

# CAN WE ALLOW LINEAR DEPENDENCIES IN THE DICTIONARY IN THE SPARSE SYNTHESIS FRAMEWORK?

*Raja Giryes and Michael Elad*

Department of Computer Science, Technion - IIT 32000, Haifa, ISRAEL.

## ABSTRACT

Signal recovery from a given set of linear measurements using a sparsity prior has been a major subject of research in recent years. In this model, the signal is assumed to have a sparse representation under a given dictionary. Most of the work dealing with this subject has focused on the reconstruction of the signal's representation as the means for recovering the signal itself. This approach forced the dictionary to be of low coherence and with no linear dependencies between its columns. Recently, a series of contributions that focus on signal recovery using the analysis model find that linear dependencies in the analysis dictionary are in fact permitted and beneficial. In this paper we show theoretically that the same holds also for signal recovery in the synthesis case for the  $\ell_0$ -synthesis minimization problem. In addition, we demonstrate empirically the relevance of our conclusions for recovering the signal using an  $\ell_1$ -relaxation.

**Index Terms**— Sparse representations, compressed sensing, analysis versus synthesis, inverse problems.

## 1. INTRODUCTION

Recovering a signal from a noisy set of linear measurements appears in many problems such as compressed sensing, image deblurring and super-resolution. In the basic setup an unknown signal  $\mathbf{x}_0 \in \mathbb{R}^d$  passes through a given linear transformation  $\mathbf{M} \in \mathbb{R}^{m \times d}$ ; an additive noise  $\mathbf{e} \in \mathbb{R}^m$  contaminates the outcome, providing a set of linear measurements

$$\mathbf{y} = \mathbf{M}\mathbf{x}_0 + \mathbf{e}. \quad (1)$$

The noise  $\mathbf{e}$  can be assumed to be either an adversarial unknown bounded noise such that  $\|\mathbf{e}\|_2 \leq \epsilon$  or a random noise with a given distribution, e.g., zero-mean white Gaussian noise. In this paper we focus on the first case. The measurement matrix  $\mathbf{M}$  is usually either rectangular with  $m < d$  or an ill-posed matrix with  $m = d$ . Thus, the information we have in the measurement vector  $\mathbf{y}$  is not enough for recovering  $\mathbf{x}$ , and additional prior knowledge is needed.

The priors considered in this work are sparsity based ones [1]. These include two main paradigms for modeling the signal – the synthesis and the analysis schemes [2]. In the synthesis framework the signal  $\mathbf{x}$  is assumed to have a sparse representation  $\boldsymbol{\alpha} \in \mathbb{R}^n$  under a given dictionary  $\mathbf{D} \in \mathbb{R}^{d \times n}$ ,

$$\mathbf{x}_0 = \mathbf{D}\boldsymbol{\alpha}_0. \quad (2)$$

Few words on our terminology: A vector  $\boldsymbol{\alpha}$  is said to be  $k$ -sparse if  $\|\boldsymbol{\alpha}\|_0 \leq k$ , where  $\|\cdot\|_0$  is the  $\ell_0$  semi-norm that counts the number of its non-zero elements, and a signal  $\mathbf{x}$  is said to have a  $k$ -sparse representation if there exists a  $k$ -sparse  $\boldsymbol{\alpha}$  such that  $\mathbf{x} = \mathbf{D}\boldsymbol{\alpha}$ . The columns of  $\mathbf{D}$  are referred to as atoms, the non-zero locations of a sparse vector  $\boldsymbol{\alpha}$  as the support  $T$ , the size of this support is  $|T|$ , and the sub-matrix of  $\mathbf{D}$  with the atoms corresponding to  $T$  as  $\mathbf{D}_T$ . Given that  $\mathbf{x}_0$  has a  $k$ -sparse representation, we can recover it from its measurement  $\mathbf{y}$  by looking for the sparsest vector  $\boldsymbol{\alpha}$  that satisfies  $\|\mathbf{y} - \mathbf{M}\mathbf{D}\boldsymbol{\alpha}\|_2 \leq \epsilon$  (In the noiseless case,  $\epsilon = 0$ ).

The analysis framework models the signal differently. Given a matrix  $\boldsymbol{\Omega} \in \mathbb{R}^{p \times d}$  – the analysis dictionary – we are interested in the multiplication  $\boldsymbol{\Omega}\mathbf{x}$ , referred to as the analysis representation. A signal  $\mathbf{x}$  is said to be  $\ell$ -cosparsity if  $\|\boldsymbol{\Omega}\mathbf{x}\|_0 \leq p - \ell$ , i.e.,  $\boldsymbol{\Omega}\mathbf{x}$  has  $\ell$  zero elements. In this case we recover the signal by looking for the most cosparsity vector  $\mathbf{x}$ , i.e., the one with the largest number of zeros in  $\boldsymbol{\Omega}\mathbf{x}$  that satisfies  $\|\mathbf{y} - \mathbf{M}\mathbf{x}\|_2 \leq \epsilon$ . More details can be found in [3, 4].

Unlike the analysis model which works in the signal domain and thus recovers it directly, the synthesis one works in the representation domain and thus recovers the signal indirectly by recovering first its representation. Because of this, most of the work that studies theoretical guarantees for the synthesis framework give bounds for the representation's reconstruction, or deal only with the case that  $\mathbf{D}$  is unitary, for which the two problems coincide [5, 6, 7, 8, 9, 10, 11, 12].

Results for more general dictionaries also exist [13, 14]. However, they require the dictionary to be highly incoherent and with no linear dependencies between small number of its columns. This requirement limits the types of dictionaries that can be used to model the signal. While this constraint is necessary for recovering the signal's representation, it is not clear that it is still required when our target is the signal. The goal of this paper is to answer this very question. We aim to show that highly-coherent and even linearly dependent

---

R. Giryes is grateful to the Azrieli Foundation for the award of an Azrieli Fellowship. This research was partly supported by the ERC Advanced grant agreement no. 320649.

atoms in  $\mathbf{D}$  still enable a reliable recovery of the signal from its measurements.

Recently, a series of works [15, 3, 4, 16, 17] showed that in the analysis framework, where we aim at recovering the signal directly, linear dependencies can be allowed in the (analysis) dictionary. These correlations were even favored as they were shown to improve the recovery performance of the different techniques analyzed in these works.

Since linear dependencies are shown to give profit in the analysis framework, it is conjectured that the requirement for an incoherent dictionary for the signal recovery in the synthesis framework is unnecessary as well. A clue for this very property is given in [15], where the reconstruction conditions are presented in terms of  $\mathbf{D}$ -RIP which is a property of the measurement matrix  $\mathbf{M}$  for the synthesis model. However, the results in [15] are derived for signals from the analysis model, thus leaving our question unresolved as of yet.

In this work we focus on the synthesis model and ask: Is it possible to have a recovery guarantee for the synthesis model for a dictionary exhibiting linear dependencies within small groups of its columns? In order to answer this question, we study the performance of the ideal  $\ell_0$ -minimization problem

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}} \|\alpha\|_0 \quad s.t. \quad \|\mathbf{y} - \mathbf{M}\mathbf{D}\alpha\|_2 \leq \epsilon. \quad (3)$$

This is the core approach for recovering a signal that is known to have the synthesis sparsity prior. The reconstructed signal in this technique is simply  $\hat{\mathbf{x}} = \mathbf{D}\hat{\alpha}$ . In the noiseless case, (3) turns to be simply

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}} \|\alpha\|_0 \quad s.t. \quad \mathbf{y} = \mathbf{M}\mathbf{D}\alpha. \quad (4)$$

In this work we study the recovery performance of (3) and (4). We first provide uniqueness conditions for the signal recovery in the noiseless case. Then we present stable reconstruction conditions for the noisy case where the noise is adversarial. This result is a particular case of the result presented in [18, 19] in which a more general form of the  $\mathbf{D}$ -RIP property was proposed giving stable recovery guarantee for signals that come from a general union of subspaces model.

The uniqueness and stability conditions that we present are generalization of previous results [20, 5, 6, 7, 8, 21] that assume  $\mathbf{D}$  to be incoherent. The contribution of this work is in providing reconstruction guarantees in the signal domain that do not pose incoherence requirement on the dictionary.

The  $\ell_0$ -minimization problem is known to be NP-hard [22] and several approximation techniques have been suggested to it [20, 5, 9, 10, 11, 12, 23, 24]. A future work should extend the existing representation recovery guarantees to the signal case. First steps in this direction have been taken for the CoSaMP algorithm in [25]. However, the theoretical study of this algorithm is impaired by the assumed existence of a near-optimal projection scheme in the signal domain, an assumption similar to the one posed in [4]. The existence of such a projection in the general case is still an open question.

In addition to the theoretical study of the  $\ell_0$ -problem, in this paper we show empirically that for the  $\ell_1$ -synthesis minimization problem, a relaxed version of (3) that replaces the  $\ell_0$ -semi-norm with an  $\ell_1$ -norm, signal recovery is at hand in the cases where  $\mathbf{D}$  is highly coherent and the representation recovery is impossible. Similar experiments has been performed in [25] for the CoSaMP algorithm.

The organization of this paper is as follows. In Section 2 we survey existing uniqueness and stability guarantees for the representation reconstruction using the  $\ell_0$ -minimization problem. In Section 3 the previously presented results are extended for the signal reconstruction case, allowing general coherent dictionaries. Empirical reconstruction results for the  $\ell_1$ -minimization problem are presented in Section 4, comparing the representation and signal recovery in the case where  $\mathbf{D}$  is highly coherent. In Section 5 the work is summarized.

## 2. BACKGROUND - GUARANTEES FOR THE REPRESENTATION RECOVERY

Two important properties that were explored for the  $\ell_0$ -problem are the uniqueness of its solution and the stability of the solution under bounded adversarial noise. In this section we survey the existing results for representation recovery. We present the results with no proofs, as they are already contained in those of Section 3 for signal recovery.

### 2.1. Uniqueness

Given a signal's representation  $\alpha_0$  with a cardinality  $k$  ( $\|\alpha_0\|_0 \leq k$ ) we are interested to know whether it is the unique solution of (4). In other words, whether there exists another representation  $\alpha_1 \neq \alpha_0$  with a cardinality at most  $k$  such that  $\mathbf{M}\mathbf{D}\alpha_0 = \mathbf{M}\mathbf{D}\alpha_1$ . An answer for this question was given in [5] using the definition of the Spark of a matrix:

**Definition 2.1 (Definition 1 in [5])** *Given a matrix  $\mathbf{A}$  we define  $\operatorname{Spark}(\mathbf{A})$  as the smallest possible number of columns from  $\mathbf{A}$  that are linearly dependent.*

This definition provides us with a sharp uniqueness condition for the representation reconstruction:

**Theorem 2.2 (Corollary 3 in [5])** *Let  $\mathbf{y} = \mathbf{M}\mathbf{D}\alpha_0$ . If  $\|\alpha_0\|_0 < \operatorname{Spark}(\mathbf{M}\mathbf{D})/2$  then  $\alpha_0$  is the unique solution of (4).*

Though this condition is sharp for finding the representation, it is not sharp at all, in terms of finding the signal itself. This can be demonstrated using the following simple example. Let us assume that  $\mathbf{D} = [\mathbf{z}, \mathbf{z}, \dots, \mathbf{z}]$ , a dictionary with columns that are a duplicate of the same atom  $\mathbf{z}$ . Clearly, the signal  $\mathbf{x}=\mathbf{z}$  can be represented by any of the atoms of  $\mathbf{D}$ , which means that it has  $n$  different sparse representations, each with cardinality 1. Thus, for any measurement matrix  $\mathbf{M}$  there is no unique solution to (4). Indeed, we have  $\operatorname{Spark}(\mathbf{M}\mathbf{D}) = 2$

and the uniqueness condition collapses to the trivial requirement  $\|\alpha_0\|_0 = 0$ . However, it is clear that if our goal is to recover the signal  $\mathbf{x}$  (i.e.  $\mathbf{D}\alpha$ ) and not its representation  $\alpha$ , then we can certainly have a uniqueness, as all the possible solutions to (4) lead to the same signal. Thus, we conclude that for the task of estimating the signal itself, the existing condition is not sharp and a better one should be explored.

## 2.2. Stability

In the case where noise exists in the measurement, the uniqueness of the solution is no longer the question since we cannot recover the original signal exactly. Instead, we ask whether the reconstruction is stable, i.e., whether the  $\ell_2$ -distance between the estimated signal's representation and the original one is proportional to the noise power  $\epsilon$ . In order to establish that, we use the restricted isometry property (RIP) [7, 8]. The RIP can be seen as an extension of the Spark, which allows noisy case analysis.

**Definition 2.3** A matrix  $\mathbf{A}$  satisfies the RIP condition with parameter  $\delta_k$  if it is the smallest value that satisfies

$$(1 - \delta_k) \|\alpha\|_2^2 \leq \|\mathbf{A}\alpha\|_2^2 \leq (1 + \delta_k) \|\alpha\|_2^2 \quad (5)$$

for any  $k$ -sparse vector  $\alpha$ .

The connection between the RIP and the Spark is the following. Given a matrix  $\mathbf{A}$ ,  $k < \text{Spark}(\mathbf{A})$  if  $\delta_k < 1$  [8]. Having the RIP it is straight forward to have a stability condition for the representation reconstruction using (3).

**Theorem 2.4** Let  $\mathbf{y} = \mathbf{M}\mathbf{D}\alpha_0 + \mathbf{e}$  where  $\|\mathbf{e}\|_2 \leq \epsilon$ ,  $\mathbf{M}\mathbf{D}$  satisfies the RIP condition with  $\delta_{2k}$  and  $\|\alpha_0\|_0 \leq k$ . If  $\delta_{2k} < 1$  then the solution  $\hat{\alpha}$  of (3) is stable. More specifically,

$$\|\alpha_0 - \hat{\alpha}\|_2 \leq \frac{2\epsilon}{\sqrt{1 - \delta_{2k}}}. \quad (6)$$

The same problem we had with the dictionary  $\mathbf{D} = [\mathbf{z}, \mathbf{z}, \dots, \mathbf{z}]$  in the uniqueness case repeats also here because  $\delta_{2k} < 1$  implies  $2k < \text{Spark}(\mathbf{M}\mathbf{D})$ .

## 3. GUARANTEES FOR SIGNAL RECOVERY

As mentioned before, though the Spark and RIP conditions for  $\mathbf{M}\mathbf{D}$  are sharp for the representation recovery, they are not designed for the signal recovery. In this section we use an extended Spark and RIP definitions that will serve better the signal recovery problem. The D-RIP [15] is used for having stable recovery conditions for the signal reconstruction. In a similar way, we propose a D-Spark property for the measurement matrix  $\mathbf{M}$ , introducing a new uniqueness condition for the signal recovery. Note that the results in this section are generalization of the ones presented in the previous section for the signal's representation. As a general guideline, by setting the measurement matrix to be  $\mathbf{M}\mathbf{D}$  and the dictionary to be the identity, the results of this section coincide with those of the previous one.

### 3.1. Uniqueness for Signal Recovery

As in the representation case, we are interested to know when we can guarantee that a signal  $\mathbf{x}_0$  with a  $k$ -sparse representation under a matrix  $\mathbf{D}$  is the unique solution of (4). In other words, whether there exist another signal  $\mathbf{x}_1 \neq \mathbf{x}_0$  with at most  $k$ -sparse representation under  $\mathbf{D}$  such that  $\mathbf{M}\mathbf{x}_0 = \mathbf{M}\mathbf{x}_1$ . For this task we introduce the D-Spark, an extension of the Spark definition.

**Definition 3.1** Given a matrix  $\mathbf{M}$  we define  $\mathbf{D}\text{-Spark}(\mathbf{M})$  as the smallest possible number of columns from  $\mathbf{D}$ , marked as  $T$ , such that  $\text{range}(\mathbf{D}_T) \cap \text{Null}(\mathbf{M}) \neq \{0\}$ .

In other words, for every set  $T$  with size  $|T| < \mathbf{D}\text{-Spark}(\mathbf{M})$  we have  $\text{range}(\mathbf{D}_T) \cap \text{Null}(\mathbf{M}) = \{0\}$ , implying that for any vector  $\mathbf{v} \in \mathbb{R}^{|T|}$ ,  $\mathbf{M}\mathbf{D}_T\mathbf{v} = 0$  if and only if  $\mathbf{D}_T\mathbf{v} = 0$ . Note that this definition coincides with the one of the Spark for  $\mathbf{D}=\mathbf{I}$ . This can be observed by noticing that in this case  $\mathbf{D}_T=\mathbf{I}_T$  and thus  $\mathbf{M}\mathbf{D}_T$  simply chooses  $T$  columns from  $\mathbf{M}$ . Thus, the above translates to the requirement that there is no subset of  $|T|$  columns in  $\mathbf{M}$  that are linearly dependent.

Having the D-Spark definition, we propose a uniqueness condition for the signal recovery.

**Theorem 3.2** Let  $\mathbf{y} = \mathbf{M}\mathbf{x}_0$  where  $\mathbf{x}_0$  has a  $k$ -sparse representation  $\alpha_0$  under  $\mathbf{D}$ . If  $k < \mathbf{D}\text{-Spark}(\mathbf{M})/2$  then  $\mathbf{x}_0 = \mathbf{D}\hat{\alpha}$  for  $\hat{\alpha}$  the minimizer of (4), implying a perfect recovery.

*Proof:* Let us assume the contrary, i.e., there exists a minimizer,  $\hat{\alpha}_1$  for (4) such that  $\mathbf{D}\hat{\alpha}_1 \neq \mathbf{D}\alpha_0$ . Let us denote the support sets of  $\alpha_0$  and  $\hat{\alpha}_1$  by  $T_0$  and  $T_1$  respectively. Since  $\alpha_0$  is a feasible solution to (4) and  $\hat{\alpha}_1$  is a minimizer, we have that  $|T_1| \leq |T_0| \leq k$ . Thus, by definition  $\mathbf{D}\hat{\alpha}_1 - \mathbf{D}\alpha_0 \in \text{range}(\mathbf{D}_{T_1 \cup T_0})$  where  $|T_1 \cup T_0| \leq 2k$ . Noticing that  $\mathbf{M}\mathbf{D}\hat{\alpha}_1 = \mathbf{M}\mathbf{D}\alpha_0$ , by the constraint of (4), we have  $\mathbf{D}\hat{\alpha}_1 - \mathbf{D}\alpha_0 \in \text{Null}(\mathbf{M})$ . This contradicts the assumption  $k < \mathbf{D}\text{-Spark}(\mathbf{M})/2$  because we get that  $\mathbf{D}\hat{\alpha}_1 - \mathbf{D}\alpha_0 \in \text{range}(\mathbf{D}_{T_1 \cup T_0}) \cap \text{Null}(\mathbf{M}) \neq \{0\}$ , which means that  $\mathbf{D}\text{-Spark}(\mathbf{M}) < |T_1 \cup T_0| \leq 2k$ .  $\square$

Unlike the uniqueness condition of the regular Spark, the one of the D-Spark allows linear dependencies within the dictionary  $\mathbf{D}$ . Looking at the example from the previous section, for the dictionary  $\mathbf{D} = [\mathbf{z}, \mathbf{z}, \dots, \mathbf{z}]$ ,  $\text{range}(\mathbf{D}_T) = \{\beta\mathbf{z}, \beta \in \mathbb{R}\}$  for any non-empty support  $T$ . Thus, the uniqueness condition turns out to be  $\mathbf{z} \notin \text{Null}(\mathbf{M})$ . This means that the family of matrices  $\mathbf{M}$  that guarantee uniqueness are the ones that have at least one row non-orthogonal to  $\mathbf{z}$ . This is far stronger compared to earlier condition as discussed in Section 2.

### 3.2. Stability for Signal Recovery

Moving to the noisy case, we seek for a generalization of the RIP that provides us with guarantees for the signal recovery. For this task we use the D-RIP, as introduced in [15].

**Definition 3.3** A matrix  $\mathbf{M}$  satisfies the **D-RIP** condition with parameter  $\delta_k^{\mathbf{D}}$  if it is the smallest value that satisfies

$$(1 - \delta_k^{\mathbf{D}}) \|\mathbf{x}\|_2^2 \leq \|\mathbf{M}\mathbf{x}\|_2^2 \leq (1 + \delta_k^{\mathbf{D}}) \|\mathbf{x}\|_2^2 \quad (7)$$

for any vector  $\mathbf{x}$  that has a  $k$ -sparse representation under  $\mathbf{D}$ , or equivalently for any  $\mathbf{x} \in \text{range}(\mathbf{D}_T)$ , where  $T$  is such that  $|T| \leq k$ .

As in the representation case, a connection between the **D-RIP** and the **D-Spark** can be established.

**Proposition 3.4** Given a matrix  $\mathbf{M}$  and sparsity  $k$ , if  $\delta_k^{\mathbf{D}} < 1$  then  $k < \mathbf{D}\text{-Spark}(\mathbf{M})$ .

*Proof:* Requiring  $\delta_k^{\mathbf{D}} < 1$  implies that for any vector  $\mathbf{x} \in \text{range}(\mathbf{D}_T)$  such that  $|T| \leq k$  it holds that  $\|\mathbf{M}\mathbf{x}\|_2 \geq (1 - \delta_k^{\mathbf{D}}) \|\mathbf{x}\|_2 > 0$ , hence  $\mathbf{M}\mathbf{x} \neq 0$ . The last is equivalent to requiring  $\text{Null}(\mathbf{M}) \cap \text{range}(\mathbf{D}_T) = \{0\}$  for any support set  $T$  such that  $|T| \leq k$ , which is exactly equivalent to  $\mathbf{D}\text{-Spark}(\mathbf{M}) > k$ .  $\square$

Having the definition of the **D-RIP** we present a stability guarantee for the signal recovery that appears in [18, 19].

**Theorem 3.5** Let  $\mathbf{y} = \mathbf{M}\mathbf{x}_0 + \mathbf{e}$  where  $\|\mathbf{e}\|_2 \leq \epsilon$ ,  $\mathbf{x}_0$  has a  $k$ -sparse representation  $\alpha_0$  under  $\mathbf{D}$ , and  $\mathbf{M}$  satisfies the **D-RIP** condition with  $\delta_{2k}^{\mathbf{D}}$ . If  $\delta_{2k}^{\mathbf{D}} < 1$  then recovering the signal using (3), where the recovered signal is  $\hat{\mathbf{x}} = \mathbf{D}\hat{\alpha}$ , is stable:

$$\|\mathbf{x}_0 - \hat{\mathbf{x}}\|_2 \leq \frac{2\epsilon}{\sqrt{1 - \delta_{2k}^{\mathbf{D}}}}. \quad (8)$$

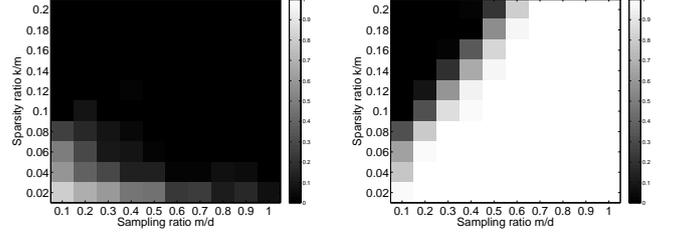
*Proof:* Since  $\alpha_0$ , the representation of  $\mathbf{x}_0$ , is a feasible solution to (3) we have that  $\|\hat{\alpha}\|_0 \leq \|\alpha_0\|_0 \leq k$ . Thus,  $\mathbf{x}_0 - \hat{\mathbf{x}}$  is a signal that has a  $2k$ -sparse representation. According to the constraint in (3) we also have  $\|\mathbf{y} - \mathbf{M}\hat{\mathbf{x}}\|_2 \leq \epsilon$ . Using the **D-RIP**, the triangle inequality, and the fact that  $\|\mathbf{y} - \mathbf{M}\mathbf{x}_0\|_2 \leq \epsilon$  as well, we get

$$\begin{aligned} \|\mathbf{x}_0 - \hat{\mathbf{x}}\|_2 &\leq \frac{1}{\sqrt{1 - \delta_{2k}^{\mathbf{D}}}} \|\mathbf{M}(\mathbf{x}_0 - \hat{\mathbf{x}})\|_2 \\ &\leq \frac{1}{\sqrt{1 - \delta_{2k}^{\mathbf{D}}}} (\|\mathbf{y} - \mathbf{M}\mathbf{x}_0\|_2 + \|\mathbf{y} - \mathbf{M}\hat{\mathbf{x}}\|_2) \leq \frac{2\epsilon}{\sqrt{1 - \delta_{2k}^{\mathbf{D}}}}, \end{aligned} \quad (9)$$

which is the stated result.  $\square$

#### 4. PRACTICAL RECONSTRUCTION

With the above observations, we turn to check the recovery performance of the  $\ell_1$ -minimization problem – a relaxed version that replaces the  $\ell_0$  with  $\ell_1$ -norm. In our experiment we consider the noiseless scenario, and since solving (4) is NP-hard [22], the  $\ell_1$  relaxation is inevitable when dealing with practical problems. We perform a simple synthetic test for



**Fig. 1.** Representation recovery rate (left) and signal recovery rate (right). Color attribute: fraction of realizations in which  $\ell_1$  minimization achieves a perfect recovery.

signals that are sparse under a dictionary which is highly coherent and with linear dependencies between its columns. We generate a dictionary  $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2]$  where  $\mathbf{D}_1, \mathbf{D}_2 \in \mathbb{R}^{d \times d}$ ,  $d=1000$ ,  $\mathbf{D}_1$  contains sparse columns with 2 non-zero entries which are 1 or  $-1$  with probability 0.5 and  $\mathbf{D}_2$  contains columns which are linear combinations of random 3 columns from  $\mathbf{D}_1$  with random zero-mean white Gaussian weights. Each entry of the measurement matrix  $\mathbf{M} \in \mathbb{R}^{m \times d}$  is distributed according to a normal Gaussian distribution, where  $m = \lfloor \gamma d \rfloor$  and  $\gamma$  is the sampling rate – a value in the range  $(0, 1]$ . We set  $k$  to be  $\lfloor \rho m \rfloor$  ( $\rho \ll 1$ ) and measure the recovery rate of the representation  $\alpha$  and the signal  $\mathbf{x}$  for various values of  $\gamma \in \{0.1, 0.2, \dots, 1\}$  and  $\rho \in \{0.02, 0.04, \dots, 0.2\}$ .

Figure 1 presents the recovery performance over 100 realizations per each parameter setting. Similar to the uniqueness results for the representation, the reconstruction fails almost always since  $\mathbf{D}$  contains many linear dependencies (note that in our case  $\text{Spark}(\mathbf{M}\mathbf{D}) = 4$ ). However, though we fail in reconstructing the representations, in most cases the recovery of the signal succeeds. Note that even in the case where  $m=d$  ( $\gamma=1$ ), we get a very low recovery rate for the representations and this is due to the non-uniqueness of any representation with cardinality beyond 1. The recovery rate on the bottom left part of the representation diagram is better than the other parts because for very small values of  $k$  the chance to choose the wrong representation decreases significantly.

#### 5. CONCLUSION

In this work we studied the  $\ell_0$ -synthesis problem for signal recovery in the case of a general dictionary  $\mathbf{D}$ . We have shown that in the case where  $\mathbf{D}$  contains linear dependencies signal recovery is still theoretically plausible. We derived theoretical conditions for uniqueness of the solution in the noiseless case and stability for the noisy case. In addition, empirical reconstruction results were presented for the  $\ell_1$ -synthesis minimization problem, demonstrating that for this problem signal recovery is achievable in the case where  $\mathbf{D}$  is highly coherent. These results motivate looking for theoretical guarantees for the signal recovery using the  $\ell_1$ -relaxation in the same spirit of those developed in the representation case [20, 5, 7], a topic that would be covered by our upcoming research.

## 6. REFERENCES

- [1] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.
- [2] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse Problems*, vol. 23, no. 3, pp. 947–968, June 2007.
- [3] S. Nam, M. Davies, M. Elad, and R. Gribonval, "The cosparsity analysis model and algorithms," *Appl. Comput. Harmon. Anal.*, vol. 34, no. 1, pp. 30 – 56, 2013.
- [4] R. Giryes, S. Nam, M. Elad, R. Gribonval, and M. E. Davies, "Greedy-like algorithms for the cosparsity analysis model," to appear in the *Special Issue in Linear Algebra and its Applications on Sparse Approximate Solution of Linear Systems*, 2013.
- [5] D. L. Donoho and M. Elad, "Optimal sparse representation in general (nonorthogonal) dictionaries via  $l_1$  minimization," *Proceedings of the National Academy of Science*, vol. 100, pp. 2197–2202, Mar 2003.
- [6] ———, "On the stability of the basis pursuit in the presence of noise," *Signal Process.*, vol. 86, no. 3, pp. 511–532, 2006.
- [7] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory.*, vol. 52, no. 12, pp. 5406 –5425, Dec. 2006.
- [8] E. Candès and T. Tao, "Decoding by linear programming," vol. 51, no. 12, pp. 4203 – 4215, dec. 2005.
- [9] D. Needell and J. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301 – 321, May 2009.
- [10] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory.*, vol. 55, no. 5, pp. 2230 –2249, May 2009.
- [11] T. Blumensath and M. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265 – 274, 2009.
- [12] S. Foucart, "Hard thresholding pursuit: an algorithm for compressive sensing," *SIAM J. Numer. Anal.*, vol. 49, no. 6, pp. 2543–2563, 2011.
- [13] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *IEEE Trans. Inf. Theory.*, vol. 54, no. 5, pp. 2210 –2219, May 2008.
- [14] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," vol. 50, no. 10, pp. 2231–2242, 2004.
- [15] E. J. Candès, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, pp. 59 – 73, 2011.
- [16] R. Rubinstein, T. Peleg, and M. Elad, "Analysis K-SVD: A dictionary-learning algorithm for the analysis sparse model," *IEEE Trans. on Signal Processing*, vol. 61, no. 3, pp. 661–677, Feb. 2013.
- [17] T. Peleg and M. Elad, "Performance guarantees of the thresholding algorithm for the cosparsity analysis model," *IEEE Trans. on Information Theory*, vol. 59, no. 3, pp. 1832–1845, Mar. 2013.
- [18] T. Blumensath and M. Davies, "Sampling theorems for signals from the union of finite-dimensional linear subspaces," *IEEE Trans. Inf. Theory.*, vol. 55, no. 4, pp. 1872 –1882, april 2009.
- [19] Y. Lu and M. Do, "A theory for sampling signals from a union of subspaces," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2334 –2345, Jun. 2008.
- [20] M. Elad and A. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases," vol. 48, no. 9, pp. 2558–2567, Sep 2002.
- [21] E. Candès, "Modern statistical estimation via oracle inequalities," *Acta Numerica*, vol. 15, pp. 257–325, 2006.
- [22] B. Natarajan, "Sparse approximate solutions to linear systems," *SIAM Journal on Computing*, vol. 24, pp. 227–234, 1995.
- [23] E. Candès and T. Tao, "The Dantzig selector: Statistical estimation when  $p$  is much larger than  $n$ ," *Annals Of Statistics*, vol. 35, p. 2313, 2007.
- [24] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, pp. 3397–3415, 1993.
- [25] M. A. Davenport, D. Needell, and M. B. Wakin, "Signal space CoSaMP for sparse recovery with redundant dictionaries," *CoRR*, vol. abs/1208.0353, 2012.