IOP Conf. Series: Earth and Environmental Science **17** (2014) 012273 doi:10.1088/1755-1315/17/1/012273

Information theoretic bounds for compressed sensing in SAR imaging

ZHANG Jingxiong [1], **YANG Ke** [1], **GUO Jianzhong** [2]

[1]School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China.

[2]School of Electronic and Electrical Engineering, Wuhan Textile University, Wuhan, China.

ke.yang.whu@gmail.com

Abstract. Compressed sensing (CS) is a new framework for sampling and reconstructing sparse signals from measurements significantly fewer than those prescribed by Nyquist rate in the Shannon sampling theorem. This new strategy, applied in various application areas including synthetic aperture radar (SAR), relies on two principles: sparsity, which is related to the signals of interest, and incoherence, which refers to the sensing modality. An important question in CS-based SAR system design concerns sampling rate necessary and sufficient for exact or approximate recovery of sparse signals. In the literature, bounds of measurements (or sampling rate) in CS have been proposed from the perspective of information theory. However, these information-theoretic bounds need to be reviewed and, if necessary, validated for CS-based SAR imaging, as there are various assumptions made in the derivations of lower and upper bounds on sub-Nyquist sampling rates, which may not hold true in CS-based SAR imaging. In this paper, information-theoretic bounds of sampling rate will be analyzed. For this, the SAR measurement system is modeled as an information channel, with channel capacity and rate-distortion characteristics evaluated to enable the determination of sampling rates required for recovery of sparse scenes. Experiments based on simulated data will be undertaken to test the theoretic bounds against empirical results about sampling rates required to achieve certain detection error probabilities.

1. Introduction

The Nyquist–Shannon sampling theorem states that perfect reconstruction of a signal is possible when the sampling frequency is greater than twice the maximum frequency of the signal being sampled. If sub-Nyquist sampling rates are used, the original signal's information may not be completely recoverable from the sampled signal[1]. Under certain conditions, compressed sensing[2-5] (or compressive sampling, CS) states that certain signals and images may be recovered from far fewer samples or measurements than that required by the Nyquist–Shannon sampling theorem.

According to the theory of CS, with an incoherent linear projection acquiring an efficient representation of a sparse or compressible signal, the signal can be reconstructed by solving an inverse problem either through linear programming or greedy pursuit[2]. Suppose that a vector $\mathbf{x}$ of length n is known to have a small number k of nonzero entries, but the values and locations of the nonzero entries are unknown and must be estimated from a set of m noisy linear projections (or samples) given by the vector $\mathbf{Y}$:

$$\mathbf{Y} = \mathbf{A}\mathbf{x} + \mathbf{b} \tag{1}$$

where $\mathbf{A}$ is a known $m \times n$ measurement matrix and $\mathbf{b}$ is additive white Gaussian noise (CH1 in [6]). The problem of determining which entries in the unknown signal $\mathbf{x}$ are nonzero is known as support recovery (sparsity pattern recovery) (CH1 in [6]).

With the development of CS theory, a new synthetic aperture radar (SAR) imaging modality [7-11] has been proposed that can obtain accurate representation of signals with much fewer samples or measurements than traditional matched filter (MF) methods. An important question in CS-based SAR system designs concerns sampling rates necessary and sufficient for exact or approximate recovery of sparse signals.

Much previous work ([1],[13-22]) has focused on necessary and sufficient conditions for exact or approximate recovery of the sparsity pattern. In the literature, bounds of measurements (or sampling rate) in CS have been proposed from perspectives including information theory. However, these information-theoretic bounds need to be reviewed and, if necessary, validated for CS-based SAR imaging, as there are various assumptions made in the derivations of lower and upper bounds on sub-Nyquist sampling rates, which may not hold true in CS-based SAR imaging.

This paper seeks to provide an overview of theoretic fundamentals that can be used to provide anchor points for determining CS sampling rates. For this, CS theorems and classic information theoretic concepts and quantities including channel capacity and Fano's inequality are described with respect to the bounds on sampling rates necessary and/or sufficient for sparsity pattern recovery. Thus, a SAR measurement system can be modeled as an information channel, with channel capacity and rate-distortion characteristics evaluated to enable the determination of sampling rates required for recovery of sparse scenes. SAR imaging is briefly described to furnish its implementation in the language of CS. This is followed by experiments based on simulated data, which were undertaken to test the theoretic bounds on sampling rates required to achieve certain detection error probabilities. Finally, some concluding remarks are given.

2. Methodology

2.1. CS and Restricted isometry property

Consider a measurement system

$$\mathbf{Y} = \mathbf{A}\mathbf{x} + \mathbf{b} \tag{2}$$

where $\mathbf{x}$ is an object we wish to reconstruct, $\mathbf{Y}$ are available measurements, $\mathbf{b}$ is noise, and $\mathbf{A}$ is a known $m \times n$ matrix with $m < n$. We define this kind of systems as underdetermined linear system, since the system has more unknowns than equations, and thus it has either no solution if $\mathbf{b}$ is not in the span of the columns of the matrix $\mathbf{A}$, or infinitely many solutions[23].

Here, we are interested in this underdetermined case with fewer equations than unknowns ($m < n$), and ask whether it is possible to reconstruct $\mathbf{x}$ with good accuracy. As such, the problem is of ill-posed but suppose now that $\mathbf{x}$ is known to be *sparse* or nearly sparse in the sense that it depends on a smaller number of unknown parameters. This premise radically changes the problem, making the search for solutions feasible[14]. In fact, it has been shown that solving the convex problem

$$\text{P2:} \qquad \min \|\mathbf{x}\|_{\ell_1} \quad \text{subject to } \|\mathbf{A}\mathbf{x} - \mathbf{Y}\|_{\ell_2} \le \mathbf{b} \tag{3}$$

will recover an unknown sparse object with an error at most proportional to the noise level provided that 1) $\mathbf{x}$ is sufficiently sparse and 2) the matrix $\mathbf{A}$ obeys a condition known as the restricted isometry property[14] introduced in Equation(4).

Theorem 1.1 (Theorem 1.1 of [16]) Define the S-restricted isometry constant[15] δ_S of $\mathbf{A}$ as the smallest quantity such that

$$(1-\delta_S)\|\mathbf{x}\|_{\ell_2}^2 \le \|\mathbf{A}\mathbf{x}\|_{\ell_2}^2 \le (1+\delta_S)\|\mathbf{x}\|_{\ell_2}^2 \tag{4}$$

Let s be such that $\delta_{3S}+3\delta_{4S}<2$ (δ_{3S} and δ_{4S} are restricted isometry constants of order $3S$ and $4S$, respectively), then for any signal $\mathbf{x_0}$ supported on T_0 (meaning the set of index of i where $\mathbf{x_0}[i]\neq 0$) with $|T_0|\leq S$ and any perturbation $\mathbf{b}$ with $\|\mathbf{b}\|_{\ell_2}\leq\varepsilon$, the solution $\mathbf{x}^*$ to (P2) obeys

$$\left\|\mathbf{x}^*-\mathbf{x}_0\right\|_{\ell_2}\leq C_S\cdot\varepsilon \tag{5}$$

where the constant C_S depends only on δ_{4S}.

If considering how many measurements are necessary to achieve the RIP, we can establish a simple lower bound if we ignore the impact of δ (restricted isometry constant) and focus only on the dimensions of the problem (n, m, and k) ([17], Section A.1).

Theorem 1.2 (Theorem of 3.5, Lemma 1.3 of [24]) Let $\mathbf{A}$ be an $m\times n$ matrix that satisfies the RIP of order $2k$ with constant $\delta_{2k}\in\left(0,\tfrac{1}{2}\right]$. Then

$$m\geq Ck\log\left(\frac{n}{k}\right) \tag{6}$$

where $C=\frac{\sqrt{2}\delta_{2k}}{1-(1+\sqrt{2})\delta_{2k}}$.

2.2. Information theoretic bounds

2.2.1. Rate distortion function and channel capacity

When an unknown signal is measured by a noisy linear measurements system as displayed in Equation (2), approximate rather than exact recovery is often what is possible as the measurements are imperfect. The rate distortion function[25] $R^{(I)}(D)$ is the lower bound for information transmission rate, and also the least average information content needed to recover the source signal within the given distortion.

$$R^{(I)}(D)=\min_{p(\hat{\mathrm{x}}|\mathrm{x}):\sum_{(\mathrm{x},\hat{\mathrm{x}})}p(\mathrm{x})p(\hat{\mathrm{x}}|\mathrm{x})d(\mathrm{x},\hat{\mathrm{x}})\leq D} I(\mathrm{X};\hat{\mathrm{X}}) \tag{7}$$

In order to determine how many measurements is necessary to recover the unknown signals, Sarvotham, Baron, and Baraniuk[20] applied information theory, by which a linear measurement system can be seen as a combination of encoder and signal transmitter while the recovery algorithm is the decoder to supply the approximation of the original signal. Obviously, the measurement equation shown in Equation (2) can be regarded as a channel which transfers the signals so that we can get information about the signals (recovery of signals) from measurements.

In [20], the amount of information that can be extracted from the CS measurements is investigated, and determined by the capacity of the measurement channel. Having upper bounded the information contained in the measurements, the authors investigate the minimum information needed to reconstruct the signal with distortion $\varepsilon(D_X)$.

Theorem 2.1 (Theorem 1 in [20]) For a signal source with rate-distortion function R(.) and measurement scheme specified above, the lower bound on the CS measurement rate $\rho=m/n$ required to obtain normalized reconstruction error $\varepsilon(D_X)$ subject to a fixed SNR is given by

$$\rho\geq\frac{2R(\xi(D_X))}{\log(1+SNR)} \tag{8}$$

where $\xi(D_X)$ is the distortion level, and SNR stands for signal to noise ratio [20].

Consider a k-sparse signal where the spikes have uniform amplitude, the lower bound on number of measurements [20] is:

$$m\geq\frac{2k\log(n/k)}{\log(1+SNR)} \tag{9}$$

2.2.2. Fano's inequality

Suppose that we know a random variable $\mathbf{Y}$ and we wish to guess the value of a correlated random variable $\mathbf{X}$, for any estimator $\hat{\mathbf{X}}$ such that $\mathbf{X} \to \mathbf{Y} \to \hat{\mathbf{X}}$ forms a Markov chain. Fano's inequality [25] relates the probability of error in guessing the random variable $\mathbf{X}$ ($P_e = \Pr\left(\mathbf{X} \neq \hat{\mathbf{X}}\right)$) to its conditional entropy $H(\mathbf{X} | \mathbf{Y})$:

$$H(P_e) + P_e \log|\chi| \geq H\left(\mathbf{X} | \hat{\mathbf{X}}\right) \geq H(\mathbf{X} | \mathbf{Y}) \tag{10}$$

where χ is the alphabet in which $\mathbf{X}$ takes on values.

Consider the issue of estimating the sparsity pattern (support) S^* of an unknown variable $\mathbf{x}$ from a measurement system

$$\mathbf{Y} = \mathbf{Ax} + \mathbf{b} \tag{11}$$

From the measurements $\mathbf{Y}$ and matrix $\mathbf{A}$, we want to obtain the estimate of S^* (i.e., $\hat{S}$). Define the probability of error

$$P_{\mathcal{E}} = \Pr\left[d\left(S^*, \hat{S}\right) > D\right] \tag{12}$$

where $d\left(S^*, \hat{S}\right)$ is the distortion between S^* and $\hat{S}$.

We observe that $S^* \to (\mathbf{Y}, \mathbf{A}) \to \hat{S}$ forms a Markov chain. From the perspective of Fano's inequality [6][25], the probability of error in reconstructing the support S^* correlates with the conditional entropy between S^* and $(\mathbf{Y}, \mathbf{A})$, and further the sampling rate m/n determined by the measurements matrix $\mathbf{A}$.

In [6](Ch3), lower bounds on the fundamental sampling rate distortion-function are derived. These bounds consist of necessary conditions which apply generally to any possible recovery algorithm.

$$\rho \geq \frac{2R(D;\kappa)}{\log(1+SNR*V_X)} \tag{13a}$$

where

$$R(D,\kappa) = \begin{cases} H(\kappa) - \kappa H(D) - (1-\kappa) H\left(\frac{\kappa D}{1-\kappa}\right), & D < 1-\kappa \\ 0, & D \geq 1-\kappa \end{cases}, \quad \kappa(\text{sparsity rate}) = k/n \tag{13b}$$

2.3. CS-based SAR

SAR can obtain a two-dimensional image of the observed scene. The conventional SAR imaging algorithm based on Matched Filter theory limits the resolution by the transmitted signal bandwidth and the antenna length [10]. When applying CS to SAR imaging, we can utilize radar measurements significantly fewer than what are required of traditional radar imaging algorithms to obtain an accurate representation for further processing. CS can eliminate the need for the Matched Filter in the radar receiver, and reduce the required receiver A/D conversion bandwidth so that CS-based SAR systems can operate at information rates that are potentially lower than Nyquist rates, which are often extremely high for fine-resolution systems [7].

Suppose a SAR system transmits a linear frequency modulated (LFM) signals as

$$s(t) = rect\left(\frac{t}{T}\right) \bullet \exp\left(j2\pi f_c t + j\pi f_{dr} t^2\right) \tag{14}$$

where f_c denotes carrier frequency, f_{dr} is LFM chirp data, t is the fast-time, T is pulse repetition time, and $rect(t)$ denotes the unit rectangular function $rect(t) = 1$ when $|t| \in T/2$. For a 2-D illuminated scene Ω with N scatters (a single scatter is denoted by P_i), the echo signal can be written as [10]:

$$s_c(t,n) = \sum_{i=1}^{N} \sigma_i \exp\left[j\pi f_{dr} \left(t - 2\frac{R(n;P_i)}{C} \right)^2 - j4\pi f_c \frac{R(n;P_i)}{C} \right] = \sum_{i=1}^{N} \sigma_i \exp\left[-j\phi_i(t,n) \right] \tag{15}$$

where n is the low-time, σ_i denotes the scattering coefficient of a scatter P_i. To easily facilitate the numerical implement, a long vector $\sigma \in C^{N\times 1}$ ($N = N_a \times N_r$) is formed by the columns of matrix $\mathbf{A}$, N_r and N_a denote sample size in range and azimuth directions, respectively .

In order to use CS, a linear measurement model of SAR should be created. Equation (15) can be written in the form of an inner product:

$$\mathbf{S} = \mathbf{\Phi}\sigma + \mathbf{n} \tag{16}$$

where

$$\sigma = \left[\sigma_1, \sigma_2, ..., \sigma_N\right]^T$$

$$\mathbf{S} = \left(s_c(t_1, n_1), s_c(t_2, n_2), ..., s_c\left(t_{N_r}, n_{N_a}\right)\right)^T$$

For a sparse scene, we assume that σ is k sparse when only k ($k \ll N$) of its coefficients is nonzero or greater than zero. According to CS theory, it is possible to recover the sparse signal with only a small number of samples of measured signal $\mathbf{S}$.

3. Experiments

In the experiment, the CS based SAR imaging system was simulated. The original signal comes from sparse binary sources[26] where the spikes have uniform amplitude 1, and then multiplied by a measurement matrix which is constructed by choosing *M* rows (*M* varies from 20 to 600 by a step of 20) from a 1000 by 1000 convolutional matrix, whose baseband is 1-D chirp signal, with bandwidth 50MHz, pulse duration $2\mu s$, sampling rate 1.2 times of Nyquist rate. The measurements are corrupted with additive Gaussian noise. To reconstruct the spike signal, a common used algorithm called LASSO[27] was used. As the original signal is complex-valued, the real and the imaginary part of the reconstituted signal and recovered support (sparsity pattern) are displayed in Figure 1(d),1(e),1(g) and 1(h), respectively.

Figure 1 shows the experiment results with a simulated CS-SAR system. For this, a total of 200 measurements of the 1-dimensional signal (a vector of length 1000) were used to reconstruct the original signal with 30 spikes. With the signal-noise ratio (SNR) set to 10 dB, the support of the spike signal was recovered with a distortion [6] (error probability) of 0.06.

Figure 2 shows the average distortion of sparsity support detection for n = 1000, SNR =10 dB, with varying k and m , where each point indicates the summed result of 10 independent trials. Simulation results for combinations of k and m ($k \in \{10, 20, 30, ...100\}$ and $m \in \{20, 40, 60, ...600\}$) are shown . As shown in Figure 2, as m is increasing, there is a transition from higher to lower distortion errors.

Overlaid on the intensity plots 2(a) 2(b) 2(c) are curve representing what would be evaluated for the lower bounds (for numbers of measurements) on the basis of Equations (6), (9), and (13a), respectively. As expected, the necessary measurement represented by the curves on the plot is less than that needed in practice. There are two reasons for this: 1) the lower bounds retrieved from RIP or information theory is a general result applicable to a few special types of measurement matrices instead of only the one used in the experiment, 2) the channel capacity and rate distortion function (tools used to get information theoretic bounds) are the limits that may be realized under ideal environments.

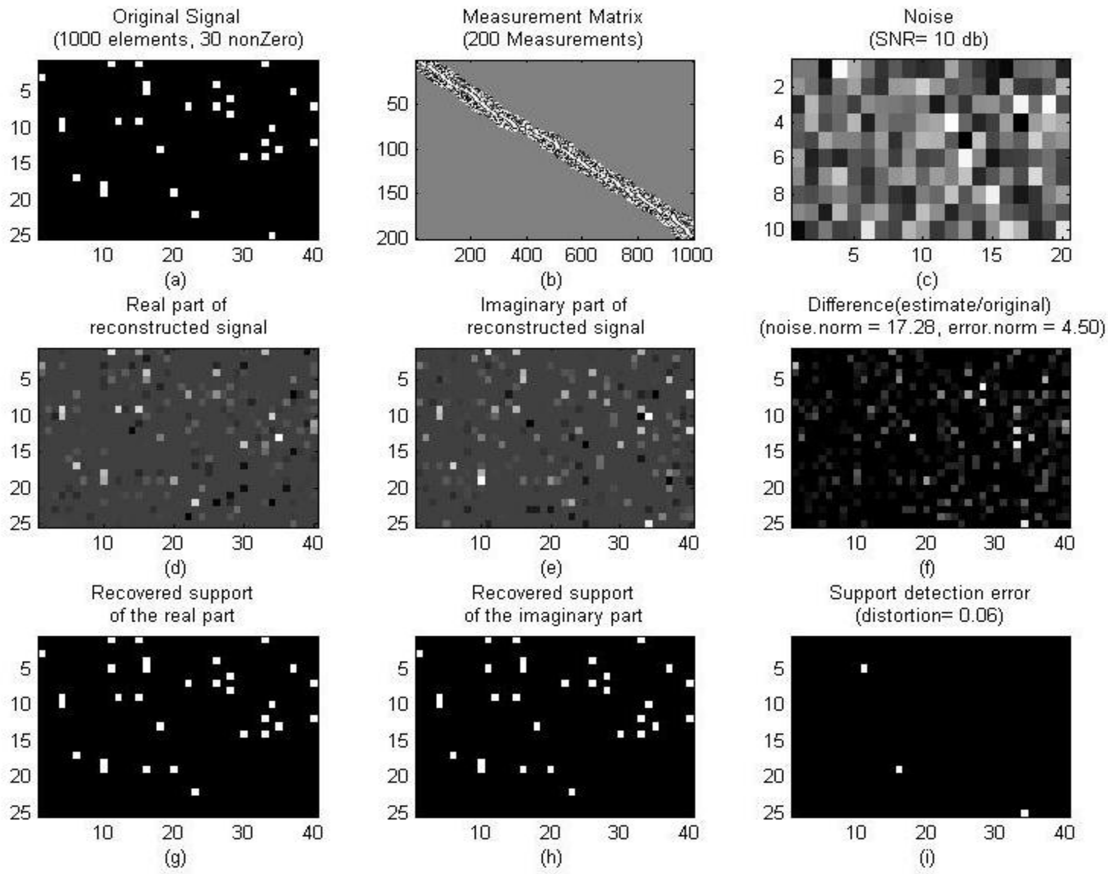

Figure 1. Simulated CS-SAR and experiments: (a) original signal, (b) measurement matrix, (c) noise, (d)-(f) reconstructed signal, (g)-(h) recovered support, and (i) support detection error.

4. Conclusion and discussion

We have considered the problem of approximate recovery of sparse signals from noisy random linear measurements, which are easily obtained by multiplying the signal with a convolutional matrix used in a SAR imaging system. Numbers of measurements required for sparsity pattern recovery were reviewed from the perspectives of CS theorems and information theory. The experiment on a simulated SAR system suggests that CS theory (random projections on a convolution matrix) is extensible to SAR imaging. Preliminary results reveal the commonality and discrepancy between theoretic and empirical results about the required sampling rates in CS-based SAR imaging system. Further research is needed to test the applicability of CS sampling rates on SAR imaging in an operational setting where continuous signal amplitudes, complex SNR and other parameters, *a priori* unknown sparsity patterns, and other complications are likely involved.

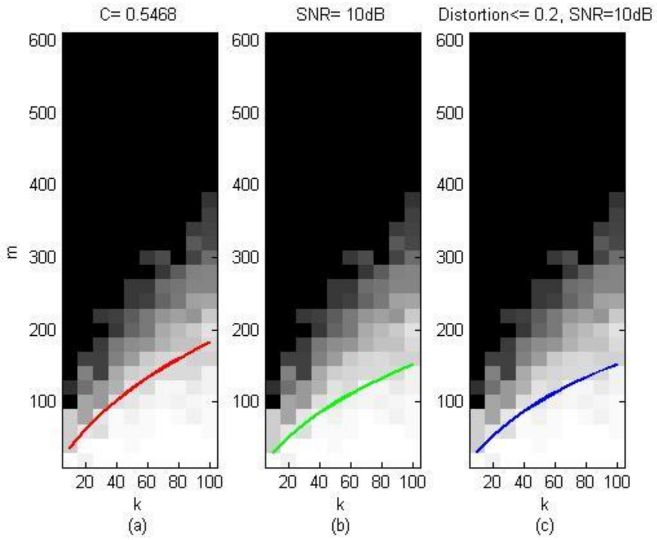

Figure 2. Errors in sparsity support detection for n= 1000 and varying values of k, m: (a) red curve representing the errors calculated by Equation (6) with C = 0.5468, (b) green curve by Equation (9) with SNR = 10 dB, and (c) blue curve by Equation (13a) with distortion <= 20% and SNR = 10 dB (the gray scale indicates 0 (black) through100 (white) percent error probability).

Acknowledgments
This work was supported by the State Key Development Program for Basic Research of China (No. 2010CB731905), and the National Natural Science Foundation of China (No. 41071286).

References

[1] Shannon, C.E. 1949. Communication in the Presence of Noise. *Proceedings of the IRE* 37(1): 10-21.
[2] Baraniuk, R. 2007. Compressive sensing. *IEEE Signal Processing Magazine*, **24**(4): 118-121.
[3] Candes, E.J. 2006. Compressive sampling. *Int. Congress of Mathmatics*, **3**: 1433-1452.
[4] Candes, E.J. and M.B Wakin. 2008. An Introduction To Compressive Sampling. *IEEE Signal Processing Magazine* **25**: 21-30.
[5] Donoho, D. 2006. Compressed sensing. *IEEE Transactions on Information Theory*, **52**(4): 1289-1306.
[6] Reeves, G. 2011. Sparsity pattern recovery in compressed sensing. PhD diss., University of California at Berkley.
[7] Baraniuk, R. and P. Steeghs. 2007. Compressive radar imaging. *Radar Conference 2007.*
[8] Huan, Y.F., Wang, J.F., Z. Tan and et al. 2011. SAR imaging based on compressed sensing. *IGARSS 2011*.
[9] Jiang, H., Zhang, B.C., Y.G. Lin and et al. 2010. Random noise SAR based on compressed sensing. *IGARSS 2010.*
[10] Wei, S.J., Zhang, X.L., Shi, J. and G. Xiang. 2010. Sparse reconstruction for SAR imaging based on compressed sensing. *Progress In Electromagnetics Research,* **109**(63-81).
[11] Zhang, Z., Zhang, B.C., Hong, W. and Y.R. Wu. 2012. Waveform design for Lq regularization based radar imaging and an approach to radar imaging with non-moving platform. *Synthetic Aperture Radar, 2012. EUSAR. 9th European Conference on*.
[12] Aeron, S. and V. Saligrama. 2010. Information Theoretic Bounds for Compressed Sensing. *IEEE Transactions on Information Theory*, **56**(10): 5111-5130.
[13] Akcakaya, M. and V. Tarokh. 2010. Shannon-Theoretic Limits on Noise Compressive Sampling. *IEEE Transactions on Information Theory*, **56**(1): 492-504

[14] Candes, E.J. 2008. The Restricted Isometry Property and Its Implications for Compressed Sensing.http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.154.8144&rep=rep1&type=pdf.

[15] Candes, E.J. and T. Tao. 2005. Decoding by linear programming. *IEEE Transactions on Information Theory*, **51**(12): 4203-4215.

[16] Candes, E.J., Romberg, J.K. and T. Tao. 2006. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, **LIX**: 1207-1223.

[17] Davenport, M.A., Duarte, M.F., Eldar, Y.C., and G. Kutyniok. 2012. Introduction to compressed sensing. In *Compressed sensing: Theory and Applications*, ed. Eldar, Y.C., and G. Kutyniok, 1-68. London: Cambridge University Press.

[18] Fletcher, A.K., S. Rangan and V.K. Goyal. 2009. Necessary and Sufficient Conditions for Sparity Pattern Recovery. *IEEE Transactions on Information Theory*, **55** (12): 5758-5772

[19] Reeves, G. and M. Gastpar. 2012. The Sampling Rate-Distortion Tradeoff for Sparsity Pattern Recovery in Compressed Sensing. *IEEE Transactions on Information Theory*, **58**(5): 3065-3092

[20] Sarvotham, S., D.Baron, and R.G. Baraniuk. 2006. Measurements vs. Bits: Compressed Sensing. *Proceedings of the 44th Allerton Conference on Communication, Control, and Computing.*

[21] Wainwright, M.J. 2009. Sharp Thresholds for High-Dimensional and Noisy Sparsity Recovery using L1 -constrained quadratic programming (Lasso). *IEEE Transactions on Information Theory*, **55**(5): 2183-2202.

[22] Wainwright, M.J. 2009. Information-Theoretic Limits on Sparsity Recovery in the High-Dimensional and Noisy setting. *IEEE Transactions on Information Theory*, **55**(12): 5728-5740.

[23] Elad, M. 2010. *Sparse and Redundant Representations: From theory to applications in signal and image processing.* New York: Springer.

[24] Davenport, M.A. 2010. Random observations on random observations: Sparse signal acquisition and processing. PhD diss., Rice University.

[25] Thomas, M.C. and J.A. Thomas. 2006. Elements of information theory, 2nd edition. Wiley.

[26] Weidmnn, C. and M. Vetterli. 2012. Rate distortion behavior of sparse sources. *IEEE Transactions on information theory*, **58**(8).

[27] Tibshirani, R. 2011. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society*, **73**(3): 273-282.

[28] Wang, W., Wainwright, M.J. and K. Ramchandran. 2010. Information-Theoretic Limits on Sparsity Recovery: Dense versus Sparse Measurement Matrice. *IEEE Transactions on Information Theory*, **56**(6): 2967-2979.