

PUSHING THE RIP PHASE TRANSITION IN COMPRESSED SENSING

Jeffrey D. Blanchard

Department of Mathematics and Statistics, Grinnell College
Grinnell, Iowa, 50112 USA
jeff@math.grinnell.edu

Andrew Thompson

School of Mathematics, University of Edinburgh
The King's Buildings, Mayfield Rd., Edinburgh EH9 3JZ, UK
a.thompson-8@sms.ed.ac.uk

ABSTRACT

We apply the asymmetric restricted isometry property (ARIP) to recent results of Cai, Wang, and Xu and formulate a two-parameter family of sufficient conditions for exact k -sparse signal recovery via ℓ^1 -minimization. We translate the sufficient conditions into the phase transition framework and apply bounds on the ARIP constants to define lower bounds on the phase transition. By selecting the parameters wisely, we determine sufficient conditions whose phase transition curves improve upon those already in the literature.

1. INTRODUCTION

In the simplest setting of compressed sensing (CS), one seeks the sparsest solution to an underdetermined system of linear equations. A predominant tool in the analysis of sparse signal recovery is the *Restricted Isometry Property* (RIP) of Candès and Tao [9]. Since the introduction of CS and the related increase in work on the analysis of recovery algorithms, the field has produced numerous RIP statements which serve as sufficient conditions for exact sparse recovery. A framework for comparison of CS results was set forth by Donoho et al. [11, 12, for example] and adapted to the RIP [1]. In this article, we apply the *asymmetric restricted isometry property* (ARIP) to recent work by Cai, Wang and Xu [6]. In doing so, and by resisting the temptation to force small support sizes on the ARIP [3], we present the sufficient condition for ℓ^1 -minimization which is provably satisfied by the largest region of Gaussian matrices. We demonstrate this by applying bounds on the ARIP constants [1] and comparing the region obtained from this analysis against the regions of previous results.

1.1 The Noiseless CS Problem and the RIP

We consider signals of dimension N and assume the target signal has no more than k -nonzero coefficients. Our target signal class is denoted $\mathcal{X}^N(k) = \{x \in \mathbb{R}^N : \|x\|_0 \leq k\}$ where $\|\cdot\|_0$ counts the nonzero entries. From $y = Ax$, we are provided n measurements of the target signal x from a linear encoder, A of size $n \times N$. Finally, the signal x is recovered from (y, A) using one of many CS decoders.

The natural decoder is a combinatorial search for the sparsest solution:

$$\min_{z \in \mathbb{R}^N} \|z\|_0 \quad \text{subject to} \quad y = Az. \quad (1)$$

If the encoder A provides distinct measurements for any two k -sparse vectors (for any $x, x' \in \mathcal{X}^N(k)$, $Ax \neq Ax'$), the solution to (1) is unique. It is now known that other more

tractable decoders can identify this unique solution under appropriate additional conditions on the encoder/decoder pair. Candès and Tao introduced the RIP [9] which permits a decoder analysis independent from the encoder and provides one form of these appropriate additional conditions for identifying the unique solution to (1). To admit the largest class of encoders which might satisfy an RIP condition, we adopt the ARIP formulation [1].

Definition 1.1 (RIP [9] and ARIP [1]). *For an $n \times N$ matrix A , the ARIP constants $L(k, n, N)$ and $U(k, n, N)$ are the smallest nonnegative numbers which satisfy*

$$(1 - L(k, n, N))\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + U(k, n, N))\|x\|_2^2 \quad (2)$$

for all $x \in \mathcal{X}^N(k)$.

The standard (symmetric) RIP constant $R(k, n, N)$ is then defined by

$$R(k, n, N) := \max\{L(k, n, N), U(k, n, N)\}. \quad (3)$$

Remark 1.2. We refer to the first entry of the RIP constant as the *support size*.

There are several decoders with sufficient conditions for exact k -sparse recovery based on the RIP. Greedy algorithm decoders with such guarantees include CoSaMP [15], Iterative Hard Thresholding [4], and Subspace Pursuit [10]. Here we focus on ℓ^1 -minimization:

$$\min_{z \in \mathbb{R}^N} \|z\|_1 \quad \text{subject to} \quad y = Az. \quad (4)$$

The ℓ^1 -decoder has been studied quite extensively in the CS setting and sufficient conditions for exact k -sparse recovery have been derived using multiple methods of analysis. In this article, we focus on the ℓ^1 decoder and the sufficient conditions based on an RIP analysis.

Remark 1.3. For simplicity, we focus exclusively on the ideal case of exact sparsity and exact measurements. Standard extensions of the RIP analysis to signals which are not exactly k -sparse or measurements corrupted by noise yield similar results. The phase transition curves in this article are upper bounds on the curves obtained for noisy measurements or compressible signals.

1.2 An Overview of RIP Conditions for ℓ^1 -minimization

The RIP was first used in CS by Candès and Tao [9] to establish a sufficient condition for recovery of a k -sparse vector. The first method of proof used the RIP to establish a separating hyperplane on the boundary of the feasible region for (4). In [8], Candès, Romberg, and Tao used a primal argument to establish a sufficient condition for successful recovery. The

JDB is an NSF International Research Fellow at the University of Edinburgh funded by the award NSF OISE 0854991.

general approach to their proof was to show that the error between the solution to (4) and the target signal x must be zero if the encoder A satisfied the RIP condition

$$\mu^{ct}(k, n, N) := \frac{1}{2}R(3k, n, N) + \frac{3}{2}R(4k, n, N) < 1. \quad (5)$$

The major step in the proof was to partition the index set in such a way that one could apply an RIP analysis. This technique was then further developed by Foucart and Lai in [14] where an ARIP analysis and an alternative partitioning of the index set resulted in the sufficient condition

$$\mu^{fl}(k, n, N) := \frac{1 + \sqrt{2}}{4} \left(\frac{1 + U(2k, n, N)}{1 - L(2k, n, N)} - 1 \right) < 1. \quad (6)$$

These theoretical results¹ are worst case guarantees rather than the average case performance observed in empirical testing. However, such theoretical guarantees are at the heart of the excitement surrounding CS and the fast decoders which (approximately) solve (1). We do not discuss average case performance nor empirical investigation, but we show the evolution of the theoretical guarantees based on the RIP.

1.3 The Phase Transition Framework

The results in Sec. 1.2 can be compared qualitatively in a variety of ways. The general approach in the literature is to simply formulate the results, using various inequalities, into statements involving a common support size of the RIP constants. Alternatively, using bounds on the ARIP constants, one may write the condition in a form that permits direct quantitative comparison via the phase transition framework advocated by Donoho. To obtain such a format, we consider the CS problem in the setting where the problem parameters (k, n, N) grow in a coordinated fashion.

Definition 1.4 (Proportional-Growth Asymptotic). *A sequence of problem sizes (k, n, N) is said to grow proportionally if, for $(\delta, \rho) \in [0, 1]^2$, $\frac{n}{N} \rightarrow \delta$ and $\frac{k}{N} \rightarrow \rho$ as $n \rightarrow \infty$.*

In the asymptotic setting, the extensive existing knowledge regarding singular values of Gaussian matrices was coupled with a large deviation analysis to obtain rather accurate² bounds on the ARIP constants [1]. We say that the matrix A is *drawn from the Gaussian ensemble* when the entries of A are taken i.i.d. from the normal distribution with mean zero and variance n^{-1} .

Theorem 1.5 (Blanchard, Cartis, Tanner [1]). *Fix $\varepsilon > 0$. Under the proportional-growth asymptotic, Definition 1.4, sample each $n \times N$ matrix A from the Gaussian ensemble. Let $\mathcal{L}(\delta, \rho)$ and $\mathcal{U}(\delta, \rho)$ be defined as in [1, Thm. 1]. Define $\mathcal{R}(\delta, \rho) = \max\{\mathcal{L}(\delta, \rho), \mathcal{U}(\delta, \rho)\}$. Then as $n \rightarrow \infty$,*

$$\text{Prob}[L(k, n, N) < \mathcal{L}(\delta, \rho) + \varepsilon] \rightarrow 1, \quad (7)$$

$$\text{Prob}[U(k, n, N) < \mathcal{U}(\delta, \rho) + \varepsilon] \rightarrow 1, \quad (8)$$

$$\text{and } \text{Prob}[R(k, n, N) < \mathcal{R}(\delta, \rho) + \varepsilon] \rightarrow 1. \quad (9)$$

Sufficient RIP conditions, such as those mentioned in Sec. 1.2 establish when the solution to (4) will coincide with

¹Equations (5) and (6) are representatives of the many RIP conditions in the literature.

²These bounds are within twice the ARIP constants observed during extensive empirical testing [1].

the solution to (1). We formalize this event as Strong ℓ^1/ℓ^0 Equivalence.

Definition 1.6 (Strong ℓ^1/ℓ^0 Equivalence). *The event $\text{StrongEquiv}(A, \ell^1)$ denotes the following property of an $n \times N$ matrix A : for every k -sparse vector x , ℓ^1 -minimization (4) exactly recovers x from the corresponding measurements $y = Ax$.*

Under the proportional-growth asymptotic there is a strictly positive function $\rho_S(\delta; \ell^1) \equiv \rho_S(\delta) > 0$ defining a region of the (δ, ρ) phase space which ensures successful recovery of every k -sparse vector $x \in \mathcal{X}^N(k)$, [9, 11, 12]. This function, $\rho_S(\delta)$, is called the *Strong phase transition function*.

Definition 1.7 (Region of Strong Equivalence). *Consider the proportional-growth asymptotic with parameters $(\delta, \rho) \in (0, 1) \times (0, 1/2)$. Draw the corresponding $n \times N$ matrices A from the Gaussian ensemble and fix $\varepsilon > 0$. Suppose that we are given a function $\rho_S(\delta)$ with the property that, whenever $0 < \rho < (1 - \varepsilon)\rho_S(\delta)$, $\text{Prob}(\text{StrongEquiv}(A, \ell^1)) \rightarrow 1$ as $n \rightarrow \infty$. We say that $\rho_S(\delta)$ bounds a region of strong equivalence.*

That is, if (δ, ρ) falls in the region of the phase space below the curve $\rho_S(\delta)$, then the probability of successfully k -sparse recovery via ℓ^1 -minimization converges to 1 when measurements are taken from the Gaussian ensemble. In this paper we apply the ARIP to a relatively recent proof technique of Cai, Wang, and Xu [6]. Then, by formulating the sufficient conditions in the phase transition framework we show that this technique yields results with larger regions of strong equivalence than previously known RIP conditions.

2. SUFFICIENT ARIP CONDITIONS FOR STRONG ℓ^1/ℓ^0 EQUIVALENCE

In a recent series of papers [5, 6, 7], Cai et al. have modified the existing RIP analysis to form an additional family of sufficient conditions on A which imply $\text{StrongEquiv}(A, \ell^1)$. While their motivation was the reduction of the support size of the RIP constants, our motivation is to attempt to close the gap between the regions of strong equivalence implied by the RIP and by more geometric, algorithm-specific analyses such as the polytope analysis of Donoho and Tanner [11, 13]. To do so, we define a function which based on an ARIP analysis following the techniques in [6].

Definition 2.1 (CWX ARIP Function). *For $\alpha, \beta > 0$ define a CWX ARIP Function by*

$$\mu_{\alpha, \beta}^{\text{CWX}}(k, n, N) := L((1 + \alpha)k, n, N) + \frac{1}{\sqrt{\beta}}\Omega(\alpha, \beta), \quad (10)$$

where

$$\Omega(\alpha, \beta) := \frac{L((1 + \alpha + \beta)k, n, N) + U((1 + \alpha + \beta)k, n, N)}{2}.$$

Definition 2.1 defines a family of CWX ARIP functions parameterized by α, β . For $\alpha, \beta \in \frac{1}{k}\mathbb{N}^+$ with $2\alpha \leq \beta \leq 4\alpha$, $\mu_{\alpha, \beta}^{\text{CWX}}(k, n, N)$ provides a sufficient condition for ℓ^1 recovery of every k -sparse signal.

Theorem 2.2. Suppose $\alpha, \beta \in \frac{1}{k}\mathbb{N}^+$ with $2\alpha \leq \beta \leq 4\alpha$ and that A is an $n \times N$ matrix with ARIP constants satisfying $\mu_{\alpha, \beta}^{\text{cwx}}(k, n, N) < 1$. Then we have $\text{StrongEquiv}(A, \ell^1)$ for k -sparse vectors.

The proof precisely follows the analysis of Cai et al. in [6] with the application of the ARIP that delivers the result in the form stated in 2.2. The general format of the proofs obtained by Cai et al. follow the original proofs of Candès et al. [8]. The index set $\{1, \dots, N\}$ is partitioned into subsets of a certain size which determine the support size of the ARIP constants. In [6], the Shifting Inequality is the main innovation permitting a flexible partitioning of the index set.

Lemma 2.3 (Shifting Inequality [6]). Suppose $q, r \in \mathbb{N}^+$ with $r \leq q \leq 3r$. If

$$c_1 \geq c_2 \geq \dots \geq c_{2r+q} \geq 0,$$

Then

$$\left(\sum_{i=r+1}^{2r+q} c_i^2 \right)^{\frac{1}{2}} \leq \frac{1}{\sqrt{r+q}} \sum_{i=1}^{r+q} c_i. \quad (11)$$

In our version of the proof, we employ the following implication of the ARIP constants which can be found in [2, Lem. 15(iv)]. For an index set $I \subset \{1, \dots, N\}$, $|I|$ denotes the cardinality of the set.

Lemma 2.4. Let $I, J \subset \{1, \dots, N\}$ with $I \cap J = \emptyset$ and $|I| + |J| = m$. Let A be an $n \times N$ matrix with ARIP constants $L(m, n, N), U(m, n, N)$. Then for any $h \in \mathbb{R}^N$,

$$|\langle Ah_I, Ah_J \rangle| \leq \frac{L(m, n, N) + U(m, n, N)}{2} \|h_I\|_2 \|h_J\|_2.$$

Proof of Theorem 2.2. Let $x \in \chi^N(k)$, $y = Ax$, and \hat{x} be the solution to (4). Define $h = |\hat{x} - x|$ and $T_0 = \text{supp}(x)$. Arrange the elements of h in decreasing order on the complement, T_0^c , and partition the complement into subsets

$$T_0^c = T_1 \cup \left(\bigcup_{j \geq 2} T_j \right)$$

where $|T_1| = \alpha k$ and $|T_j| = \beta k$ for $j \geq 2$. Now, for $j \geq 2$, further partition each set into the first αk largest entries and the last $(\beta - \alpha)k$ entries:

$$T_j = T_{j1} \cup T_{j2}; \quad j \geq 2$$

with $|T_{j1}| = \alpha k$ and $|T_{j2}| = (\beta - \alpha)k$. For notational brevity, let $T_{01} = T_0 \cup T_1$, $T_{12} = T_1$. Finally, h_{T_j} is the set of coefficients in h indexed by T_j . By design, $\alpha k \leq (\beta - \alpha)k \leq 3\alpha k$ so we may apply Lem. 2.3 to these sets. For $j \geq 2$,

$$\begin{aligned} \|h_j\|_2 &= \left(\sum_{i \in T_j} h_i^2 \right)^{\frac{1}{2}} \leq \frac{1}{\sqrt{\beta k}} \left(\sum_{i \in T_{(j-1)2}} h_i + \sum_{i \in T_{j1}} h_i \right) \\ &= \frac{1}{\sqrt{\beta k}} \left(\|h_{T_{(j-1)2}}\|_1 + \|h_{T_{j1}}\|_1 \right). \end{aligned} \quad (12)$$

So, using (12), we can write

$$\begin{aligned} \sum_{j \geq 2} \|h_j\|_2 &\leq \frac{1}{\sqrt{\beta k}} \sum_{j \geq 2} \left(\|h_{T_{(j-1)2}}\|_1 + \|h_{T_{j1}}\|_1 \right) \\ &= \frac{1}{\sqrt{\beta k}} \sum_{j \geq 1} \|h_{T_j}\|_1 = \frac{1}{\sqrt{\beta k}} \|h_{T_0^c}\|_1. \end{aligned} \quad (13)$$

From [8], we use $\|h_{T_0^c}\|_1 \leq \|h_{T_0}\|_1$, the Cauchy-Schwarz inequality, and the fact that $\|h_{T_0}\|_2 \leq \|h_{T_{01}}\|_2$ to see that

$$\sum_{j \geq 2} \|h_j\|_2 \leq \frac{1}{\sqrt{\beta}} \|h_{T_{01}}\|_2. \quad (14)$$

Finally, we use this partitioning and (14) to apply the ARIP constants. Since $Ax = A\hat{x}$, $Ah = 0$ and so

$$\begin{aligned} 0 &= |\langle Ah, Ah_{T_{01}} \rangle| = \left| \|Ah_{T_{01}}\|_2^2 + \sum_{j \geq 2} \langle Ah_{T_j}, Ah_{T_{01}} \rangle \right| \\ &\geq [1 - L((1 + \alpha)k, n, N)] \|h_{T_{01}}\|_2^2 \\ &\quad - \Omega(\alpha, \beta) \|h_{T_{01}}\|_2 \sum_{j \geq 2} \|h_{T_j}\|_2 \\ &\geq [1 - L((1 + \alpha)k, n, N)] \|h_{T_{01}}\|_2^2 - \Omega(\alpha, \beta) \frac{1}{\sqrt{\beta}} \|h_{T_{01}}\|_2^2 \\ &= \left[1 - \mu_{\alpha, \beta}^{\text{cwx}}(k, n, N) \right] \|h_{T_{01}}\|_2^2, \end{aligned} \quad (15)$$

where the sequence of inequalities relies on Def. 1.1, Lem. 2.4, and (14). Thus, if $\mu_{\alpha, \beta}^{\text{cwx}}(k, n, N) < 1$, (15) implies that $h_{T_{01}} = 0$ and therefore $\hat{x} = x$. \square

Theorem 2.2 provides a family of sufficient condition for exact k -sparse signal recovery in the form of the CWX ARIP functions $\mu_{\alpha, \beta}^{\text{cwx}}(k, n, N)$. This family of functions leads to two natural questions. First, when is it possible to satisfy the condition $\mu_{\alpha, \beta}^{\text{cwx}}(k, n, N) < 1$ and for which values of α, β will this condition be easiest to satisfy? Second, do any of the functions $\mu_{\alpha, \beta}^{\text{cwx}}(k, n, N)$ lead to regions of strong equivalence which are larger than the regions of strong equivalence determined by the either $\mu^{\text{crt}}(k, n, N)$ or $\mu^{\text{fl}}(k, n, N)$ defined, respectively, in (5) and (6).

To answer these question, it will be necessary to define a function which bounds the regions of strong equivalence. We apply the bounds from Thm. 1.5.

Definition 2.5 (CWX Region of $\text{StrongEquiv}(A, \ell^1)$). Define

$$\mu_{\alpha, \beta}^{\text{cwx}}(\delta, \rho) := \mathcal{L}(\delta, (1 + \alpha)\rho) + \frac{\mathcal{L}(\delta, (1 + \alpha + \beta)\rho) + \mathcal{U}(\delta, (1 + \alpha + \beta)\rho)}{2\sqrt{\beta}}$$

and $\rho_S^{\text{cwx}}(\delta; \alpha, \beta)$ as the solution to $\mu_{\alpha, \beta}^{\text{cwx}}(\delta, \rho) = 1$.

Theorem 2.6. Consider the proportional growth asymptotic with parameters $(\delta, \rho) \in (0, 1) \times (0, 1/2)$. Draw the corresponding $n \times N$ matrices A from the Gaussian ensemble. Fix $\varepsilon > 0$. If $\rho < (1 - \varepsilon)\rho_S^{\text{cwx}}(\delta; \alpha, \beta)$, then $\text{Prob}(\text{StrongEquiv}(A, \ell^1)) \rightarrow 1$ as $n \rightarrow \infty$.

Therefore the function $\rho_S^{\text{cwx}}(\delta; \alpha, \beta)$ bounds a region of strong equivalence for ℓ^1 -minimization.

Proof. By [2, Lem. 16], with overwhelming probability on the draw of A from the Gaussian ensemble $\mu_{\alpha, \beta}^{\text{cwx}}(k, n, N) <$

$\mu_{\alpha,\beta}^{cwx}(\delta, (1+\varepsilon)\rho)$ and $\mu_{\alpha,\beta}^{cwx}(\rho, \delta)$ is strictly increasing in ρ . Thus, by [2, Lem. 17], if $\rho < (1-\varepsilon)\rho_S^{cwx}(\delta; \alpha, \beta)$, then $\mu_{\alpha,\beta}^{cwx}(\delta, (1+\varepsilon)\rho) < 1$. Therefore, with overwhelming probability the hypotheses of Thm. 2.2 are satisfied and we have $StrongEquiv(A, \ell^1)$ for all k -sparse vectors. \square

In the next section, we present a heuristic argument which leads to a choice of α, β that provides the largest region of strong equivalence while permitting every problem instance (k, n, N) . Then we compare the regions of strong equivalence for matrices drawn from the Gaussian ensemble and associated to each of the three conditions $\mu_{\alpha,\beta}^{cwx}(k, n, N), \mu^{fl}(k, n, N), \mu^{crt}(k, n, N) < 1$.

3. PUSHING THE RIP REGION OF STRONG EQUIVALENCE

As discussed in [3], the general approach in the literature is to use the parameters α, β in $\mu_{\alpha,\beta}^{cwx}(k, n, N)$ to obtain a statement with RIP constants with the smallest possible support size. However, this does not produce the largest region of strong equivalence. Here we attempt to intelligently choose α and β to obtain the largest region of strong equivalence. Because no single choice of α, β results in a function $\rho_S^{cwx}(\delta; \alpha, \beta)$ which is largest for all values of δ , a heuristic argument for the choice of α, β is reasonable. First, we fix the problem instance (k, n, N) and observe that $\mu_{\alpha,\beta}^{cwx}(k, n, N)$ consists of two terms,

$$L((1+\alpha)k, n, N) \quad \text{and}$$

$$\frac{1}{2\sqrt{\beta}} (L((1+\alpha+\beta)k, n, N) + U((1+\alpha+\beta)k, n, N)).$$

The first term depends only on alpha and the second term has the important weight $1/(2\sqrt{\beta})$. Since the RIP constants are nondecreasing, we balance increasing the support sizes of the RIP constants with impact of the weight $1/(2\sqrt{\beta})$. As α has no effect on the weight and only plays the role of altering the support size of the ARIP constants, we take the minimum value of α in order to keep the support sizes of the RIP constants as small as possible. Therefore, we must choose $\alpha = \beta/4$ according to the hypotheses of Thm. 2.2. Then, we can rewrite the sufficient condition as $\mu_{\beta}^{cwx}(k, n, N) < 1$ for

$$\mu_{\beta}^{cwx}(k, n, N) := \mu_{\frac{1}{4}\beta, \beta}^{cwx}(k, n, N) =$$

$$L((1+\frac{1}{4}\beta)k, n, N) + \frac{L((1+\frac{5}{4}\beta)k, n, N) + U((1+\frac{5}{4}\beta)k, n, N)}{2\sqrt{\beta}}.$$

Now, the condition is dependent only on β and one must balance the growth rates of the ARIP constants as the support size grows against the benefits of increasing β to force a smaller weight on the second part of the condition with larger support sizes. This is highly dependent on the matrix ensemble from which A is drawn. Define $\rho_S^{cwx}(\delta; \beta)$ as the solution to $\mu_{\beta}^{cwx}(\delta, \rho) = 1$.

Choosing $\beta \in 4\mathbb{N}^+$ ensures that $\mu_{\beta}^{cwx}(k, n, N)$ applies to every problem instance (k, n, N) . It is possible that selecting $\beta \notin 4\mathbb{N}^+$ may result in larger regions of Strong Equivalence. For the Gaussian ensemble, such changes have a minimal effect on the magnitude of the phase transition curve. From $4\mathbb{N}^+$, $\beta = 4$ defines the phase transition function with the largest magnitude: $\rho_S^{cwx}(\delta; 4)$ is the red curve in Fig. 1.

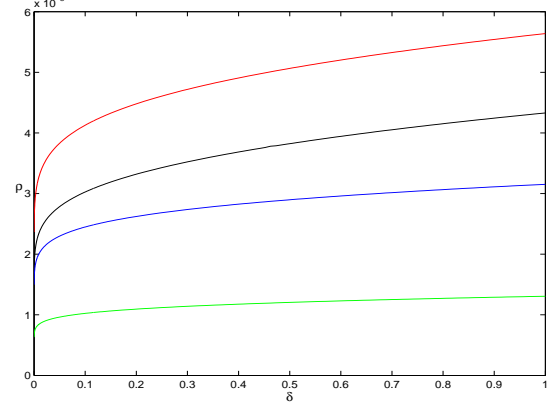


Figure 1: *Regions of Strong Equivalence*: The region below the curve $\rho_S(\delta)$ is a region of $StrongEquiv(A, \ell^1)$ for A drawn from the Gaussian ensemble; $(\rho_S^{cwx}(\delta; 4)$, red), $(\rho_S^{bt}(\delta)$, black), $(\rho_S^{fl}(\delta)$, blue), $(\rho_S(\delta)$, green).

Figure 1 also displays three other regions of strong equivalence defined by RIP results from the literature. First, $\rho_S^{crt}(\delta)$, displayed as the green curve in Fig. 1, bounds the region of strong equivalence determined by Candès et al. [8] and stated here as (5). The blue curve is defined by $\rho_S^{fl}(\delta)$ which is the region of strong equivalence associated with Foucart and Lai's RIP condition (6). Previously, the highest phase transition curves were represented by the following condition from [3] which are a generalization of (5); Let

$$\mu^{bt}(\delta, \rho) := \frac{11\mathcal{L}(\delta, 12\rho) + \mathcal{U}(\delta, 11\rho)}{10}, \quad (16)$$

and define $\rho_S^{bt}(\delta)$ as the solution to $\mu^{bt}(\delta, \rho) = 1$. Then $\rho_S^{bt}(\delta)$ bounds the region of strong equivalence and is displayed as the black curve in Fig. 1. Figure 1 gives a straightforward means of quantitatively stating that the method of analysis in [6] yields larger regions of strong equivalence. This figure also displays the advances over time and the improvements obtained by incorporating new partitioning techniques and the ARIP in the analysis.

The choice of $\beta = 4$ is not arbitrary. By plotting other regions of strong equivalence, it is clear that $\rho_S^{cwx}(\delta; 4)$ bounds the largest region of strong equivalence for $\beta \in 4\mathbb{N}^+$. If we are willing to let β vary continuously, it is possible to push the strong phase transition curve slightly higher. However, no other choice of β will result in a major improvement for Gaussian matrices. When $\delta \rightarrow 0$, the improvement approaches 2%, but even for $\delta > 0.1$ the improvement is never even 1% (or 5.7×10^{-5}). Figure 2 provides the improvement ratio for 3 alternative choices of β , which further demonstrates that no single choice of β is optimal.

Another interesting use of the phase transition curve is the extraction of a constant of proportionality for the required number of measurements implied by a given sufficient condition. Namely, by taking the inverse of a curve that bounds a region of strong equivalence, we can determine a constant $C = (\rho_S(n/N))^{-1}$ such that if $n > Ck$ then the problem instance (k, n, N) falls in the region of strong equivalence for Gaussian matrices. Thus the curves $(\rho_S(\delta))^{-1}$ de-

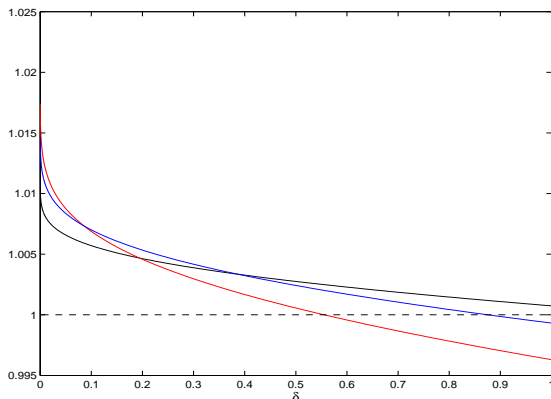


Figure 2: *Improvement Ratios for Selected β* : The ratio $\rho_S^{cwx}(\delta, \beta) / \rho_S^{cwx}(\delta; 4)$ for ($\beta = 3$, red), ($\beta = 3.25$, blue), ($\beta = 3.5$, black).

scribe the smallest constant of proportionality for which a particular theorem guarantees strong equivalence with overwhelming probability on the draw of A . Figure 3 shows the inverse of the red curve in Fig. 1 defined by $\rho_S^{cwx}(\delta; 4)$. In [9], Candès and Tao developed the first RIP based sufficient condition and determined that for Gaussian matrices, if $n/N = 1/2$, then $n > 2174k$ was a sufficient number of measurements for $StrongEquiv(A, \ell^1)$ with A from the Gaussian ensemble. Since that seminal paper, the introduction of new proof techniques, the ARIP analysis, and the improved bounds of Thm. 1.5 have dramatically reduced this constant of proportionality. Using $\rho_S^{cwx}(\delta; 4)$, if $n/N = 1/2$, then $n > 198k$ is a sufficient number of measurements to ensure the problem instance falls in the region of strong equivalence. However, there is still a gap from the necessary and sufficient conditions determined by Donoho and Tanner with a geometric algorithm specific analysis of the ℓ^1 decoder. Their polytope analysis tells us that if $n/N = 1/2$, the actual required number of measurements is $n > 11k$ to guarantee $StrongEquiv(A, \ell^1)$ with overwhelming probability on the draw of A from the Gaussian ensemble.

REFERENCES

- [1] J. D. Blanchard, C. Cartis, and J. Tanner. Compressed sensing: How sharp is the restricted isometry property? *SIAM Review*, in press, 2010.
- [2] J. D. Blanchard, C. Cartis, J. Tanner, and A. Thompson. Phase transitions for greedy sparse approximation algorithms. submitted, 2009.
- [3] J. D. Blanchard and A. Thompson. On support sizes of restricted isometry constants. *Appl. Comput. Harmon. Anal.*, in press, 2010.
- [4] T. Blumensath and M. E. Davies. Iterative hard thresholding for compressed sensing. *Appl. Comput. Harmon. Anal.*, 27(3):265–274, 2009.
- [5] T. T. Cai, L. Wang, and G. Xu. New bounds for restricted isometry constants. submitted, online arXiv:0911.1564v1, 2009.

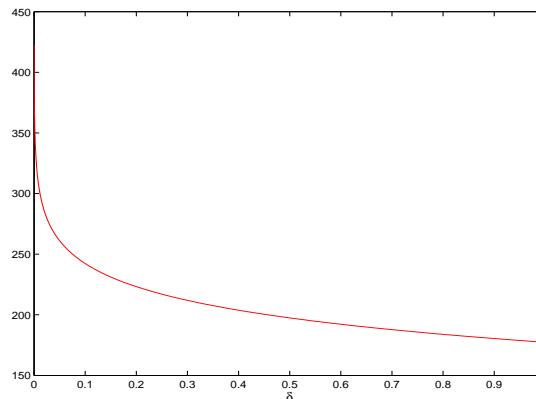


Figure 3: *Proportionality Constants*: the curve $1/\rho_S^{cwx}(\delta; 4)$ which provides the constant of proportionality between the required number of measurements and the sparsity level.

- [6] T. T. Cai, L. Wang, and G. Xu. Shifting inequality and recovery of sparse signals. *IEEE Trans. Sig. Proc.*, 2009. in press, doi:10.1109/TSP.2009.2034936.
- [7] T. T. Cai, G. Xu, and J. Zhang. On recovery of sparse signals via ℓ_1 minimization. *IEEE Trans. Inform. Theory*, 55(7):3388–3397, 2009.
- [8] E. J. Candès, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.*, 59(8):1207–1223, 2006.
- [9] E. J. Candès and T. Tao. Decoding by linear programming. *IEEE Trans. Inform. Theory*, 51(12):4203–4215, 2005.
- [10] W. Dai and O. Milenkovic. Subspace pursuit for compressive sensing signal reconstruction. submitted, arXiv:0803.0811v3, 2008.
- [11] D. L. Donoho. Neighborly polytopes and sparse solution of underdetermined linear equations. Technical Report, Department of Statistics, Stanford University, 2004.
- [12] D. L. Donoho and J. Tanner. Sparse nonnegative solution of underdetermined linear equations by linear programming. *Proc. Natl. Acad. Sci. USA*, 102(27):9446–9451 (electronic), 2005.
- [13] D. L. Donoho and J. Tanner. Sparse nonnegative solutions of underdetermined linear equations by linear programming. *Proc. Natl. Acad. Sci. USA*, 102(27):9446–9451, 2005.
- [14] S. Foucart and M.-J. Lai. Sparsest solutions of underdetermined linear systems via ℓ_q -minimization for $0 < q \leq 1$. *Appl. Comput. Harmon. Anal.*, 26(3):395–407, 2009.
- [15] D. Needell and J. Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Appl. Comp. Harm. Anal.*, 26(3):301–321, 2009.