

A null space property approach to compressed sensing with frames

Xuemei Chen

Department of Mathematics
University of Maryland, College Park
Email: xuemeic@math.umd.edu

Haichao Wang

Department of Mathematics
U C Davis
Email: hchwang@ucdavis.edu

Rongrong Wang

Department of Mathematics
University of Maryland, College Park
Email: rongwang@math.umd.edu

Abstract—An interesting topic in compressive sensing concerns problems of sensing and recovering signals with sparse representations in a dictionary. In this note, we study conditions of sensing matrices A for the ℓ^1 -synthesis method to accurately recover sparse, or nearly sparse signals in a given dictionary D . In particular, we propose a dictionary based null space property (D -NSP) which, to the best of our knowledge, is the first sufficient and necessary condition for the success of the ℓ^1 recovery. This new property is then utilized to detect some of those dictionaries whose sparse families cannot be compressed universally. Moreover, when the dictionary is of full spark, we show that AD being NSP, which is well-known to be only sufficient for stable recovery via ℓ^1 -synthesis method, is necessary as well.

I. INTRODUCTION

Compressed sensing concerns the problem of recovering a sparse signal $x_0 \in \mathbb{C}^d$ from its undersampled linear measurements $y = Ax_0 \in \mathbb{C}^m$, where the number of measurements m is usually much less than the ambient dimension d . A vector is said to be k -sparse if it has at most k nonzero entries. The following linear optimization algorithm, also known as the Basis Pursuit, can reconstruct x_0 efficiently from a perturbed observation $y = Ax_0 + w$ where $\|w\|_2 \leq \epsilon$ [8][4]:

$$\hat{x} = \arg \min_{x \in \mathbb{R}^d} \|x\|_1, \quad \text{subject to } \|y - Ax\|_2 \leq \epsilon. \quad (1)$$

A primary task of compressed sensing is to choose appropriate sensing matrix A in order to achieve good performance of (1). A matrix A is said to have the *Restricted isometry property (RIP)* with order k if

$$(1 - \delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta)\|x\|_2^2 \quad (2)$$

for any k -sparse vectors x . RIP is shown to provide stable reconstruction of approximately sparse signals via (1) [5][8]. Moreover, many random matrices satisfy RIP with high probability [6], [14]. A matrix A is said to have the *Null space property of order k (k -NSP)* if

$$\forall v \in \ker A \setminus \{0\}, \quad \forall |T| \leq k, \quad \|v_T\|_1 < \|v_{T^c}\|_1.$$

NSP is known as a characterization of uniqueness of problem (1) when there is no noise [10]. It has also been proven that the NSP matrices admit a similar stability result as RIP does except that the constants may be larger [1].

A recent direction of interest in compressed sensing concerns problems where signals are sparse in an overcomplete

dictionary D instead of a basis, see [3], [13], [10], [11], [1], [12], [9]. This is motivated by the widespread use of overcomplete dictionaries in signal processing and data analysis. Many signals naturally possess sparse frame coefficients, such as images consisting of curves (curvelet frame). In addition, the greater flexibility and stability of frames make them preferable for practical purposes in order to compensate the imperfectness of measurements. In this setting, the signal $x_0 \in \mathbb{C}^d$ can be represented as $x_0 = Dz_0$, where z_0 is k -sparse and D is a $d \times n$ matrix with $n \geq d$. The columns of D may be thought of as an overcomplete frame or dictionary for \mathbb{C}^d . The linear measurements are $y = Ax_0$.

A natural way to recover x_0 from y is first solving

$$\hat{z} = \arg \min_{z \in \mathbb{R}^n} \|z\|_1, \quad \text{subject to } y = ADz. \quad (3)$$

for the sparse coefficients \hat{z} , then synthesizing it to obtain \hat{x} , i.e., $\hat{x} = D\hat{z}$. The resulting method is therefore called ℓ^1 -synthesis or synthesis based method [11], [13]. Since we are only seeking the recovery of x_0 , we say the ℓ^1 -synthesis method (3) is *successful* when every minimizer \hat{z} of (3) satisfies $D\hat{z} = x_0$.

In the case when the measurements are perturbed, we naturally solve the following:

$$\hat{z} = \arg \min_{z \in \mathbb{R}^n} \|z\|_1, \quad \text{subject to } \|y - ADz\| \leq \epsilon. \quad (4)$$

The work in [13] established conditions on A and D to make the compound AD satisfy RIP. However, as pointed in [3], [11], forcing AD to satisfy RIP or even the weaker NSP implies the exact recovery of both z_0 and x_0 , which is unnecessary if we only care about obtaining a good estimate of x_0 . In particular, if D is perfectly correlated (has two identical columns), then there are infinitely many minimizers of (3) that may be assigned to \hat{z} , but all of them lead to the true signal x_0 . It seems reasonable to expect that similar result may hold in the case of highly correlated dictionaries, since they are only a small perturbation away from the perfectly correlated ones.

A. Overview and main results

In this paper, we generalize the ordinary null space property to the dictionary case (D -NSP), and prove in Theorem II.1 that this new condition is equivalent to the accurate recovery

of sparse signals in dictionaries via ℓ^1 -synthesis. Moreover, a stability result is given in Theorem III.1. To the best of our knowledge, these results are the first characterization of compressed sensing with dictionaries via ℓ^1 -synthesis approach.

Section IV studies more properties of D -NSP, and shows that A has D -NSP is equivalent to AD has NSP as long as D is of full spark (every d columns of D are linearly independent). As a consequence, under the full spark assumption, the ℓ^1 -synthesis method cannot accurately recover the signals without accurate recoveries of their sparse representations, therefore an incoherent dictionary is needed under this circumstance.

All proofs of the theorems presented can be found in [7], while some proofs are provided here.

II. A SUFFICIENT AND NECESSARY CONDITION FOR NOISELESS SPARSE RECOVERY

In this section, we develop a sufficient and necessary condition for the success of ℓ^1 -synthesis method (3). We show that the following property on A is a necessary and sufficient condition for successfully recovering all signals in $D\Sigma_s$ via (3), where $D\Sigma_k = \{x : \exists z, \text{ such that } x = Dz, \|z\|_0 \leq k\}$ is the set of signals that have k -sparse representations in D .

Definition 1 (Null space property of a dictionary D (D -NSP)). Fix a dictionary $D \in \mathbb{C}^{d,n}$, a matrix $A \in \mathbb{C}^{m,d}$ is said to satisfy the D -NSP of order k (k - D -NSP) if for any index set T with $|T| \leq k$, and any $v \in D^{-1}(\ker A \setminus \{0\})$, there exists $u \in \ker D$, such that

$$\|v_T + u\|_1 < \|v_{T^c}\|_1. \quad (5)$$

Theorem II.1. D -NSP is a necessary and sufficient condition for ℓ^1 -synthesis (3) to successfully recover all signals in the set $D\Sigma_k$.

Proof: Necessary part. We need to show that, if from measurements taken by a sensing matrix A , ℓ^1 -synthesis is successful in recovering all signals in $D\Sigma_k$, then A must be k - D -NSP.

For any $v \in D^{-1}(\ker A \setminus \{0\})$ and any index set T with $|T| = k$, we define $x_0 = Dv_T$ be a signal in $D\Sigma_k$, $y = Ax_0$ be its measurements, and let \hat{x} , \hat{z} be the reconstructed signal and its coefficients from y via (3). If ℓ^1 -synthesis is successful for all signals in $D\Sigma_k$, then we must have $\hat{x} = x_0$, and so $\hat{z} = v_T + u$ with some $u \in \ker D$.

Observe that $v_T - v$ is also feasible to (3), but it is not a minimizer since it cannot be represented in the form of $v_T + u$ with any $u \in \ker D$. Therefore, its ℓ_1 norm is strictly greater than that of \hat{z} :

$$\|v_T + u\|_1 < \|v_T - v\|_1 = \|v_{T^c}\|_1,$$

implying A is k - D -NSP.

Sufficient part. Assuming A is k - D -NSP, we will show that the ℓ_1 synthesis can recover all signals $x \in D\Sigma_k$ from $y = Ax$. Suppose to the contrary that there exists an $x_0 = Dz_0 \in D\Sigma_k$, such that its reconstruction $\hat{x} = D\hat{z}$ is wrong. Then we must have $v := z_0 - \hat{z} \in D^{-1}(\ker A \setminus \{0\})$. Let T be the support of

z_0 , by D -NSP, therefore there exists a $u \in \ker D$, such that $\|v_T + u\|_1 < \|v_{T^c}\|_1$, i.e., $\|z_0 - \hat{z} + u\|_1 < \|\hat{z}_{T^c}\|_1$. Hence, $\|z_0 + u\|_1 \leq \|z_0 - \hat{z} + u\|_1 + \|\hat{z}\|_1 < \|\hat{z}_{T^c}\|_1 + \|\hat{z}\|_1 = \|\hat{z}\|_1$.

This is a contradicts to the assumption that \hat{z} is a minimizer. \square

Notice when D is the canonical basis of \mathbb{C}^d , the D -NSP is reduced to the normal NSP with the same order. In other words, D -NSP is a generalization of NSP for the dictionary case. It is, however, a nontrivial generalization.

The intuition of D -NSP rises from the fact that we are only interested in recovering x_0 instead of the representation z_0 . As long as the minimizer \hat{z} lies in the affine plane $z_0 + \ker D$, our reconstruction is a success.

III. D -NSP BASED STABILITY ANALYSIS

It is known that the NSP is a sufficient and necessary condition not only for the sparse and noiseless recovery, but also for compressible signals with noisy measurement [1], [15]. However, the stability analysis of NSP [1] cannot be easily generalized to our case because essentially we need the function $f(v) = (\|v_{T^c}\|_1 - \|v_T + u\|_1) / \|Dv\|_2$ to be bounded away from zero. In the basis case, we have knowledge of $f(v)$ on a compact set, and consequently the extreme value theorem can be applied to prove the existence of a positive lower bound. In our case we do not have a compact set, therefore other constructions to overcome this difficulty is necessary.

Definition 2 (Strong null space property of a dictionary D (D -SNSP)). A sensing matrix A is said to have the strong null space property with respect to D of order k (k - D -SNSP) if for any index set T with $|T| \leq k$, and any $v \in \ker(AD)$, there exists $u \in \ker D$, such that

$$\|v_{T^c}\|_1 - \|v_T + u\|_1 \geq c\|Dv\|_2 \quad (6)$$

D -SNSP is a stronger assumption than D -NSP by definition. We prove that under this assumption, the ℓ^1 -synthesis recovery is stable with respect to perturbations on the measurement vector y .

Theorem III.1. If A is k - D -SNSP, then any solution \hat{z} of problem (4) satisfies

$$\|D\hat{z} - x_0\|_2 \leq C_1\sigma_k(z_0) + C_2\epsilon.$$

where $\sigma_k(z_0)$ denotes the ℓ^1 residue of the best k -term approximation to z_0 , C_1, C_2 are constants dependent on n , the c in (6), the minimum singular values of A and D , but not on x_0 .

Proof: Let $x_0 = Dz_0$ with z_0 being an k -sparse representation of x_0 . Let $h = D(\hat{z} - z_0)$, and decompose it as $h = Dw + \eta$ where $Dw \in \ker A$, $\eta \in \ker A^\perp$. It is easy to show that $\|\eta\|_2 \leq \frac{1}{\nu_A}\|Ah\|_2 \leq \frac{2\epsilon}{\nu_A}$ with ν_A being the smallest singular value of A .

Define $\xi = D^T(DD^T)^{-1}\eta$, then $\eta = D\xi$, and

$$\|\xi\|_2 \leq \frac{1}{\nu_D}\|\eta\|_2 \leq \frac{2}{\nu_A\nu_D}\epsilon. \quad (7)$$

Moreover, by our setting, $D(\hat{z} - z_0) = h = D(w + \xi)$, and therefore $\hat{z} - z_0 = w + \xi + u_1$ with some $u_1 \in \ker D$.

Let $v = w + u_1$, then $\hat{z} - z_0 = v + \xi$ and $v \in \ker(AD)$. By the assumption of D -SNSP, there exists a $u \in \ker D$ such that (6) holds for u and v . Therefore,

$$\begin{aligned} & \|v + z_{0,T}\|_1 - \|-u + z_{0,T}\|_1 \\ &= \|v_{T^c}\|_1 + \|v_T + z_{0,T}\|_1 - \|-u_T + z_{0,T}\|_1 - \|u_{T^c}\|_1 \\ &\geq \|v_{T^c}\|_1 - \|v_T + u_T\|_1 - \|u_{T^c}\|_1 \\ &= \|v_{T^c}\|_1 - \|v_T + u\|_1 \geq c\|Dv\|_2 \end{aligned} \quad (8)$$

On the other hand, from the fact that \hat{z} is a minimizer, we have

$$\begin{aligned} & \|-u + z_{0,T}\|_1 + \|z_{0,T^c}\|_1 \geq \|-u + z_0\|_1 = \|\hat{z}\|_1 \\ & \geq \|v + z_0 + \xi\|_1 \geq \|v + z_0\|_1 - \|\xi\|_1 \\ & \geq \|v + z_{0,T}\|_1 - \|z_{0,T^c}\|_1 - \|\xi\|_1. \end{aligned}$$

Rearrange the above inequality, we will obtain

$$\|v + z_{0,T}\|_1 - \|-u + z_{0,T}\|_1 \leq 2\|z_{0,T^c}\|_1 + \|\xi\|_1. \quad (9)$$

Combining (8) and (9), we get

$$\|Dv\|_2 \leq \frac{2}{c}\|z_{0,T^c}\|_1 + \frac{1}{c}\|\xi\|_1 \leq \frac{2}{c}\|z_{0,T^c}\|_1 + \frac{\sqrt{n}}{c}\|\xi\|_2 \quad (10)$$

In the end, using (10) and (7),

$$\begin{aligned} \|h\|_2 &= \|Dv + D\xi\|_2 = \|Dv + \eta\|_2 \leq \|Dv\|_2 + \|\eta\|_2 \\ &\leq \frac{2}{c}\|z_{0,T^c}\|_1 + \frac{\sqrt{n}}{c}\|\xi\|_2 + \frac{1}{\nu_A}2\epsilon \\ &\leq \frac{2}{c}\|z_{0,T^c}\|_1 + \frac{2\sqrt{n}}{c\nu_A\nu_D}\epsilon + \frac{1}{\nu_A}2\epsilon. \end{aligned}$$

□

It is natural to ask how much stronger this new assumption is than D -NSP. We address this question partially in the next section.

IV. A FURTHER STUDY OF D -NSP AND ADMISSIBLE DICTIONARIES

This section explores the two assumptions D -NSP and D -SNSP further for the purpose of answering the following important questions: What kind of dictionaries will allow sensing matrices A with few measurements to satisfy D -NSP? How to find those sensing matrices given a dictionary?

We call a $d \times n$ dictionary D k -admissible if there exists a measurement matrix $A \in \mathbb{C}^{m,d}$ with $m < d$ such that A is k - D -NSP. We call D inadmissible if D is not k -admissible for any $k \geq 2$. Intuitively speaking, D is not k -admissible means that $D\Sigma_k$ cannot be universally compressed by any linear matrix A .

The following proposition shows that adding repeated columns to the dictionary D will not affect admissibility. This is quite intuitive since we do not change the set $D\Sigma_k$ during this procedure, and we only care about recovering the signal x_0 rather than the representation z_0 .

Proposition IV.1. *Let $D \in \mathbb{C}^{d,n}$, and let I be any index set $I \subset \{1, \dots, n\}$. Define $\tilde{D} = [D, D_I]$, then for any sensing matrix $A \in \mathbb{C}^{m,n}$, we have A is D -NSP if and only if A is \tilde{D} -NSP.*

Proposition IV.1 states that a perfectly correlated dictionary D does not get in the way of the reconstruction of signals. It is only natural to ask whether this is still the case for a highly coherent dictionary. We answer this question partially by showing a class of highly correlated dictionaries is inadmissible. Moreover, equivalent conditions of D -NSP is given in Section IV-B under the assumption that D is of full spark.

A. A Class of inadmissible matrices

The following theorem constructs a class of inadmissible matrices with a one dimensional kernel.

Theorem IV.2. *Given an orthonormal basis $\Phi = [\phi_1, \dots, \phi_d]$.*

Let $H = \bigcup_{j=1}^d \text{span}\{\phi_i\}_{i=1, i \neq j}^d$ be the union of the hyperplanes spanned by every combination of $d - 1$ columns of Φ . Then there exists a small constant r_0 such that for every $v \in B(\phi_1, r_0) \setminus H$ where $B(\phi_1, r_0)$ is the ball centered at ϕ_1 with radius r_0 , $D = [\Phi, v] \in \mathbb{C}^{d,d+1}$ is inadmissible.

We need the following lemma for the proof of this Theorem.

Lemma IV.3. *Suppose D is a $d \times (d + 1)$ dictionary. If there exist $T \subset \{1, \dots, d + 1\}$ with $|T| \geq 2$ such that any vector $u \in \ker D \setminus \{0\}$ satisfies*

1. $\|u_T\|_1 > \|u_{T^c}\|_1$, and
2. $T^c \subset \text{supp}(u)$,

Then D cannot be $|T|$ -admissible.

For any vector $w \in \mathbb{C}^n$, we define $\|w\|_{\min} = \min_{1 \leq i \leq n} \{|w_i| \neq 0\}$ to be the minimum magnitude in w .

Proof: Assume that the dictionary D defined in Lemma IV.3 is $|T|$ -admissible, we will show how this leads to a contradiction.

Since D is admissible, then there exists at least one A that is k - D -NSP. Pick one of them, and fix a $v_0 \in D^{-1}(\ker(A) \setminus \{0\})$. Define $\alpha = 2\|v_0\|_{\infty} / \|u\|_{\min}$. Now that $v_0 + \alpha u, -v_0 + \alpha u \in D^{-1}(\ker(A) \setminus \{0\})$, by the definition of D -NSP, there exist $c_1, c_2 \in \mathbb{C}$ such that

$$\|v_T + \alpha u_T - c_1 u\|_1 < \|v_{T^c} + \alpha u_{T^c}\|_1, \quad (11)$$

and

$$\|-v_T + \alpha u_T - c_2 u\|_1 < \|-v_{T^c} + \alpha u_{T^c}\|_1. \quad (12)$$

Therefore,

$$2\alpha\|u_{T^c}\|_1 \quad (13)$$

$$= \|v_{T^c} + \alpha u_{T^c}\|_1 + \|-v_{T^c} + \alpha u_{T^c}\|_1 \quad (14)$$

$$> \|v_T + \alpha u_T - c_1 u\|_1 + \|-v_T + \alpha u_T - c_2 u\|_1 \quad (15)$$

$$= \|v_T + (\alpha - c_1)u_T\|_1 + |c_1|\|u_{T^c}\|_1$$

$$+ \|-v_T + (\alpha - c_2)u_T\|_1 + |c_2|\|u_{T^c}\|_1$$

$$\geq |2\alpha - c_1 - c_2|\|u_T\|_1 + (|c_1| + |c_2|)\|u_{T^c}\|_1, \quad (16)$$

where (14) follows from our assumption on α and Assumption 2, while (15) from adding (11) and (12). Combining (13) and (16) to get

$$\|u_T\|_1 < \|u_{T^c}\|_1.$$

This is a contradiction to Assumption 1 of Lemma IV.3. \square

Proof of Theorem IV.2: Notice that $\ker(D) = \text{span}\{u\}$ with $u = (a^T, -1)$. Let T be an index set with $|T| \geq 2$ such that $\{1, n+1\} \in T$. First, since $v \notin H$, then $\langle v, \phi_i \rangle \neq 0$ for $i = 1, \dots, d$. This means that all coordinates of u are nonzero, so Assumption 2 of Lemma IV.3 holds. Second, we can pick r_0 small enough such that whenever $v \in B(\phi_1, r)$, it holds $\|u_T\|_1 > \|u_{T^c}\|_1$, so Assumption 1 is satisfied.

Applying Lemma IV.3 completes the proof. \square

We have constructed an example of inadmissible dictionaries of special sizes: $d \times (d+1)$. The following proposition asserts that this dictionary can be used to generate inadmissible dictionaries of arbitrary dimension by adding appropriate columns to it.

Proposition IV.4. *If $D = [B, v]$ where B is a full rank $d \times (n-1)$ matrix and $v = B\alpha$ with $\|\alpha\|_1 \leq 1$, then A has D -NSP implies that A has B -NSP with the same order k .*

B. The relation between D -NSP and NSP

It is obvious that AD satisfies NSP implies A satisfies D -NSP, which explains why imposing RIP or incoherence conditions on AD could be too strong and unnecessary. To explore how much room there is between these two conditions can possibly answer the question whether we can allow highly coherent dictionaries or not, since AD being NSP will inevitably leads to the incoherence of D . Surprisingly enough, we show that whenever D is of full spark, these two conditions are equivalent.

A dictionary is of full spark means every d columns of this matrix are linearly independent.

Theorem IV.5. *The following conditions are equivalent under the assumption that D is of full spark,*

- A is k - D -NSP;
- AD is k -NSP;
- A is k - D -SNSP;
- For any $v \in \ker AD$, there exists a u such that

$$\|v_T + u\|_1 < \|v_{T^c}\|_1.$$

Remark IV.1. *We comment that full spark is not a strong assumption on matrices. In fact, full spark matrices is dense in the space of matrices [2], and a large class of full spark Harmonic frames is also constructed in [2].*

Remark IV.2. *Earlier we mentioned that we only care about recovering the signals x and allow the recovery of their representations z to be wrong. Theorem IV.5 tells us that when the dictionary is of full spark this requirement is actually not any looser than requiring both signals and their representations to be recovered. In spite of being negative, this result is quite*

important, since it has been largely thought that the opposite is true.

Like the RIP, NSP is essentially an incoherence property of a matrix. Hence a highly coherent dictionary D cannot be NSP, nor can the composite AD be, because whichever vector in $\ker D$ that fails to satisfy NSP, is also contained in $\ker(AD)$. Consequently, the equivalence of the first two items in Theorem IV.5 implies that if a highly coherent D is also full spark, then it must be inadmissible.

Perfectly coherent dictionaries are not full spark, so they can be and many of them are indeed admissible (Proposition IV.1). However, if these dictionaries are perturbed a little bit, then no matter how small the perturbations are, with probability one, they will turn into highly coherent and full spark dictionaries and therefore become inadmissible. We conclude that admissibility is not stable with respect to perturbations.

ACKNOWLEDGMENT

This research has been supported in part by Laboratory for Telecommunications Science (LTS) and by Defense Threat Reduction Agency HDTRA1-13-1-0015.

REFERENCES

- [1] Akram Aldroubi, Xuemei Chen, and Alexander M. Powell. Perturbations of measurement matrices and dictionaries in compressed sensing. *Appl. Comput. Harmon. Anal.*, 33(2):282–291, 2012.
- [2] Boris Alexeev, Jameson Cahill, and Dustin G. Mixon. Full Spark Frames. *J. Fourier Anal. Appl.*, 18(6):1167–1194, 2012.
- [3] E. Candes, Y. C. Eldar, D. Needell, and P. Randall. Compressed sensing with coherent and redundant dictionaries. *Applied and Computational Harmonic Analysis*, 31(1):59–73, 2010.
- [4] E. Candes, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.*, 59:1207–1223, 2006.
- [5] E. Candes and T. Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 51(12):4203–4215, December 2005.
- [6] E. Candes and T. Tao. Near optimal signal recovery from random projections and universal encoding strategies. *IEEE Transactions on Information Theory*, 52:5406–5425, 2006.
- [7] Xuemei Chen, Haichao Wang, and Rongrong Wang. A null space analysis of the ℓ_1 -synthesis method in frame-based compressed sensing. *in preparation*, 2013.
- [8] S. Foucart and M. Lai. Sparsest solutions of underdetermined linear systems via l_q -minimization for $0 < q \leq 1$. *Applied and Computational Harmonic Analysis*, 26(3):395–407, 2009.
- [9] R. Giryas and M. Elad. Can we allow linear dependencies in the dictionary in the sparse synthesis framework? *to appear in ICASSP*, 2013.
- [10] R. Gribonval and M. Nielsen. Highly sparse representations from dictionaries are unique and independent of the sparseness measure. *Applied and Computational Harmonic Analysis*, 22(3):335–355, May 2007.
- [11] Shidong Li, Tiebin Mi, and Yulong Liu. Performance analysis of ℓ_1 -synthesis with coherent frames. <http://arxiv.org/abs/1202.2223>, 2012.
- [12] D. Needell M. A. Davenport and M. B. Wakin. Signal space cosamp for sparse recovery with redundant dictionaries. *arXiv preprint arXiv:1208.0353*, 2012.
- [13] H. Rauhut, K. Schnass, and P. Vandergheynst. Compressed sensing and redundant dictionaries. *IEEE Transactions on Information Theory*, 54(5):2210–2219, 2008.
- [14] Mark Rudelson and Roman Vershynin. On sparse reconstruction from fourier and gaussian measurements. *Communications on Pure and Applied Mathematics*, 61:1025–1045, 2008.
- [15] Qiyu Sun. Sparse approximation property and stable recovery of sparse signals from noisy measurements. *IEEE Trans. Signal Process.*, 59(10):5086–5090, 2011.