Sparsity-undersampling tradeoff of compressed sensing in the complex domain

2 authors:

Zai Yang
Nanjing University of Science and Technology
38 PUBLICATIONS   1,321 CITATIONS

Cishen Zhang
Swinburne University of Technology
227 PUBLICATIONS   2,781 CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Stochastic Computation (SC) View project

SPARSITY-UNDERSAMPLING TRADEOFF OF COMPRESSED SENSING IN THE COMPLEX DOMAIN

Zai Yang and Cishen Zhang

School of Electrical and Electronic Engineering, Nanyang Tech. Univ., Singapore, 639798
yang0248@e.ntu.edu.sg, ecszhang@ntu.edu.sg

ABSTRACT

In this paper, recently developed ONE-L1 algorithms for compressed sensing are applied to complex-valued signals and sampling matrices. The optimal and iterative solution of ONE-L1 algorithms enables empirical investigation and evaluation of the sparsity-undersampling tradeoff of ℓ_1 minimization of complex-valued signals. A remarkable finding is that, not only there exists a sharp phase transition for the complex case determining the behavior of the sparsity-undersampling tradeoff, but also this phase transition is different and superior to that for the real case, providing a significantly improved success phase in the transition plane.

Index Terms— Compressed sensing, complex signal, ℓ_1 minimization, ONE-L1 algorithms, phase transition.

1. INTRODUCTION

Compressed sensing (CS) has made significant impact to information theory and signal processing since the pioneering works of Candès, Romberg and Tao [1] and Donoho [2]. It aims at reconstructing a signal from fewer linear measurements under a sparsity condition, which can be satisfied in most signals in practical applications, i.e. most signals are compressible. Compared with the conventional linear reconstruction, a nonlinear scheme is used in CS, i.e. to solve the convex optimization problem

$$(P_1) \qquad \min_x \|x\|_1, \text{subject to } Ax = b,$$

where $x \in \mathbb{C}^N$ is the sparse signal to recover, $A \in \mathbb{C}^{n \times N}$ is the sampling matrix and $b \in \mathbb{C}^n$ is the sample data vector, with the signal length N, sample size n and $n < N$, respectively.

The theory of CS has been mainly focused on how aggressively a signal can be undersampled, or under what condition a sparse signal can be exactly reconstructed from undersampled data by solving (P_1). The most precise undersampling theorem is given in [3] based on combinatorial geometry for real-valued A and x where A is random with i.i.d. Gaussian entries. It asserts that the behavior of sparsity-undersampling tradeoff is precisely controlled by the sampling ratio $\delta = n/N$ and sparsity ratio $\rho = k/N$ with k being the number of nonzero entries of x. Accordingly, there exists a curve, called the phase transition, which divides the plane of (δ, ρ) into two phases. Moreover, the precise expression of this phase transition has also been calculated [3].

To the best of our knowledge, the existing phase transition theory only deals with the real-valued case where the sparse signal and sampling matrix are both in the real domain. So far there have been few results studying the phase transition for CS in the more general complex domain. In this paper we investigate the sparsity-undersampling tradeoff for CS in the complex domain in the term of the phase transition. Its significance is twofold: first, it discovers a sharp phase transition in the complex domain so to complement the theory of phase transition; and next, it can provide a fundamental support to numerous applications involving complex data and signal processing, such as magnetic resonance imaging (MRI), radar and terahertz imaging [4, 5, 6].

In this paper, we show that the recently developed orthonormal expansion ℓ_1-minimization (ONE-L1) algorithms [7] for CS by the authors and collaborators are also applicable to the complex-valued case. The algorithms can compute the optimal solution of (P_1) with low computational complexities and can thus be used to empirically explore the sparsity-undersampling tradeoff of ℓ_1 minimization of complex-valued signals. A remarkable finding is that, not only there exists a sharp phase transition for the complex case determining the behavior of the sparsity-undersampling tradeoff, but also this phase transition is different and superior to that for the real case, providing a significantly improved area of the success phase in the transition plane. It shows that a random complex-valued sparse signal is easier to reconstruct than a real-valued one, in the sense that less samples are needed. Moreover, the universality of phase transitions [8] also holds for the complex-valued case.

2. ONE-L1 ALGORITHMS

This section first retrospects the main results in [7] on ONE-L1 algorithms for solving the ℓ_1 minimization problem (P_1), where x, A and b are all in the real domain, the sampling matrix A is assumed to be partial-orthonormal, i.e. A is composed of a subset of rows of an orthonormal matrix. A novel

technique called orthonormal expansion method is introduced to reformulate (P_1) as

$$(P_1^o) \quad \begin{array}{l} \min_{(x,p)} \|x\|_1 \\ \text{subject to } \Phi x = p \text{ and } \Gamma(p) = b, \end{array}$$

where orthonormal matrix $\Phi = \begin{bmatrix} A \\ B \end{bmatrix}$, $\Gamma(p)$ is an operator projecting the vector p onto its first n entries.

Exact ONE-L1 algorithm is proposed to exactly solve (P_1^o) based on augmented Lagrange multiplier (ALM) method. A relaxed version called relaxed ONE-L1 algorithm further speeds up exact ONE-L1 via updating the variables in the inner loop iteration only once. Numerical simulation shows that relaxed ONE-L1 algorithm is optimal in the sparsity-undersampling tradeoff under reasonable parameter settings. On the other hand, the relaxed ONE-L1 algorithm is equivalent to the following iteration:

$$x_{t+1} = S_{\mu_t^{-1}} \left(x_t + A' z_t \right),$$

$$z_t = b - A \left[\left(1 + \frac{\mu_{t-1}}{\mu_t} \right) x_t - \frac{\mu_{t-1}}{\mu_t} x_{t-1} \right] + \frac{\mu_{t-1}}{\mu_t} z_{t-1},$$

starting from $x_t = 0$ and $z_t = 0$ as $t \leq 0$. Hence, it is of iterative soft thresholding type (with some modifications) so can perform very fast computing. The relaxed ONE-L1 algorithm, in the worst case, has a computational complexity of $O\left(N \log^2 N\right)$ for some fast operator A.

For $w \in \mathbb{C}$, the soft thresholding of w with threshold $\varepsilon \in \mathbb{R}^+$ is defined as

$$S_\varepsilon(w) = \text{sgn}(w) \cdot (|w| - \varepsilon)^+,$$

where $(\cdot)^+ = \max(\cdot, 0)$ and $\text{sgn}(w) = \begin{cases} w/|w|, & w \neq 0; \\ 0, & w = 0. \end{cases}$ Using the definition above, ONE-L1 algorithms can be extended to the complex case, after minor and straightforward modifications.

3. EMPIRICAL SPARSITY-UNDERSAMPLING TRADEOFF

3.1. Phase Transition

Phase transition is a measure of the sparsity-undersampling tradeoff. For the real-valued signal x and sampling matrix A with A having random i.i.d. Gaussian entries, the sparsity-undersampling tradeoff of (P_1) is controlled by the sampling ratio δ and sparsity ratio ρ, as the signal length $N \to \infty$ [3], i.e. the plane of (δ, ρ) is divided by a curve, called the phase transition, into two phases, a 'success' phase where (P_1) successfully recovers the sparse signal and a 'failure' phase where the original signal cannot be recovered by solving (P_1), both with an overwhelming probability.

For real-valued x and A, three different expressions of the phase transition of (P_1) have been given based on different methods, including combinatorial geometry [3], null space method [10] and state evolution [11]. The three theoretically calculated results agree with each other although derived from different methods. We call this phase transition for the real-valued case as the real phase transition of (P_1). In theory, the phase transition is defined for the case as $N \to \infty$. Numerical simulations show that the observed phase transition coincides with the theoretical calculation even for modestly large N, e.g. $N = 1000$ [7, 8], where the finite-N phase transition is defined as the value of ρ at which the original signal is successfully reconstructed with the probability of 50%. Moreover, it is observed that the assumption of Gaussianity of A can be relaxed in practice, known as the universality of phase transitions [8].

While the existing results on the phase transition are only for the real-valued x and A, the questions about whether there exists a sharp phase transition for the complex-valued x and A and whether the complex phase transition, if it exists, is different from or the same as the real phase transition are still open. We now explore answers to these questions.

3.2. Method of Estimating Phase Transition

ONE-L1 algorithms exactly solve (P_1) and, hence, achieve the optimal sparsity-undersampling tradeoff of (P_1). In the mean time, ONE-L1 are also iterative soft thresholding type algorithms and can perform very fast computing. Hence, they can provide an efficient tool for empirically exploring the phase transition of (P_1).

The implementations of ONE-L1 algorithms are the same as that of the real versions in [7]. We fix $r > 1$ and let $\mu_{t+1} = r \cdot \mu_t$. The parameter r is set to $r = 1 + \delta$ in the exact ONE-L1 algorithm, while $r = \min(1 + 0.04\delta, 1.02)$ is chosen in relaxed ONE-L1. The success of recovering the original signal can be stated if the relative root mean squared error (relative RMSE) $\|\hat{x} - x^o\|_2 / \|x^o\|_2 < 10^{-4}$, where x^o and $\hat{x}$ are the original and recovered signal, respectively. Meanwhile, the failure of ONE-L1 algorithms solving (P_1) is stated if $\|\hat{x}\|_1 \geq (1 + 10^{-5}) \|x^o\|_1$ and $\|\hat{x} - x^o\|_2 / \|x^o\|_2 \geq 10^{-4}$.

Following from the the estimation of phase transition in [8, 11, 7], we first set a matrix ensemble, e.g. Gaussian, and dimension N. Then a grid of (δ, ρ) is generated in the plane $[0, 1] \times [0, 1]$, where equispaced $\delta \in \{0.02, 0.05, \cdots, 0.98\}$ and equispaced $\rho \in \{\rho_R(\delta) + 0.01(i - 21) : i = 1, 2, \cdots, 41\}$ with respect to each δ with $\rho_R(\delta)$ denoting the real phase transition. For each combination of (δ, ρ), $M = 20$ random problem instances are generated and solved with $n = \lceil \delta N \rceil$ and $k = \lceil \rho n \rceil$. The number of success among M instances is recorded. After data acquisition, a generalized linear modal (GLM) with a logistic link is used to estimate the phase transition as that in [7, 11].

3.3. Phase Transition of Partial-Fourier Sampling

In this subsection, we explore the sparsity-undersampling tradeoff of (P_1) with partial-Fourier sampling. Four values
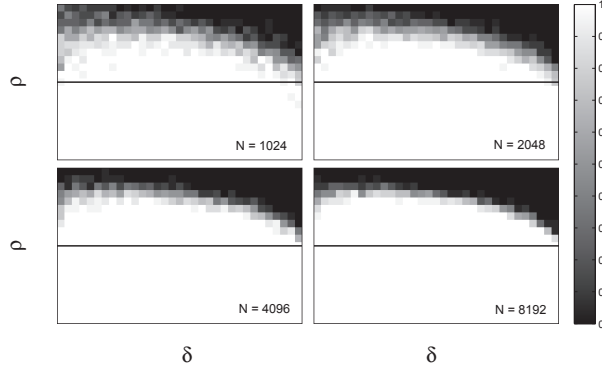
Fig. 1. Observed success rates of partial-Fourier sampling. Four images show results of relaxed ONE-L1 with $N = 1024$, 2048, 4096 and 8192, respectively. Sparsity ratio ρ is from $\rho_R(\delta) - 0.1$ (bottom) to $\rho_R(\delta) + 0.1$ (top) with an interval 0.01 with respect to each δ (here only the middle half of experiment results is shown). The middle line in each image refers to the real phase transition $\rho_R(\delta)$ of (P_1).

Fig. 2. Observed phase transitions of partial-Fourier sampling. Note that all observed phase transitions of partial-Fourier sampling by ONE-L1 algorithms coincide and are superior to the real phase transition.

of signal length N are considered, including 1024, 2048, 4096 and 8192. When $N = 1024$, both exact ONE-L1 and relaxed ONE-L1 are used to estimated the phase transition of (P_1). Relaxed ONE-L1 is used for other N. Few failures of solving (P_1) occur when using relaxed ONE-L1. See time consumptions and numbers of failures in Table 1.
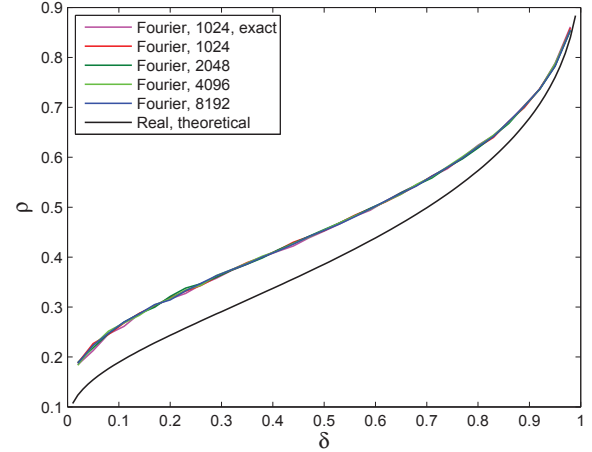
The observed success rates of our experiments are shown in Fig. 1, where the phase-transition performance can be observed. The phase transitions occur at about the same location and larger signal length N can provide clearer phase transition, which is consistent with the behavior of the real phase transition of (P_1).

Fig. 2 presents the observed phase transitions of partial-Fourier sampling. The five observed phase transitions of (P_1), estimated respectively by exact ONE-L1 with $N = 1024$ and relaxed ONE-L1 with $N = 1024$, 2048, 4096 and 8192, coincide with each other and are superior to the real phase transition of (P_1).

Based on our experimental results, a remarkable finding is stated as follows.

Finding 1 *For the partial-Fourier sampling and large dimension N, we observe*

I. *The ℓ_1 minimization approach exhibits phase transition in the plane of (δ, ρ), and larger N can result in smaller transition width.*

II. *The complex phase transition of ℓ_1 minimization is superior to the real phase transition with considerably enlarged success phase.*

Table 1. Failure Numbers and CPU Time Consumptions

Ensemble	Dim. N	Method	Failure number			Time (hour)
			$\delta = 0.02$	$\delta = 0.05$	Others	
Fourier	1024	exact	0(480)[1]	0(340)	0(8180)	6.26
Fourier	1024	relaxed	14(440)	4(340)	0(8280)	0.327
Fourier	2048	relaxed	1(320)	1(320)	0(7460)	0.594
Fourier	4096	relaxed	4(240)	0(260)	0(6980)	1.01
Fourier	8192	relaxed	1(260)	0(240)	0(6660)	2.15
Gaussian	1000	relaxed	16(420)	0(340)	0(8240)	5.56
Bernoulli	1000	relaxed	9(420)	2(360)	0(8340)	5.60
Ternary	1000	relaxed	9(360)	2(280)	0(8240)	5.65

[1] In the brackets is the corresponding number of problem instances solved.

Remark 1 *Finding 1 says that it is easier to reconstruct a complex-valued signal than to reconstruct a real signal in the sense that less samples are required. An intuitive explanation of this finding is as follows. Suppose that there are k nonzero entries in the complex signal x. It seems that we need to recover $2k$ variables from $2n$ samples when treating a complex number as two real numbers. But there is a constraint for this case, i.e. the $2k$ nonzero entries are, in fact, k pairs, or the freedom of the support of x is only k. This constrained freedom makes the complex signal essentially easier to recover.*

3.4. Observed Universality of Phase Transitions

The observed universality of phase transitions of (P_1) for the real-valued case has been discussed in [8] and the same result is also stated in [11, 7]. In this subsection, we examine whether the same property holds for the complex phase transitions of (P_1). Apart from the Fourier ensemble, three other complex matrix ensembles are considered with signal length $N = 1000$, including Gaussian, Bernoulli and Ternary.
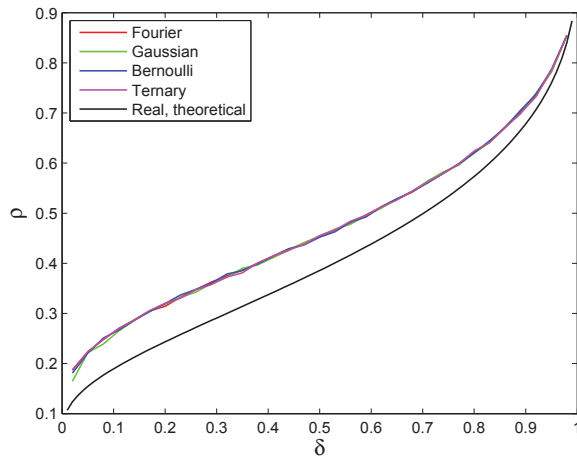
Fig. 3. Observed universality of phase transitions of (P_1) for the complex-valued case. The signal length $N = 8192$ for partial-Fourier matrix ensemble and $N = 1000$ for the other three cases.

All random matrices have i.i.d. real and imaginary parts following the corresponding distributions. Bernoulli refers to equally likely being 0 or 1, and Ternary is equally likely to be $-1, 0$ or 1.

Few failures in solving (P_1) occurred in our experiment, which can be seen in Table 1. Fig. 3 presents the observed phase transitions of (P_1) with different matrix ensembles. Like the universality of phase transitions of (P_1) for the real-valued case, we have the following finding.

Finding 2 *For large dimension N and a number of complex-valued matrix ensembles, the phase transitions of (P_1) coincide with that of (P_1) with partial-Fourier sampling.*

4. CONCLUSION

We have empirically explored the sparsity-undersampling tradeoff of ℓ_1 minimization with complex-valued signals and complex-valued matrix ensembles. It is found that similar properties hold as the real-valued case, such as the existence of phase transition and the observed universality across different matrix ensembles. A remarkable result is that the complex phase transition is superior to the real phase transition of (P_1) with considerably enlarged success phase, so to reconstruct a complex-valued sparse signal requires less samples and is easier. A rigorous mathematical derivation of this phase transition is still under investigation.

5. REFERENCES

[1] E.J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Info. Theo.*, vol. 52, no. 2, pp. 489–509, 2006.

[2] D.L. Donoho, "Compressed sensing," *IEEE Trans. Info. Theo.*, vol. 52, no. 4, pp. 1289–1306, 2006.

[3] D.L. Donoho and J. Tanner, "Counting the faces of randomly-projected hypercubes and orthants, with applications," *Discrete and Computational Geometry*, vol. 43, no. 3, pp. 522–541, 2010.

[4] W.L. Chan, K. Charan, D. Takhar, K.F. Kelly, R.G. Baraniuk, and D.M. Mittleman, "A single-pixel terahertz imaging system based on compressed sensing," *Appl. Phys. Lett.*, vol. 93, no. 12, pp. 121105, 2009.

[5] M. Lustig, D. Donoho, and J.M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[6] M.A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *Signal Processing, IEEE Transactions on*, vol. 57, no. 6, pp. 2275–2284, 2009.

[7] Z. Yang, C. Zhang, J. Deng, and W. Lu, "Orthonormal expansion ℓ_1-minimization algorithms for compressed sensing," *submitted for publication*, 2010.

[8] D. Donoho and J. Tanner, "Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing," *Phil. Trans. Royal Soci. A*, vol. 367, no. 1906, pp. 4273–4293, 2009.

[9] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure and Applied Math.*, vol. 57, no. 11, pp. 1413–1457, 2004.

[10] M. Stojnic, "Various thresholds for l1-optimization in compressed sensing," *arXiv:0907.3666*, 2009.

[11] D.L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *PNAS*, vol. 106, no. 45, pp. 18914–18919, 2009.