_{j^{[i]}} c_{j^{[i]}}^{[i]}$ ▷ Calculate the new residual
7: **end while**
return $x \leftarrow x^{[i]}$

It solves the aforementioned (Eq. 28) formulation of the original problem, where the objective function $||\boldsymbol{y} - A\boldsymbol{x}||_2^2$ tried to be minimized by the iteratingly built estimation. It calculates the signal estimation by adding the newly found correlated elements to form the vector $\boldsymbol{x}$, one element per iteration. However, MP may select the same column vectors if the column vectors of matrix A are coherent.

2.7.1.2 Orthogonal Matching Pursuit (OMP)

Shortly after the obvious convergence issues of MP got noticed, Pati has improved upon the idea of MP by introducing the *Orthogonal Matching Pursuit* (OMP) (Pati et al., 1993). The algorithm has been slightly changed to project $\boldsymbol{y}$ onto the matrix $\boldsymbol{A}$ with the selected indices at each iteration, and therefore does not only just count the inner product of a column of $\boldsymbol{A}$**.** This leads to the optimization where the already selected indices gets to be accounted for. The residual vector $\boldsymbol{r}$ (where $\boldsymbol{r} = \boldsymbol{y}$ at initialization) is projected orthogonally onto the already selected column set Ω of A, then the signal $\boldsymbol{x}$ and the residual $\boldsymbol{r}$ gets updated, resulting in with a more robust algorithm.

Algorithm 2 Orthogonal Matching Pursuit (OMP)

Input: Measurement vector $\boldsymbol{y}$, Measurement matrix A
Output: Sparse vector $\boldsymbol{x}$
Initialize: $\boldsymbol{x}^{[0]} \leftarrow 0, \boldsymbol{r}^{[0]} \leftarrow \boldsymbol{y}, i \leftarrow 0$, $\Omega \leftarrow [\,]$

1: **while** halting criterion false **do**
2: $i \leftarrow i + 1$
3: $c^{[i]} \leftarrow A^T r^{[i-1]}$ ▷ Compute the correlation
4: $j^{[i]} \leftarrow \underset{j}{\operatorname{argmax}} |c_j^{[i]}|$ ▷ Find the index of the most correlated
5: $\Omega^{[i]} \leftarrow \Omega^{[i-1]} \cup j^{[i]}$ ▷ Update the support set
6: $j^{[i]} \leftarrow \underset{j}{\arg\min} ||A_\Omega x^{[i]} - y||$ ▷ Solve least squares
7: $r^{[i]} \leftarrow y - Ax - c_{j^{[i]}}^{[i]}$ ▷ Calculate the new residual
8: **end while**
return $x \leftarrow x^{[i]}$

Although its computational complexity is somewhat higher than the MP algorithm, it still remains performative against convex relaxation. Nevertheless, OMP having greater computational complexity, and iterative nature, makes it impractical against large-scale operation. In order to combat this issue, several algorithms have been proposed that follow and develop upon the OMP algorithm. Namely Stagewise Orthogonal Matching Pursuit (StOMP) by Donoho in 2006 (D. L. Donoho, 2006b), Regularized Orthogonal Matching Pursuit (ROMP) by Needell in 2007 (Needell & Vershynin, 2010), and Orthogonal Multimatching Pursuit (OMMP) by Liu in 2010 (Liu & Temlyakov, 2010). One common property shared by among all these algorithms, is their multiple-element operation, whereas OMP was able to update one element at a time. This thesis will not go further into details of these algorithms, but the reader is advised to read the brilliant PhD Thesis of Needell (Needell, 2009), and the articles (Needell & Vershynin, 2009)(D. L. Donoho, 2006b)(Liu & Temlyakov, 2010) for further information on these methods.

2.7.1.3 Compressive Sampling Matching Pursuit (CoSaMP)

CoSaMP is first introduced by Needell and Tropp in 2008 (Tropp & Needell, 2010). Although OMP fixed a great deal of drawback of the MPs flaws, the area was still

relatively newborn and there was a lot of room for progress in the algorithmic sense. ROMP, improved upon the convex relaxation and greedy approaches. It has the advantages of a recovery guarantee (as BP does in the convex relaxation), and the fast nature of greedy approaches (Needell, 2009). Moreover, CoSaMP takes this improvement further, by reducing the error tolerances that is caused by stronger RIP property of ROMP.

CoSaMP is similar in a way with the aforementioned StOMP, ROMP, OMMP due to its multiple element operation. However, it is mildly different, in a way that it refines the selection of support elements at each iteration. This allows for index correction at later iterations. It starts by calculating the correlation product between the measurement vector $\boldsymbol{y}$ and the measurement matrix A, and calls it the signal proxy, unlike others that call the same step residual. Then, selects $2K$ elements that are of the highest value. selecting the mostly correlated $2K$ elements from this signal proxy. The support locations of the signal proxy, and the estimated signal vector $\boldsymbol{x}$ from the previous round are merged at every iteration. Through midway, a signal estimation is calculated again by least squares with this newly updated support set. Lastly, only the K number of elements get selected, and rest discarded. This signal is feeded back to the algorithm to form a new signal proxy to iterate further, unless a stopping criterion is reached.

Algorithm 3 Compressive Sampling Matching Pursuit

Input: Measurement vector $\boldsymbol{y}$, Measurement Matrix A,Sparsity level K
Output: Sparse vector $\boldsymbol{x}$
Initialize: $\boldsymbol{x}^{[0]} \leftarrow 0, \boldsymbol{r} \leftarrow \boldsymbol{y}, i \leftarrow 0, \Omega_{[0]} \leftarrow [\,]$

while halting criterion false **do**
  $i \leftarrow i + 1$
  $c \leftarrow A^T r$ ▷ Compute the correlation
  $\Omega^{[i]} \leftarrow \text{supp}\,(c_{2K})$ ▷ Index the largest 2K supports
  $W^{[i]} \leftarrow \Omega \cup \text{supp}\,(x^{[i-1]})$ ▷ Update the support set
  $x^{[i]} \leftarrow \underset{x,\text{supp}(x)\subset W^{[i]}}{\arg\min} ||A_\Omega x^{[i-1]} - y||_2$ ▷ Solve Least Squares
  $x^{[i]} \leftarrow x_K$ ▷ Prune everything but max K entries
  $r \leftarrow y - Ax^{[i]}$ ▷ Update the residual
end while
  return $x \leftarrow x^{[i]}$

Another algorithm by Dai (Dai & Milenkovic, 2009) that came out around the same time in 2008 as CoSaMP, where the strategies are very similar, is called *Subspace Pursuit* (SP). However, they calculate the signal estimation differently at the last stage, whereas CoSaMP uses a transition variable midway and keeps only the elements according to the support set, SP solves for a least squares second time. In the following studies, it has been shown that RICs $\delta_{3K} < 0.4859$ for SP and $\delta_{4K} < 0.5$ for CoSaMP suffice to guarantee a s-sparse recovery (Song et al., 2013).

2.7.2 Iterative Thresholding Algorithms

Convex optimization and greedy approaches are the two main pathways to solve for the main problem (Eq. 21). Whereas convex optimization is more convenient for recovery guarantees and error bounds but heavier by means of computation, greedy approaches are more computationally feasible, so they are faster in general, but somewhat lacks the recovery guarantee and require a lot more measurements than convex relaxation algorithms. There is another family of algorithms in greedy approaches called iterative thresholding algorithms or iterative shrinkage algorithms besides the pursuit type algorithms. This class of algorithms, although still being greedy, somewhat possess similar guarantees like the convex relaxation approaches. These algorithms follow a very similar scheme to try to solve the (Eq. 32) formulation of the problem, and can be shown as,

$$\boldsymbol{x}^{[i+1]} = {}^{H,S}T_K\left(\boldsymbol{x}^{[i]} + f\left(\boldsymbol{x}^{[i]}\right)\right) \qquad \text{(Eq. 33)}$$

where ${}^{H,S}T_K$ is a thresholding operator that operates on a vector **x** with K-nonzero elements, the subscript indicates hard thresholding for ${}^{H}T_K$ or soft thresholding for ${}^{S}T_K$, and $f(.)$ is some function that operates on the vector **x** at i^{th} iteration.

2.7.2.1 Iterative Hard Thresholding (IHT)

IHT algorithm, as stated by Eldar (Eldar & Kutyniok, 2012), is introduced for CS by Blumensath (Blumensath & Davies, 2008a) and Portilla (Portilla, 2009) seperately. The thresholding operator, essentially keeps K-largest entries of the vector that it is operationg on, and zeroes out the rest. Hence, the general formulation in (Eq. 34) takes the form,

$$\boldsymbol{x}^{[i+1]} = T_H\left(\boldsymbol{x}^{[i]} + \mathrm{A}^{\mathrm{T}}(\boldsymbol{y} - A\boldsymbol{x}^{[i]})\right) \qquad \text{(Eq. 34)}$$

and if M_i is the set of K-largest entries of $\boldsymbol{x}$ by its absolute value at i^{th} iteration, where,

$$T_H(x) = \begin{cases} \boldsymbol{x}_i & if\ |\boldsymbol{x}_i| \epsilon |\mathcal{M}_i| \\ 0 & otherwise \end{cases} \qquad \text{(Eq. 35)}$$

The algorithm is based for solving the (Eq. 35). In order to solve it, the algorithm calculates the gradient of the objective function at each iteration and adjusts oppositely. The objective function from (Eq. 32),

$$J(x) = \frac{1}{2}||\boldsymbol{y} - A\boldsymbol{x}||_2^2 \qquad \text{(Eq. 36)}$$

and its gradient,

$$-\nabla J(\boldsymbol{x}) = A^T(\boldsymbol{y} - A\boldsymbol{x}) \qquad \text{(Eq. 37)}$$

can be recognized in the pseudocode and the (Eq. 34) for estimating the signal.

Algorithm 4 Iterative Hard Thresholding (IHT)

Input: Measurement vector $\boldsymbol{y}$, Measurement Matrix A, Sparsity level K
Output: Sparse vector $\boldsymbol{x}$
Initialize: $\boldsymbol{x}^{[0]} \leftarrow 0, i \leftarrow 0$

1: **while** halting criterion false **do**
2: $i \leftarrow i + 1$
3: $x^{[i]} \leftarrow x^{[i-1]} + A^T(y - Ax^{[i-1]})$
4: $x^{[i]} \leftarrow {}^{H}T_K^{[i-1]}(x^{[i]})$
5: **end while**
return $x \leftarrow x^{[i]}$

Inherently, a thresholding operator acting upon the resulting vector at each iteration also makes sure the constraint is fulfilled. It has been later proven by Blumensath that RIC of $\delta_{2k} < \frac{1}{\sqrt{8}}$ guarantees an exact recovery of k-sparse signal, which is similiar to that of BP (Blumensath & Davies, 2008b). Moreover, the algorithm is guaranteed to converge where the measurement matrix A has its Euclidian norm $||\boldsymbol{A}||_2 < 1$.

2.7.2.2 Iterative Soft Thresholding (IST/ISTA)

Fornasier state that several authors independently suggested IST algorithm (Fornasier, 2010). Iterative soft thresholding is motivated to solve the unconstrained form of the system, described in its Lagrangian form as (Eq. 28),

$$\mathcal{L}(\boldsymbol{x}, \lambda) = \min_{\boldsymbol{x} \in R^N} ||\boldsymbol{y} - A\boldsymbol{x}||_2^2 + \lambda \big|\big|\boldsymbol{x}\big|\big|_1 \qquad \text{(Eq. 38)}$$

The problem is then restated as (Fornasier, 2010),

$$\min_{\boldsymbol{x} \in R^N} ||\boldsymbol{y} - A\boldsymbol{x}||_2^2 + 2\tau \big|\big|\boldsymbol{x}\big|\big|_1 \qquad \text{(Eq. 39)}$$

where $\lambda = 2\tau$. In order to solve the unconstrained system above, Landweber iteration step (Eq. 34) is utilized which regularizes the ill-posed problem

$$\boldsymbol{x}^{[i+1]} = T_S\left(\boldsymbol{x}^{[i]} + \mathrm{A}^{\mathrm{T}}\boldsymbol{y} - A^T A\boldsymbol{x}^{[i]}\right) \qquad \text{(Eq. 40)}$$

$$T_S(\boldsymbol{x}) = \begin{cases} \boldsymbol{x} - \tau & \boldsymbol{x} > \tau \\ 0 & |\boldsymbol{x}| \leq \tau \\ \boldsymbol{x} + \tau & \boldsymbol{x} < \tau \end{cases} \qquad \text{(Eq. 41)}$$

where T_S is defined as solft-thresholding operator which acts elementwise. It is proven to converge to the objective function, required that $\big|\big|A\big|\big| < 1$ (Daubechies et al., 2003), as it does also in IHT.

Algorithm 5 Iterative Soft Thresholding (IST)

Input: Measurement vector $\boldsymbol{y}$, Measurement Matrix A,
Output: Sparse vector $\boldsymbol{x}$
Initialize: $\boldsymbol{x}^{[0]} \leftarrow 0, i \leftarrow 0$

1: **while** halting criterion false **do**
2: $i \leftarrow i + 1$
3: $x^{[i]} \leftarrow x^{[i-1]} + A^T(y - Ax^{[i-1]})$
4: $x^{[i]} \leftarrow ST^{[i-1]}(x^{[i]})$
5: **end while**
return $x \leftarrow x^{[i]}$

It is also possible to add an adaptive descent parameter to ensure convergency. Thus, the Landweber iterative step in (Eq. 40) can be modified as,

$$\boldsymbol{x}^{[i+1]} = T_S\left(\boldsymbol{x}^{[i]} + \beta^{[i]}\mathrm{A}^{\mathrm{T}}(\boldsymbol{y} - A\boldsymbol{x}^{[i]})\right) \qquad \text{(Eq. 42)}$$

where β can used as a tuning parameter.

CHAPTER 3

METHOD

This section will include the methodology of the proposed study. First, the acquisition setup for the fiber speckles will be explained. Second, the sensing matrix classes will be introduced and the utilization of the aforementioned speckles for imaging will be illustrated. At last, the performance metrics to be used will be described.

3.1 Introduction

Until 1980s, communications had to rely mostly on line wires which were susceptible to considerable loss. In order to deal with this issue, coaxial cables were to be used, however, this would still render them trapped in short distance. Upon the emergence of fiber optics, communications field saw tremendous development, by the fiber optics remarkable performance concerning the bandwidth and loss. Moreover, fibers could support multiple channels, called modes, which allow them to transfer numerous spatial information, in which case it is referred as multimode fiber (MMF). This surge in 1980s, also benefited a lot of other fields that saw the advantages and applicability, hence enjoyed the vast research and declining cost of fiber optics. One such research area is imaging.

Since then, various attempts have been made to exploit the fiber optics for imaging (Yariv, 1976) (Fischer & Sternklar, 1985) and still are being made (Papadopoulos et al., 2012) (Čižmár & Dholakia, 2012) (Plöschner et al., 2015) (Amitonova & de Boer, 2020).

This chapter describes the methodology, from speckle acquisition to, in the end, using these speckles to image a 2D scene. The scheme, as hinted before, uses SPC structure.

3.2 Fiber Speckle Acquisition Scheme

When an image is given as input at the one end of MMF, the light is coupled to the multiple spatial modes of the fiber. This in turn, results in a seemingly unrelated random pattern as an output at the other end. The main reason for this, stems from the fact that different modes, allow for different velocities (Fedor Mitschke, 2016), and referred as *modal dispersion*. This results with the input image getting its amplitude and phase mixed (Caramazza et al., 2019). Furthermore, any environmental changes have considerable perturbation on the output as well. Therefore, the resulting output at then other end shows up as scrambled, and called as *speckles*. This phenomenon is not exclusively observed for fibers. Speckle formation can be found in any coherent imaging modality, from ultrasound to tomography, especially the most obvious are in systems that involve lasers (Duncan & Kirkpatrick, 2008). At the entrance of the MMF, the image is transmitted along fiber as E-field,

$$E_{in} = Ae^{i\Phi} \qquad \text{(Eq. 43)}$$

where A is the amplitude, Φ is the phase and E_{in} is the E-field at the input. It reaches out to the exit of the MFF, and captured on CCD as an intensity distribution as,

$$I_{out} = |E_{out}|^2 = |T * E_{in}|^2 \qquad \text{(Eq. 44)}$$

where I_{out} is the output intensity, E_{out} is the E-field at the output and T is the transmission matrix of the MMF.

The experimental setup to generate speckles at the output by providing everyday imagery at the input, is depicted in Fig. 7 as below.

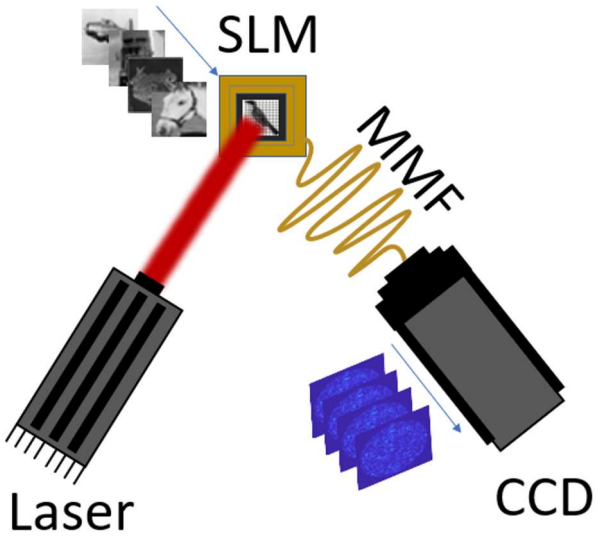


Figure 7. Schematics of the speckle acquisition setup

Fig. 7 shows a diode laser of 671 nm is used as a coherent light source. The laser beam then illuminates a spatial light modulator (SLM) of 1280x768 pixels. A directory of everyday images consisting of 1000 images from Cifar-10 (Krizhevsky, 2009) are fed and written on the SLM sequentially as 32x32 greyscale images. The image-set is consisting of mostly by animal and vehicle pictures. A set of sample between the input images and the corresponding output speckles are given below.

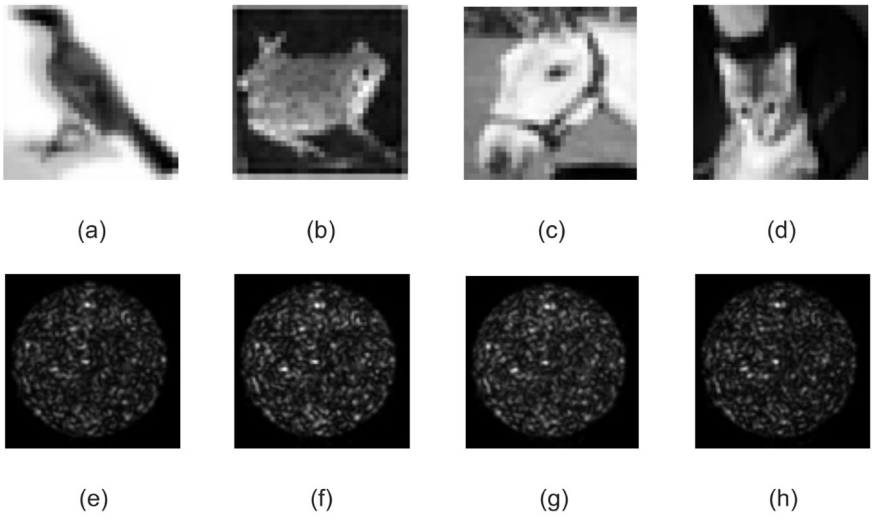


Figure 8. a-d) Example random animal themed images from the image dataset e-h) corresponding fiber speckle outputs

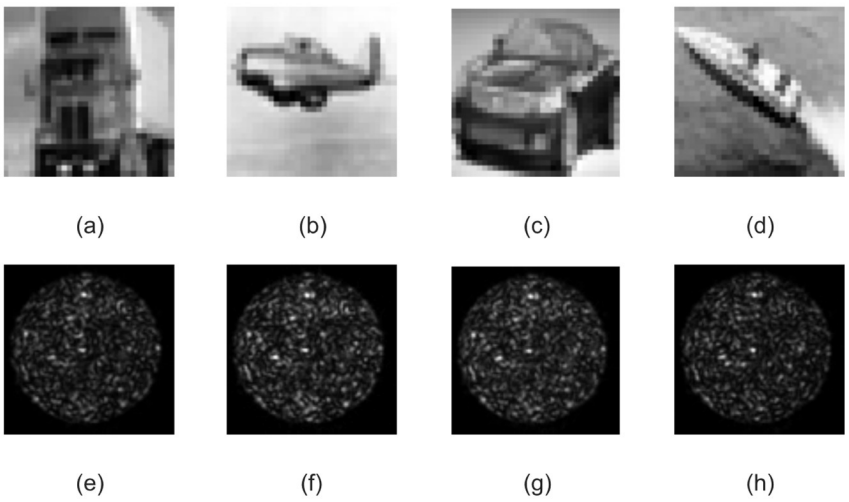


Figure 9. a-d) Example random vehicle themed images from the image dataset e-h) corresponding fiber speckle outputs

The images are written by modulating the amplitude and phase of the incident laser light. The images are then conveyed by MMF of 2 meters, which was set safe and secure from outside perturbations. The MMF has diameter of 105 μm, and numerical aperture (NA) of 0.22 . The resulting speckle patterns formed at the output end, are then acquired by a CCD camera of resolution 1024x768 pixels, performing in visible range. The corresponding speckle patterns are all collected, and archived accordingly.

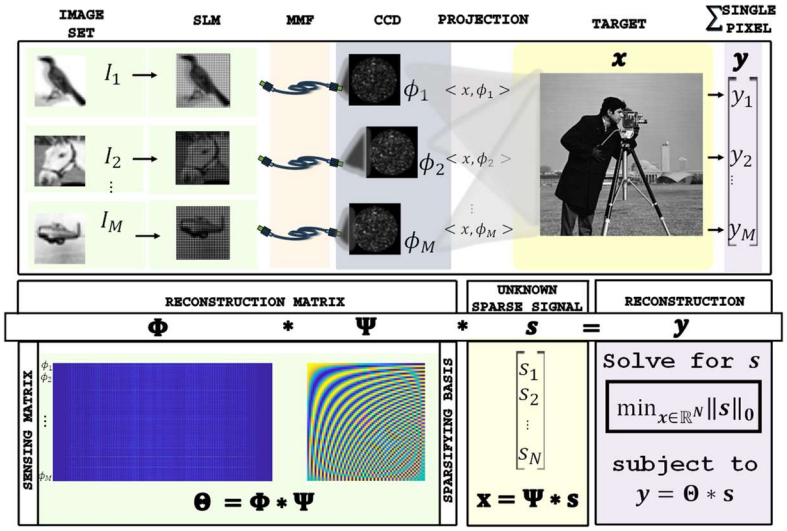


Figure 10. The general scheme of the experiment illustrated

Fig. 10 illustrates the framework of the general progression of the study. The Image set feeds the SLM where I_M denotes the images. After passing through MMF, CCD collects the speckles as ϕ_M. Then, the measurement $\boldsymbol{y}_i$ is calculated that corresponds to the inner product between the scene $\boldsymbol{x}$ and the speckle pattern ϕ_i. This is equivalent to projection of speckle patterns onto the scene, then acquisition by a single-pixel detector (SPD). Proceedingly, the measurement vector $\boldsymbol{y}$ is built according to these projection calculations. Constructing the sensing matrix Φ from the projection speckle patterns ψ_i, and choosing a proper sparsifying basis for the task Ψ, enables the calculation of the reconstruction matrix Θ. At last, an algorithm is chosen to solve the approximation or relaxation for the original ℓ_0-problem, in the pursuit of solving for the sparse coefficient vector $\boldsymbol{S}$.

3.3 Sensing Matrix Construction

There are two important aspects to consider in CS scheme. First, choosing or building a measurement matrix that complies with the conditions of the CS theory. Second, choosing a reconstruction algorithm that best fits the problem at hand. One can classify sensing matrices into two groups, deterministic and random matrices. In general, random matrices often have small RIC values and more robust. However, deterministic matrices are difficult to design as they are NP-Hard to check for the RIP condition. Aside from random matrices, most commonly the employed structured illumination bases are Hadamard, speckle (i.e., Laser speckles (Devaux et al., 2016), (Ferri et al., 2005), Fourier, Haar, LeGall and Daubechies (Angelo et al., 2018).

3.3.1 Random Matrices

The realization of randomness into signal processing for sparse modelling context, came just before the formal introduction of the CS. The great contribution came from Candes, Romberg and Tao in June 2004 (E. Candes et al., 2005) as acknowledged

by Donoho in his CS paper of September 2004 (D. Donoho, 2006). The striking results include the random matrices with independent and identically distributed entries are proven to be hold the RIP property with very high probability, and with a measurement amount relation as $N > M \geq O(Klog\left(\frac{N}{K}\right))$.

The sensing matrix Φ is drawn from a Gaussian such that its entries are independent normal variables, with zero mean and $1/M$ variance.

$$\phi_{i,j} \sim \mathcal{N}(0, \frac{1}{M}) \qquad \text{(Eq. 45)}$$

where the entries $\phi_{i,j}$ are independent realizations of Gaussian random variables (Baraniuk et al., 2006).

Bernoulli drawn sensing matrix Φ can also be created with a similar strategy as,

$$\phi_{i,j} := \begin{cases} +\frac{1}{\sqrt{M}} & with\ probability\ \frac{1}{2} \\ -\frac{1}{\sqrt{M}} & with\ probability\ \frac{1}{2} \end{cases} \qquad \text{(Eq. 46)}$$

One can often find these random matrices turned to binary in order to be usable in the CS context. If a matrix is drawn from a Bernoulli distribution, where it has its entries $\{-1, +1\}$, it can be found to be referred as Rademacher matrix or signed Bernoulli matrix. One can also encounter the value set converted to $\{0, +1\}$ as it is more convenient to use for some devices. Application wise, a lot of devices are binary friendly such as DMD and Bernoulli distribution is favored over Gaussian due to its binary base which can be represented in 1 bit. Considering hardware/optics implementation friendliness, sensing matrix consisting of set of values $\{-1, 0, +1\}$ is essential (Do et al., 2012). Unfortunately, this reduces the RIP condition of a matrix (Chandar, 2008). Nevertheless, it does not necessarily mean that using these random matrices are fruitless. The recovery is still achievable, however with weaker conditions.

3.3.2 Deterministic Matrices

Although random matrices are robust, and well defined in terms of recovery, randomness puts a heavy weight on the computational side. As the dimensions increase, the required storage becomes several gigabytes as Gaussian and Bernoulli matrices computational storage cost is $O(MN)$ (Calderbank et al., 2010). Furthermore, the needed processing power reaches the order of giga-flops, even at a few hundreds of dimensions. Due to this issue, deterministic matrices become more appealing as they are more machine friendly. Tao has posted about the need for structured matrices in CS around 2007, making it explicitly known (Tao, 2007) to the following community.

Deterministic matrices are easily implemented, and highly reproducible compared to the random matrices. Hadamard basis has already been in use since 1960s, especially in space communication protocols for image transfer (Beser, 1994) and regarded as being the option with one of the best reconstruction accuracies compared to the other matrices (Zhuoran et al., 2013). In the context of single pixel imaging with CS, Hadamard transform is often used to create deterministic measurement matrix as it is a simpler version of the Fourier transform that does not involve multiplication or division. A self-adjoint orthogonal Hadamard matrix can be put as,

$$H_0 = 1 \qquad \text{(Eq. 47)}$$

$$H_q = \frac{1}{\sqrt{2}}\begin{pmatrix} H_{q-1} & H_{q-1} \\ H_{q-1} & -H_{q-1} \end{pmatrix} \qquad \text{(Eq. 48)}$$

where $N = 2^q$ and q is a positive integer. Moreover, by replacing the negative values in the Hadamard matrix results with a matrix called *semi-Hadamard matrix* (Zhang et al., 2010). Furthermore, the Hadamard matrices are often not used directly, but they are rather reordered in a specific strategy to get better performances. Although the matrices have the same number of ones and zeros, their rearrangement allows for a better computation time. The often-encountered strategy called *Walsh-Hadamard matrix* encourages the frequency ordering, which also have a fast transformation and

gives the best results in most of the cases (Zhuoran et al., 2013). Author encourages further interested readers to the recent article about comparison of Hadamard orderings by Vaz *et al.* (Vaz et al., 2020).

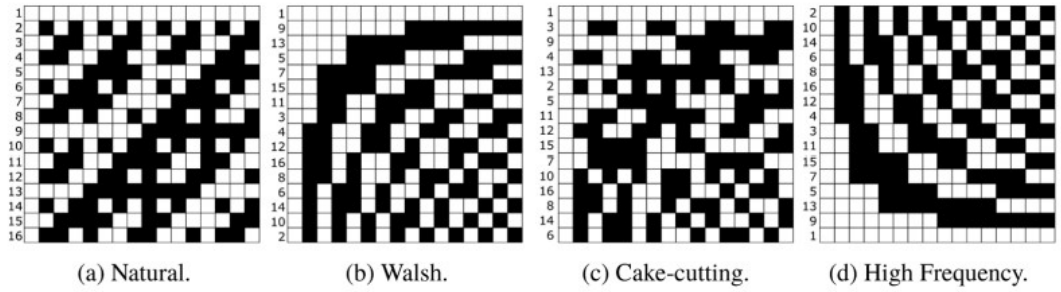


Figure 11. Hadamard orderings (Vaz et al., 2020)

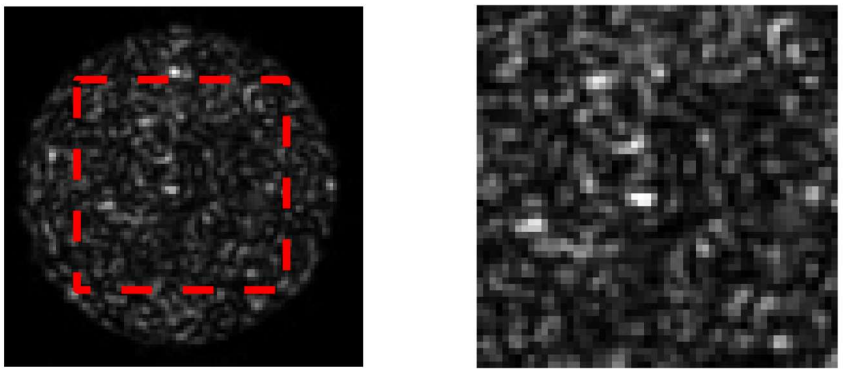
Partial Fourier Ensembles are another popular deterministic approach to constructing a sensing matrix. This class of matrices enjoy the same fast transformation availability as Walsh-Hadamard matrices. Fast Fourier Transform algorithm and its inverse can be utilized to benefit computational complexity. The partial Fourier Ensembles also enjoy a similar guarantee for unique recovery. For exact reconstruction, the required measurement amount first given by Candes *et al.* (E. Candes et al., 2005) as $M = \mathrm{O}(\mathrm{k}\log^6 N)$, then improved by Rudelson and Vershynin as $M = O(k\log^4 N)$ (Rudelson & Vershynin, 2008). Also, the memory cost for a Fourier matrix is given as $O(M\log N)$ (Calderbank et al., 2010). More detailed recovery conditions on Fourier Enbsembles are given in the articles by Candes *et al.* (E. Candes et al., 2004) and Rudelson *et al.* (Rudelson & Vershynin, 2006). The discrete Fourier matrix $\mathcal{F}$ is constructed with the below formulation as,

$$\mathcal{F}_{j,k} = \frac{1}{\sqrt{N}} e^{\frac{2\pi(j-1)(k-1)}{N}} \qquad \text{(Eq. 49)}$$

where $j,k \in [N]$. In CS context, the sensing matrix of size $M \times N$ $(M < N)$ is constructed by choosing M rows from the previously mentioned full-rank matrices at random. Hence, the new matrix is often referred as *partial*.

The last broad class of matrix is, as one of the key questions of this thesis, *incoherent ensemble*. Choosing any orthonormal matrix of size $N \times N$ matrix that is incoherent

enough to use as sampling projections, in other words, qualifies as a good sensing matrix. Here, the question is whether the fiber speckle patterns that are produced by every-day still imagery from Cifar-10 image set, can qualify as a sensing matrix that is capable of reconstruction of a 2D scene with CS. Two sets of sensing matrices Φ are created to examine the discrete edge effects. First, the raw speckle size of 96×128 that is recorded through CCD camera is reduced to 96×96. Downsizing medially, clears the image from obvious linearly dependent column vectors which contain no valuable information. Second, the resulting 96×96 image involves a circular speckle and outside of the circle is blank (See Figure 11a). In order to increase incoherence, another set of sensing matrix from speckles of size 56×56 is constructed (Figure 11b).



(a) (b)

Figure 12. a) Full-size speckle with rectangular area showing b) second set of speckles that shows the cropped section of the full image in (a)

By using the speckle patterns, the created sensing matrices Φ are shown in color scaled form in Fig. 12. Each row ϕ_j of the sensing matrix Φ corresponds to a unique speckle pattern. The scaled color enables to observe the intensity distributions in the speckle patterns more conveniently, compared to grayscale. One can observe the redundant columns of the sensing matrix Φ that is created with the full sized speckles of size 96×96 in Fig. 12a. These redundant columns at the beginning and the end are caused by the rectangular speckle image having circular speckle distribution.

Moreover, a vertical line trend can be seen through the whole image. Even around the center, a similar intensity level can be observed for most of the images. Meaning, regardless of the image, the speckles tend to share similar speckle intensity at the same points.

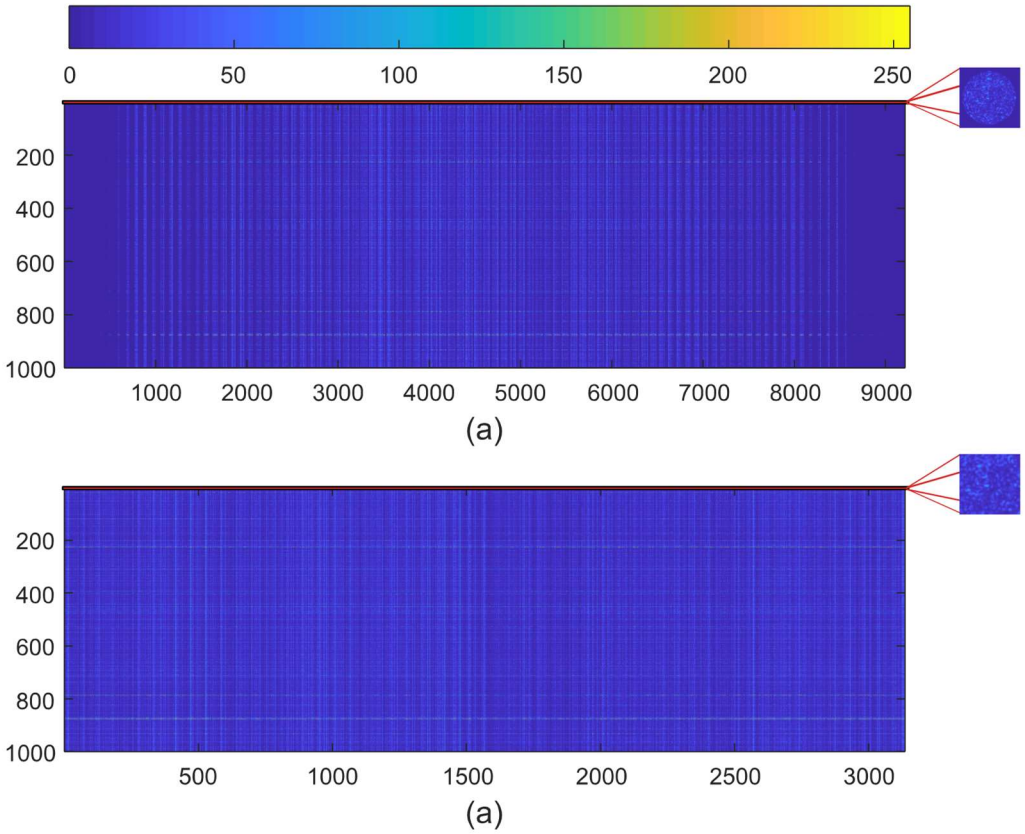


Figure 13. Sensing matrices utilizing a) full-size speckles b) partial speckles

3.4 Speckle Projections

In the context of SPI, the structured projections are collected by (SPD). A single projection that falls onto the 2D scene, gives rise to a single measurement reading on a SPD. This single measurement is the inner product of the projection and the image. A 2D image of size $N \times N$ is converted into a single vector by stacking the columns and denoted as $\boldsymbol{x} \in \mathbb{R}^{N^2 \times 1}$. Further, the speckle pattern is denoted by $\boldsymbol{\phi}_j \in \mathbb{R}^{1 \times N^2}$ for the j^{th} row. This can be shown as,

$$y_j = < \boldsymbol{x}, \phi_j > = \phi_j^T \boldsymbol{x} \qquad \text{(Eq. 50)}$$

where a single intensity measurement y_j is collected. Provided with sequential illumination, the above expression becomes,

$$\boldsymbol{y} = \Phi \mathbf{x} + \boldsymbol{\epsilon} \qquad \text{(Eq. 51)}$$

where $\boldsymbol{\epsilon}$ is the noise vector. The signal $\boldsymbol{x}$ is assumed to be sparse in some specified basis $\Psi \in \mathbb{R}^{N \times N}$. The projected patterns need to be inchorent with the basis Ψ.

$$\boldsymbol{y} = \Phi \Psi \mathbf{s} \qquad \text{(Eq. 52)}$$

where $\boldsymbol{s} \in \mathbb{R}^{N^2 \times 1}$ is the underlying sparse coefficient vector. One of the advantages of SPC with CS is, after the measurement vector $\boldsymbol{y} \in \mathbb{R}^{M \times 1}$ acquired, reconstruction is made computationally afterwards. One does not have to choose basis, or the reconstruction algorithm right away at the acquisition phase. This allows for the reconstruction of $\mathbf{x}$ from previously taken data $\boldsymbol{y}$, with the new state of the art reconstruction algorithms, or newly realized a better choice of basis.

The measurement vector $\boldsymbol{y}$ is constructed with the equations above, by calculating the inner products between the speckle patterns and the image. The common test pictures from image processing community are used. Afterwards, the reconstruction matrix $\Theta = \Phi\Psi$ is built with DCT chosen as the basis Ψ. Among most common choices of bases DCT, Fourier and Wavelets; DCT is fairly straightforwad and provides a relatively good performances for the imaging purposes. The reconstruction algorithms that are mentioned in the previous sections are applied. Namely, ℓ_1 and ℓ_2 minimization as convex relaxations, and OMP, CoSaMP, IST as the greedy alternatives. ℓ_1-magic package has been used to solve the ℓ_1-norm minimization problem. The reconstruced images have their resolution limited according to the fiber speckle resolutions. 56×56 and 96×96 images are reconstructed with the two sensing matrices are given in Fig. 13. Measurement amount of 600, 800, 1000 are trialed and their performances compared.

3.5 Image Quality Metrics

Starting from the 1950s, growing technological advancements such as microprocessors and Silicon technology, has led digital imaging to its peak. Moreover, last couple of decades has a great deal of developments in artificial intelligence (AI). Subfields of AI, such as machine vision and computer vision further ignited the progress of digital imaging.

In order to assess the quality of the reconstruction, there are several metrics that is used in the field. One of the most basic metrics in use is the mean squared error (MSE) as,

$$MSE(\boldsymbol{x}, \hat{\boldsymbol{x}}) = E((\boldsymbol{x} - \hat{\boldsymbol{x}})^2) \qquad \text{(Eq. 53)}$$

where $E(\cdot)$ is the expected value operator. However, this metric utilizes a squared term, which amplifies larger errors compared to the small ones. In order to make it the same units with the original signal, the MSE value is square rooted, and the new metric is called root mean squared error (RMSE). It describes how much spread the reconstructed values are compared to the their true positions or original signal. RMSE is formulated as,

$$RMSE(x, \hat{x}) = \sqrt{E((\boldsymbol{x} - \hat{\boldsymbol{x}})^2)} \qquad \text{(Eq. 54)}$$

Another one of the most common metrics is called peak signal-to-noise ratio. It is often encountered in compression of media and describes the ratio of maximum possible power of the signal and the noise present. PSNR is given as,

$$PSNR(\boldsymbol{x}, \hat{\boldsymbol{x}}) = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right) \qquad \text{(Eq. 55)}$$

where MAX is the maximum possible pixel value. In this thesis, the images are calculated as grayscale 8 bits, which have the corresponding MAX value of 255. Another metric that has been proposed (Wang et al., 2004) relatively recent called structural similarity index measure (SSIM), which takes luminance (l), contrast (c) and structure (s) into account for comparison,

$$SSIM(x, \hat{x}) = [l(x, \hat{x})]^{\alpha} \cdot [c(x, \hat{x})]^{\beta} \cdot [s(x, \hat{x})]^{\gamma} \quad \text{(Eq. 56)}$$

where $\alpha, \beta, \gamma > 0$ are tunable parameters. This metric differs from the aforementioned methods, it has a normalized range. Whereas the RMSE and PSNR have different range values for comparison, the SSIM has a fixed range of $[0,1]$, so the perfect score (i.e. $SSIM = 1$) is only attained when the two signals are perfectly equal (i.e. $\widehat{\boldsymbol{x}} = \boldsymbol{x}$).

CHAPTER 4

RESULTS AND DISCUSSIONS

4.1 Introduction

This chapter gives the simulation results for CS reconstruction, using fiber speckles as projections onto the conventional 2D test images.

4.2 Reconstruction

The measurement vector $\boldsymbol{y} \in \mathbb{R}^{M\times 1}$ is acquired by the sensing matrix projections $\phi_j \in \mathbb{R}^{1\times N}$. Reconstructions are performed with the algortihms in Section 2.6. Convex relaxation algortihms ℓ_1 minimization and ℓ_2 minimization are trialed for performance. For applying ℓ_1 minimization, the MATLAB package called ℓ_1-magic(E. J. Candès, 2005) , created by Candes and Romberg is utilized to solve (Eq. 23) with the parameters $iter = 50$ and $pdtol = 10^{-3}$. Least squares solution is employed for the ℓ_2-minimization. The standard described OMP and CoSaMP algorithms are used for pursuits. Lastly, ISTA is trialed for greedy iterative class as described in (Bayram, 2016).

4.2.1 Different Speckle Size Reconstruction Comparison

Initially, the inherent sparsity value for the signal is chosen as $K = M/4 = 250$. The resolution of the reconstructed image is dependent on the spatial resolution of the speckle projection. Pursuits are iterated K times, whereas the thresholding is iterated at specified times. All three (OMP, CoSaMP and ISTA) are iterated 250 times for the below reconstructions in Fig. 14 and Fig. 15, and regularizer parameter

chosen as $\lambda = 1$ for ISTA. Fig. 14 shows the reconstruction results for 96×96 speckle projections.

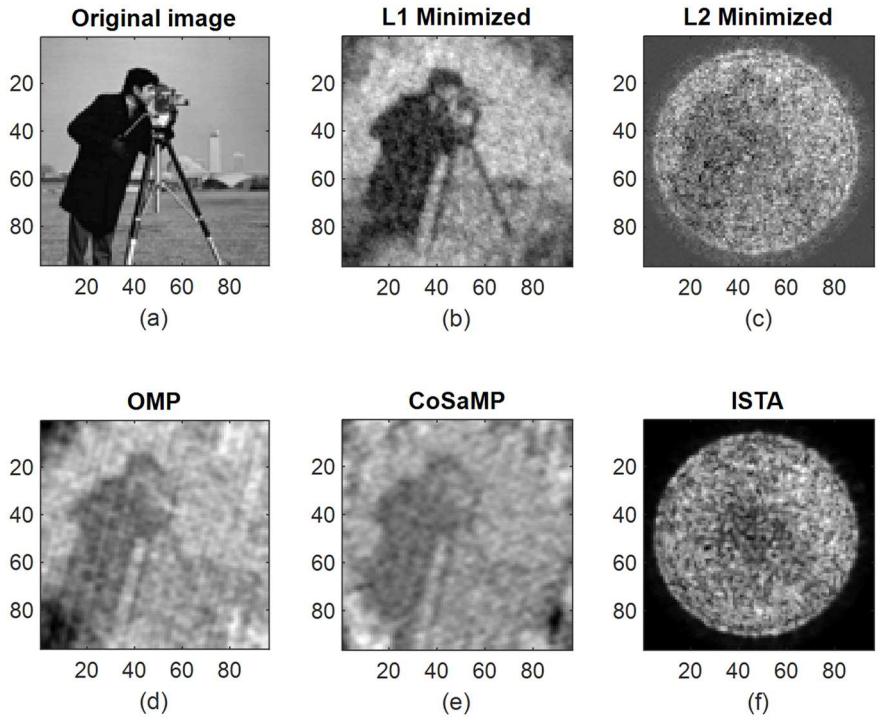


Figure 14. Reconstruction results for cameraman image created with 96x96 speckle patterns (N = 9216, M = 1000, K = 250)

Table 1. Performance metrics for Fig. 14

	ℓ_1 Min	**ℓ_2 Min**	**OMP**	**CoSaMP**	**ISTA**
RMSE	39.86	91.89	71.82	64.72	101.17
PSNR	16.12	8.87	11.01	11.91	8.02
SSIM	0.222	0.081	0.157	0.164	0.005

The reconstruction results in Fig. 14 show that ℓ_1-minimization gives the best result compared to the other strategies followed. Since ℓ_2 is a power keeping norm, the reconstruction dispersed the energy at the center around, and nearly zeroed out all the information outside of the circular ring. The pursuit algorithms expectedly proceeded by the correlation, which explains their partial success. Adding the

column vectors to reconstruct the image, resulted in a more distributed, noncircular formation. CoSaMP gave better performance by pruning the solution vector along the algoritm at each step. Hence, OMP seems to have more texture of high frequency, which is non-existent in the CoSaMP reconstruction. ISTA seemingly underperformed. The sensing matrix choice is important for thresholding algorithms and $||\Phi||_2 < 1$ is required for convergence to a solution. Even though the RMSE and PSNR values are close to the ℓ_2-min algorithm, SSIM shows ℓ_2-min performs much better by mean of structural similarity with the original. This can also be verified by visual inspection of the two.

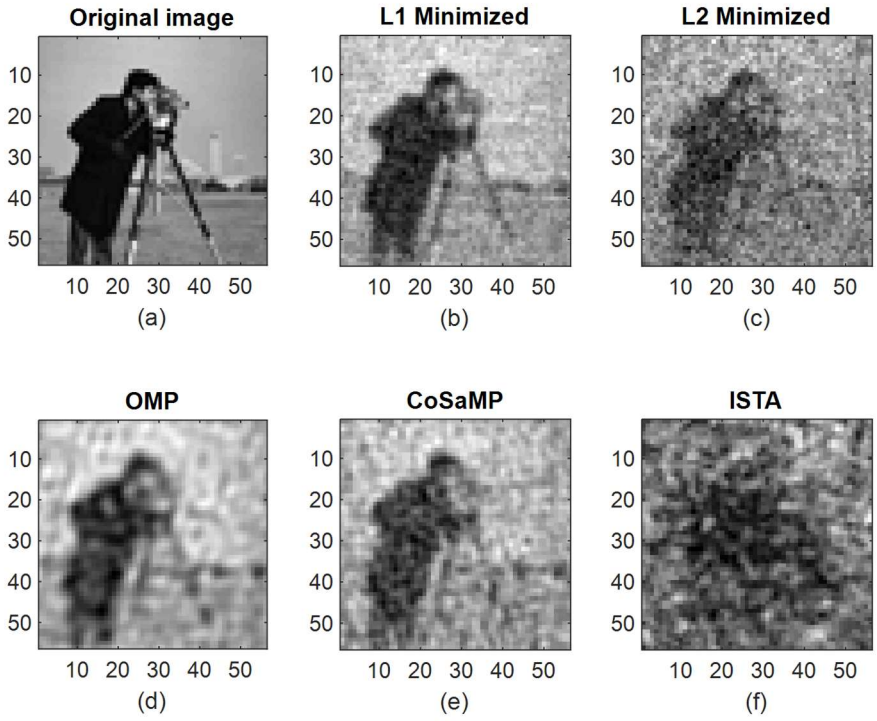


Figure 15. Reconstruction results for cameraman image created with 56x56 speckle patterns (N = 3136, M = 1000, K = 250)

Table 2. Performance metrics for Fig. 15

	ℓ_1 Min	**ℓ_2 Min**	**OMP**	**CoSaMP**	**ISTA**
RMSE	20.07	35.49	24.28	30.05	54.51
PSNR	22.08	17.13	20.43	18.57	13.40
SSIM	0.408	0.232	0.353	0.294	0.001

One can deduce from the Table 2, the reconstruction by 56×56 speckle projections, overall benefited all the figures except the one by ISTA. The sparsity value K and the measurement amount M are kept the same but N has decreased. At the heart of the CS theory, $M \sim O(K \log N)$ is sustained, hence the required measurement amount M also decreased. Yet, the recoveries took place with the same M value. In other words, the requirements for a succesfull recovery got lowered, and obtaining more than minimum measurement amounts improved overall almost all the images. Furthermore, by only taking the centeral region for projection, helped to remove monotonous parts. These monotonous parts were causing the sensing matrix Φ to be more coherent. Removing these, has led to an increased incoherency among the projection vectors of the sensing matrix Φ and the image. Almost all the reconstructions are doubled in quality, except ISTA. Thresholding and shrinkage algorithms already require high amounts of iterations. Morever, they try to keep the highest values. The downside of the iterative algorithms as mentioned in Section 2.6, they are mostly oblivious to the reconstruction. The PSNR has improved and the RMSE has reduced, which in general indicate improvement. However, their blind structure do not take the relatability in the process of reconstruction.

4.2.2 Reconstruction Performance with Different Measurement Amounts

The reconstruction performance is directly affected by the measurement amount, as it tries to reconstruct the full size image with lesser information. Surprisingly, losing

almost 7% of the total information, does not seem to affect the visible image quality besides CoSaMP.

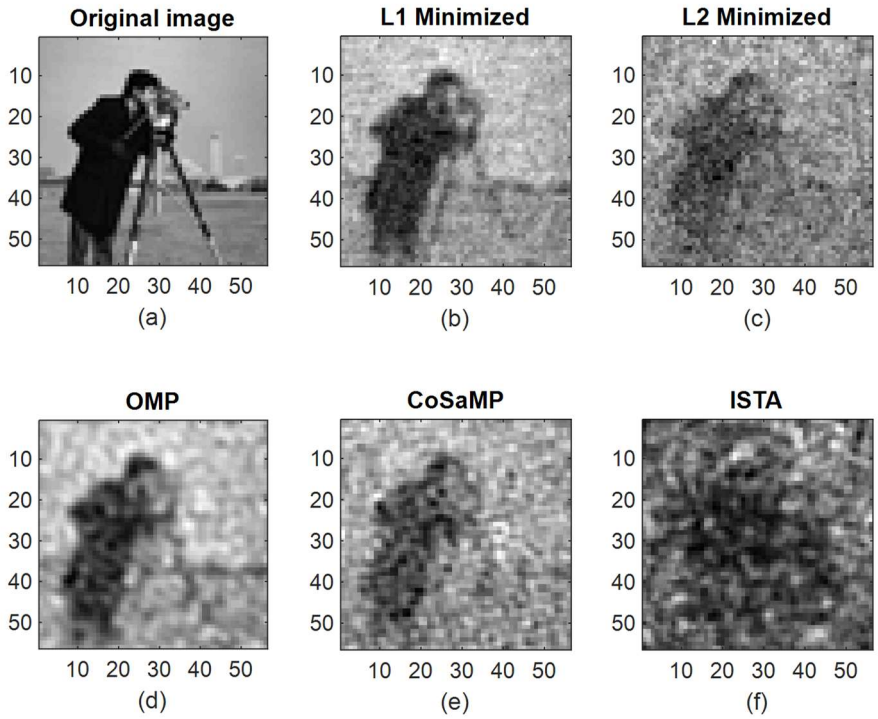


Figure 16. Reconstruction results for cameraman image created with 56x56 speckle patterns (N = 3136, M = 800, K = 250)

Table 3. Performance metrics for Fig. 16

	ℓ_1 Min	**ℓ_2 Min**	**OMP**	**CoSaMP**	**ISTA**
RMSE	22.76	38.23	26.53	40.20	55.46
PSNR	20.99	16.48	19.65	16.05	13.25
SSIM	0.361	0.200	0.303	0.221	0.004

The most affected algorithm by the measurement amount M seems to be CoSaMP. By going from $M = 800$ to $M = 600$, the reconstruction got obsolete. This may be caused by how close the approximity of the values M and K become. Since the algorithm works by selecting $2K$ number of most correlated values, and then pruning

the rest, it got stuck with the same pool of values. Since $2K$ and M are close values, there is not much new information pool to better the reconstruction.

The convex relaxation method of ℓ_1-minimization, still outperforms the other algorithms and gives a very recognizable reconstruction even at the levels of 20% of the available information.

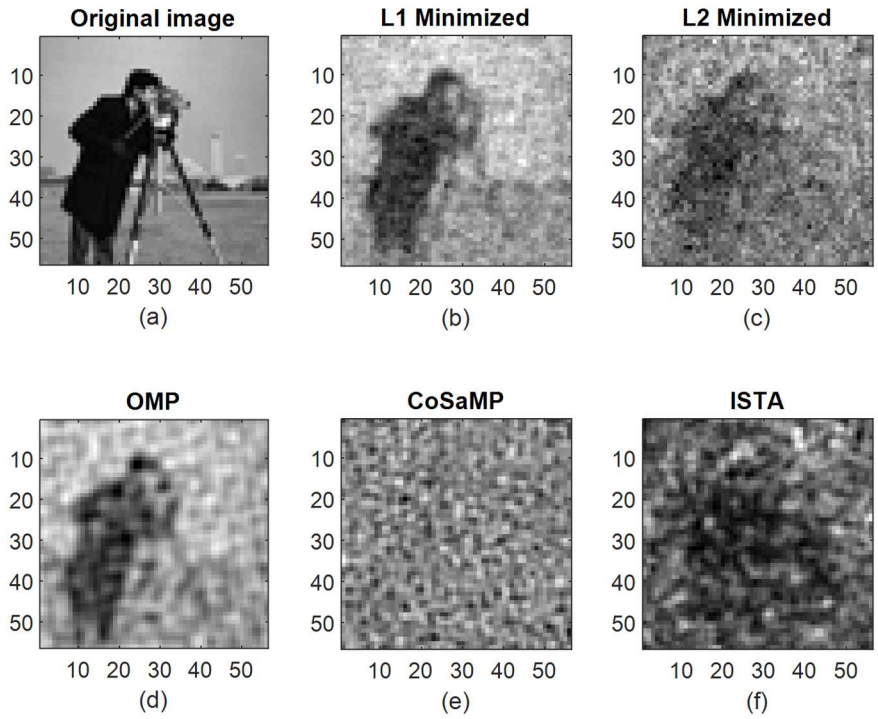


Figure 17. Reconstruction results for cameraman image created with 56x56 speckle patterns (N = 3136, M = 600, K = 250)

Table 4. Performance metrics for Fig. 17

	ℓ_1 Min	**ℓ_2 Min**	**OMP**	**CoSaMP**	**ISTA**
RMSE	25.55	41.55	30.49	353.65	56.001
PSNR	19.98	15.76	18.45	-2.84	13.17
SSIM	0.306	0.165	0.277	0.008	0.006

CHAPTER 5

CONCLUSION

One of the main aims of this study was to investigation and validation of the proof-of-concept for the imaging 2D with optical fiber optics using CS. Among various sensing matrix strategies including random and deterministic, the applicability of the fiber speckles has been explored. Various reconstruction algortihms are studied, and applied to reconstruct within the SPI scheme. These reconstructions later compared with different image quality metrics.

The setup has been built and improved to acquire the fiber optic speckles of everyday still imagery after passing through 2 m MMF. These still imagery are produced by illuminating the related SLM pattern by a laser of 671 nm. These speckles are then acquired via a CCD camera.

The accumulated speckles are then utilized to form a sensing matrix. The measurement is determined by the product between the projected speckles and the 2D image. Several methods are trialed to investigate the better reconstruction algorith. Among these algorithms, convex relaxation prevailed to be the reconstruction with the most fidelity and quality.Several metrics including RMSE, PSNR and SSIM are employed to compare, and ℓ_1 minimization is especially highlighted as the algorithm with the best results. The computational complexity of the convex methods are repeatedly emphasised.

This work has enabled to show a succesfull demonstration of the fiber optic SPI. It has been recently experimentally shown by utilizing the modal structure of the MMF fiber, and raster scanning the input facet CS can be used to reconstruct a scene (Amitonova & de Boer, 2020). However, without raster scanning the input face and directly exposing with different patterns still allow for a decent recovery of the output facet. The findings in this study shows, although the sensing matrix structure

shows the excitation of the same vertical vectors the recovery is attainable even without prioritizing the excitation or illumination of the whole modal structure at the end of the fiber.

The utilization and the applicability of the fiber optic speckles in SPC and CS context has been shown. It can be further supported with the 2D experimental data. Different bases can be trialed to explore the existance of a better reconstruction matrix. Precise calculations that use advanced SPDs can further improve the image quality and timing. Moreover, the same SPI scheme with fibers can be utilized to achieve depth reconstruction, similar to the CS LiDAR applications with structured illumination. Moreover, using two or more SPDs at different positions can provide reflection and 3D imaging. Filters or prisms can be utilized for color imaging at endoscopic scale. Ghost imaging can be realized to further explore the imaging possibilities of speckles.

REFERENCES

Ament, S. (2017). *CS6220 Lecture Notes: Basis Pursuit De-noising, LASSO, and Compressed Sensing*.

Amitonova, L. v, & de Boer, J. F. (2020). Endo-microscopy beyond the Abbe and Nyquist limits. *Light: Science & Applications*, *9*(1), 81. https://doi.org/10.1038/s41377-020-0308-x

Angelo, J. P., Chen, S.-J. K., Ochoa, M., Sunar, U., Gioux, S., & Intes, X. (2018). Review of structured light in diffuse optical imaging. *Journal of Biomedical Optics*, *24*(7), 71602. https://doi.org/10.1117/1.JBO.24.7.071602

Bandeira, A. S., Dobriban, E., Mixon, D. G., & Sawin, W. F. (2013). Certifying the Restricted Isometry Property is Hard. *IEEE Transactions on Information Theory*, *59*(6), 3448–3450. https://doi.org/10.1109/TIT.2013.2248414

Baraniuk, R., Davenport, M. A., DeVore, R. A., & Wakin, M. B. (2006). *The Johnson-Lindenstrauss Lemma Meets Compressed Sensing*.

Baraniuk, R., & Steeghs, P. (2007). Compressive Radar Imaging. *2007 IEEE Radar Conference*, 128–133. https://doi.org/10.1109/RADAR.2007.374203

Bayram, I. (2016). On the Convergence of the Iterative Shrinkage/Thresholding Algorithm With a Weakly Convex Penalty. *IEEE Transactions on Signal Processing*, *64*, 1597–1608.

Beser, N. (1994). *Space data compression standards.*

Blumensath, T., & Davies, M. (2008a). Iterative Thresholding for Sparse Approximations. *Journal of Fourier Analysis and Applications*, *14*, 629–654. https://doi.org/10.1007/s00041-008-9035-z

Blumensath, T., & Davies, M. E. (2008b). Iterative Hard Thresholding for Compressed Sensing. *ArXiv*, *abs/0805.0510*.

Calderbank, A. R., Howard, S. D., & Jafarpour, S. (2010). Construction of a Large Class of Deterministic Sensing Matrices That Satisfy a Statistical Isometry Property. *IEEE Journal of Selected Topics in Signal Processing*, *4*, 358–374.

Candès, E. (2006). Compressive sampling. *Proceedings Oh the International Congress of Mathematicians, Vol. 3, 2006-01-01, ISBN 978-3-03719-022-7, Pags. 1433-1452*, *3*.

Candès, E. J. (2005). *11-magic : Recovery of sparse signals via convex programming*.

Candès, E. J., Eldar, Y. C., Needell, D., & Randall, P. (2011). Compressed sensing with coherent and redundant dictionaries. *Applied and Computational Harmonic Analysis*, *31*(1), 59–73. https://doi.org/https://doi.org/10.1016/j.acha.2010.10.002

Candes, E. J., & Tao, T. (2006). Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies? *IEEE Transactions on Information Theory*, *52*(12), 5406–5425. https://doi.org/10.1109/TIT.2006.885507

Candes, E. J., & Wakin, M. B. (2008). An introduction to compressive sampling: A sensing/sampling paradigm that goes against the common knowledge in data acquisition. *IEEE Signal Processing Magazine*, *25*(2), 21–30. https://doi.org/10.1109/MSP.2007.914731

Candès, E., & Romberg, J. (2005a). *-magic : Recovery of Sparse Signals via Convex Programming 1 Seven problems*.

Candès, E., & Romberg, J. (2005b). *Practical Signal Recovery from Random Projections*.

Candes, E., Romberg, J., & Tao, T. (2004). *Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information*.

Candes, E., Romberg, J., & Tao, T. (2005). *Stable Signal Recovery from Incomplete and Inaccurate Measurements*.

Candes, E., & Tao, T. (2004). *Decoding by Linear Programming*.

Caramazza, P., Moran, O., Murray-Smith, R., & Faccio, D. (2019). Transmission of natural scene images through a multimode fibre. *Nature Communications*, *10*(1), 2029. https://doi.org/10.1038/s41467-019-10057-8

Chan, W. L., Charan, K., Takhar, D., Kelly, K. F., Baraniuk, R. G., & Mittleman, D. M. (2008). A single-pixel terahertz imaging system based on compressed sensing. *Applied Physics Letters*, *93*(12), 121105. https://doi.org/10.1063/1.2989126

Chandar, V. (2008). *A Negative Result Concerning Explicit Matrices With The Restricted Isometry Property*.

Chen, S. S., Donoho, D. L., & Saunders, M. A. (1998). Atomic Decomposition by Basis Pursuit. *SIAM Journal on Scientific Computing*, *20*(1), 33–61. https://doi.org/10.1137/S1064827596304010

Čižmár, T., & Dholakia, K. (2012). Exploiting multimode waveguides for pure fibre-based imaging. *Nature Communications*, *3*(1), 1027. https://doi.org/10.1038/ncomms2024

Dai, W., & Milenkovic, O. (2009). Subspace Pursuit for Compressive Sensing Signal Reconstruction. *IEEE Transactions on Information Theory*, *55*, 2230–2249.

Daubechies, I., Defrise, M., & Mol, C. de. (2003). An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, *57*.

Daubechies, I., Fornasier, M., & Loris, I. (2008) Accelerated Projected Gradient Method for Linear Inverse Problems with Sparsity Constraints. *Journal of Fourier Analysis and Applications*, *14, 764-792*.

Devaux, F., Huy, K. P., Denis, S., Lantz, E., & Moreau, P.-A. (2016). Temporal ghost imaging with pseudo-thermal speckle light. *Journal of Optics*, *19*(2), 24001. https://doi.org/10.1088/2040-8986/aa5328

Do, T. T., Gan, L., Nguyen, N. H., & Tran, T. D. (2012). Fast and Efficient Compressive Sensing Using Structurally Random Matrices. *IEEE Transactions on Signal Processing*, *60*, 139–154.

Donoho, D. (2006). Compressed Sensing. *Information Theory, IEEE Transactions On*, *52*, 1289–1306. https://doi.org/10.1109/TIT.2006.871582

Donoho, D. L. (2004). *Compressed Sensing*.

Donoho, D. L. (2006a). Compressed sensing. *IEEE Transactions on Information Theory*, *52*(4), 1289–1306. https://doi.org/10.1109/TIT.2006.871582

Donoho, D. L. (2006b). *Sparse Solution Of Underdetermined Linear Equations By Stagewise Orthogonal Matching Pursuit*.

Duarte, M., Davenport, M., Wakin, M. B., & Baraniuk, R. (2006). Sparse Signal Detection from Incoherent Projections. *Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference On*, *3*, III–III. https://doi.org/10.1109/ICASSP.2006.1660651

Duncan, D., & Kirkpatrick, S. (2008, August). *Algorithms for simulation of speckle (laser and otherwise)*.

Edgar, M. P., Johnson, S., Phillips, D. B., & Padgett, M. J. (2017). Real-time computational photon-counting LiDAR. *Optical Engineering*, *57*(3), 31304. https://doi.org/10.1117/1.OE.57.3.031304

Eldar, Y., & Kutyniok, G. (2012). *Compressed Sensing: Theory and Applications*. https://doi.org/10.1017/CBO9780511794308

Farebrother Richard William. (2013). *L1-Norm and L∞-Norm Estimation* (1st ed.). Springer Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-36300-9_1,

Fedor Mitschke. (2016). *Fiber Optics* (2nd ed.). Springer Berlin. https://doi.org/https://doi.org/10.1007/978-3-662-52764-1

Feng, L., Benkert, T., Block, K. T., Sodickson, D. K., Otazo, R., & Chandarana, H. (2017). Compressed sensing for body MRI. *Journal of Magnetic Resonance Imaging*, *45*(4), 966–987. https://doi.org/https://doi.org/10.1002/jmri.25547

Ferri, F., Magatti, D., Gatti, A., Bache, M., Brambilla, E., & Lugiato, L. A. (2005). High-Resolution Ghost Image and Ghost Diffraction Experiments with Thermal Light. *Physical Review Letters*, *94*, 183602. https://doi.org/10.1103/PhysRevLett.94.183602

Fischer, B., & Sternklar, S. (1985). Image transmission and interferometry with multimode fibers using self-pumped phase conjugation. *Applied Physics Letters*, *46*(2), 113–114. https://doi.org/10.1063/1.95703

Forman, C., Wetzl, J., Hayes, C., & Schmidt, M. (2016). Compressed Sensing: a Paradigm Shift in MRI. *66*, 8–13. www.siemens.com/magnetom-world

Fornasier, M. (2010). Theoretical Foundations and Numerical Methods for Sparse Recovery. *Radon Series on Computational and Applied Mathematics*.

Foucart, S., & Rauhut, H. (2013). *A Mathematical Introduction to Compressive Sensing* (1st ed.). Birkhäuser. https://doi.org/https://doi.org/10.1007/978-0-8176-4948-7

Gantz, J. F., Chute, C., Manfrediz, A., Minton, S., Reinsel, D., Schlichting, W., & Toncheva, A. (2008). *The Diverse and Exploding Digital Universe*. https://www.ifap.ru/library/book268.pdf

Gover, A., Lee, C. P., & Yariv, A. (1976). Direct Transmission of Pictorial Information in MM Optical Fibers. *Journal of Optical Society of America*, *66*(4), 306–311.

Harding, B. J., & Milla, M. (2013). Radar imaging with compressed sensing. *Radio Science*, *48*(5), 582–588. https://doi.org/https://doi.org/10.1002/rds.20063

Herman, M. A., & Strohmer, T. (2009). High-Resolution Radar via Compressed Sensing. *IEEE Transactions on Signal Processing*, *57*, 2275–2284.

Herman, M., & Strohmer, T. (2008). Compressed sensing radar. *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, 1509–1512. https://doi.org/10.1109/ICASSP.2008.4517908

Howland, G. A., Dixon, P. B., & Howell, J. C. (2011). Photon-counting compressive sensing laser radar for 3D imaging. *Appl. Opt.*, *50*(31), 5917–5920. https://doi.org/10.1364/AO.50.005917

Howland, G. A., Zerom, P., Boyd, R. W., & Howell, J. C. (2011). Compressive Sensing LIDAR for 3D Imaging. *CLEO:2011 - Laser Applications to Photonic Applications*, CMG3. https://doi.org/10.1364/CLEO_SI.2011.CMG3

Imaging Technoogy News. (2017, February 21). *FDA Clears Compressed Sensing MRI Acceleration Technology From Siemens Healthineers*.

Jaspan, O. N., Fleysher, R., & Lipton, M. L. (2015). Compressed sensing MRI: a review of the clinical literature. *The British Journal of Radiology*, *88*(1056), 20150487. https://doi.org/10.1259/bjr.20150487

Krizhevsky, A. (2009). *Learning Multiple Layers of Features from Tiny Images*.

Kutz, J. N. (2013). *Data-Driven Modeling & Scientific Computation* (1st ed.). OUP.

Lau, R. C., & Woodward, T. K. (2014). A new approach to apply compressive sensing to LIDAR sensing. In F. Ahmad (Ed.), *Compressive Sensing III* (Vol. 9109, p. 91090U). SPIE. https://doi.org/10.1117/12.2058777

Lee, K., Li, Y., Jin, K. H., & Ye, J. C. (2016). Unified Theory for Recovery of Sparse Signals in a General Transform Domain. *CoRR*, *abs/1612.09565*. http://arxiv.org/abs/1612.09565

Liu, E., & Temlyakov, V. (2010). The Orthogonal Super Greedy Algorithm and Applications in Compressed Sensing. *IEEE Transactions on Information Theory - TIT*, *58*. https://doi.org/10.1109/TIT.2011.2177632

Lustig, M., Donoho, D. L., Santos, J. M., & Pauly, J. M. (2008). Compressed Sensing MRI. *IEEE Signal Processing Magazine*, *25*(2), 72–82. https://doi.org/10.1109/MSP.2007.914728

Lustig, M., Donoho, D., & Pauly, J. M. (2007). Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, *58*(6), 1182–1195. https://doi.org/https://doi.org/10.1002/mrm.21391

Mallat, S., & Zhang, Z. (1994). Matching Pursuit with Time-Frequency Dictionaries. *Signal Processing, IEEE Transactions On*, *41*, 3397–3415. https://doi.org/10.1109/78.258082

Marco Duarte. (2009). *Compressed sensing for signal ensembles* [Doctor of Philosophy]. Rice University.

Needell, D. (2009). *Topics in Compressed Sensing* [Doctor of Philosophy]. University of California .

Needell, D., & Vershynin, R. (2009). Uniform Uncertainty Principle and Signal Recovery via Regularized Orthogonal Matching Pursuit. *Foundations of Computational Mathematics*, *9*, 317–334.

Needell, D., & Vershynin, R. (2010). Signal Recovery From Incomplete and Inaccurate Measurements Via Regularized Orthogonal Matching Pursuit. *IEEE Journal of Selected Topics in Signal Processing*, *4*, 310–316.

Papadopoulos, I. N., Farahi, S., Moser, C., & Psaltis, D. (2012). Focusing and scanning light through a multimode optical fiber using digital phase conjugation. *Opt. Express*, *20*(10), 10583–10590. https://doi.org/10.1364/OE.20.010583

Pati, Y. C., Rezaiifar, R., & Krishnaprasad, P. S. (1993). Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, 40–44 vol.1. https://doi.org/10.1109/ACSSC.1993.342465

Philips. (n.d.). *Philips Compressed SENSE saves you valuable time*. Retrieved August 19, 2022, from https://www.philips.com.tr/healthcare/resources/landing/the-next-mr-wave/compressed-sense

Plöschner, M., Tyc, T., & Čižmár, T. (2015). Seeing through chaos in multimode fibres. *Nature Photonics*, *9*(8), 529–535. https://doi.org/10.1038/nphoton.2015.112

Portilla, J. (2009). Image restoration through l0 analysis-based sparse optimization in tight frames. *2009 16th IEEE International Conference on Image Processing (ICIP)*, 3909–3912. https://doi.org/10.1109/ICIP.2009.5413975

Potter, L. C., Ertin, E., Parker, J. T., & Cetin, M. (2010). Sparsity and Compressed Sensing in Radar Imaging. *Proceedings of the IEEE*, *98*(6), 1006–1020. https://doi.org/10.1109/JPROC.2009.2037526

Razavy, M. (2020). *An Introduction to Inverse Problems in Physics*. World Scientific Publishing Company. https://doi.org/10.1142/11860

Richards, M. A. (Mark A.), Scheer, J., & Holm, W. A. (2010). *Principles of modern radar. Volume I, Basic principles: Vol. Vol. I*. SciTech Pub.

Rish, I., & Grabarnik, G. (2014). *Sparse Modeling: Theory, Algorithms, and Applications* (1st ed.). CRC Press. https://doi.org/https://doi.org/10.1201/b17758

Romberg, J. (2007). *COMPRESSIVE SENSING BY RANDOM CONVOLUTION*.

Rudelson, M., & Vershynin, R. (2006). *Sparse reconstruction by convex relaxation: Fourier and Gaussian measurements*. 207–212. https://doi.org/10.1109/CISS.2006.286463

Rudelson, M., & Vershynin, R. (2008). On sparse reconstruction from Fourier and Gaussian measurements. *Communications on Pure and Applied Mathematics*, *61*(8), 1025–1045. https://doi.org/https://doi.org/10.1002/cpa.20227

Santosa, F., Symes, W. W., & Raggio, G. (1987). Inversion of band-limited reflection seismograms using stacking velocities as constraints. *Inverse Problems*, *3*(3), 477–499. https://doi.org/10.1088/0266-5611/3/3/015

Shannon, C. E. (1949). Communication Theory of Secrecy Systems*. *Bell System Technical Journal*, *28*(4), 656–715. https://doi.org/https://doi.org/10.1002/j.1538-7305.1949.tb00928.x

Siemens-Healthineers. (n.d.). *Compressed Sensing Cardiac Sine*. Retrieved August 19, 2022, from https://www.siemens-healthineers.com/magnetic-resonance-imaging/options-and-upgrades/clinical-applications/compressed-sensing-cardiac-cine

Song, C.-B., Xia, S.-T., & Liu, X.-J. (2013). Improved Analyses for SP and CoSaMP Algorithms in Terms of Restricted Isometry Constants. *CoRR*, *abs/1309.6073*. http://arxiv.org/abs/1309.6073

Takhar, D., Laska, J., Baron, D., Wakin, M., Duarte, M. F., Sarvotham, S., Baraniuk, R., & Kelly, K. (2006). A Single Pixel Camera Based on White-noise Compressed Sensing. *Frontiers in Optics*, FWN3. https://doi.org/10.1364/FIO.2006.FWN3

Takhar, D., Laska, J. N., Wakin, M. B., Duarte, M. F., Baron, D., Sarvotham, S., Kelly, K. F., & Baraniuk, R. G. (2006). A new compressive imaging camera architecture using optical-domain compression. In C. A. Bouman, E. L. Miller, & I. Pollak (Eds.), *Computational Imaging IV* (Vol. 6065, p. 606509). SPIE. https://doi.org/10.1117/12.659602

Tao, T. (2007). *Open question: deterministic UUP matrices*. https://terrytao.wordpress.com/2007/07/02/open-question-deterministic-uup-matrices/

Taubman, D. S., & Marcellin, M. W. (2002). *JPEG2000 Image Compression Fundamentals, Standards and Practice* (1st ed., Vol. 642). Springer New York, NY. https://doi.org/https://doi.org/10.1007/978-1-4615-0799-

Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288. https://doi.org/https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Tropp, J. A., & Needell, D. (2010). CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *ArXiv*, *abs/0803.2392*.

Vauhkonen Marko and Tarvainen, T. and L. T. (2016). Inverse Problems. In S. Pohjolainen (Ed.), *Mathematical Modelling* (pp. 207–227). Springer International Publishing. https://doi.org/10.1007/978-3-319-27836-0_12

Vaz, P., Amaral, D., Ferreira, L., Morgado, M., & Cardoso, J. (2020). Image quality of compressive single pixel imaging using different Hadamard orderings. *Optics Express*, *28*. https://doi.org/10.1364/OE.387612

Wallace, G. K. (1992). The JPEG still picture compression standard. *IEEE Transactions on Consumer Electronics*, *38*(1), xviii–xxxiv. https://doi.org/10.1109/30.125072

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*(4), 600–612. https://doi.org/10.1109/TIP.2003.819861

Yang, J., Jin, T., Xiao, C., & Huang, X. (2019). Compressed Sensing Radar Imaging: Fundamentals, Challenges, and Advances. *Sensors*, *19*(14). https://doi.org/10.3390/s19143100

Yariv, A. (1976). On transmission and recovery of three-dimensional image information in optical waveguides*. *Journal of Optical Society of America*, *66*(4), 301–306.

Yin, P., Lou, Y., He, Q., & Xin, J. (2015). Minimization of L_1-2 for Compressed Sensing. *SIAM Journal on Scientific Computing*. https://doi.org/10.1137/140952363

Yoon, Y.-S., & Amin, M. G. (2008). Compressed sensing technique for high-resolution radar imaging. In I. Kadar (Ed.), *Signal Processing, Sensor Fusion, and Target Recognition XVII* (Vol. 6968, p. 69681A). SPIE. https://doi.org/10.1117/12.777175

Zhang, G., Jiao, S., Xu, X., & Wang, L. (2010). Compressed sensing and reconstruction with Bernoulli matrices. *2010 IEEE International Conference on Information and Automation, ICIA 2010*. https://doi.org/10.1109/ICINFA.2010.5512379

Zhuoran, C., Honglin, Z., Min, J., Gang, W., & Jingshi, S. (2013). An improved Hadamard measurement matrix based on Walsh code for compressive sensing. *2013 9th International Conference on Information, Communications & Signal Processing*, 1–4. https://doi.org/10.1109/ICICS.2013.6782833