On Partial Sparse Recovery

A. S. Bandeira, K. Scheinberg, L. N. Vicente

Abstract—We consider the problem of recovering a partially sparse solution of an underdetermined system of linear equations by minimizing the ℓ_1-norm of the part of the solution vector which is known to be sparse. Such a problem is closely related to a classical problem in Compressed Sensing where the ℓ_1-norm of the whole solution vector is minimized. We introduce analogues of restricted isometry and null space properties for the recovery of partially sparse vectors and show that these new properties are implied by their original counterparts. We show also how to extend recovery under noisy measurements to the partially sparse case.

Index Terms—Partial sparse recovery, compressed sensing, ℓ_1-minimization, Sparse quadratic polynomial interpolation.

I. INTRODUCTION

$\mathbf{I}$N Compressed Sensing one is interested in recovering a sparse solution $\bar{x} \in \mathbb{R}^N$ of an underdetermined system of the form $y = A\bar{x}$, given a vector $y \in \mathbb{R}^k$ and a matrix $A \in \mathbb{R}^{k \times N}$ with far fewer rows than columns ($k \ll N$). A direct approach is to minimize the number of non-zero components of x, i.e., the ℓ_0-*norm* of x (which is defined as $\|u\|_0 = |\{i : u_i \neq 0\}|$ but, strictly speaking, is not a norm),

$$\min \|x\|_0 \quad \text{s.t.} \quad Ax = y. \tag{1}$$

Since (1) is known to be NP-Hard, a tractable approximation is commonly considered which is obtained by substituting the non-convex ℓ_0-*norm* by a convex approximation. Recent results indicate that the ℓ_1-norm can serve as such an approximation (see [1] for a survey on some of this material). Hence (1) is replaced by the following optimization problem

$$\min \|x\|_1 \quad \text{s.t.} \quad Ax = y. \tag{2}$$

Note that (2) is equivalent to a linear program and thus is much easier to solve than (1).

In this paper we consider the case (see [2], [3], [4]) when it is known a priori that the solution vector consists of two parts, one of which is expected to be dense, in other words we have $x = (x_1, x_2)$, where $x_1 \in \mathbb{R}^{N-r}$ is sparse and $x_2 \in \mathbb{R}^r$ is possibly dense. A natural generalization of problem (2) to this setting of partially sparse recovery is given by

$$\min \|x_1\|_1 \quad \text{s.t.} \quad A_1 x_1 + A_2 x_2 = y, \tag{3}$$

ASB is with the Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA (ajsb@math.princeton.edu).

KS is with the Department of Industrial and Systems Engineering, Lehigh University, Harold S. Mohler Laboratory, 200 West Packer Avenue, Bethlehem, PA 18015-1582, USA (katyas@lehigh.edu). Support for this author is provided by AFOSR under grant FA9550-11-1-0239 and by NSF under grant DMS 10-16571.

LNV is with CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal (lnv@mat.uc.pt). Support for this author was provided by FCT under grant PTDC/MAT/098214/2008.

where $A = (A_1, A_2)$, $A_1 \in \mathbb{R}^{k \times (N-r)}$, and $A_2 \in \mathbb{R}^{k \times r}$. We will refer to this setting as *partially sparse recovery of size* $N - r$. One of the key applications of partially sparse recovery is image reconstruction [2] but they also arise naturally in sparse Hessian recovery [5].

Vaswani and Lu [2] gave a first sufficient condition for partially sparse recovery. Later, Friedlander et al. [3] proposed a weaker sufficient condition and covered the extension to the noisy case. After obtaining our results we were directed to the work of Jacques [4] who addressed the noisy case, deriving another sufficient condition for partially sparse recovery. His conditions guarantee the same recovery as ours but, as far as we can tell, are not the simple extensions of the NSP and RIP properties. The conditions in [2], [3], [4] are somewhat weaker than the known restricted isometry property for general sparse recovery, which is natural since the case of partial sparsity can be considered as a case of general sparsity where part of the support of the solution is known in advance.

The contribution of our paper is to introduce the analogues of restricted isometry and null space properties for the case of partial sparsity. We prove that these new properties are sufficient for partially sparse recovery (including the noisy case) and are implied by the original conditions of fully sparse recovery. We show that it is possible to guarantee recovery of a partially sparse signal using Gaussian random matrices with the number of measurements an order smaller than the one necessary for general recovery.

A. Notation

We will use the following notation in this paper. $[N]$ denotes the set of integers $\{1, \dots, N\}$, and $[N]^{(s)}$ denotes the set of all subsets of $[N]$ of cardinality $s \leq N$. If A is a matrix, then by $\mathcal{N}(A)$ and $\mathcal{R}(A)$ we denote the null and range spaces of A, respectively. We say that a vector x is $s-$sparse if at most s components of x are non-zero. This is also denoted by $\|x\|_0 \leq s$. Given $v \in \mathbb{R}^N$ and $S \in [N]$, $v_S \in \mathbb{R}^N$ denotes a vector defined by $(v_S)_i = v_i$, $i \in S$ and $(v_S)_i = 0$, $i \notin S$.

II. SPARSE RECOVERY IN COMPRESSED SENSING

One of the main questions addressed by Compressed Sensing is under what conditions on the matrix A can every sparse vector $\bar{x}$ be recovered by solving problem (2) given A and the right hand side $y = A\bar{x}$. The next definition is a well known characterization of such matrices (see, e.g., [6], [7]).

Definition 2.1 (Null Space Property): The matrix $A \in \mathbb{R}^{k \times N}$ is said to satisfy the Null Space Property (NSP) of order s if, for every $v \in \mathcal{N}(A) \setminus \{0\}$ and for every $S \in [N]^{(s)}$, one has

$$\|v_S\|_1 < \frac{1}{2}\|v\|_1. \tag{4}$$

It is well known that NSP is a necessary and sufficient condition for the recovery of an s-sparse vector $\bar{x}$ (see [8]).

Theorem 2.1: The matrix A satisfies the Null Space Property of order s if and only if, for every $s-$sparse vector $\bar{x}$, problem (2) with $y = A\bar{x}$ has an unique solution and it is given by $x = \bar{x}$.

It is difficult to analyze whether NSP is satisfied. On the other hand, the *Restricted Isometry Property* (RIP), introduced in [9], is considerably more useful and insightful, although it provides only sufficient conditions for recovery with (2). We present below the definition of the *RIP Constant*.

Definition 2.2 (Restricted Isometry Property Constant): One says that $\delta_s > 0$ is the Restricted Isometry Property Constant, or *RIP* constant, of order s of the matrix $A \in \mathbb{R}^{k \times N}$ if δ_s is the smallest positive real number such that:

$$(1 - \delta_s) \|x\|_2^2 \ \leq \ \|Ax\|_2^2 \ \leq \ (1 + \delta_s) \|x\|_2^2 \qquad (5)$$

for every $s-$sparse vector x.

The following theorem (see, e.g., [10]) provides a useful sufficient condition for successful recovery by (2).

Theorem 2.2: [10] Let $A \in \mathbb{R}^{k \times N}$ and $2s < k$. If $\delta_{2s} < \sqrt{2} - 1$, where δ_{2s} is the RIP constant of A of order $2s$, then, for every $s-$sparse vector $\bar{x}$, problem (2) with $y = A\bar{x}$ has an unique solution and it is given by $x = \bar{x}$.

It is known that RIP is satisfied with some probability if the entries of the matrix are randomly generated (see, e.g., [11]) according to some distribution such as a sub-Gaussian. However, it is in general computationally hard to check whether it is satisfied by a certain realization matrix [12], and it is still an open problem to find such matrices deterministically when the underlying system is highly underdetermined (see [13]).

III. PARTIAL SPARSE RECOVERY

In this section we consider the following extension of the NSP to the case of partially sparse recovery.

Definition 3.1 (Partial Null Space Property): We say that $A = (A_1, A_2)$ satisfies the Null Space Property (NSP) of order $s - r$ for partially sparse recovery of size $N - r$ with $r \leq s$ if A_2 is full column rank ($\mathcal{N}(A_2) = \{0\}$) and for every $v_1 \in \mathbb{R}^{N-r} \setminus \{0\}$ such that $A_1 v_1 \in \mathcal{R}(A_2)$ and every $S \in [N - r]^{(s-r)}$, we have

$$\|(v_1)_S\|_1 \ < \ \frac{1}{2} \|v_1\|_1. \qquad (6)$$

Note that when $r = 0$, the partial NSP naturally reduces to the NSP in Definition 2.1. Wang and Yin [14] have suggested a stronger NSP adapted to a setting where it is not known the location of the partial support.

The new property is a necessary and sufficient condition for any solution of (3) with $y = A\bar{x}$ to satisfy $x = \bar{x}$ if $\bar{x}_1$ is appropriately sparse.

Theorem 3.1: The matrix $A = (A_1, A_2)$ satisfies the Null Space Property of order $s - r$ for Partially Sparse Recovery of size $N - r$ if and only if for every $\bar{x} = (\bar{x}_1, \bar{x}_2)$ such that $\bar{x}_1 \in \mathbb{R}^{N-r}$ is $(s-r)-$sparse and $\bar{x}_2 \in \mathbb{R}^r$, problem (3) with $y = A\bar{x}$ has an unique solution and it is given by $(x_1, x_2) = (\bar{x}_1, \bar{x}_2)$.

Proof: The proof follows the steps of the proof of [8, Theorem 2.3] with appropriate modifications. Let us assume first that for any vector $(\bar{x}_1, \bar{x}_2) \in \mathbb{R}^N$, where $\bar{x}_1$ is an $(s - r)-$sparse vector and $\bar{x}_2 \in \mathbb{R}^r$, the minimizer (x_1, x_2) of $\|x_1\|_1$ subject to $A_1 x_1 + A_2 x_2 = A\bar{x}$ satisfies $x_1 = \bar{x}_1$. Consider any $v_1 \neq 0$ such that $A_1 v_1 \in \mathcal{R}(A_2)$. Then consider minimizing $\|x_1\|_1$ subject to $A_1 x_1 + A_2 x_2 = A_1 (v_1)_S + A_2 v_2$ for any $v_2 \in \mathbb{R}^r$ and for any $S \in [N - r]^{(s-r)}$. By the assumption, the corresponding minimizer (x_1, x_2) satisfies $x_1 = (v_1)_S$. Since $A_1 v_1 \in \mathcal{R}(A_2)$, there exists u_2 such that $A_1(-(v_1)_{S^c}) + A_2 u_2 = A_1 (v_1)_S + A_2 v_2$. As $-(v_1)_{S^c} \neq (v_1)_S$, $(-(v_1)_{S^c}, u_2)$ is not the minimizer of $\|x_1\|_1$ subject to $A_1 x_1 + A_2 x_2 = A_1 (v_1)_S + A_2 v_2$, hence, $\|(v_1)_{S^c}\|_1 > \|(v_1)_S\|_1$ and (6) holds.

Let us now assume that A satisfies the NSP of order $s-r$ for partially sparse recovery of size $N - r$ (Definition 3.1). Then, given a vector $(\bar{x}_1, \bar{x}_2) \in \mathbb{R}^N$, where $\bar{x}_1$ is $(s - r)-$sparse and $\bar{x}_2 \in \mathbb{R}^r$, and a vector $(u_1, u_2) \in \mathbb{R}^N$ with $u_1 \neq \bar{x}_1$ and satisfying $A_1 u_1 + A_2 u_2 = A_1 \bar{x}_1 + A_2 \bar{x}_2$, consider $(v_1, v_2) = ((\bar{x}_1 - u_1), (\bar{x}_2 - u_2)) \in \mathcal{N}(A)$, which implies $A_1 v_1 \in \mathcal{R}(A_2)$ and $v_1 \neq 0$. Thus, setting S to be the support of $\bar{x}$, one has that

$$\begin{aligned}
\|\bar{x}_1\|_1 & \leq & \|\bar{x}_1 - (u_1)_S\|_1 + \|(u_1)_S\|_1 \\
& = & \|(v_1)_S\|_1 + \|(u_1)_S\|_1 < \|(v_1)_{S^c}\|_1 + \|(u_1)_S\|_1 \\
& = & \|-(u_1)_{S^c}\|_1 + \|(u_1)_S\|_1 = \|u_1\|_1,
\end{aligned}$$

(the strict inequality coming from (6)), guaranteeing that all solutions (x_1, x_2) of (3) with $y = A\bar{x}$ satisfy $x_1 = \bar{x}_1$.

It remains to note that $x_2 = \bar{x}_2$ is uniquely determined by solving $A_2 x_2 = y - A_1 \bar{x}_1$ if and only if A_2 is full column rank. ∎

We now define an extension of the RIP to the partially sparse recovery setting. For this purpose, let $A = (A_1, A_2)$ be as considered above, under the assumption that A_2 has full column rank. Let

$$\mathcal{P} \ = \ I - A_2 \left(A_2^\top A_2 \right)^{-1} A_2^\top \qquad (7)$$

be the matrix of the orthogonal projection from $\mathbb{R}^N$ onto $\mathcal{R}(A_2)^\perp$. Then, the problem of recovering $(\bar{x}_1, \bar{x}_2)$, where $\bar{x}_1$ is an $(s - r)-$sparse vector satisfying $A_1 \bar{x}_1 + A_2 \bar{x}_2 = y$, can be stated as the problem of recovering an $(s - r)-$sparse vector $x_1 = \bar{x}_1$ satisfying $(\mathcal{P} A_1) x_1 = \mathcal{P} y$ and then recovering $x_2 = \bar{x}_2$ satisfying $A_2 x_2 = y - A_1 \bar{x}_1$. The solution of the resulting linear system in the second step exists and is unique given that A_2 has full column rank and $(\mathcal{P} A_1)\bar{x}_1 = \mathcal{P} y$. Note that the first step is now reduced to the classical setting of Compressed Sensing. This motivates the following definition of RIP for partially sparse recovery.

Definition 3.2 (Partial RIP): We say that $\delta_{s-r}^r > 0$ is the Partial Restricted Isometry Property Constant of order $s - r$ of the matrix $A = (A_1, A_2) \in \mathbb{R}^{k \times N}$, for recovery of size $N - r$ with $r \leq s$, if A_2 is full column rank and δ_{s-r}^r is the RIP constant of order $s - r$ (see Definition 2.2) of the matrix $\mathcal{P} A_1$, where $\mathcal{P}$ is given by (7).

Again, when $r = 0$ the Partial RIP reduces to the RIP of Definition 2.2. We also note that, given a matrix $A =$

$(A_1, A_2) \in \mathbb{R}^{k \times N}$ with Partial RIP constant $\delta^r_{2(s-r)}$ of order $2(s-r)$ for recovery of size $N-r$, satisfying $\delta^r_{2(s-r)} < \sqrt{2}-1$, Theorems 2.1 and 2.2, guarantee that $\mathcal{P}A_1$ satisfies the NSP of order $s-r$. Thus, given $\bar{x} = (\bar{x}_1, \bar{x}_2)$ such that $\bar{x}_1 \in \mathbb{R}^{N-r}$ is $(s-r)$−sparse and $\bar{x}_2 \in \mathbb{R}^r$, $\bar{x}_1$ can be recovered by minimizing the ℓ_1-norm of x_1 subject to $(\mathcal{P}A_1)x_1 = \mathcal{P}A\bar{x}$ and, recalling that A_2 is full-column rank, $x_2 = \bar{x}_2$ is uniquely determined by $A_2x_2 = y - A_1\bar{x}_1$. (In particular, this implies that A satisfies the NSP of order $s-r$ for partially sparse recovery of size $N-r$.)

IV. PARTIALLY SPARSE RECOVERY IMPLIED BY FULLY SPARSE RECOVERY CONDITIONS

We are now interested in showing that partially sparse recovery is achievable under the conditions which guarantee fully sparse recovery. In particular we will show that the NSP and RIP imply, respectively, the partial NSP and the partial RIP. We first establish the relationship between the corresponding null space properties.

Theorem 4.1: If a given matrix A satisfies the NSP of order s then it satisfies the NSP for partially sparse recovery of order $s-r$ for any $r \leq s$.

Proof: Let $A = (A_1, A_2)$ satisfy the NSP of order s. First we note that since $r \leq s$, the NSP implies that A_2 is full column rank. Let $v_1 \in \mathbb{R}^{N-r}$ be a non-zero vector such that $A_1v_1 \in \mathcal{R}(A_2)$ and let $T \in [N-r]^{(s-r)}$.

Since there exists v_2 such that $A_1v_1 + A_2v_2 = 0$, we have that $v = (v_1, v_2) \in \mathcal{N}(A) \setminus \{0\}$, and therefore by setting $S = T \cup ([N] \setminus [N-r])$ and by using the NSP, $\|(v_1)_T\|_1 + \|v_2\|_1 = \|v_S\|_1 < \frac{1}{2}\|v\|_1 = \frac{1}{2}\|v_1\|_1 + \frac{1}{2}\|v_2\|_1$. Thus, $\|(v_1)_T\|_1 \leq \|(v_1)_T\|_1 + \frac{1}{2}\|v_2\|_1 \leq \frac{1}{2}\|v_1\|_1$, and A satisfies the NSP of order $s-r$ for partially sparse recovery of size $N-r$. ∎

Partial RIP is also implied by RIP without the change in the RIP constant value.

Theorem 4.2: Let $\delta_s > 0$ and $A = (A_1, A_2)$ satisfy the following property: For every $(s-r)$-sparse vector $x_1 \in \mathbb{R}^{N-r}$ and $x_2 \in \mathbb{R}^r$ we have

$$(1-\delta_s)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta_s)\|x\|_2^2, \tag{8}$$

where $x = (x_1, x_2)$. Then A satisfies partial RIP of order $s-r$ with $\delta^r_{s-r} = \delta$ for partially sparse recovery of size $N-r$, for any $r \leq s$.

Proof: First we note that setting $x_1 = 0$ implies that A_2 is full column rank. Consider now any given $(s-r)$−sparse vector $x_1 \in \mathbb{R}^{N-r}$. Now, by setting $x_2 = -\left(A_2^\top A_2\right)^{-1} A_2^\top A_1 x_1$, one obtains $(1-\delta_s)\|x_1\|_2^2 \leq (1-\delta_s)\left(\|x_1\|_2^2 + \|x_2\|_2^2\right) \leq \|A_1x_1 + A_2x_2\|_2^2 = \|\mathcal{P}A_1x_1\|^2$. On the other hand, the choice $x_2 = 0$ provides $\|\mathcal{P}A_1x_1\|_2^2 \leq \|A_1x_1\|_2^2 \leq (1+\delta_s)\|x_1\|_2^2$. We have thus arrived at the conditions of Definition 3.2. ∎

Corollary 4.1: Let $A = (A_1, A_2)$ satisfy the RIP of order s with the RIP constant δ_s. Then A satisfies partial RIP of order $s-r$ with $\delta^r_{s-r} = \delta_s$ for partially sparse recovery of size $N-r$, for any $r \leq s$.

V. PARTIAL (AND TOTAL) COMPRESSIBILITY RECOVERY WITH NOISY MEASUREMENTS

In most realistic applications the observed measurement vector y often contains noise and the true signal vector $\bar{x}$ is

not sparse but rather compressible, meaning that most components are very small but not necessarily zero. It is known, however, that Compressed Sensing is robust to noise and can approximately recover compressible vectors. This statement is formalized in the following theorem taken from [10].

Theorem 5.1: Assume that the matrix $A \in \mathbb{R}^{k \times N}$ satisfies RIP with the RIP constant δ_{2s} such that $\delta_{2s} < \sqrt{2} - 1$. For any $\bar{x} \in \mathbb{R}^N$, let noisy measurements $y = A\bar{x} + \epsilon$ be given satisfying $\|\epsilon\|_2 \leq \eta$. Let $x^\#$ be a solution of

$$\min_{x \in \mathbb{R}^N} \|x\|_1 \quad \text{s.t.} \quad \|Ax - y\|_2 \leq \eta. \tag{9}$$

Then

$$\|x^\# - \bar{x}\|_2 \leq c\eta + d\frac{\sigma_s(\bar{x})_1}{\sqrt{s}}, \tag{10}$$

for constants c, d only depending on the RIP constant, and where $\sigma_s(\bar{x})_1 = \min_{x:\|x\|_0 \leq s} \|x - \bar{x}\|_1$.

The following theorem provides an analogous result for the partially sparse recovery setting introduced in Section III.

Theorem 5.2: Assume that the matrix $A = (A_1, A_2) \in \mathbb{R}^{k \times N}$ satisfies partial RIP of order $2(s-r)$ for recovery of size $N-r$ with the RIP constant $\delta^r_{2(s-r)} < \sqrt{2} - 1$. For any $\bar{x} = (\bar{x}_1, \bar{x}_2) \in \mathbb{R}^N$, let noisy measurements $y = A\bar{x} + \epsilon$ be given satisfying $\|\epsilon\|_2 \leq \eta$. Let $x^* = (x_1^*, x_2^*)$ be a solution of

$$\min_{x=(x_1,x_2) \in \mathbb{R}^N} \|x_1\|_1 \quad \text{s.t.} \quad \|Ax - y\|_2 \leq \eta. \tag{11}$$

Then

$$\|x_1^* - \bar{x}_1\|_2 \leq c\eta + d\frac{\sigma_{s-r}(\bar{x}_1)_1}{\sqrt{s-r}}, \tag{12}$$

and

$$\|x_2^* - \bar{x}_2\|_2 \leq C_2\left(2\eta + C_1\left(c\eta + d\frac{\sigma_{s-r}(\bar{x}_1)_1}{\sqrt{s-r}}\right)\right), \tag{13}$$

for constants c, d only depending on $\delta^r_{2(s-r)}$, and where C_1 and C_2 are given by $C_1 = \|A_1\|_2$, and $C_2 = \|A_2^\dagger\|_2$, (Since A_2 is full column rank recall that $A_2^\dagger = (A_2^\top A_2)^{-1}A_2^\top$ and $C_2 > 0$.)

Proof: From Theorem 4.2, the matrix $\mathcal{P}A_1$, where $\mathcal{P}$ is given by (7), satisfies the condition of Theorem 5.1. Thus, since $\mathcal{P}$ is a projection matrix, $\|\mathcal{P}A_1\bar{x}_1 - \mathcal{P}y\| = \|\mathcal{P}A\bar{x} - \mathcal{P}y\| \leq \|A\bar{x} - y\| \leq \eta$, and a solution $x_1^\#$ of

$$\min_{x_1 \in \mathbb{R}^{N-r}} \|x_1\|_1 \quad \text{s.t.} \quad \|\mathcal{P}A_1x_1 - \mathcal{P}y\|_2 \leq \eta, \tag{14}$$

satisfies

$$\|x_1^\# - \bar{x}_1\|_2 \leq c\eta + d\frac{\sigma_{s-r}(x_1)_1}{\sqrt{s-r}}. \tag{15}$$

Now, we will prove that the solutions of problems (11) and (14) coincide in their x_1 parts, completing thus the proof of (12). Let (x_1^*, x_2^*) be a feasible point of (11). Again, since $\mathcal{P}$ is a projection matrix, we obtain that

$$\begin{aligned}\|\mathcal{P}A_1x_1^* - \mathcal{P}y\|_2 &= \|\mathcal{P}(A_1x_1^* + A_2x_2^* - y)\|_2 \\ &\leq \|A_1x_1^* + A_2x_2^* - y\|_2 \leq \eta,\end{aligned}$$

which proves that x_1^* is a feasible point of (14). Now let $x_1^\#$ be a feasible point of (14). Since $I - \mathcal{P}$ projects (orthogonally) onto the column space of A_2 there must exist an $x_2^\#$ such that $A_2x_2^\# = (I-\mathcal{P})(y-A_1x_1^\#)$, and then $\|A_1x_1^\# + A_2x_2^\# - y\|_2 =$

$\|\mathcal{P}A_1 x_1^{\#} - \mathcal{P}y\|_2 \leq \eta$. Therefore $(x_1^{\#}, x_2^{\#})$ is a feasible point of (11). Hence we have proved that, any solution of problem (11) is also a solution of problem (14), and the inequality (12) results directly from (15).

We now use this inequality to bound the error on the reconstruction of $\bar{x}_2$. Since both $\bar{x}$ and x^* satisfy the measurements constraints $\|Ax - y\|_2 \leq \eta$ we have that $\|A_1(\bar{x}_1 - x_1) + A_2(\bar{x}_2 - x_2)\|_2 \leq 2\eta$, and thus $\|A_2(x_2^* - \bar{x}_2)\|_2 \leq 2\eta + \|A_1(x_1^* - \bar{x}_1)\|_2$. Using the definitions of C_1 and C_2 we have $\|x_2^* - \bar{x}_2\|_2 \leq C_2 (2\eta + C_1\|x_1^* - \bar{x}_1\|_2)$, and the result (13) follows from bounding $\|x_1^* - \bar{x}_1\|_2$ by (12) in this last inequality. ∎

The condition on the matrix A imposed in the previous theorem involved only its partial RIP constant. In the next proposition we describe how one can bound the constants C_1 and C_2 in terms of the RIP constant of A (the proof is simple and is omitted, see also [15]).

Proposition 5.1: Consider the RIP constant δ_s of order s of $A = (A_1, A_2) \in \mathbb{R}^{k \times N}$. The constants C_1 and C_2 of Theorem 5.2 satisfy $C_1 \leq \sqrt{1 + \delta_s}$ and $C_2 \leq \frac{1}{\sqrt{1 - \delta_s}}$.

VI. MATRICES WITH PARTIAL RIP

In this section we investigate regimes of N, s, and k for which random Gaussian matrices satisfy partial RIP. Similar results can be obtained for other families of random matrices, like sub-Gaussian or Bernoulli matrices.

Theorem 6.1: Let $0 < \delta < 1$ and $r \leq s$. Let $A = (A_1, A_2)$ with $A_1 \in \mathbb{R}^{k \times (N-r)}$ and $A_2 \in \mathbb{R}^{k \times r}$ have independent Gaussian entries with variance $1/k$. Then, as long as

$$k > \frac{2 \times 48}{3\delta^2 - \delta^3} \left((s - r) \log \left(\frac{N - r}{s - r} e \right) + s \log \left(\frac{12}{\delta} \right) \right), \ (16)$$

$A = (A_1, A_2)$ satisfies partial RIP of order $s - r$ with $\delta_{s-r}^r \leq \delta$ for partially sparse recovery of size $N - r$, with high probability.

Proof: Given a particular sparsity pattern, the probability that (8) does not hold is (see [11, Lemma 5.1])

$$\leq 2 (12/\delta)^s e^{-\left(\frac{\delta^2}{16} - \frac{\delta^3}{48}\right)k}.$$

There are $\binom{N-r}{s-r} \leq \left(\frac{N-r}{s-r}e\right)^{s-r}$ different sparsity patterns (see, e.g., [11]). Let $\mathcal{P}$ denote the probability that $A = (A_1, A_2)$ does not satisfy the partial RIP of order $s - r$ with $\delta_{s-r}^r = \delta$ for partially sparse recovery of size $N-r$. For this to happen, (8) has to fail for at least one sparsity pattern, setting $\beta = \frac{\delta^2}{16} - \frac{\delta^3}{48}$ and using a union bound

$$\begin{aligned} \mathcal{P} &\leq e^{(s-r)\log\left(\frac{N-r}{s-r}e\right)} 2 \left(\frac{12}{\delta}\right)^s e^{-\beta k} \\ &\leq 2e^{\left((s-r)\log\left(\frac{N-r}{s-r}e\right)+s\log\left(\frac{12}{\delta}\right)-\beta k\right)} \\ &\leq 2e^{-\beta\left[k-\frac{1}{\beta}\left((s-r)\log\left(\frac{N-r}{s-r}e\right)+s\log\left(\frac{12}{\delta}\right)\right)\right]} \\ &\leq 2e^{-\left[(s-r)\log\left(\frac{N-r}{s-r}e\right)+s\log\left(\frac{12}{\delta}\right)\right]}, \\ &\leq 2 \left(\frac{N-r}{s-r}e\right)^{-(s-r)} \left(\frac{12}{\delta}\right)^{-s}, \end{aligned}$$

where the second to last inequality was obtained using (16). It is easy to see that either $(e(N-r)/(s-r))^{-(s-r)}$ or $\left(\frac{12}{\delta}\right)^{-s}$ goes to zero polynomially with N, thus $\mathcal{P} \leq \mathcal{O}\left(N^{-\mathcal{O}(1)}\right)$ ∎

Note that the condition (16) can be asymptotically smaller than the one found in the classical case $r = 0$. If, e.g., $s - r = \mathcal{O}(1)$ then (16) just requires $k = \mathcal{O}(s + \log(N - r))$ instead of the classical $k = \mathcal{O}(s \log(N/s))$.

VII. CONCLUDING REMARKS

In some applications of Compressed Sensing one may be interested in a sparse (or compressible) vector whose support is partially known in advance. In such a setting we show that one can consider the ℓ_1-minimization of the part of the vector for which the support is not known. We have shown that such a sparse recovery can be then ensured under conditions that are potentially weaker than those assumed for the full approach. We have explored this feature to show that it is possible to guarantee partial sparse recovery (with Gaussian random matrices) for an order of measurements below the one necessary for general recovery.

REFERENCES

[1] E. J. Candès, "Compressive sampling," *Proceedings of the International Congress of Mathematicians Madrid 2006*, vol. Vol. III, 2006.
[2] N. Vaswani and W. Lu, "Modified-CS: Modifying compressive sensing for problems with partially known support," *IEEE Trans. Signal Process.*, vol. 58, pp. 4595–4607, 2010.
[3] M. P. Friedlander, H. Mansour, R. Saab, and O. Yilmaz, "Recovering compressively sampled signals using partial support information," *IEEE Trans. Inform. Theory*, vol. 58, pp. 1122–1134, 2012.
[4] L. Jacques, "A short note on compressed sensing with partially known signal support," *Signal Processing*, vol. 90, pp. 3308–3312, 2010.
[5] A. S. Bandeira, K. Scheinberg, and L. N. Vicente, "Computation of sparse low degree interpolating polynomials and their application to derivative-free optimization," *Math. Program.*, vol. 134, pp. 223–257, 2012.
[6] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best k-term approximation," *J. Amer. Math. Soc.*, vol. 22, pp. 211–231, 2009.
[7] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decompositions," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2845–2862, 2001.
[8] H. Rauhut, "Compressive sensing and structured random matrices," in *Theoretical Foundations and Numerical Methods for Sparse Recovery*, ser. Radon Series Comp. Appl. Math., M. Fornasier, Ed., 2010, pp. 1–92.
[9] E. Candès and T. Tao, "Near optimal signal recovery from random projections: universal encoding strategies?" *IEEE Trans. Inform. Theory*, vol. 52, pp. 5406–5425, 2006.
[10] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Compte Rendus de l'Academie des Sciences, Paris, Serie I*, vol. 589–592, 2009.
[11] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constr. Approx.*, vol. 28, pp. 253–263, 2008.
[12] A. S. Bandeira, E. Dobriban, D. G. Mixon, and W. F. Sawin, "Certifying the restricted isometry property is hard," *arXiv:1204.1580 [math.FA]*, 2012.
[13] A. S. Bandeira, M. Fickus, D. G. Mixon, and P. Wong, "The road to deterministic matrices with the restricted isometry property," *submitted, available online: arXiv:1202.1234 [math.FA]*, 2012.
[14] Y. Wang and W. Yin, "Compressed sensing via iterative support detection," *CoRR*, vol. abs/0909.4359, 2009.
[15] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, pp. 301–321, 2008.