DICTIONARY-BASED RECONSTRUCTION OF THE CYCLIC AUTOCORRELATION VIA ℓ_1-MINIMIZATION FOR CYCLOSTATIONARY SPECTRUM SENSING

Andreas Bollig and Rudolf Mathar

Institute for Theoretical Information Technology
RWTH Aachen University
D-52074 Aachen, Germany
Email: {bollig, mathar}@ti.rwth-aachen.de

ABSTRACT

One of the main enablers of dynamic spectrum access is fast and reliable spectrum sensing. Acquiring the occupation status of a spectral band can be accomplished in different ways, one of which is called cyclostationary spectrum sensing. The aforementioned method exploits the prior knowledge of periodicities inherent in most man-made signals for the purpose of detecting their presence in a set of sample data. One prerequisite for the detection is the knowledge of the signal's cyclic autocorrelation (CA), which can be estimated from a finite amount of time-domain samples. This work introduces a new method for estimating the CA using a very small amount of time-domain samples, i.e. a short observation time. This is accomplished by modeling the desired CA vector using a custom dictionary describing its known properties and recovering it by solving a convex optimization problem.

***Index Terms*—** compressive sampling, convex optimization, cyclic autocorrelation, spectrum sensing

1. INTRODUCTION

In the recent past, the use of wireless communication technologies has increased dramatically. Despite technological advancements aimed at boosting the achievable utilization of the given radio spectrum, this lead to the advent of spectral scarcity in certain bands. In stark contrast to that, other bands making up a large part of the spectral band suitable for radio communication are idle either most of the time or in most geographical regions as a result of traditional licensing policies [1]. To tackle this problem, the use of dynamic spectrum access (DSA) has been proposed as a part of a new paradigm for wireless communication centered around the cognitive radio (CR) [2]. The idea is to let intelligent transceivers dynamically allocate bandwidth to a secondary system when the licensee of said band, i.e. the primary system, is not occupying it. The main technology enabling reliable DSA is spectrum sensing, i.e. before utilizing some spectral band, the secondary system has to make sure that it is unoccupied in order to not interfere with the primary system's usage of its licensed band.

One way of probing a band's occupancy status is to test for the presence of cyclostationarity. This method exploits the fact that most man-made signals vary periodically with time [3] and can thus be characterized as cyclostationary. Although the data contained in a modulated signal may be a purely stationary random process, the coupling with sine wave carriers, pulse trains, repeating, spreading, hopping sequences and cyclic prefixes going along with its modulation causes a built-in periodicity [4]. This underlying periodicity can be exploited for detecting the presence of a signal despite contamination with noise. One method for detecting the presence of cyclostationarity in a signal is the *time-domain test* as detailed in [5].

A prerequisite of said *time-domain test* is the cyclic autocorrelation of the observed signal. The CA function of a signal $x(t)$ is given by [3]:

$$R_x^\alpha(\tau) = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} x(t+\tau/2)x^*(t-\tau/2)e^{-j2\pi\alpha t}dt. \tag{1}$$

For purely stationary signals, it holds that $R_x^\alpha(\tau) = 0$ for all $\alpha \neq 0$, while for cyclostationary signals $R_x^\alpha(\tau) \neq 0$ for some $\alpha \neq 0$. The α with non-zero CA coefficients are called cycle frequencies. The set of cycle frequencies caused by one of potentially multiple incommensurate second-order periodicities in a cyclostationary signal comprises the periodicity's fundamental cycle frequency (the reciprocal of the fundamental period) as well as its harmonics (integer multiples). Since the CA is zero on its whole support except the set of cycle frequencies and $\alpha = 0$, it can be called sparse. The CA's sparsity can be taken advantage of for the purpose of estimating it given a short observation time.

Sparsity has been exploited for signal recovery for a long time [6]. However, the topic has seen enormous development in the last decade, leading to a new sampling-paradigm called compressive sampling (CS) [7] [8]. A discrete signal which is sparse in some domain carries a lot less information than

This work was partly supported by the Deutsche Forschungsgemeinschaft (DFG) project CoSeCo (grant MA 1184/18-1) and by the UMIC research cluster of the RWTH Aachen University.

suggested by its dimension. However, when the signal is observed in some other domain, its low information load may not be directly visible so that the number of samples that needs to be taken in order to acquire the signal depends on the signal's dimension instead of the amount of information it carries. Applying CS techniques, the signal to be acquired can be recovered from a small amount of samples by solving a convex optimization problem. Practically, this means that in contrast to traditional methods for estimating the CA, employing a solution based on CS will reduce the necessary observation time. Apart form cyclostationary spectrum sensing, the CS theory has seen utilization in different branches of spectrum sensing as for example in energy detection [9] [10].

The performance of estimating the CA via compressive sampling has been investigated in [11]. The algorithm developed in said paper will be used as a benchmark against which this paper's contribution will be compared. Another example of employing CS for the recovery of the CA is found in [12].

The contribution of this paper lies in developing an algorithm for reconstructing the CA of an observed signal with a high accuracy while further diminishing the minimum viable observation time compared to [11]. Our algorithm is based on the central idea behind CS. However, in addition to the sparsity property, it exploits further prior knowledge about the signal.

The remainder of this paper is structured as follows. Section 2 introduces the signal model as well as the traditional CA estimation method. In Section 3 the estimation algorithm from [11] and the new algorithm are presented and the intuition behind their design is given. The numerical simulation of the three approaches as well as the interpretation of the results is given in Section 4. Section 5 concludes the paper.

2. ANALYTICAL MODEL AND TRADITIONAL CA ESTIMATION METHOD

In order to decide about a spectral band's occupancy status, a CR receiver observes the time-domain signal $x(t)$ in the baseband. The signal is sampled uniformly with a sampling period T_e, resulting in the sample vector $\mathbf{x}_t \in \mathbb{C}^N$, where

$$\mathbf{x}_t = [x(0), x(T_e), ..., x((N-1)T_e)]^{\mathrm{T}}. \tag{2}$$

Due to the nature of man-made signals, $\mathbf{x}_t$ is a discrete zero-mean (almost) cyclostationary process [5]. To detect the presence of periodicity in the sampled band, the CR needs to run a detection algorithm on the signal's CA. For this purpose, the CA can be obtained in multiple ways, one of which is the traditional unbiased estimator as used in [5]. It is given by

$$\hat{R}_x^{\alpha}(\tau) = \frac{1}{N}\sum_{t=0}^{N-1} x(tT_e)x^*(tT_e+\tau)e^{-j2\pi\alpha\tau}. \tag{3}$$

With $\alpha \in \{\frac{i}{NT_e}|_{i=0}^{N-1}\}$ and choosing the delay τ to be an integer multiple of the sampling period, we define the CA vector as

$$\hat{\mathbf{r}}_x^{\tau_0} = [\hat{R}_x^0(\tau_0), ..., \hat{R}_x^{\frac{N-1}{NT_e}}(\tau_0)]^{\mathrm{T}}. \tag{4}$$

This vector is a scaled discrete Fourier transform (DFT) of the delay-product $\mathbf{y}_{\tau_0} = \mathbf{x}_t \circ \mathbf{x}_{t+\tau_0}$, where $\circ$ denotes the component-wise multiplication. Thus, the CA vector can also be written as

$$\hat{\mathbf{r}}_x^{\tau_0} = \frac{1}{N}\mathbf{F}\mathbf{y}_{\tau_0}, \tag{5}$$

where $\mathbf{F}$ denotes the DFT matrix.

This work uses an artificially generated BPSK signal with symbol length T_s for evaluating the performance of the CA estimation algorithms. The fundamental cycle frequency of the built-in periodicity of this signal is $\frac{1}{T_s}$. A theoretical expression for the CA of a BPSK signal is given in [3]. For a rectangular transmission window its absolute value reduces to

$$|\tilde{R}_x^{\alpha}(\tau)| = \begin{cases} 0 & \text{for } \alpha \neq \frac{k}{T_s}, k \in \mathbb{Z} \\ \frac{|T_s-\tau|}{T_s}\frac{\sin(\pi\alpha(T_s-\tau))}{\pi\alpha(T_s-\tau)} & \text{otherwise.} \end{cases} \tag{6}$$

Its value at the relevant frequencies α are arranged into a vector in the following way as to match the format of the DFT matrix.

$$\tilde{\mathbf{r}}_x^{\tau_0}[i] = \begin{cases} \left|\tilde{R}_x^{\frac{i}{NT_e}}(\tau_0)\right| & \text{for } i \in \{0, ..., \frac{N}{2}\}, \\ \left|\tilde{R}_x^{\frac{i-N}{NT_e}}(\tau_0)\right| & \text{for } i \in \{\frac{N}{2}+1, ..., N-1\}. \end{cases} \tag{7}$$

In the present work, this term will be used as a theoretical reference for the CA estimation algorithms.

3. OPTIMIZATION PROBLEMS

3.1. Basic CS Approach

To estimate the N-element CA vector, the traditional estimator (3) needs the knowledge of all N elements of $\mathbf{y}_{\tau_0}$. In contrast to that, the CA estimation algorithm presented in [11] can recover the CA vector from the first $n \ll N$ elements of $\mathbf{y}_{\tau_0}$, such that the observation time necessary for reliable detection is drastically reduced. This improvement is facilitated by the employment of CS, i.e. the exploitation of the CA's sparse property. The CA vector's sparsity is apparent since it holds non-zero entries only for $\alpha = 0$ as well as a built-in periodicity's fundamental cycle frequency and its harmonics.

The n known entries of $\mathbf{y}_{\tau_0}$ define a set of constraints given by the following under-determined system of equations

$$N\mathbf{M}\mathbf{F}^{-1}\hat{\mathbf{r}}_x^{\tau_0} = \tilde{\mathbf{y}}_{\tau_0}, \tag{8}$$

where $\mathbf{M}$ contains the first n lines of the $N \times N$ identity matrix, $\mathbf{F}^{-1}$ is the IDFT matrix of according size and

$$\tilde{\mathbf{y}}_{\tau_0} = \mathbf{M}\mathbf{y}_{\tau_0}. \tag{9}$$

Intuitively, this means that the CA vector that is to be reconstructed has to match the n samples which have been observed. The constraints span a set of possible solutions for $\mathbf{y}_{\tau_0}$ giving rise to the under-determined linear inverse problem of finding the right CA vector in the set. Applying the prior knowledge about the vector's sparsity, the sparsest vector from the set , i.e. the one with the lowest ℓ_0-"norm" ($\|\cdot\|_{\ell_0}$) should be picked as the solution, where $\|\cdot\|_{\ell_0}$ is defined as the number of non-zero entries in a vector. However, this is impractical since the problem at hand is NP-hard. Making use of the recent developments in the field of sparse recovery, i.e. CS, the problem can be solved by picking the vector with the lowest ℓ_1-norm instead of the lowest ℓ_0-"norm". This approach is viable since *for most large under-determined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution* [13]. The resulting *convex* optimization problem is given by

$$\begin{aligned} &\underset{\hat{\mathbf{r}}_x^{\tau_0}}{\text{minimize}} \;\; \|\hat{\mathbf{r}}_x^{\tau_0}\|_{\ell_1} \\ &\text{subject to} \;\; N\mathbf{M}\mathbf{F}^{-1}\hat{\mathbf{r}}_x^{\tau_0} = \tilde{\mathbf{y}}_{\tau_0}. \end{aligned} \tag{10}$$

3.2. CA Dictionary Approach

Compared to the basic CS approach a more detailed model of the CA will be developed, allowing us to further reduce the required amount of sensing data while keeping the reconstruction accuracy high. To accomplish this, the desired CA vector is divided into a sum of three solution vectors, i.e. $\hat{\mathbf{r}}_x^{\tau_0} = \mathbf{e}+\mathbf{f}+\mathbf{g}$ resulting in the following *convex* optimization problem

$$\begin{aligned} &\underset{\mathbf{e},\mathbf{f},\mathbf{g},\mathbf{z}}{\text{minimize}} \;\; \|\mathbf{D}^{-1}\mathbf{z}\|_{\ell_1} + \beta_a\|\mathbf{z}\|_{\ell_1} + \beta_b\|\mathbf{f}\|_{\ell_1} \\ &\text{subject to} \;\; N\mathbf{M}\mathbf{F}^{-1}(\mathbf{e}+\mathbf{f}+\mathbf{g}) = \tilde{\mathbf{y}}_{\tau_0} \\ &\qquad\qquad\;\; \mathbf{C}\mathbf{z} \geq |\mathbf{g}|. \end{aligned} \tag{11}$$

Subsequently, an intuition for the inner workings of this optimization problem will be given. The vector $\mathbf{e} = [\tilde{e}, 0, ..., 0]^{\mathrm{T}}$ represents the conventional autocorrelation function. Since we cannot make any statement about its value from the knowledge about the signal's cyclostationarity, it is not part of the objective, meaning that we don't discriminate possible solution vectors based on their $\tilde{e}$. It only needs to match the observed samples which is enforced by the first constraint.

For each periodicity that appears in the signal, the CA vector shows non-zero entries at the periodicities' respective fundamental cycle frequency and at its harmonics. This is modeled by constructing a dictionary $\mathbf{D}$ describing the structure of the CA vector. Each column of $\mathbf{D} \in \{0,1\}^{\frac{N}{2}\times\frac{N}{2}}$ represents one of the possible cycle frequencies contained in the set $\alpha \in \{\frac{i}{NT_e}|_{i=1}^{N/2}\}$. For simplicity, this set is chosen such that the frequencies contained in it hit center frequencies of the CA's DFT bins. An entry of the dictionary covers elements 1 to $\frac{N}{2}$ of $\hat{\mathbf{r}}_x^{\tau_0}$ which is indexed from 0 to $N-1$. The dictionary is built according to the following rule

$$[\mathbf{D}]_{ij} = \begin{cases} 1 & \text{if } i \bmod j = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

An example of size 6 is given by

$$\mathbf{D}_6 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \tag{13}$$

Note, that the dictionary exhibits a full rank independent of its size and can thus be inverted.

The vector $\mathbf{z}$ is directly modeled with $\mathbf{D}$. Indeed, it is a sum of weighted dictionary entries. Given the special sparse property of the CA vector which is reflected in the dictionary, the term $\mathbf{D}^{-1}\mathbf{z}$, i.e. the vector $\mathbf{z}$ mapped to the dictionary domain, exhibits one non-zero entry for each of possibly multiple incommensurate periodicities in the observed signal. Obviously this is desired to be sparse leading to the first term in the objective. Sampling at a rate higher than the symbol rate of the signal and expecting a low number of incommensurate periodicities in the signal, $\mathbf{z}$ is sparse too, giving rise to the second term in the objective.

The connection between the model (represented by the dictionary) and the observed data is the vector $\mathbf{g}$. Since the CA vector is complex-valued ($\hat{\mathbf{r}}_x^{\tau_0} \in \mathbb{C}^N$), its non-zero elements have both an amplitude and a phase. However, the dictionary is real and can only describe the CA's amplitude. Thus, the second constraint makes a connection between $\mathbf{z}$ and the absolute value of $\mathbf{g}$ instead of $\mathbf{g}$'s actual value. Furthermore, the CA is symmetric to its DC component, which is why $\mathbf{z} \in \mathbb{R}^{\frac{N}{2}}$ only describes the positive frequencies. In contrast to that, $\mathbf{g} \in \mathbb{C}^N$ describes the whole CA vector. The matrix $\mathbf{C} \in \{0,1\}^{N\times\frac{N}{2}}$ mirrors $\mathbf{z}$ as to fit the format of $\hat{\mathbf{r}}_x^{\tau_0}$ and $\mathbf{g}$. For $\mathbf{z}$ to actually represent the amplitude of $\mathbf{g}$, the second constraint would have to read $\mathbf{C}\mathbf{z} = |\mathbf{g}|$, which would lead to the optimization problem not being convex. Thus, the equality has to be relaxed to the present inequality.

The non-zero elements in $\hat{\mathbf{r}}_x^{\tau_0}$ stemming from a periodicity in the observed signal, i.e. the one at the fundamental cycle frequency as well as its harmonics, can have different amplitudes. This is not taken care of by the dictionary, as all entries of a dictionary word are weighted by the same factor. Therefore, the residuum vector $\mathbf{f}$ is introduced. Naturally, $\mathbf{f}$ should be (nearly) sparse, which is reflected in the third term of the objective.

4. NUMERICAL EVALUATION AND ANALYSIS OF THE RESULTS

To compare the three methods for reconstructing the CA vector presented in the preceding sections, we investigate the

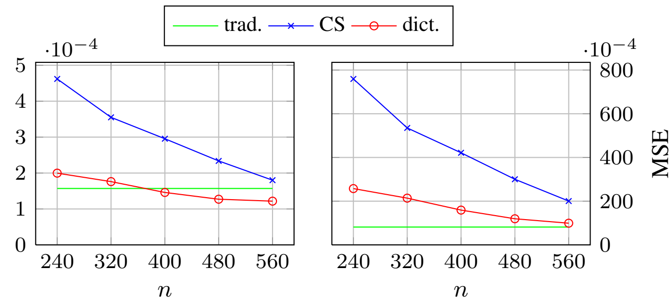

Fig. 1. MSE of recovery, no noise in the signal; left: overall, right: peaks

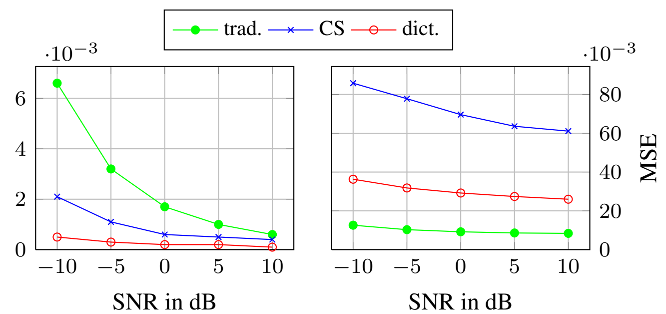

Fig. 2. MSE of recovery from 240 samples, different SNRs; left: overall, right: peaks

mean square error (MSE) of the respective recovered CA vectors to the theoretical curve (7). The three algorithms are run on a BPSK signal generated with the parameters given in table 1. The convex optimization problems are solved employing the CVX software package [14].

The sequence of symbols is drawn randomly from a uniform distribution and the noise is distributed according to a zero-mean Gaussian distribution. The variance of the noise is defined by the respective SNR, where the SNR is defined as the signal energy of all $(N + \tau_0)$ observed time-domain samples divided by the noise energy of all observed samples.

Table 1. Scenario parameters

Parameter	Symbol	Value(s)
Size of the CA vector	N	2000
# of known delay-product entries	n	$\{240, 320, 400, 480, 560\}$
Time delay	τ_0	$2\,T_e$
BPSK symbol length	T_s	$8\,T_e$
Signal to noise ratio	SNR	$\{-10, -5, 0, 5, 10\}$dB
Regularization parameters	β_a	0.25
	β_b	1

In the following, the overall MSE and the MSE at the peaks are defined as follows

$$\begin{aligned} \text{MSE}_{\text{overall}} &= \frac{1}{N}\sum_{i=0}^{N-1}\left(\left|\hat{\mathbf{r}}_x^{\tau_0}[i]\right| - \tilde{\mathbf{r}}_x^{\tau_0}[i]\right)^2 \\ \text{MSE}_{\text{peaks}} &= \frac{1}{n_{\text{peaks}}}\sum_{j\in\mathcal{P}}\left(\left|\hat{\mathbf{r}}_x^{\tau_0}[j]\right| - \tilde{\mathbf{r}}_x^{\tau_0}[j]\right)^2, \end{aligned} \tag{14}$$

where the set $\mathcal{P}$ contains all indices of the frequency bins representing cycle frequencies, their harmonics and the DC component.

Figure 1 shows the overall MSE as well as the MSE at the peaks resulting from applying the three presented methods, namely the traditional CA estimation, the basic compressive sampling approach and the custom dictionary based approach to the problem of estimating the CA vector from time-domain samples which are not contaminated by noise. While the traditional method estimates the N-element CA vector from all of the N elements of the delay-product vector $\mathbf{y}_{\tau_0}$, the basic CS approach and the CA dictionary approach recover the CA vector from the first $n \ll N$ elements of $\mathbf{y}_{\tau_0}$, where n is labeled on the x-axis. As can be seen in the plot, the sparse recovery approaches manage to match the performance of the traditional estimator with a drastically reduced number of samples. The additional knowledge used in the CA dictionary approach leads to a better error performance compared to the basic CS approach, which only exploits the CA vector's sparsity. Due to its strong assumption of the CA vector matching some dictionary entries, the CA dictionary approach favors solution vectors with very little estimation noise between the peaks and thus takes less samples to outperform the traditional method regarding the overall MSE than the MSE at the peaks.

While the CA should be independent of stationary noise, its estimation from a finite number of samples is not. This can be observed in figure 2, where the overall MSE as well as the MSE at the peaks produced by the three respective approaches is plotted over the ratio of the signal energy to the energy of additional white Gaussian noise. For the noisy case, the equality constraints of both optimization problems have been replaced by an ℓ_2-norm ball constraint. Here, the effect of the sparsity assumption made by the two sparse recovery approaches can be observed, since given this noisy scenario, the traditional method performs worst in the overall MSE due to its higher estimation noise between the peaks although it uses more than 8 times as much samples for the estimation. Concerning the MSE at the peaks, the sparse recovery methods cannot match the error performance of the traditional approach given the low number of samples. In all cases and parameter sets investigated, the CA dictionary approach produces a smaller average error than the basic CS approach.

5. CONCLUSION

In this work, a new approach for estimating the CA vector has been introduced. It makes extensive usage of prior knowledge about the vector to be reconstructed in order to diminish the number of samples necessary for reliably reconstructing the CA. Due to its more detailed model of the CA it outperforms the basic CS approach introduced in [11].

6. REFERENCES

[1] Spectrum Policy Task Force Federal Communications Commission, "Report of the spectrum efficiency working group," Tech. Rep., November 2002.

[2] Simon Haykin, "Cognitive radio: Brain-empowered wireless communications," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 2, pp. 201 – 220, February 2005.

[3] William A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals," *Signal Processing Magazine, IEEE*, vol. 8, no. 2, pp. 14 –36, April 1991.

[4] William A. Gardner, "Signal interception: a unifying theoretical framework for feature detection," *Communications, IEEE Transactions on*, vol. 36, no. 8, pp. 897 –906, aug 1988.

[5] Amod V. Dandawate and Georgios B. Giannakis, "Statistical tests for presence of cyclostationarity," *Signal Processing, IEEE Transactions on*, vol. 42, no. 9, pp. 2355 –2369, September 1994.

[6] David L. Donoho, "Scanning the technology," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 910–912, June 2010.

[7] David L. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[8] Emmanuel J. Candès, Justin Romberg, and Terence Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[9] Zhi Tian, "Compressed wideband sensing in cooperative cognitive radio networks," in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, Nov. 2008, pp. 1–5.

[10] Andreas Bollig, Steven Corroy, and Rudolf Mathar, "Joint sparse spectrum reconstruction and information fusion via l1-minimization," in *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*, may 2012, pp. 1 –5.

[11] Ziad Khalaf and Jacques Palicot, "Exploiting sparse property of the cyclic autocorrelation function for cyclostationary process compressed sensing," in *Proceedings of the 7th Workshop on Software Radios*, Karlsruhe, Germany, January 2012, p. 6 p.

[12] Zhi Tian, "Cyclic feature based wideband spectrum sensing using compressive sampling," in *Communications (ICC), 2011 IEEE International Conference on*, June 2011, pp. 1–5.

[13] David L. Donoho, "For most large underdetermined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution," *Comm. Pure Appl. Math*, vol. 59, pp. 797–829, September 2004.

[14] Michael Grant and Stephen Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," Apr. 2011.