

A Necessary and Sufficient Condition for Generalized Demixing

Chun-Yen Kuo, Gang-Xuan Lin, and Chun-Shien Lu

Abstract—Demixing is the problem of identifying multiple structured signals from a superimposed observation. This work analyzes a general framework, based on convex optimization, for solving demixing problems. We present a new solution to determine whether or not a specific convex optimization problem built for generalized demixing is successful. This solution also creates a way to estimate the probability of success by the approximate kinematic formula.

Index Terms—Compressive sensing, conic geometry, convex optimization, ℓ_1 -minimization, sparse signal recovery.

I. INTRODUCTION

ACCORDING to the theory of convex analysis, convex cones have been exploited to express the optimal conditions for a convex optimization problem [6]. In particular, Amelunxen *et al.* [1] presented the necessary and sufficient conditions for the problems of basis pursuit (BP) and demixing to be successful.

Let $x_0 \in \mathbb{R}^n$ be an unknown k -sparse vector with k nonzero entries in a certain domain, let A be an $m \times n$ random matrix whose entries are independent standard normal variables, and let $z = Ax_0 \in \mathbb{R}^m$ be the measurement vector obtained via random transformation by A . In regard to the basis pursuit (BP) problem, which is defined as

$$(\text{BP}) \left\{ \begin{array}{l} \text{minimize } \|x\|_1 \\ \text{subject to } z = Ax, \end{array} \right. \quad (\text{I.1})$$

a convex optimization method was proposed in Chen *et al.* [4] to solve the sparse signal recovery problem in the context of compressive sensing [5] when $m < n$.

To explore whether BP has a unique optimal solution, Amelunxen *et al.* [1] started from the concept of the conic integral.

Definition 1.1. (descent cone): [1] The descent cone $\mathcal{D}(f, x)$ of a proper convex function $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ at the point $x \in \mathbb{R}^n$ is the conical hull of the perturbations that do not increase f near x .

$$\mathcal{D}(f, x) := \bigcup_{\tau > 0} \{y \in \mathbb{R}^n : f(x + \tau y) \leq f(x)\}. \quad (\text{I.2})$$

Manuscript received March 22, 2015; revised June 01, 2015; accepted July 13, 2015. Date of publication July 16, 2015; date of current version July 21, 2015. This work was supported by MOST Grants 102-2221-E-001-002-MY2 and 102-2221-E-001-022-MY2. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Negar Kiyavash.

The authors are with Institute of Information Science, Academia Sinica, Taipei, Taiwan, R.O.C. (e-mail: r00221004@ntu.edu.tw; spybeiman@gmail.com; lcs@iis.sinica.edu.tw).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LSP.2015.2457403

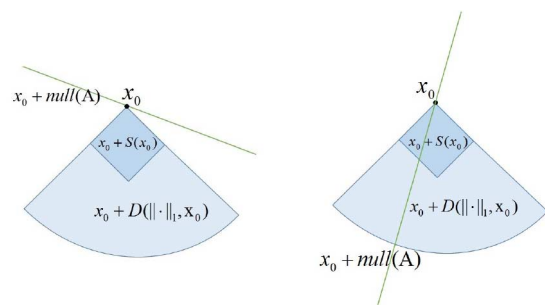


Fig. 1. The optimality condition for the BP problem. [Left] BP succeeds. [Right] BP fails. $S(x_0) = \{y \in \mathbb{R}^d : \|x_0 + y\|_1 \leq \|x_0\|_1\}$.

We say that problem BP, defined in Eq. (I.1), succeeds when it has a unique minimizer \hat{x} that coincides with the true unknown, that is, $\hat{x} = x_0$. To characterize when the BP problem succeeds, Amelunxen *et al.* presented the primal optimality condition as:

$$\text{null}(A) \cap \mathcal{D}(\|\cdot\|_1, x_0) = \{0\} \quad (\text{I.3})$$

in terms of the descent cone [1] (*cf.*, [3] and [7]), where $\text{null}(A)$ denotes null space of A . The optimality condition for the BP problem is also illustrated in Fig. 1.

Amelunxen *et al.* [1] also explored the demixing problem (sparse + sparse) characterized as

$$z = x_0 + Uy_0, \quad (\text{I.4})$$

where $U \in \mathbb{R}^{n \times n}$ is a known orthogonal matrix, x_0 is itself sparse, and y_0 is sparse with respect to U . The optimization problem of recovering signals x_0 and y_0 is formally defined as follows, which we call the demixing problem (DP) in short:

$$(\text{DP}) \left\{ \begin{array}{l} \text{minimize } \|x\|_1 \\ \text{subject to } \|y\|_1 \leq \|y_0\|_1 \text{ and } z = x + Uy. \end{array} \right. \quad (\text{I.5})$$

They propose the primal optimality condition (also illustrated in Fig. 2) as

$$\mathcal{D}(\|\cdot\|_1, x_0) \cap -U\mathcal{D}(\|\cdot\|_1, y_0) = \{0\} \quad (\text{I.6})$$

to characterize whether (x_0, y_0) is the unique minimizer to problem (DP).

The authors in [1] also estimated the probabilities of success of problem (BP) and problem (DP) with Gaussian random sensing matrices by the approximate kinematic formula. They derived the probability¹ using convex (descent) cones. Note

¹Nevertheless, the authors still fail to calculate the actual probabilities. In fact, they only derive the bounds of probabilities that involve the calculation of statistical dimension. Unfortunately, up to now, the statistical dimension still cannot be calculated accurately.

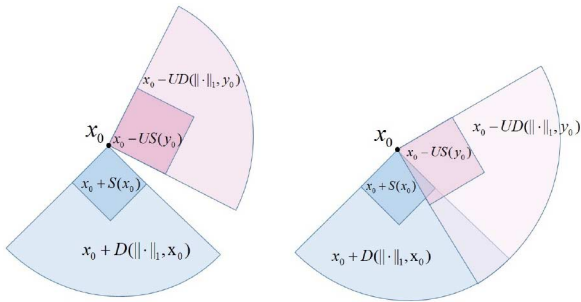


Fig. 2. The optimality condition for (DP) problem. [Left] problem (DP) succeeds. [Right] problem (DP) fails.

that, as shown in Fig. 1 and Fig. 2, the affine ℓ_1 balls $S(\cdot)$ are defined as $S(a) = \{x : \|x + a\|_1 \leq \|a\|_1\}$. Also note that $\text{cone}(S(a)) = \mathcal{D}(\|\cdot\|_1, a)$, where $\text{cone}(S(a))$ is a conical hull of $S(a)$.

In Sections II and III, we generalize the demixing problem specified in Eq. (I.4), set the corresponding optimization problem to recover the signals of such generalized demixing model, and explore its necessary and sufficient condition for successful demixing.

II. MOTIVATION AND PROBLEM DEFINITION

The demixing problem we discuss in this paper refers to the extraction of two informative signals from a single observation. We consider a more general model for a mixed observation $z \in \mathbb{R}^m$, which takes the form

$$z = Ax_0 + By_0, \quad (\text{II.1})$$

where $x_0 \in \mathbb{R}^{n_1}$ and $y_0 \in \mathbb{R}^{n_2}$ are the unknown informative signals that we wish to find and where the matrices $A \in \mathbb{R}^{m \times n_1}$ and $B \in \mathbb{R}^{m \times n_2}$ are arbitrary linear operators (not necessarily $m \leq n_1$ or n_2). We assume that all elements appearing in Eq. (II.1) are known, except for x_0 and y_0 . The broad applications of the general model in Eq. (II.1) can be found in [8] (and the references therein).

It should be noted that: (1) if y_0 in Eq. (II.1) is set to zero, then the generalized demixing model is degenerated to BP; (2) the demixing model in [1] is a special case of Eq. (II.1) if A is set to an identity matrix and B is enforced to be an orthogonal matrix; (3) our generalized demixing model has more freedom in the sense of dimension than that in [2] because A and B can be arbitrarily selected. Moreover, the two components x_0 and y_0 in our generalized model are permitted to have different lengths.

III. MAIN RESULT

The ground truths, x_0 and y_0 , in Eq. (II.1) are approximated by solving the convex optimization problem, which we call the generalized demixing problem (GDP):

$$(\text{GDP}) \begin{cases} \text{minimize } \|x\|_1 \\ \text{subject to } \|y\|_1 \leq \|y_0\|_1 \text{ and } z = Ax + By. \end{cases} \quad (\text{III.1})$$

We say problem (GDP) succeeds provided (x_0, y_0) is the unique optimal solution to GDP. Our goal in this paper is to characterize when the problem (GDP) succeeds.

Theorem III.1: The problem (GDP) has a unique minimizer (\hat{x}, \hat{y}) to coincide with (x_0, y_0) **if and only if**

$$\begin{cases} \text{null}(A) \cap S(x_0) = \{0\}, \\ \text{null}(B) \cap S(y_0) = \{0\}, \\ -AS(x_0) \cap BS(y_0) = \{0\}. \end{cases} \quad (\text{III.2})$$

Proof: First, we assume that the problem (GDP) succeeds in having a unique minimizer (\hat{x}, \hat{y}) to coincide with (x_0, y_0) .

1) Claim: $\text{null}(A) \cap S(x_0) = \{0\}$.

Given $h_1 \in \text{null}(A) \cap S(x_0)$, we have $Ah_1 = 0$. By letting $(x', y') = (x_0 + h_1, y_0)$, it follows that $z = Ax' + By'$ and $\|y'\|_1 \leq \|y_0\|_1$, which means that the point (x', y') is a feasible point of problem (GDP). On the other hand, since $h_1 \in S(x_0)$, we have $\|x_0 + h_1\|_1 = \|x'\|_1 \leq \|x_0\|_1$. By the fact that the problem (GDP) is assumed to have a unique minimizer (x_0, y_0) , we conclude that $h_1 = 0$.

2) Claim: $\text{null}(B) \cap S(y_0) = \{0\}$.

Given $h_2 \in \text{null}(B) \cap S(y_0)$, we have $Bh_2 = 0$ and $\|y_0 + h_2\|_1 \leq \|y_0\|_1$. By letting $(x'', y'') = (x_0, y_0 + h_2)$, it follows that $z = Ax'' + By''$ and $\|y''\|_1 \leq \|y_0\|_1$. Thus, $h_2 = 0$, otherwise $(x'', y'') \neq (x_0, y_0)$ will be another minimizer to problem (GDP).

3) Claim: $-AS(x_0) \cap BS(y_0) = \{0\}$.

Given $s \in -AS(x_0) \cap BS(y_0)$, there exist $\bar{x} \in S(x_0)$ and $\bar{y} \in S(y_0)$ to satisfy $-A\bar{x} = B\bar{y} = s$, $\|x_0 + \bar{x}\|_1 \leq \|x_0\|_1$, and $\|y_0 + \bar{y}\|_1 \leq \|y_0\|_1$. By letting $(x''', y''') = (x_0 + \bar{x}, y_0 + \bar{y})$, it follows that $z = Ax''' + By'''$ and $\|y'''\|_1 \leq \|y_0\|_1$, which means that the point (x''', y''') is a feasible point of problem (GDP). On the other hand, since $\|x'''\|_1 \leq \|x_0\|_1$, (x''', y''') is also an optimal solution. By the fact that the problem (GDP) is assumed to have a unique minimizer (x_0, y_0) , we conclude that $(x''', y''') = (x_0, y_0)$; therefore, $\bar{x} = 0_{\mathbb{R}^{n_1}}$, $\bar{y} = 0_{\mathbb{R}^{n_2}}$, and $s = -A\bar{x} = 0_{\mathbb{R}^m}$.

Conversely, we suppose the point (x_0, y_0) satisfies Eq. (III.2). Letting (x^*, y^*) be a feasible point of problem (GDP), we show that either $\|x^*\|_1 > \|x_0\|_1$ or $(x^*, y^*) = (x_0, y_0)$.

Let $h_1 = x^* - x_0$ and $h_2 = y^* - y_0$. Since (x^*, y^*) is feasible in problem (GDP), $z = Ax^* + By^* = A(x_0 + h_1) + B(y_0 + h_2) = z + Ah_1 + Bh_2$, which implies $-Ah_1 = Bh_2$. If $\|x^*\|_1 > \|x_0\|_1$, then we are done. So, we may assume $\|x^*\|_1 \leq \|x_0\|_1$, which means $h_1 \in S(x_0)$. Moreover, $\|y^*\|_1 \leq \|y_0\|_1$ implies $h_2 \in S(y_0)$. Then, we see that $-Ah_1 = Bh_2 \in -AS(x_0) \cap BS(y_0) = \{0\}$, namely, $h_1 \in \text{null}(A)$ and $h_2 \in \text{null}(B)$. Thus, we have $h_1 \in \text{null}(A) \cap S(x_0) = \{0\}$ and $h_2 \in \text{null}(B) \cap S(y_0) = \{0\}$, which means $(x^*, y^*) = (x_0, y_0)$ and we complete the proof. \square

We know that $S(x_0)$ and $S(y_0)$ are the affine ℓ_1 -balls of the points x_0 and y_0 , respectively. Nevertheless, Eq. (III.2) is the formula, consisting of null spaces of sensing matrices and affine ℓ_1 -balls. Indeed, we can relax the affine ℓ_1 -ball to be its conical hull, such as $\mathcal{D}(\|\cdot\|_1, x_0) = \text{cone}(S(x_0))$ and $\mathcal{D}(\|\cdot\|_1, y_0) = \text{cone}(S(y_0))$, and attain the following result.

Corollary III.1: The problem (GDP) has a unique minimizer (\hat{x}, \hat{y}) that coincides with (x_0, y_0) **if and only if**

$$\begin{cases} \text{null}(A) \cap \mathcal{D}(\|\cdot\|_1, x_0) = \{0\}, \\ \text{null}(B) \cap \mathcal{D}(\|\cdot\|_1, y_0) = \{0\}, \\ -A\mathcal{D}(\|\cdot\|_1, x_0) \cap B\mathcal{D}(\|\cdot\|_1, y_0) = \{0\}. \end{cases} \quad (\text{III.3})$$

We emphasize again that, if x_0 and y_0 have the same length, as in Eq. (I.4), then $\mathcal{D}(\|\cdot\|_1, x_0)$ and $\mathcal{D}(\|\cdot\|_1, y_0)$ will reside

in the same linear space and their intersection can be geometrically visible, as shown in Fig. 2. Nevertheless, since matrices A and B have arbitrary dimensions in our model, their geometrical interaction cannot be observed simply. Thus, we argue that the derivation of a necessary and sufficient condition via combination of all of the cones is significantly different from the standard problems [1], [2].

IV. SIMULATIONS AND VERIFICATIONS

We conduct simulations to verify the consistency between Theorem III.1 and GDP.

A. Verification Procedures

The verification steps for practical sparse signal recovery based on Eq. (III.1) are described as follows.

- 1) Construct the vectors $x_0 \in \mathbb{R}^{n_1}$ and $y_0 \in \mathbb{R}^{n_2}$ with k_1 and k_2 nonzero entries, respectively. The locations of the nonzero entries are selected at random; such a nonzero entry equals ± 1 with equal probability.
- 2) Draw two standard normal matrices $A \in \mathbb{R}^{m \times n_1}$ and $B \in \mathbb{R}^{m \times n_2}$ before capturing the sample $z = Ax_0 + By_0$.
- 3) Solve problem (GDP) to obtain an optimal solution (\hat{x}_1, \hat{y}_1) .
- 4) Declare successful demixing if $\|\hat{x}_1 - x_0\|_2 \leq 10^{-5}$.
In addition, the verification steps for theoretical recovery based on Theorem III.1 are described as follows.
- 5) Solve $\min \|x_0 + x\|_1$ subject to $Ax = 0$ to obtain an optimal point \hat{x}_2 .
- 6) Solve $\min \|y_0 + y\|_1$ subject to $By = 0$ to obtain an optimal point \hat{y}_2 .
- 7) Solve $\min \|x_0 + x\|_1$ subject to $Ax + By = 0$ and $\|y_0 + y\|_1 \leq \|y_0\|_1$ to obtain a pair of optimal points (\hat{x}_3, \hat{y}_3) .
- 8) Declare success in Theorem III.1 if ℓ_2 -norms of $\hat{x}_2, \hat{y}_2, \hat{x}_3,$ and \hat{y}_3 are all smaller than or equal to 10^{-5} .

B. Simulation Setting and Results

In our simulations, n_1 and n_2 were the signal dimensions for signals x_0 and y_0 , respectively. Their sparsities, k_1 and k_2 , ranged from 1 to n_1 and 1 to n_2 , respectively.

First, we let $n_1 = n_2 = 100$ and $k_1 = k_2$. Under this circumstance, the simulation results for both the demixing problems in Eq. (III.1) and Theorem III.1 are illustrated in Fig. 3, where the x -axis denotes the sparsity k and the y -axis denotes the number m of measurements. We can see that the performance of these two seem to be identical, and we notice that the smaller k is, the easier it is for sparse signal recovery to succeed.

Second, we considered $n_1 \neq n_2$, where $n_1 = 100$ and $n_2 = 160$. Again, k_1 and k_2 ranged from 1 to n_1 and 1 to n_2 , respectively. By additionally considering a varying number of measurements, the visualization of recovery results, unlike Fig. 3, will be multidimensional. So, we chose different numbers of measurements with $10 \leq m \leq 100$ in the simulations to ease observations. The recovery result at each pair of k_1 and k_2 for each measurement rate ($\frac{m}{n}$) was obtained by averaging from 100 trials. In sum, the simulation results reveal that, if each optimal solution in Steps (5)–(7) is zero, then the point, x_0 and y_0 , satisfies Eq. (III.2), and *vice versa*. That is to say, we can check if x_0 and y_0 satisfy Eq. (III.2) by solving these three optimization problems in Steps (5)–(7).

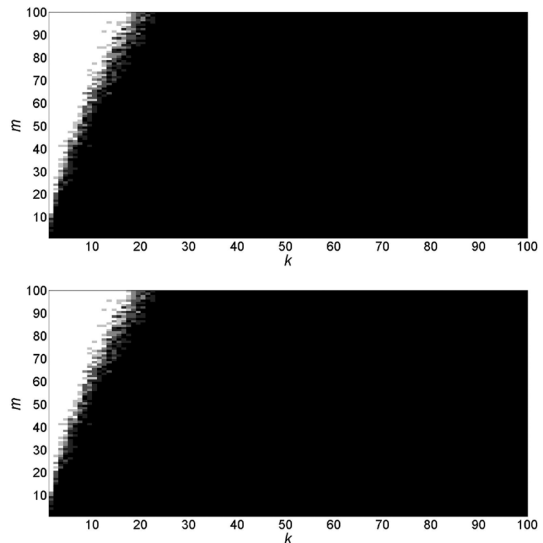


Fig. 3. Phase transitions for demixing problems: [Top] Practical recovery of two sparse vectors based on Eq. (III.1) and [Bottom] theoretic recovery based on Theorem (III.1). In each figure, the heat map indicates the empirical probability of success (Black = 0%; White = 100%).

C. Proof of Feasibility of Our Verification

Now, we show why the above verification is feasible. We say that $\hat{x}_2, \hat{y}_2, \hat{x}_3,$ and \hat{y}_3 obtained from Steps (5)–(7) are all zero vectors if and only if Eq. (III.2) in Theorem III.1 holds. We will validate this claim in the following.

Definition IV.1: Two cones C and D are said to **touch** if they share a ray but are weakly separable by a hyperplane.

Fact 1 [9, pp. 258-260]

Let C and D be closed and convex cones such that both C and $D \neq \{0\}$. Then,

$$\mathbb{P}\{QC \text{ touches } D\} = 0,$$

where Q is a random rotation.

Lemma IV.1: Steps (5)–(8) constitute a complete verification to (III.2) in Theorem III.1.

Proof: We want to prove that Steps (5)–(8) form a valid verification for Eq. (III.2). First, we assume the point (x_0, y_0) satisfies Eq. (III.2).

A1 Claim: \hat{x}_2 in Step (5) is zero.

Since \hat{x}_2 is an optimal solution to the problem in Step (5), we have $\|x_0 + \hat{x}_2\|_1 \leq \|x_0\|_1$, which implies $\hat{x}_2 \in S(x_0)$, and we have $A\hat{x}_2 = 0$, which is followed by $\hat{x}_2 \in \text{null}(A)$. That is $\hat{x}_2 \in \text{null}(A) \cap S(x_0) = \{0\}$; hence, $\hat{x}_2 = 0$.

A2 Claim: \hat{y}_2 in Step (6) is zero.

The proof is similar to the one in A1.

A3 Claim: (\hat{x}_3, \hat{y}_3) in Step (7) is zero.

Since (\hat{x}_3, \hat{y}_3) is an optimal solution to the problem in Step (7), we have $\|y_0 + \hat{y}_3\|_1 \leq \|y_0\|_1$, which means $\hat{y}_3 \in S(y_0)$; $\|x_0 + \hat{x}_3\|_1 \leq \|x_0\|_1$, which says that $\hat{x}_3 \in S(x_0)$; and $A\hat{x}_3 + B\hat{y}_3 = 0$, which implies $-A\hat{x}_3 = B\hat{y}_3 \in -AS(x_0) \cap BS(y_0) = \{0\}$. Thus, $\hat{x}_3 \in \text{null}(A)$ and $\hat{y}_3 \in \text{null}(B)$, then $\hat{x}_3 \in \text{null}(A) \cap S(x_0) = \{0\}$ and $\hat{y}_3 \in \text{null}(B) \cap S(y_0) = \{0\}$, and we come to the conclusion that $(\hat{x}_3, \hat{y}_3) = (0, 0)$.

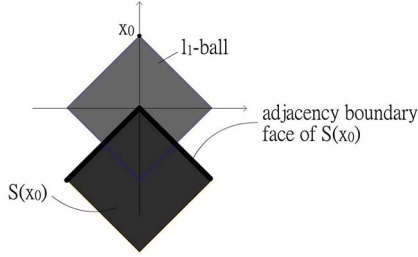


Fig. 4. Adjacency boundary face of $S(x_0)$.

On the other hand, suppose that the optimal solutions \hat{x}_2, \hat{y}_2 , and (\hat{x}_3, \hat{y}_3) corresponding to the minimization problems in Steps (5), (6), and (7), respectively, are all zeros.

B1 Claim: $null(A) \cap S(x_0) = \{0\}$.

Given $x^* \in null(A) \cap S(x_0)$, we have $x^* \in S(x_0)$, meaning that $\|x_0 + x^*\|_1 \leq \|x_0\|_1$. Furthermore, we also have $x^* \in null(A)$, which implies that x^* is a feasible point of the problem in Step (5). Due to the fact that

$$\|x_0\|_1 = \|x_0 + \hat{x}_2\|_1 \leq \|x_0 + x\|_1 \quad \forall x \in null(A),$$

we have $\|x_0 + x^*\|_1 \geq \|x_0\|_1$. Thus, $\|x_0 + x^*\|_1 = \|x_0\|_1$, which means x^* belongs to the adjacency boundary face $\partial_*(S(x_0))$ of $S(x_0)$ at x_0 , where the adjacency boundary face $\partial_*(\cdot) = \partial(\cdot) \cap \partial(\text{cone}(\cdot))$ is the intersection of the boundary of itself and the boundary of its conical hull (as shown in Fig. 4). Therefore, $null(A)$ touches $\mathcal{D}(\|\cdot\|_1, x_0)$ or $null(A) \cap S(x_0) = \{0\}$.

By Fact 1, we may assume that “ $null(A)$ touches $\mathcal{D}(\|\cdot\|_1, x_0)$ ” never happens. So, we conclude that “ $null(A) \cap S(x_0) = \{0\}$ ”.

B2 Claim: $null(B) \cap S(y_0) = \{0\}$. The proof is similar to the one in B1.

B3 Claim: $-AS(x_0) \cap BS(y_0) = \{0\}$.

Given $s^* \in -AS(x_0) \cap BS(y_0)$, there exist $x^* \in S(x_0)$ and $y^* \in S(y_0)$ such that $s^* = -Ax^* = By^*$. Since $y^* \in S(y_0)$, we have $\|y_0 + y^*\|_1 \leq \|y_0\|_1$, together with the fact that $Ax^* + By^* = 0$, the point (x^*, y^*) is a feasible point of the problem in Step (7).

Since $(\hat{x}_3, \hat{y}_3) = (0, 0)$ is an optimal solution to the problem in Step (7), we have

$$\|x_0\|_1 = \|x_0 + \hat{x}_3\|_1 \leq \|x_0 + x^*\|_1.$$

Moreover, $x^* \in S(x_0)$ means $\|x_0 + x^*\|_1 \leq \|x_0\|_1$. Thus, $\|x_0 + x^*\|_1 = \|x_0\|_1$, i.e., $x^* \in \partial_*(S(x_0))$ and $s^* \in \partial_*(-AS(x_0))$. Therefore,

$$-AS(x_0) \cap BS(y_0) \subseteq \partial_*(-AS(x_0)),$$

which means $-AD(\|\cdot\|_1, x_0)$ touches $BD(\|\cdot\|_1, y_0)$ or $-AS(x_0) \cap BS(y_0) = \{0\}$. Due to Fact 1, we may assume that “ $-AD(\|\cdot\|_1, x_0)$ touches $BD(\|\cdot\|_1, y_0)$ ” never happens. So, we conclude that “ $-AS(x_0) \cap BS(y_0) = \{0\}$ ”. \square

V. FUTURE WORK

We plan to employ Corollary III.1 to estimate the probability of success under some assumptions by the approximate kinematic formula from [1].

Theorem V.1 (Approximate kinematic formula): Fix a tolerance $\eta \in (0, 1)$. Let C and K be convex cones in \mathbb{R}^n , and draw a random orthogonal basis $Q \in \mathbb{R}^{d \times d}$. Then,

$$\delta(C) + \delta(K) \leq d - a_\eta \sqrt{d} \Rightarrow \mathbb{P}\{C \cap QK \neq \{0\}\} \leq \eta$$

$$\delta(C) + \delta(K) \geq d + a_\eta \sqrt{d} \Rightarrow \mathbb{P}\{C \cap QK \neq \{0\}\} \geq 1 - \eta,$$

where $a_\eta := \sqrt{8 \log(4/\eta)}$ and δ means the statistical dimension.

Definition V.1 (Statistical dimension): Let $C \subseteq \mathbb{R}^d$ be a closed convex cone. Define the Euclidean projection $\Pi_C : \mathbb{R}^d \rightarrow C$ onto C by

$$\Pi_C(x) := \arg \min_{y \in C} \|y - x\|^2.$$

The statistical dimension $\delta(C)$ of C is defined as:

$$\delta(C) := \mathbb{E}_g[\|\Pi_C(g)\|^2],$$

where $g \sim N(0, I)$ is a standard Gaussian vector.

For the generalized demixing model proposed in this paper, we suppose $A \in \mathbb{R}^{m \times n_1}$ and $B \in \mathbb{R}^{m \times n_2}$ have independent standard normal entries and we let $z = Ax_0 + By_0$. For the compressive sensing demixing, we may assume $m < n_1, m < n_2$, and both A and B have full rank. Then, we can derive:

$$\begin{cases} m \geq \delta(\mathcal{D}(\|\cdot\|_1, x_0)) + a_{\eta_1} \sqrt{n_1}, \\ m \geq \delta(\mathcal{D}(\|\cdot\|_1, y_0)) + a_{\eta_2} \sqrt{n_2}, \\ m \geq \delta(AD(\|\cdot\|_1, x_0)) + \delta(BD(\|\cdot\|_1, y_0)) + a_{\eta_3} \sqrt{m}, \end{cases} \quad (\text{V.1})$$

which implies

$$\begin{cases} \mathbb{P}\{null(A) \cap \mathcal{D}(\|\cdot\|_1, x_0) = \{0\}\} \geq 1 - \eta_1, \\ \mathbb{P}\{null(B) \cap \mathcal{D}(\|\cdot\|_1, y_0) = \{0\}\} \geq 1 - \eta_2, \\ \mathbb{P}\{-AD(\|\cdot\|_1, x_0) \cap BD(\|\cdot\|_1, y_0) = \{0\}\} \geq 1 - \eta_3. \end{cases}$$

On the other hand, we also have

$$\begin{cases} m \leq \delta(\mathcal{D}(\|\cdot\|_1, x_0)) - a_{\eta_1} \sqrt{n_1}, \\ m \leq \delta(\mathcal{D}(\|\cdot\|_1, y_0)) - a_{\eta_2} \sqrt{n_2}, \\ m \leq \delta(AD(\|\cdot\|_1, x_0)) + \delta(BD(\|\cdot\|_1, y_0)) - a_{\eta_3} \sqrt{m}, \end{cases} \quad (\text{V.2})$$

which implies

$$\begin{cases} \mathbb{P}\{null(A) \cap \mathcal{D}(\|\cdot\|_1, x_0) = \{0\}\} \leq \eta_1, \\ \mathbb{P}\{null(B) \cap \mathcal{D}(\|\cdot\|_1, y_0) = \{0\}\} \leq \eta_2, \\ \mathbb{P}\{-AD(\|\cdot\|_1, x_0) \cap BD(\|\cdot\|_1, y_0) = \{0\}\} \leq \eta_3. \end{cases}$$

Apparently, if the number m of measurements is large enough, then successful sparse recovery can be achieved. On the other hand, failed recovery is possible due to an insufficient number of measurements. Nevertheless, if we want to realize the above derived results, computation of the statistical dimensions of $AD(\|\cdot\|_1, x_0)$ and $BD(\|\cdot\|_1, y_0)$, as indicated in Eqs. (V.1) and (V.2), will be an unavoidable difficulty.

VI. CONCLUSION

Our major contribution in this paper is to derive the necessary and sufficient condition for a successful generalized demixing problem. There is an issue worth mentioning, i.e., Amelunxen *et al.* have evaluated an upper bound and a lower bound of the probability of successful recovery for the demixing problem (DP). The reason we did not do that is due to the known unavoidable difficulty raised by the generalized model (GDP problem), that is, “**How to compute the statistical dimension of a descent cone operated by a linear operator?**” We believe that, if this open problem can be solved, we will complete the generalized demixing problem with Gaussian random measurements.

REFERENCES

- [1] D. Amelunxen, M. Lotz, M. B. McCoy, and J. A. Tropp, "Living on the edge: A geometric theory of phase transitions in convex optimization," *arXiv preprint arXiv:1303.6672*, 2013.
- [2] M. B. McCoy and J. A. Tropp, "The achievable performance of convex demixing," *arXiv preprint arXiv:1309.7478*, 2013.
- [3] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Found. Comput. Math.*, vol. 12, no. 6, pp. 805–849, 2012.
- [4] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001.
- [5] D. L. Donoho, "Compressed sensing," *IEEE Trans. Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [6] J.-B. Hiriart-Urruty and C. Lemarchal, "Convex analysis and minimization algorithms," in *I, volume 305 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Berlin, Germany: Springer-Verlag, 1993, Fundamentals.
- [7] M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," *Commun. Pure Appl. Math.*, vol. 61, no. 8, pp. 1025–1045, 2008.
- [8] C. Studer and R. G. Baraniuk, "Stable restoration and separation of approximately sparse signals," *Appl. Comput. Harmon. Anal.*, vol. 37, pp. 12–35, 2014.
- [9] R. Schneider and W. Weil, *Stochastic and Integral Geometry*, ser. Series in Statistics: Probability and its Applications. Springer. Berlin, Germany: Springer, 2008.
- [10] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Reprinted from SIAM," *J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.