

Research Article

Linear Program Relaxation of Sparse Nonnegative Recovery in Compressive Sensing Microarrays

Linxia Qin, Naihua Xiu, Lingchen Kong, and Yu Li

Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Linxia Qin, lxqin.echo@163.com

Received 23 May 2012; Accepted 30 July 2012

Academic Editor: Huafeng Liu

Copyright © 2012 Linxia Qin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Compressive sensing microarrays (CSM) are DNA-based sensors that operate using group testing and compressive sensing principles. Mathematically, one can cast the CSM as sparse nonnegative recovery (SNR) which is to find the sparsest solutions subjected to an underdetermined system of linear equations and nonnegative restriction. In this paper, we discuss the l_1 relaxation of the SNR. By defining nonnegative restricted isometry/orthogonality constants, we give a nonnegative restricted property condition which guarantees that the SNR and the l_1 relaxation share the common unique solution. Besides, we show that any solution to the SNR must be one of the extreme points of the underlying feasible set.

1. Introduction

Nowadays, with the rapid development of molecular biology techniques, scientists use compressive sensing microarrays to collect the gene expression changes of patients suffer from specific diseases and test a lot of different drugs on cells genetically to look for medicine being able to change the abnormal gene expression [1, 2]. A DNA microarray is a collection of microscopic DNA spots attached to a solid surface. Each DNA spot contains a string of specific DNA sequences, known as probes. These can be a short section of a gene or other DNA element that are used to hybridize an organism's genetic sample under high-stringency conditions. Probe-target hybridization is usually detected and quantified by detection of chemiluminescence-labeled targets to infer the genetic makeup in the test sample.

Although the number of DNA sequences is extremely large, not all agents are expected to be present in a significant concentration at a given time and location. In traditional microarrays, this results in many inactive probes during sensing. On the other hand, we are often interested in only a small quantity of certain harmful biological agents. Therefore, it is important to not just detect the presence of agents in a sample but also estimate the concentrations with which they are present.

Assume that there are m spots and n labeled targets, and we have far fewer spots than target agents such that $m \ll n$. Mathematically, one can represent the DNA concentration of each organism as an element in a vector $x \in \mathbb{R}^n$ and the measurements as $b \in \mathbb{R}^m$. For $1 \leq i \leq m$ and $1 \leq j \leq n$, the probe at spot i hybridizes to target j with probability a_{ij} . The target j occurs in the tested DNA sample with concentration x_j , which is clearly nonnegative. Denoting by $A := (a_{ij})_{m \times n}$, the process of DNA microarrays leads to the *sparse nonnegative recovery* (SNR) which is to find the sparsest solutions subjected to an underdetermined system of linear equations and nonnegative constraints, with the mathematical model as follows:

$$\begin{aligned} \min & \|x\|_0, \\ \text{s.t.} & Ax = b, \quad x \geq 0, \end{aligned} \tag{P_0}$$

where the variable vector $x \in \mathbb{R}^n$, $\|x\|_0$ denotes the number of the nonzero entries of x , $A \in \mathbb{R}^{m \times n}$ is the measurement matrix with full row rank, $m \ll n$, and $b \in \mathbb{R}^m$.

SNR can be regarded as a special case of the sparse recovery, which is related to program $\min\{\|x\|_0 \mid Ax = b\}$. This program has sparked the significant concern and rapid development in recent years [3–5] owing to its wide applications. However, with the nonnegativity prior information about the object to be recovered in various applications

such as CSM, solutions on (P_0) tend to be closer to the actual situations and lead to substantial improvements in the image reconstruction. Moreover, with the nonnegative constraints, the feasible set becomes a polyhedral set instead of an affine subspace. This will bring us essential hardness in projecting on the feasible set. Thus, (P_0) is more likely difficult to solve. Therefore, SNR deserves specific study.

Problem (P_0) has been shown to be NP-hard [6, 7] in general from the perspective of computational complexity. One popular approach is to reconstruct the vector via the l_1 relaxation, which refers to

$$\begin{aligned} \min \|x\|_1, \\ \text{s.t. } Ax = b, \quad x \geq 0. \end{aligned} \quad (P_1)$$

Since (P_1) is a standard linear program, it is easy to solve. An important issue is how to guarantee the equivalence of (P_0) and (P_1) in the sense that they have the same unique k -sparse solution under some conditions. Here, we call a vector x k -sparse if the number of its nonzero entries is no more than k . There has been some increasing interest and activity in this area; see, for example, [8–14]. Donoho and Tanner [9] firstly proposed that (P_0) and (P_1) share the common k -sparse unique solution if the polytope AT is outwardly k -neighborly, where T is the standard simplex in \mathbb{R}^n . Zhang [13] proved that (P_0) and (P_1) share the common k -sparse unique solution if the null space of A is strictly half k -balance. Juditsky et al. [11] developed several different necessary and sufficient conditions for the (P_0) - (P_1) equivalence in the case of general type sign restrictions, including the nonnegative constraints as its special case. When the feasible set of (P_0) is a singleton, the unknown can be recovered by optimizing any objective function over this constraint set, and (P_0) and (P_1) definitely get the same unique solution. In this case, Bruckstein et al. [8] got the uniqueness of the feasible solution under a sufficient condition that A has a rowspan intersecting the positive orthant. Furthermore, Wang et al. [14] proved that the above sufficient condition is also necessary to the uniqueness of the feasible solution. Donoho and Tanner [10] proved that the underlying feasible set is a singleton if and only if the polytope $A\mathbb{R}_+^n$ and \mathbb{R}_+^n have the same number of k -faces. Khajehnejad et al. [12] gave another equivalent condition of the uniqueness property by characterizing the support size of vectors in the null space of A .

For the l_1 relaxation of sparse recovery, one of the most significant conditions is the restricted isometry property (RIP), named by Candès and Tao [15] with the groundbreaking work of Donoho et al. [16, 17]. However, to the best of our knowledge, the nonnegative case of RIP has not been investigated. This paper will deal with this issue. We begin with investigating the solution property of SNR and show that any solution to the SNR must be one extreme point of its feasible set in Section 2. We prove in Section 3 the nonemptiness and boundedness of the solution set of (P_1) and show that any solution of (P_1) could be stated as the convex combination of its optimal extreme points. In Section 4, by defining the nonnegative restricted isometry/orthogonality constants, we derive a sufficient

condition for exact recovery of the sparsest nonnegative image/signal via the linear program relaxation.

Now we give some notations used in the text. We use $\text{sol}(\cdot)$ and $v(\cdot)$ to denote the solution set and optimal value of problem (\cdot) . The $e_i \in \mathbb{R}^n$ would be the vector with only the i th entry 1 and the rest all 0. $e^l \in \mathbb{R}^l$ is the vector with each entry equal to 1; we also use e to demonstrate that $e^n \in \mathbb{R}^n$ for short. The $a_i \in \mathbb{R}^m$ for $i = 1, \dots, n$ denote the column vectors of the matrix A and $A_I = (a_i)_{i \in I \subset \{1, \dots, n\}}$. For any $x \in \mathbb{R}^n$, x_i is the i th component and $I(x)$ is the support set of x ; that is, $I(x) = \{i \mid x_i \neq 0, i = 1, \dots, n\}$. For any subset $T \subset \{1, \dots, n\}$, T^c denotes the complement set of T out of $\{1, \dots, n\}$.

2. Solution Property

Throughout the paper we assume that

$$\mathcal{S} := \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\} \neq \emptyset. \quad (1)$$

Apparently, \mathcal{S} is a polyhedral set in \mathbb{R}^n . According to the representation theorem, any $x \in \mathcal{S}$ could be represented as follows:

$$x = \sum_{i=1}^t \lambda_i \bar{x}^{(i)} + \sum_{j=1}^q \mu_j d^{(j)}, \quad (2)$$

where $\sum_{i=1}^t \lambda_i = 1$, $\lambda_i \geq 0$, $i = 1, \dots, t$, and $\{\bar{x}^{(i)} \in \mathbb{R}^n \mid i = 1, 2, \dots, t\}$ are the extreme point set of \mathcal{S} ; $\mu_j \geq 0$, $j = 1, \dots, q$, and $\{d^{(j)} \in \mathbb{R}^n \mid j = 1, 2, \dots, q\}$ are the extreme direction set of \mathcal{S} . Apparently, $A\bar{x}^{(i)} = b$, $i = 1, \dots, t$; $Ad^{(j)} = 0$, $d^{(j)} \geq 0$, $j = 1, \dots, q$.

Define subsets of \mathbb{R}^n as follows:

$$\begin{aligned} S_0 &= \text{sub}\{0\}, \\ S_1 &= \left\{ \bigcup_{i=1}^n \text{sub}\{e_i\} \right\} \setminus S_0, \\ S_2 &= \left\{ \bigcup_{i_1, i_2=1}^n \text{sub}\{e_{i_1}, e_{i_2}\} \right\} \setminus \bigcup_{j=0}^1 S_j, \\ &\vdots \\ S_r &= \left\{ \bigcup_{i_1, \dots, i_r=1}^n \text{sub}\{e_{i_1}, \dots, e_{i_r}\} \right\} \setminus \bigcup_{j=0}^{r-1} S_j, \\ &\vdots \\ S_n &= \left\{ \text{sub}\{e_1, \dots, e_n\} \right\} \setminus \bigcup_{j=0}^{n-1} S_j, \end{aligned} \quad (3)$$

where $\text{sub}\{e_{i_1}, \dots, e_{i_k}\}$ denotes the subspace spanned by the vectors e_{i_1}, \dots, e_{i_k} , $k = 1, \dots, n$. Clearly, $\{S_1, S_2, \dots, S_n\}$ forms a partition of \mathbb{R}^n ; that is, $\bigcup_{j=1}^n S_j = \mathbb{R}^n$ and $S_i \cap S_j = \emptyset$, $i \neq j$. Moreover, $\|x\|_0 = r$ for any $x \in S_r$. Along with the nonemptiness of \mathcal{S} , it is easy to see that \mathcal{S} must intersect one

of these sets, hence $v(P_0) = \min\{r \in \mathbb{R} \mid \mathcal{S} \cap S_r \neq \emptyset\}$ and $\text{sol}(P_0) \neq \emptyset$. Furthermore, we have the following result for the optimal value of (P_0) .

Lemma 1. *Assume that $v(P_0) = k$, one must have $k \leq m$.*

Proof. Suppose that the conclusion is not true; that is, there is $x^* \in \text{sol}(P_0)$ with $\|x^*\|_0 = k > m$. Without loss of generality, let $x_i^* > 0, i = 1, \dots, k$. We get

$$x_1^* a_1 + \dots + x_k^* a_k = b. \quad (4)$$

Meanwhile, $\text{rank}(A) = m < k$. Thus, $\{a_i \in \mathbb{R}^m \mid i = 1, \dots, k\}$ must be linearly dependent; that is, there exist d_1, \dots, d_k , not all zero, such that

$$d_1 a_1 + \dots + d_k a_k = 0. \quad (5)$$

Assume that $d_1 > 0$. By denoting $d = (d_1, \dots, d_k, 0, \dots, 0)^T \in \mathbb{R}^n$, we get $Ad = 0$. Taking $\delta = \min\{x_i^*/d_i \mid d_i > 0, i = 1, \dots, k\}$, it holds that

$$\begin{aligned} A(x^* - \delta d) &= Ax^* - \delta Ad = b, \\ x^* - \delta d &\geq 0, \\ \|x^* - \delta d\|_0 &\leq k - 1. \end{aligned} \quad (6)$$

This is a contradiction with $x^* \in \text{sol}(P_0)$. We complete the proof. \square

To characterize property of the solution set $\text{sol}(P_0)$, we need the next lemma. In particular, this brand new result will play a key role in proposing the sufficient condition of the uniqueness of $\text{sol}(P_0)$ in Section 4.

Lemma 2. *Any two distinct solutions of (P_0) must have different support sets.*

Proof. Assume for contradiction that x^* and \tilde{x} are two different solutions of (P_0) , $x^* \neq \tilde{x}$. If $I(x^*) = I(\tilde{x})$, we have $x_i^* > 0, \tilde{x}_i > 0$, for all $i \in I(x^*)$. Set $\lambda = \min\{\min_{i \in I(x^*)} \{x_i^*/\tilde{x}_i\}, \min_{i \in I(x^*)} \{\tilde{x}_i/x_i^*\}\}$. Since $x^* \neq \tilde{x}$, it must hold $\lambda < 1$. When $\lambda = x_{i_0}^*/\tilde{x}_{i_0}$, take

$$\hat{x} = \frac{1}{1-\lambda} x^* - \frac{\lambda}{1-\lambda} \tilde{x}. \quad (7)$$

It is easy to see that

$$\hat{x}_i = \frac{1}{1-\lambda} x_i^* - \frac{\lambda}{1-\lambda} \tilde{x}_i \in \begin{cases} \{0\}, & i = i_0, \\ \mathbb{R}_+, & i \in I(x^*) \setminus \{i_0\}. \end{cases} \quad (8)$$

Thus, we have $A\hat{x} = b, \hat{x} \geq 0$ and $\|\hat{x}\|_0 \leq \|x^*\|_0 - 1$. This is a contradiction with the optimality of x^* . When $\lambda = \tilde{x}_{j_0}/x_{j_0}^*$, just taking $\hat{x} = (1/(1-\lambda))\tilde{x} - (\lambda/(1-\lambda))x^*$ instead, we get the contradiction by a similar way. The proof is completed. \square

Now we are in a position to give the main theorem in this section.

Theorem 3. *Any solution of (P_0) must be one of the extreme points of \mathcal{S} .*

Proof. Given any solution $x^* \in \text{sol}(P_0)$ with representation

$$x^* = \sum_{i=1}^t \lambda_i^* \bar{x}^{(i)} + \sum_{j=1}^q \mu_j^* d^{(j)}, \quad (9)$$

where $\sum_{i=1}^t \lambda_i^* = 1, \lambda_i^* \geq 0, \bar{x}^{(i)} \geq 0, A\bar{x}^{(i)} = b$, for all $i = 1, \dots, t; \mu_j^* \geq 0, d^{(j)} \geq 0, Ad^{(j)} = 0$, for all $j = 1, \dots, q$. We only need to prove that

$$\begin{aligned} \sum_{i=1}^t \lambda_i^* &= 1, \quad \lambda_i^* \in \{0, 1\}, \quad i = 1, \dots, t, \\ \mu_j^* &= 0, \quad j = 1, \dots, q, \end{aligned} \quad (10)$$

in (9). To this end, we have the following three steps.

Firstly, we claim that in (9),

$$\begin{aligned} I(\bar{x}^{(i)}) &= I(x^*), \quad i \in \{i \mid \lambda_i^* > 0, i = 1, \dots, t\}, \\ I(d^{(j)}) &\subset I(x^*), \quad j \in \{j \mid \mu_j^* > 0, j = 1, \dots, q\}. \end{aligned} \quad (11)$$

According to the fact that $\bar{x}^{(i)} \geq 0, i = 1, \dots, t$, and $\mu_j^* \geq 0, d^{(j)} \geq 0, j = 1, \dots, q$, one has for any $l \in I(\bar{x}^{(i)}), i \in \{i \mid \lambda_i^* > 0, i = 1, 2, \dots, t\}$,

$$x_l^* = \sum_{i=1}^t \lambda_i^* \bar{x}_l^{(i)} + \sum_{j=1}^q \mu_j^* d_l^{(j)} > 0, \quad (12)$$

which means that $l \in I(x^*)$. This implies $I(\bar{x}^{(i)}) \subset I(x^*)$. Similarly, we get $I(d^{(j)}) \subset I(x^*)$. On the contrary, from the optimality of x^* and (9), we know that $I(x^*) \subset I(\bar{x}^{(i)})$.

Secondly, we will show that $\mu_j^* = 0, j = 1, \dots, q$. If this is not true, there is an index j_0 such that $\mu_{j_0}^* > 0$, then $\mu_{j_0}^* d^{(j_0)}$ has at least one positive component. Denote $d = \sum_{j=1}^q \mu_j^* d^{(j)}$, so $Ad = 0$ and $\{d_l > 0 \mid l = 1, \dots, n\} \neq \emptyset$. Noting that $\bar{x}^{(i)} \geq 0$, (11) implies the fact that $I(\sum_{i=1}^t \lambda_i^* \bar{x}^{(i)}) = I(x^*), I(d) \subset I(x^*)$. Take $\delta = \min\{(\sum_{i=1}^t \lambda_i^* \bar{x}^{(i)})/d_l \mid d_l > 0, l = 1, \dots, n\}$, and

$$\tilde{x} = \sum_{i=1}^t \lambda_i^* \bar{x}^{(i)} - \delta d. \quad (13)$$

Without loss of generality, set $\delta = (\sum_{i=1}^t \lambda_i^* \bar{x}^{(i)})_{l_0}/d_{l_0}$. It is easy to verify that

$$\tilde{x}_j = \left(\sum_{i=1}^t \lambda_i^* \bar{x}^{(i)} \right)_j - \delta d_j \in \begin{cases} \{0\}, & j = l_0 \\ \mathbb{R}_+, & i \in I(x^*) \setminus \{l_0\}, \end{cases} \quad (14)$$

hence

$$\begin{aligned} A\tilde{x} &= b, \\ \tilde{x} &\geq 0, \end{aligned} \quad (15)$$

$$\|\tilde{x}\|_0 \leq \|x^*\|_0 - 1,$$

which is a contradiction with the optimality of x^* .

Thirdly, we will prove that $\sum_{i=1}^t \lambda_i^* = 1$, $\lambda_i^* \in \{0, 1\}$, $i = 1, \dots, t$. Suppose that there exist $\lambda_1^* > 0$, $\lambda_2^* > 0$, and $\lambda_1^* + \lambda_2^* = 1$ and two different extreme points of \mathcal{S} , say $\bar{x}^{(1)}, \bar{x}^{(2)}$, such that

$$x^* = \lambda_1^* \bar{x}^{(1)} + \lambda_2^* \bar{x}^{(2)}. \quad (16)$$

Based on (11), we have $I(\bar{x}^{(1)}) = I(\bar{x}^{(2)}) = I(x^*)$, hence $\bar{x}^{(1)} \in \text{sol}(P_0)$, $\bar{x}^{(2)} \in \text{sol}(P_0)$. Nevertheless, by Lemma 2, this is impossible. Hence, we show that

$$x^* \in \{\bar{x}^{(1)}, \bar{x}^{(2)}, \dots, \bar{x}^{(t)}\}. \quad (17)$$

We complete the argument. \square

Theorem 3 tells us that each solution of (P_0) lies in the extreme point set of \mathcal{S} . Here is a concrete example.

Example 4. Let $A = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$, $b = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$. Obviously, the solution set and optimal value of (P_0) are

$$\text{sol}(P_0) = \{(2, 1, 0)\}, \quad v(P_0) = 2, \quad (18)$$

respectively. In this case, $(2, 1, 0)$ is the only extreme point of \mathcal{S} . While the solution set and optimal value of $(P) := \min\{\|x\|_0 \mid Ax = b\}$ are

$$\text{sol}(P) = \{(0, 0, -1)\}, \quad v(P) = 1, \quad (19)$$

respectively.

At the end of this section, we consider the l_p ($0 < p < 1$) relaxation of (P_0)

$$\begin{aligned} \min \quad & \|x\|_p^p = \sum_{i=1}^n x_i^p, \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0. \end{aligned} \quad (P_p)$$

Clearly, (P_p) is a concave relaxation of (P_0) . For the program (P_p) , Ge et al. [7] derived the useful result as the following.

Lemma 5. *The set of all extreme points of \mathcal{S} is exactly the set of all local minimizers to (P_p) .*

This lemma implies that any global solution of l_p relaxation must be one of its extreme points. From Theorem 3 and Lemma 5, we immediately draw a new proposition.

Proposition 6. *For any $p \in (0, 1)$, there exists an extreme point of \mathcal{S} that is both an exact solution of (P_0) and a local minimizer of (P_p) .*

This is different from the result of Fung and Mangasarian in [18], where they showed that for sufficiently small $\bar{p} \in (0, 1)$, there exists an extreme point \bar{x} of the polyhedral set \mathcal{T} , obtained by lifting the set \mathcal{S} , such that \bar{x} is an exact solution of (P_0) and a global solution of the $l_{\bar{p}}$ relaxation.

3. Linear Program Relaxation

Consider the linear program relaxation (P_1) . Since the linear objective function $\langle e, x \rangle$ is bounded below over the feasible set, based on the Frank-Wolfe theorem, the minimum of (P_1) is attainable. Among all the extreme points of \mathcal{S} , $\{\bar{x}^{(i)} \in \mathbb{R}^n \mid i = 1, \dots, t\}$, we call $\bar{x}^{(i)}$ an optimal extreme point if it also meets $\langle e, \bar{x}^{(i)} \rangle = v(P_1)$.

Proposition 7. *Any $x^* \in \text{sol}(P_1)$ could be stated as the convex combination of optimal extreme points of (P_1) . Hence, $\text{sol}(P_1)$ is bounded.*

Proof. Given any $x^* \in \text{sol}(P_0)$ with representation (9). If there is $i_0 \in \{1, \dots, t\}$ such that $\bar{x}^{(i_0)}$ is not an optimal extreme point of (P_1) and $\lambda_{i_0}^* > 0$, we have $\langle e, \bar{x}^{(i_0)} \rangle > v(P_1)$, hence

$$\begin{aligned} v(P_1) &= \langle e, x^* \rangle \\ &= \sum_{i=1}^t \lambda_i^* \langle e, \bar{x}^{(i)} \rangle + \sum_{j=1}^q \mu_j^* \langle e, d^{(j)} \rangle \\ &\geq \sum_{i=1}^t \lambda_i^* \langle e, \bar{x}^{(i)} \rangle \\ &> \sum_{i=1}^t \lambda_i^* v(P_1) = v(P_1), \end{aligned} \quad (20)$$

which is a contradiction.

Similarly, if there is $j_0 \in \{1, \dots, q\}$ such that $\mu_{j_0}^* > 0$, we have $\mu_{j_0}^* \langle e, d^{(j_0)} \rangle > 0$, hence

$$\begin{aligned} v(P_1) &= \langle e, x^* \rangle \\ &= \sum_{i=1}^t \lambda_i^* \langle e, \bar{x}^{(i)} \rangle + \sum_{j=1}^q \mu_j^* \langle e, d^{(j)} \rangle \\ &> \sum_{i=1}^t \lambda_i^* \langle e, \bar{x}^{(i)} \rangle \\ &\geq \sum_{i=1}^t \lambda_i^* v(P_1) = v(P_1), \end{aligned} \quad (21)$$

which is a contradiction. This completes the proof. \square

From the above proposition, we know that linear program (P_1) has at least one optimal extreme point. Thus, we could use simplex method or interior point method to solve (P_1) .

4. Nonnegative Restricted Property

In the framework of l_1 relaxation, a significant problem is how to guarantee the exact recovery of sparse image/signal via the l_1 relaxation. One of the most important qualifications is the restricted isometry property; see [15]. Recall that

the k -restricted isometry constants (RIC) δ_k is the smallest scalar satisfying

$$(1 - \delta_k)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_k)\|x\|_2^2, \quad \forall \|x\|_0 \leq k. \quad (22)$$

Similarly, the k, k' -restricted orthogonality constants (ROC) $\theta_{k,k'}$ for $k + k' \leq n$ is defined as the smallest scalar satisfying

$$\begin{aligned} |\langle Ax, Ay \rangle| &\leq \theta_{k,k'} \|x\|_2 \|y\|_2, \\ \forall \|x\|_0 &\leq k; \forall \|y\|_0 \leq k', \end{aligned} \quad (23)$$

where x and y have disjoint support sets. The RIC δ_k and ROC $\theta_{k,k'}$ measure how close each submatrix of A with certain cardinality is behaving like an orthonormal system. Under some restricted isometry property, one can get the sparse recovery via its l_1 relaxation. Nevertheless, for the nonnegative case, the sparse recovery may maintain new characterizations. Above all, we define NRIC and NROC.

Definition 8. Let $A \in \mathbb{R}^{m \times n}$. We define the nonnegative k -restricted isometry constants (NRIC) δ_k^+ as the smallest number satisfying

$$\begin{aligned} (1 - \delta_k^+) \|x\|_2^2 &\leq \|Ax\|_2^2 \leq (1 + \delta_k^+) \|x\|_2^2, \\ \forall x \geq 0, \|x\|_0 &\leq k. \end{aligned} \quad (24)$$

Similarly, we define the nonnegative k, k' -restricted orthogonality constants (NROC) $\theta_{k,k'}^+$ for $k + k' \leq n$ as the smallest number satisfying

$$\begin{aligned} |\langle Ax, Ay \rangle| &\leq \theta_{k,k'}^+ \|x\|_2 \|y\|_2, \\ \forall x \geq 0, \|x\|_0 &\leq k, \forall y \geq 0, \|y\|_0 \leq k', \end{aligned} \quad (25)$$

with $I(x)$ and $I(y)$ being disjoint sets.

Clearly,

$$\delta_k^+ \leq \delta_k, \quad \theta_{k,k'}^+ \leq \theta_{k,k'}. \quad (26)$$

Moreover, the numbers δ_k^+ and $\theta_{k,k'}^+$ are nondecreasing in k, k' .

By employing the projections of vectors in the null space of A to \mathbb{R}_+^n , we now provide a sufficient condition to determine a solution of (P_0) .

Theorem 9. Suppose that $k \geq 1$ is such that $\delta_k^+ + \theta_{k,k-1}^+ < 1$ and $x^* \in \mathcal{S}$ with $\|x^*\|_0 = k$. Then, x^* is a solution of (P_0) .

Proof. We complete the proof by contradiction. If this is not true, there exists \bar{x} such that $A\bar{x} = b$, $\bar{x} \geq 0$, $\|\bar{x}\|_0 \leq k - 1$. Set $h = x^* - \bar{x}$. Clearly, $Ah = 0$. Take $h = h^+ - h^-$ with h^+ and h^- being the projections of h and $-h$ to \mathbb{R}_+^n , respectively. We have $h^+ \geq 0$, $h^- \geq 0$, $\langle h^+, h^- \rangle = 0$. In particular,

$$\begin{aligned} (h^+)_i &= \begin{cases} (x_i^* - \bar{x}_i)_+, & i \in I(x^*), \\ 0, & \text{else,} \end{cases} \\ (h^-)_i &= \begin{cases} (x_i^* - \bar{x}_i)_-, & i \in I(\bar{x}), \\ 0, & \text{else.} \end{cases} \end{aligned} \quad (27)$$

Therefore,

$$\begin{aligned} 1 &\leq \|h\|_0 \leq \|x^*\|_0 + \|\bar{x}\|_0 \leq 2k - 1, \\ 1 &\leq \|h^+\|_0 \leq k, \\ 0 &\leq \|h^-\|_0 \leq k - 1. \end{aligned} \quad (28)$$

Thus, we get

$$\begin{aligned} 0 &= \|Ah\|_2^2 \\ &= \|Ah^+ - Ah^-\|_2^2 \\ &= \|Ah^+\|_2^2 + \|Ah^-\|_2^2 - 2\langle Ah^+, Ah^- \rangle \\ &\geq (1 - \delta_k^+) \|h^+\|_2^2 + (1 - \delta_{k-1}^+) \|h^-\|_2^2 \\ &\quad - \theta_{k,k-1}^+ (\|h^+\|_2^2 + \|h^-\|_2^2) \\ &\geq (1 - \delta_k^+ - \theta_{k,k-1}^+) \|h^+\|_2^2 \\ &\quad + (1 - \delta_{k-1}^+ - \theta_{k,k-1}^+) \|h^-\|_2^2 \\ &> 0, \end{aligned} \quad (29)$$

in which the first inequality is due to (24), (25), and the fact that $2ab \leq a^2 + b^2$, and the last inequality is because of the assumption of $\delta_k^+ + \theta_{k,k-1}^+ < 1$ and the monotonicity of δ_k^+ in k . This is a contradiction. Therefore, $x^* \in \text{sol}(P_0)$. \square

With the special result that any two solutions of (P_0) have different support sets, we next derive a sufficient condition on the uniqueness of solution to (P_0) .

Theorem 10. Suppose that $k \geq 1$ is such that $\delta_k^+ + \theta_{k,k}^+ < 1$ and $x^* \in \mathcal{S}$ with $\|x^*\|_0 = k$. Then, x^* is the unique solution of (P_0) .

Proof. Since $\delta_k^+ + \theta_{k,k}^+ < 1$ implies $\delta_k^+ + \theta_{k,k-1}^+ < 1$, we know that $x^* \in \text{sol}(P_0)$. Now we just need to verify that x^* is the unique solution of (P_0) . Assume that this is not true; that is, there is another solution $\bar{x} \neq x^*$. According to Lemma 2, it must hold $I(\bar{x}) \neq I(x^*)$. Take $h = x^* - \bar{x}$. By the argument similar to that in the proof of Theorem 9, we get

$$\begin{aligned} 2 &\leq \|h\|_0 \leq \|x^*\|_0 + \|\bar{x}\|_0 \leq 2k, \\ 1 &\leq \|h^+\|_0 \leq k, \\ 1 &\leq \|h^-\|_0 \leq k, \end{aligned} \quad (30)$$

and the contradiction. We conclude the proof. \square

Now we are ready to give the main result of this paper, which is called the nonnegative restricted property.

Theorem 11. Assume that $k \geq 1$ is such that

$$\delta_{2k}^+ + (\sqrt{2} + 1)\theta_{k,2k}^+ < 1, \quad (31)$$

and $x^* \in \mathcal{S}$ with $\|x^*\|_0 = k$. Then x^* is exactly the common unique minimizer of (P_0) and (P_1) .

Proof. Since (31) implies $\delta_k^+ + \theta_{k,k}^+ < 1$, $\text{sol}(P_0) = \{x^*\}$ by Theorem 10.

Suppose that \bar{x} is a solution of (P_1) . Take $h = \bar{x} - x^*$. To get $\text{sol}(P_1) = \{x^*\}$, it suffices to verify that $h = 0$. The proof includes three steps, the first two steps are parallel to that in [19], in the third step, we utilize the technique of projecting the null space of A on \mathbb{R}_+^n ; for details, see (42) and the argument around it.

Firstly, we introduce a partition of $\{1, \dots, n\}$. Let T_0 be the support set of x^* , T_1 the index set including the first k large components of \bar{x} in T_0^c , T_2 the index set including the next k large components of \bar{x} in $T_0^c \setminus T_1$, and so on. Thus,

$$\begin{aligned} |T_j| &= k, \quad j = 0, 1, \dots, \left\lfloor \frac{n}{k} \right\rfloor - 1, \\ |T_j| &\leq k, \quad j = \left\lfloor \frac{n}{k} \right\rfloor. \end{aligned} \quad (32)$$

Moreover, for any $j = 0, 1, \dots, \lfloor n/k \rfloor$, we define

$$(h_{T_j})_i = \begin{cases} h_i, & i \in T_j, \\ 0, & \text{else,} \end{cases} \quad (33)$$

which is exactly that for any $j = 1, \dots, \lfloor n/k \rfloor$

$$(h_{T_j})_i = \begin{cases} \bar{x}_i - x_i^*, & i \in T_0, \\ \bar{x}_i, & i \in T_j, \\ 0, & \text{else.} \end{cases} \quad (34)$$

Therefore, $Ah = 0, h = \sum_{j=0}^{\lfloor n/k \rfloor} h_{T_j}$, and

$$\begin{aligned} h_{T_j} &\geq 0, \quad j = 1, \dots, \left\lfloor \frac{n}{k} \right\rfloor, \\ \|h_{T_j}\|_0 &\leq k, \quad j = 0, 1, \dots, \left\lfloor \frac{n}{k} \right\rfloor. \end{aligned} \quad (35)$$

Next, we show that $\|h_{(T_0 \cup T_1)^c}\|_2$ is bounded by $\|h_{T_0}\|_1$. Note that for each $j \geq 2$,

$$\|h_{T_j}\|_2 \leq \sqrt{k} \|h_{T_j}\|_\infty \leq \frac{1}{\sqrt{k}} \|h_{T_{j-1}}\|_1, \quad (36)$$

where the second inequality is because of the monotonicity of h_i on T_0^c , and

$$\sum_{j=2}^{\lfloor n/k \rfloor} \|h_{T_j}\|_2 \leq \frac{1}{\sqrt{k}} \sum_{j=1}^{\lfloor n/k \rfloor - 1} \|h_{T_j}\|_1 \leq \frac{1}{\sqrt{k}} \|h_{T_0^c}\|_1. \quad (37)$$

This gives

$$\|h_{(T_0 \cup T_1)^c}\|_2 = \left\| \sum_{j=2}^{\lfloor n/k \rfloor} h_{T_j} \right\|_2 \leq \sum_{j=2}^{\lfloor n/k \rfloor} \|h_{T_j}\|_2 \leq \frac{1}{\sqrt{k}} \|h_{T_0^c}\|_1. \quad (38)$$

In fact,

$$\begin{aligned} \|x_{T_0}^*\|_1 &= \|x^*\|_1 \geq \|\bar{x}\|_1 = \|x^* + h\|_1 \\ &= \|x_{T_0}^* + h_{T_0}\|_1 + \|x_{T_0^c}^* + h_{T_0^c}\|_1 \\ &\geq \|x_{T_0}^*\|_1 - \|h_{T_0}\|_1 + \|h_{T_0^c}\|_1 - \|x_{T_0^c}^*\|_1 \\ &= \|x_{T_0}^*\|_1 - \|h_{T_0}\|_1 + \|h_{T_0^c}\|_1, \end{aligned} \quad (39)$$

which implies

$$\|h_{T_0^c}\|_1 \leq \|h_{T_0}\|_1. \quad (40)$$

By applying (38) and (40), we have

$$\|h_{(T_0 \cup T_1)^c}\|_2 \leq \frac{1}{\sqrt{k}} \|h_{T_0^c}\|_1 \leq \frac{1}{\sqrt{k}} \|h_{T_0}\|_1. \quad (41)$$

Finally, we show that $\|h_{T_0 \cup T_1}\|_2 = 0$. By utilizing the projection $h = h^+ - h^-$, where h^+ and h^- are projections of h and $-h$ on \mathbb{R}_+^n , we have $h_{(T_0 \cup T_1)} = h_{(T_0 \cup T_1)}^+ - h_{(T_0 \cup T_1)}^-$. Moreover,

$$(h_{(T_0 \cup T_1)}^+)_i = \begin{cases} (\bar{x}_i - x_i^*)_+, & i \in T_0, \\ \bar{x}_i, & i \in T_1, \\ 0, & \text{else,} \end{cases} \quad (42)$$

$$(h_{(T_0 \cup T_1)}^-)_i = \begin{cases} (\bar{x}_i - x_i^*)_-, & i \in T_0, \\ 0, & i \in T_1, \\ 0, & \text{else.} \end{cases}$$

It is easy to see

$$\begin{aligned} \|h_{(T_0 \cup T_1)}^+\|_0 &\leq 2k, \\ \|h_{(T_0 \cup T_1)}^-\|_0 &\leq k. \end{aligned} \quad (43)$$

From $Ah = 0$, we compute

$$\begin{aligned} \|Ah_{(T_0 \cup T_1)}\|_2^2 &= \langle Ah_{(T_0 \cup T_1)}, Ah_{(T_0 \cup T_1)} \rangle \\ &= - \sum_{j=2}^{\lfloor n/k \rfloor} \langle Ah_{(T_0 \cup T_1)}, Ah_{T_j} \rangle. \end{aligned} \quad (44)$$

On one hand, based on the definition of NROC, (35), and (43), for $j \geq 2$,

$$\begin{aligned} & - \langle Ah_{(T_0 \cup T_1)}, Ah_{T_j} \rangle \\ &= - \langle Ah_{(T_0 \cup T_1)}^+, Ah_{T_j} \rangle + \langle Ah_{(T_0 \cup T_1)}^-, Ah_{T_j} \rangle \\ &\leq | \langle Ah_{(T_0 \cup T_1)}^+, Ah_{T_j} \rangle | + | \langle Ah_{(T_0 \cup T_1)}^-, Ah_{T_j} \rangle | \\ &\leq \theta_{k,2k}^+ \|h_{(T_0 \cup T_1)}^+\|_2 \cdot \|h_{T_j}\|_2 \\ &\quad + \theta_{k,k}^+ \|h_{(T_0 \cup T_1)}^-\|_2 \cdot \|h_{T_j}\|_2 \\ &\leq \theta_{k,2k}^+ \|h_{T_j}\|_2 \cdot (\|h_{(T_0 \cup T_1)}^+\|_2 + \|h_{(T_0 \cup T_1)}^-\|_2) \\ &\leq \sqrt{2} \theta_{k,2k}^+ \|h_{T_j}\|_2 \cdot \|h_{(T_0 \cup T_1)}\|_2, \end{aligned} \quad (45)$$

where the last inequality is by the fact that $(a + b)^2 \leq 2(a^2 + b^2)$. On the other hand, together with the definition of NRIC and (43), one has

$$\begin{aligned}
& \|Ah_{(T_0 \cup T_1)}\|_2^2 \\
&= \left\| Ah_{(T_0 \cup T_1)}^+ - Ah_{(T_0 \cup T_1)}^- \right\|_2^2 \\
&= \left\| Ah_{(T_0 \cup T_1)}^+ \right\|_2^2 + \left\| Ah_{(T_0 \cup T_1)}^- \right\|_2^2 \\
&\quad - 2 \left\langle Ah_{(T_0 \cup T_1)}^+, Ah_{(T_0 \cup T_1)}^- \right\rangle \\
&\geq (1 - \delta_{2k}^+) \left\| h_{(T_0 \cup T_1)}^+ \right\|_2^2 + (1 - \delta_k^+) \left\| h_{(T_0 \cup T_1)}^- \right\|_2^2 \\
&\quad - 2\theta_{k,2k}^+ \left\| h_{(T_0 \cup T_1)}^+ \right\|_2 \cdot \left\| h_{(T_0 \cup T_1)}^- \right\|_2 \quad (46) \\
&\geq (1 - \delta_{2k}^+) \left\| h_{(T_0 \cup T_1)}^+ \right\|_2^2 + (1 - \delta_k^+) \left\| h_{(T_0 \cup T_1)}^- \right\|_2^2 \\
&\quad - \theta_{k,2k}^+ \left(\left\| h_{(T_0 \cup T_1)}^+ \right\|_2^2 + \left\| h_{(T_0 \cup T_1)}^- \right\|_2^2 \right) \\
&= (1 - \delta_{2k}^+ - \theta_{k,2k}^+) \left\| h_{(T_0 \cup T_1)}^+ \right\|_2^2 \\
&\quad + (1 - \delta_k^+ - \theta_{k,2k}^+) \left\| h_{(T_0 \cup T_1)}^- \right\|_2^2 \\
&\geq (1 - \delta_{2k}^+ - \theta_{k,2k}^+) \left\| h_{(T_0 \cup T_1)} \right\|_2^2.
\end{aligned}$$

Therefore, we compute

$$\begin{aligned}
& (1 - \delta_{2k}^+ - \theta_{k,2k}^+) \left\| h_{(T_0 \cup T_1)} \right\|_2^2 \\
&\leq \sqrt{2}\theta_{k,2k}^+ \left\| h_{(T_0 \cup T_1)} \right\|_2 \\
&\quad \cdot \sum_{j=2}^{\lfloor n/k \rfloor} \left\| h_{T_j} \right\|_2 \quad (\text{by (44), (45), and (46)}) \\
&\leq \sqrt{2}\theta_{k,2k}^+ \left\| h_{(T_0 \cup T_1)} \right\|_2 \cdot \frac{1}{\sqrt{k}} \left\| h_{T_0} \right\|_1 \quad (\text{by (37)}) \quad (47) \\
&\leq \sqrt{2}\theta_{k,2k}^+ \left\| h_{(T_0 \cup T_1)} \right\|_2 \cdot \frac{1}{\sqrt{k}} \left\| h_{T_0} \right\|_1 \quad (\text{by (40)}) \\
&\leq \sqrt{2}\theta_{k,2k}^+ \left\| h_{(T_0 \cup T_1)} \right\|_2 \cdot \left\| h_{T_0} \right\|_2 \\
&\leq \sqrt{2}\theta_{k,2k}^+ \left\| h_{(T_0 \cup T_1)} \right\|_2^2,
\end{aligned}$$

where the forth inequality is from the fact that $\|h_{T_0}\|_1^2 \leq k\|h_{T_0}\|_2^2$. Then the assumption $\delta_{2k}^+ + (\sqrt{2} + 1)\theta_{k,2k}^+ < 1$ forces

$$\left\| h_{(T_0 \cup T_1)} \right\|_2^2 = 0. \quad (48)$$

Thus,

$$h_{(T_0 \cup T_1)} = 0. \quad (49)$$

Therefore, we get $h_{(T_0 \cup T_1)^c} = 0$ by (41), hence $\|h\| = 0$. This is exactly what we want. We complete the proof. \square

5. Conclusion

In this paper, we have derived a nonnegative restricted property condition, which ensures the exact recovery of sparse nonnegative image/signal via the linear program relaxation. Since the NRIC and NROC are defined in \mathbb{R}_+^n , there may be more types of measurement matrices satisfying the nonnegative restricted property than that in the case of RIP, regardless of random matrices or deterministic matrices. As a byproduct of the main result, we have investigated the solution property of the sparse nonnegative recovery and shown that any solution of (P_0) must be one of the extreme points of its feasible set. However, it is not clear whether a given extreme point of the feasible set is a solution to (P_0) . This can serve as a target for future work.

Acknowledgments

The work was supported in part by the National Basic Research Program of China (2010CB732501) and the National Natural Science Foundation of China (11171018).

References

- [1] W. Dai, M. A. Sheikh, O. Milenkovic, and R. G. Baraniuk, "Compressive sensing DNA microarrays," *Eurasip Journal on Bioinformatics and Systems Biology*, vol. 2009, Article ID 162824, 12 pages, 2009.
- [2] M. A. Sheikh, S. Sarvotham, O. Milenkovic, and R. G. Baraniuk, "DNA array decoding from nonlinear measurements by belief propagation," in *Proceedings of the IEEE/SP 14th WorkShoP on Statistical Signal Processing (SSP '07)*, pp. 215–219, IEEE, Washington, DC, USA, August 2007.
- [3] 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.
- [4] E. Candès, "Compressive sampling," in *Proceedings of the Congress of Mathematics*, vol. 3, pp. 1433–1452, Madrid, Spain, 2006.
- [5] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [6] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM Journal on Computing*, vol. 24, no. 2, pp. 227–234, 1995.
- [7] D. Ge, X. Jiang, and Y. Ye, "A note on the complexity of l_p minimization," *Mathematical Programming*, vol. 129, no. 2, pp. 285–299, 2011.
- [8] A. M. Bruckstein, M. Elad, and M. Zibulevsky, "On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 4813–4820, 2008.
- [9] D. L. Donoho and J. Tanner, "Sparse nonnegative solution of underdetermined linear equations by linear programming," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 27, pp. 9446–9451, 2005.
- [10] D. L. Donoho and J. Tanner, "Counting the faces of randomly-projected hypercubes and orthants, with applications," *Discrete and Computational Geometry*, vol. 43, no. 3, pp. 522–541, 2010.
- [11] A. Juditsky, F. Killınç Karzan, and A. Nemirovski, "Verifiable conditions of ℓ_1 -recovery for sparse signals with sign

- restrictions,” *Mathematical Programming B*, vol. 127, no. 1, pp. 89–122, 2011.
- [12] M. A. Khajehnejad, A. G. Dimakis, W. Xu, and B. Hassibi, “Sparse recovery of nonnegative signals with minimal expansion,” *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 196–208, 2011.
- [13] Y. Zhang, “A simple proof for the recoverability of l_1 -minimization (II): the nonnegative case,” Tech. Rep. TR05-10, Department of Computational and Applied Mathematical, Rice University, Houston, Tex, USA, 2005.
- [14] M. Wang, W. Xu, and A. Tang, “A unique “nonnegative” solution to an underdetermined system: from vectors to matrices,” *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1007–1016, 2011.
- [15] E. J. Candès and T. Tao, “Decoding by linear programming,” *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.
- [16] S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic decomposition by basis pursuit,” *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.
- [17] D. L. Donoho and X. Huo, “Uncertainty principles and ideal atomic decomposition,” *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2845–2862, 2001.
- [18] G. M. Fung and O. L. Mangasarian, “Equivalence of minimal l_0 - and l_p -norm solutions of linear equalities, inequalities and linear programs for sufficiently small p ,” *Journal of Optimization Theory and Applications*, vol. 151, no. 1, pp. 1–10, 2011.
- [19] E. J. Candès, “The restricted isometry property and its implications for compressed sensing,” *Comptes Rendus Mathématique*, vol. 346, no. 9-10, pp. 589–592, 2008.