Surpassing the Theoretical 1-Norm Phase Transition in Compressive Sensing by Tuning the Smoothed L0 Algorithm

Oxvig, Christian Schou; Pedersen, Patrick Steffen; Arildsen, Thomas; Larsen, Torben

Published in:
Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on

DOI (link to publication from Publisher):
10.1109/ICASSP.2013.6638820

Publication date:
2013

Document Version
Accepted author manuscript, peer reviewed version

SURPASSING THE THEORETICAL 1-NORM PHASE TRANSITION IN COMPRESSIVE SENSING BY TUNING THE SMOOTHED L0 ALGORITHM

Christian Schou Oxvig, Patrick Steffen Pedersen, Thomas Arildsen, and Torben Larsen

Department of Electronic Systems, Faculty of Engineering and Science
Aalborg University, DK-9220 Aalborg, Denmark

ABSTRACT

Reconstruction of an undersampled signal is at the root of compressive sensing: when is an algorithm capable of reconstructing the signal? what quality is achievable? and how much time does reconstruction require? We have considered the worst-case performance of the smoothed ℓ_0 norm reconstruction algorithm in a noiseless setup. Through an empirical tuning of its parameters, we have improved the phase transition (capabilities) of the algorithm for fixed quality and required time. In this paper, we present simulation results that show a phase transition surpassing that of the theoretical ℓ_1 approach: the proposed modified algorithm obtains 1-norm phase transition with greatly reduced required computation time.

Index Terms— Signal Reconstruction, Compressed Sensing, Smoothing Methods, Iterative Algorithms

1. INTRODUCTION

The Compressive Sensing (CS) signal acquisition paradigm asserts that certain signals sampled far below their Nyquist frequencies can be successfully recovered, if they are sparse in some dictionary [1]. For example, the Fourier dictionary may be used with frequency sparse signals. Hence, CS combines the usual sample-and-then-compress setup into a single efficient step. However, CS requires acquired signals to be reconstructed which, in the noiseless case, entails a non-convex optimisation problem of the form [2]:

$$\begin{aligned} &\text{minimise} && ||\hat{\mathbf{x}}||_0 \\ &\text{subject to} && \mathbf{y} = \mathbf{A}\hat{\mathbf{x}} \end{aligned} \tag{1}$$

where $\hat{\mathbf{x}} \in \mathbb{R}^{N\times 1}$ is the reconstructed vector, $\mathbf{A} \in \mathbb{R}^{n\times N}$ is a known measurement matrix, and $\mathbf{y} \in \mathbb{R}^{n\times 1}$ is the measured vector with $n \ll N$. In the CS context, n is the number of samples sensed whereas N is the number of samples in the reconstructed signal. We take $||\hat{\mathbf{x}}||_0$ to denote the ℓ_0 pseudo norm [3], i.e. the number of non-zero entries in $\hat{\mathbf{x}}$.

Authors may be contacted at cssf-authors@googlegroups.com

This work was partially financed by The Danish Council for Strategic Research under grant number 09-067056 and by the Danish Center for Scientific Computing.

Several approaches to solving the problem in (1) have been proposed. One approach substitutes the ℓ_1 norm for the ℓ_0 norm thereby relaxing the problem to a linear program (LP) and enabling the use of existing LP solvers [2]. Another approach "reverses" the problem by minimising $||\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}||_2^2$ subject to some sparsity-enforcing constraint [4]. These so-called greedy algorithms include Matching Pursuit [5], Iterative Hard Thresholding [6], Orthogonal Matching Pursuit [7], and Compressive Sampling Matching Pursuit [8].

Regardless of the approach, reconstruction algorithms generally provide a trade-off between reconstruction quality (e.g. SNR), reconstruction capabilities, and required reconstruction time. A standard measure of reconstruction capabilities is the phase transition measure [9] (see [4] and [10] for examples of its use). Phase transitions evaluate the probability of successful reconstruction versus the indeterminacy $\delta = n/N$ of the constraints $\mathbf{y} = \mathbf{A}\hat{\mathbf{x}}$ and the true sparsity of $\hat{\mathbf{x}}$. In terms of phase transition, the ℓ_1 approach is superior to greedy algorithms for fixed reconstruction quality [4]; whereas in terms of required reconstruction time, greedy algorithms are superior to state-of-the-art LP solvers used with the ℓ_1 approach [11]. Potentially, any algorithm could be improved by focusing on either reconstruction capabilities, quality, or time.

In an attempt to match the phase transition of the ℓ_1 approach while reducing the required reconstruction time, the smoothed ℓ_0 norm (SL0) approach approximates the ℓ_0 norm in (1) with a continuous function [12]. We have found that, for fixed reconstruction quality and time, the phase transition of SL0 may be dramatically improved by selecting the parameters in a way that exploits the known indeterminacy of the constraints in (1). In this paper, we present a modified SL0 algorithm and provide a set of empirically determined recommended parameters. Furthermore, we support our claim of a dramatically improved phase transition by results from an extensive set of simulations.

The paper is organised as follows. In Section 2, we restate the SL0 algorithm and present our proposed modification. Section 3 describes the setup used for simulations while Section 4 provides the simulation results. A discussion of the results is given in Section 5 followed by conclusions in Section 6. Finally, Section 7 reviews relations to prior work.

2. SMOOTHED ℓ_0 NORM

SL0 approximates the ℓ_0 norm using the continuous Gaussian function $f_\sigma(x) = \exp\left(-x^2/(2\cdot\sigma^2)\right)$ [12]. The accuracy of this approximation is controlled by the parameter $\sigma \in \mathbb{R}_+$ as

$$\lim_{\sigma\to 0}\left(1 - f_\sigma(x)\right) = ||x||_0 \quad \text{for } x \in \mathbb{R} \tag{2}$$

Using this approximation in (1) yields the problem:

$$\begin{aligned} \text{minimise} \quad & \sum_{i=1}^{N}\left(1 - f_\sigma(\hat{x}_i)\right) \\ \text{subject to} \quad & \mathbf{y} = \mathbf{A}\hat{\mathbf{x}} \end{aligned} \tag{3}$$

The SL0 approach to solving the problem in (1) is to solve the problem in (3) for each σ in a sequence (σ_i) abiding by two underlying thoughts: 1) The sequence should be such that local minima are avoided. 2) The value of σ should gradually decrease to increase the accuracy of the approximation in (3). The suggested sequence is the geometric sequence where $\sigma_i = \sigma_{\text{start}} \cdot \sigma_{\text{up}}^i$ for i from 0 to the largest integer where $\sigma_i > \sigma_{\text{min}}$. For each σ, the problem in (3) is solved by repeatedly taking an unconstrained gradient step and projecting $\hat{\mathbf{x}}$ back onto the feasible set.

The SL0 algorithm from [12] is stated in Algorithm 1. Notation-wise, $\mathbf{A}^\dagger$ denotes the Moore-Penrose pseudo-inverse of the matrix $\mathbf{A}$, $\mathbf{x} \circ \mathbf{y}$ denotes the Hadamard product of the vectors $\mathbf{x}$ and $\mathbf{y}$, and

$$\begin{aligned} \exp(\mathbf{x}) &= [\exp(x_1) \ \ldots \ \exp(x_\text{N})]^T \in \mathbb{R}^{N\times 1} \\ \max|\mathbf{x}| &= \max\{|x_1|, \ldots, |x_\text{N}|\} \in \mathbb{R} \end{aligned}$$

The best choice of σ_{up} and σ_{min} is problem-dependent (with $\sigma_{\text{up}} \in [0.5\,;1[$ typically) whereas $\sigma_{\text{start}} = 2\cdot\max|\hat{\mathbf{x}}|$, iteration-count $L = 3$, and step-size $\mu = 1$ are problem-independent recommendations given in [12].

For fixed reconstruction quality and time, we have attempted to improve the reconstruction capabilities of Algorithm 1. Specifically, we have carried out an extensive empirical analysis using the simulation framework presented in Section 3, with the objective of finding the parameter values that provide the best phase transition on a worst-case problem [4]. Through the analysis, we have made three overall empirically based observations which have given rise to successive modifications of Algorithm 1. The three resulting phase transition curves are shown in Figure 1 alongside the phase transition curve of Algorithm 1 and the theoretical ℓ_1 curve presented in [9][1]. In brief, the three observations and modifications are:

1) A larger σ_{start} is required for smaller δ whereas a smaller σ_{start} is allowable for larger δ: an inversely proportional relation between σ_{start} and δ is a good model. In the algorithm, we modify (σ_i) by introducing the relation $\sigma_{\text{start}} = 1/(2.75\cdot\delta)\cdot\max|\hat{\mathbf{x}}|$, increasing the value of σ_{up} to 0.7, and letting the value of σ_{min} remain unchanged.

2) A larger L is required as σ decreases, i.e. as the accuracy of the Gaussian approximation increases. In the algorithm, we initially let $L = 2$ and multiply this with $L_{\text{up}} = 2$ after each update of σ.

3) The recommended μ is likely to make $\hat{\mathbf{x}}$ converge to local minima when $\sigma_i \approx \sigma_{\text{start}}$: making μ a step-function betters this. In the algorithm, we replace the constant step-size μ with a variable step-size:

$$\mu(\sigma_i, i) = \begin{cases} \mu_{\text{start}} & i < 4 \ \vee \ \sigma_i > \sigma_{\text{thres}} \\ \mu_{\text{end}} & \text{otherwise} \end{cases}$$

where $\sigma_{\text{thres}} = 0.75\cdot\max\left|\mathbf{A}^\dagger\mathbf{y}\right|$, $\mu_{\text{end}} = 1.5$, and

$$\mu_{\text{start}} = \begin{cases} 0.05 & \delta \le 1/2 \\ 0.001 & \text{otherwise} \end{cases}$$

To allow a larger value of σ_{up} and L without increasing the reconstruction time, we terminate the inner loop when the algorithm is sufficiently close to convergence. That is, we choose to terminate the inner loop when the relative change $||\hat{\mathbf{x}} - \hat{\mathbf{x}}_{\text{prev}}||_2$ falls below $\sigma\cdot\epsilon$ where $\epsilon = 0.01$. The resulting modified SL0 algorithm is presented in Algorithm 2.

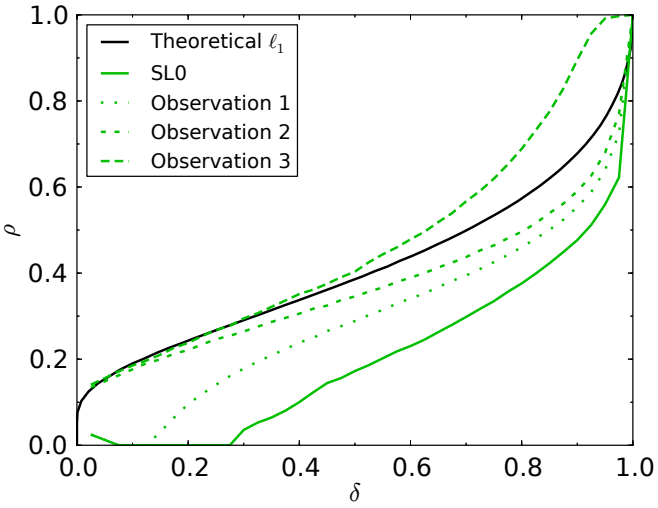

Fig. 1: Phase transition curve of the SL0 algorithm alongside the phase transition curves resulting from three overall observations. The theoretical ℓ_1 curve is included for reference.

3. SIMULATION FRAMEWORK

We evaluate the reconstruction capabilities of an algorithm by use of the phase transition measure [9]. Hence, let k denote the number of non-zero entries in the true vector $\mathbf{x}$. Define

[1] Tabulated values of this phase transition is available at `http://ecos.maths.ed.ac.uk/polytopes.shtml`

Algorithm 1 - Smoothed ℓ_0 norm (SL0)

init: $\sigma_{\text{up}} = 0.5$, $\sigma_{\text{min}} = 0.01$, $L = 3$, $\mu = 1$
$\hat{\mathbf{x}} \leftarrow \mathbf{A}^{\dagger}\mathbf{y}$
$\sigma \leftarrow 2 \cdot \max|\hat{\mathbf{x}}|$
while $\sigma > \sigma_{\text{min}}$ **do**
  for $j = 1 \ldots L$ **do**
    $\mathbf{d} \leftarrow \hat{\mathbf{x}} \circ \exp\big(-(\hat{\mathbf{x}} \circ \hat{\mathbf{x}})/(2 \cdot \sigma^2)\big)$
    $\hat{\mathbf{x}} \leftarrow \hat{\mathbf{x}} - \mu \cdot \mathbf{d}$ # Gradient step
    $\hat{\mathbf{x}} \leftarrow \hat{\mathbf{x}} - \mathbf{A}^{\dagger}(\mathbf{A}\hat{\mathbf{x}} - \mathbf{y})$ # Projection
  end for
  $\sigma \leftarrow \sigma \cdot \sigma_{\text{up}}$
end while

the measure of indeterminacy $\delta = n/N$ and the generalised measure of sparsity (density) $\rho = k/n$. Given a success criterion, the probability of reconstruction is then evaluated on the phase space $\delta, \rho \in [0, 1]$. In general, reconstruction is easier for larger δ and smaller ρ; it then gets more difficult when decreasing δ and when increasing ρ. The phase space is, consequently, separated by the phase transition curve into a phase where reconstruction is likely and one where it is unlikely.

The probability of reconstruction versus ρ for fixed N and δ may be modelled by a logistic regression [4], [13]. Thus, we adopt the approach from [4]: estimate the location of the phase transition curve through logistic regression with the reconstruction success criterion given by:

$$\frac{||\hat{\mathbf{x}} - \mathbf{x}||_2^2}{||\mathbf{x}||_2^2} < 10^{-4} \tag{4}$$

where $\hat{\mathbf{x}}$ and $\mathbf{x}$ are the reconstructed and true vectors, respectively. If the criterion is not met, the attempted reconstruction is considered unsuccessful. That is, the attempted reconstruction cannot be considered indeterminate.

For the simulations, we fix $N = 800$ as proposed in [4]. We then attempt reconstruction on a 4000 points uniform grid (δ, ρ) in the phase space. The grid is specified by:

$$\delta \in \{0.025, 0.05, \ldots, 1.00\} \tag{5}$$

$$\rho \in \{0.01, 0.02, \ldots, 1.00\} \tag{6}$$

For each point in the grid, we evaluate 100 different draws of $\mathbf{A}$ and $\mathbf{x}$. This number may be justified as follows. Let the outcome of an attempted reconstruction, X, in a given point be a Bernoulli random variable with probability p of success, i.e. $X \sim \text{Ber}(p)$. The number of successful reconstructions Y out of M attempts in that point is Binomially distributed, i.e. $Y \sim \text{Bin}(M, p)$. Using a normal approximation of Y, a worst case (at $p = 0.5$) 95 % confidence interval of length $l = 0.2$ then results in an upper bound for the required number of attempts (see e.g. [14]), i.e. $M \leq 1.96^2/l^2 \approx 100$. We then assume that this is a sufficient precision for the logistic regression (which takes 100 points on the ρ-axis into account) to determine the phase transition curve.

Algorithm 2 - Modified smoothed ℓ_0 norm (SL0-mod)

init: $\sigma_{\text{up}} = 0.7$, $\sigma_{\text{min}} = 0.01$, $L = 2$, $L_{\text{up}} = 2$, $\epsilon = 0.01$, $i = 0$
$\hat{\mathbf{x}} \leftarrow \mathbf{A}^{\dagger}\mathbf{y}$
$\sigma_i \leftarrow 1/(2.75 \cdot \delta) \cdot \max|\hat{\mathbf{x}}|$
while $\sigma_i > \sigma_{\text{min}}$ **do**
  for $j = 0 \ldots L$ **do**
    $\hat{\mathbf{x}}_{\text{prev}} \leftarrow \hat{\mathbf{x}}$
    $\mathbf{d} \leftarrow \hat{\mathbf{x}} \circ \exp\big(-(\hat{\mathbf{x}} \circ \hat{\mathbf{x}})/(2 \cdot \sigma_i^2)\big)$
    $\hat{\mathbf{x}} \leftarrow \hat{\mathbf{x}} - \mu(\sigma_i, i) \cdot \mathbf{d}$
    $\hat{\mathbf{x}} \leftarrow \hat{\mathbf{x}} - \mathbf{A}^{\dagger}(\mathbf{A}\hat{\mathbf{x}} - \mathbf{y})$
    if $||\hat{\mathbf{x}} - \hat{\mathbf{x}}_{\text{prev}}||_2 < \sigma_i \cdot \epsilon$ **then**
      break
    end if
  end for
  $\sigma_{i+1} \leftarrow \sigma_i \cdot \sigma_{\text{up}}$; $L \leftarrow L \cdot L_{\text{up}}$; $i \leftarrow i + 1$
end while

We have chosen the distributions of $\mathbf{A}$ and $\mathbf{x}$ (known as the problem suite) to be those that are most difficult to obtain good phase transitions on [4]: we draw $\mathbf{A}$ from the Uniform Spherical Ensamble (USE) [4] and draw the non-zero entries in $\mathbf{x}$ from the Rademacher distribution, i.e. $\{-1, 1\}$ with equal probability. To simulate our modified algorithm on a problem suite which the algorithm has not been optimised for, we have included the problem suite with $\mathbf{A}$ drawn from USE and the non-zero entries in $\mathbf{x}$ drawn from the zero-mean, unit-variance Gaussian distribution.

For the experiments we used an Intel Core i7 970 6-core 3.2 GHz based PC with 24 GiB DDR3 RAM running 64-bit Ubuntu 12.04 LTS Linux. Our implementation[2] of the tested algorithms is in Python. The simulations were carried out in double precision using the Enthought Python Distribution (EPD) 7.2-2 (64-bit). To validate our simulation results, we also simulated the IHT algorithm presented in [4].

4. EXPERIMENTAL RESULTS

Three algorithms have been simulated: 1) SL0 which is the algorithm stated in Algorithm 1. 2) SL0-mod which is the modified algorithm stated in Algorithm 2. 3) IHT which is the Iterative Hard Thresholding algorithm described in [4]. The experimental results are presented in Figures 2 and 3. Figure 2 shows the phase transitions for Rademacher distributed non-zero entries in $\mathbf{x}$ whereas Figure 3 shows the phase transitions for zero-mean, unit-variance Gaussian non-zero entries in $\mathbf{x}$. In both figures, the theoretical ℓ_1 curve from [9] is included for reference. To validate our simulation results, Figure 2 also shows the IHT phase transition curve points listed in [4].

[2]The source is available at:
http://dx.doi.org/10.5278/VBN/MISC/D4U3SF7B

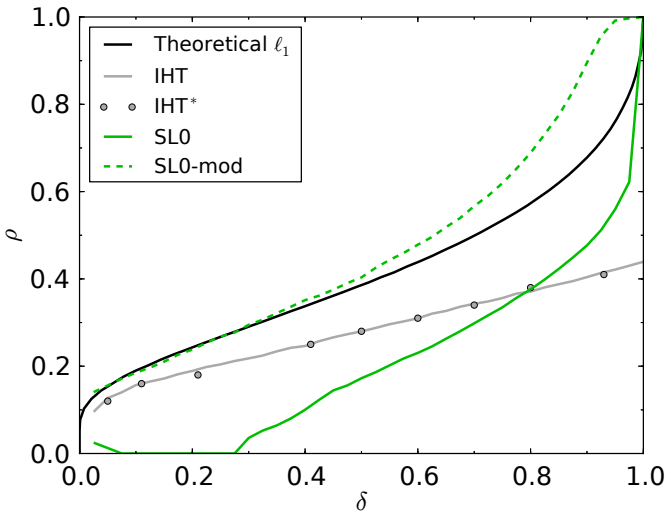

Fig. 2: The phase transition curves for the original SL0 algorithm and our modified SL0 algorithm. The theoretical ℓ_1 curve and the phase transition curve for the IHT algorithm are included for reference. IHT* is from table II in [4]. Problem suite: USE/Rademacher.

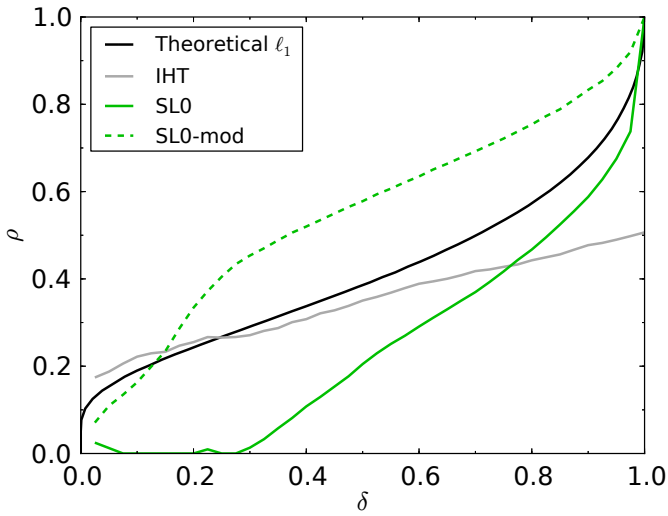

Fig. 3: The phase transition curves for the original SL0 algorithm and our modified SL0 algorithm. The theoretical ℓ_1 curve and the phase transition curve for the IHT algorithm are included for reference. Problem suite: USE/Gaussian.

5. DISCUSSION

We have attempted to improve the reconstruction capabilities of the SL0 algorithm without impairing the reconstruction quality and time. We claim that the reconstruction quality is fixed since the phase transition measure uses the static reconstruction success criterion in (4). Furthermore, we claim that the reconstruction time is unimpaired based on the reconstruction times of successful reconstructions, t_s, measured in the experimental results; mean$(t_\mathrm{s}) = 71\,\mathrm{ms}$ and std$(t_\mathrm{s}) = 27\,\mathrm{ms}$ for SL0-mod whereas mean$(t_\mathrm{s}) = 78\,\mathrm{ms}$ and std$(t_\mathrm{s}) = 23\,\mathrm{ms}$ for SL0.

The phase transition curves in Figures 2 and 3 show that the reconstruction capabilities have been improved significantly. However, there exist other attempts at optimising the SL0 algorithm such as the attempt which produced the phase transitions presented in [10]. For Rademacher non-zero entries in $\mathbf{x}$, our proposed algorithm provides significantly improved phase transition. However, for Gaussian non-zero entries in $\mathbf{x}$, our proposed algorithm provides a slightly poorer phase transition for $\delta < 0.25$. It should be noted that the focus of [10] is not on the SL0 algorithm and, consequently, it is unclear how the algorithm has been modified.

This simple comparison stresses the importance of defining the optimisation criteria and selecting a set of problem suites. Whereas both the optimisation criteria and selected problem suites of [10] are unknown, we have emphasised reconstruction capabilities and tuned for the USE/Rademacher problem suite used in Figure 2. In that respect, our proposed algorithm outperforms the original algorithm as well as the modified algorithm used in [10]. To achieve similar results for another problem suite, the algorithm should be tuned for that problem suite.

6. CONCLUSIONS

We have obtained a significantly improved phase transition for the SL0 algorithm without impairing its reconstruction quality and time. Moreover, our modified SL0 algorithm is applicable to any reconstruction problem since no assumptions are made about the signals involved. A large set of simulations based on the USE/Rademacher problem suite have shown that the phase transition curve of our modified SL0 algorithm surpasses the theoretical ℓ_1-curve. A very significant result since USE/Rademacher is a worst case problem suite. This result addresses those compressive sensing reconstruction problems, requiring ℓ_1 reconstruction capabilities, which have previously been rendered infeasible by the required reconstruction time of ℓ_1 based methods: the proposed modified SL0 algorithm has a required reconstruction time comparable to that of greedy algorithms while surpassing the theoretical 1-norm phase transition.

7. RELATION TO PRIOR WORK

The present work outlines a modification to the compressive sensing reconstruction algorithm "smoothed ℓ_0 norm" (SL0) studied in [12], [15], and [16]. The modification exploits the known problem indeterminacy in an attempt at optimising the reconstruction capabilities of SL0 when evaluated in the phase transition framework discussed in [4]. The result is a highly improved empirical phase transition for fixed reconstruction quality and time. While a phase transition curve has been presented for SL0 in [10] (with an unknown set of parameters), our obtained phase transition is superior and even surpasses that of the theoretical ℓ_1-curve given in [9].

8. REFERENCES

[1] Emmanuel J. Candès and Michael B. Wakin, "An Introduction To Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[2] David L. Donoho, "For Most Large Underdetermined Systems of Linear Equations the Minimal 1-norm Solution is also the Sparsest Solution," Tech. Rep. 2004-9, Department of Statistics, Stanford University, USA, Sept. 2004.

[3] David L. Donoho, "Compressed Sensing," *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[4] Arian Maleki and David L. Donoho, "Optimally Tuned Iterative Reconstruction Algorithms for Compressed Sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 330–341, Apr. 2010.

[5] Stéphane G. Mallat and Zhifeng Zhang, "Matching Pursuits With Time-Frequency Dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, Mar. 1993.

[6] Thomas Blumensath and Mike E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[7] Joel A. Tropp and Anna C. Gilbert, "Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[8] D. Needell and J.A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[9] David L. Donoho and Jared Tanner, "Precise Undersampling Theorems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 913–924, June 2010.

[10] Bob L. Sturm, Mads G. Christensen, and Rémi Gribonval, "Cyclic Pure Greedy Algorithms for Recovering Compressively Sampled Sparse Signals," in *45th IEEE Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove (CA), USA, Nov. 6–9, 2011, pp. 1143–1147.

[11] Thomas Blumensath and Mike E. Davies, "Normalized Iterative Hard Thresholding: Guaranteed Stability and Performance," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 298–309, Apr. 2010.

[12] G. Hosein Mohimani, Massoud Babaie-Zadeh, and Christian Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed l0 norm," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 289–301, Jan. 2009.

[13] David Donoho and Jared Tanner, "Observed Universality of Phase Transitions in High-Dimensional Geometry, with Implications for Modern Data Analysis and Signal Processing," *Phil. Trans. R. Soc. A*, vol. 367, no. 1906, pp. 4273–4293, Nov. 2009.

[14] Sheldon M. Ross, *Introduction to Probability and Statistics for Engineers and Scientists*, Elsevier Academic Press, Burlington, MA 01803, USA, fourth edition, 2009.

[15] G. Hosein Mohimani, Massoud Babaie-Zadeh, and Christian Jutten, "Complexed-Valued Sparse Representation Based on Smoothed l0 Norm," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)*, Las Vegas, Nevada, USA, Mar. 31 – Apr. 4, 2008, pp. 3881–3884.

[16] Armin Eftekhari, Massoud Babaie-Zadeh, Christian Jutten, and Hamid Abrishami Moghaddam, "Robust-SL0 For Stable Sparse Representation In Noisy Settings," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009)*, Taipei, Taiwan, Apr. 19–24, 2009, pp. 3433–3436.