S-semigoodness for Low-Rank Semidefinite Matrix Recovery

Lingchen Kong*, Jie Sun† and Naihua Xiu‡

January 14, 2013

Abstract

We extend and characterize the concept of s-semigoodness for a sensing matrix in sparse nonnegative recovery (proposed by Juditsky , Karzan and Nemirovski [Math Program, 2011]) to the linear transformations in low-rank semidefinite matrix recovery. We show that s-semigoodness is not only a necessary and sufficient condition for exact s-rank semidefinite matrix recovery by a semidefinite program, but also provides a stable recovery under some conditions. We also show that both s-semigoodness and semiNSP are equivalent.

Keywords: low-rank semidefinite matrix recovery, unitary property, necessary and sufficient condition, s-semigoodness, exact and stable recovery

1 Introduction

This paper deals with the *low-rank (positive) semidefinite matrix recovery*, which is the problem of recovering a low-rank matrix from a certain number of linear measurements when the the matrix is known to be low-rank and positive semidefinite. Mathematically, it is a *rank minimization problem with the (positive) semidefinite matrix constraints* as follows:

$$\min \ \operatorname{rank}(X), \quad \text{s.t. } \ \mathcal{A}X = b, \ X \succeq 0 \tag{1}$$

where $X \in S^n$ is the matrix variable, $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ is a linear transformation (measurement ensemble) and $b \in \mathbb{R}^m$ (known measurement), $X \succeq 0$ says that X is positive semidefinite, i.e., $X \in \mathbb{S}^n_+$. Here $\mathbb{S}^n$ denotes the space of all $n \times n$ real symmetric matrices with the inner product $\langle X, Y \rangle := \operatorname{Trace}(XY)$ and $\mathbb{S}^n_+$ is the set of all positive semidefinite symmetric matrices. When the variable matrix is diagonal, the low-rank semidefinite matrix recovery reduces the sparse nonnegative recovery (SNR), which is of great practical interest since nonnegative signals are naturally used in image processing, DNA microarrays, network monitoring, hidden Markov

* Lingchen Kong, Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044, P. R. China (e-mail: konglchen@126.com)

† Jie Sun, Department of Decision Sciences, School of Business and Risk Management Institute, National University of Singapore, Republic of Singapore (e-mail: jsun@nus.edu.sg)

‡ Naihua Xiu, Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044, P. R. China (e-mail: nhxiu@bjtu.edu.cn)

models, and so on. The model of SNR is a generalization of the well-known l_0 *norm minimization problem with nonnegative constraints*:

$$\min \|x\|_0 \ \text{ s.t. } \ Ax = b, \ x \geq 0,$$

where $\|x\|_0$ is l_0 *norm* of the vector $x \in \mathbb{R}^n$, i.e., the number of nonzero entries in x (this is not a true norm, as $\|\cdot\|_0$ is not positive homogeneous), $A \in \mathbb{R}^{m\times n}$. This is just the model of the compressed sensing (CS) with nonnegative constraints, while CS has obtained rapid developments in recent years after the fundamental and pioneering work by Donoho [6], Candés, Romberg and Tao [4, 5]. It is NP-hard in general even without the positive semidefinite constraints [1, 9, 10, 19].

A popular approach for solving problem (1) in the systems and control community is to minimize the trace of a positive semidefinite matrix variable instead of its rank (see, e.g., [17]). This yields the convex relaxation of problem (1), i.e., semidefinite program (SDP) relaxation:

$$\min \ \text{tr}(X) \quad \text{s.t. } \ \mathcal{A}X = b, \ X \succeq 0, \tag{2}$$

where $\text{tr}(X)$ is the trace of matrix X. In the vector case, the above problem becomes the linear program (LP) relaxation of SNR

$$\min \sum_{i=1}^{n} x_i \ \text{ s.t. } \ Ax = b, \ x \geq 0. \tag{3}$$

Thus, the above LP problem is the l_1 norm minimization problem with nonnegative constraints.

In control, statistics, signal and image processing, econometrics, quantum information, and many other fields, many applications can be formulated as problem (1). Below we list some of them.

The Feedback Synthesis Problem Feedback synthesis problem is of importance in control and system theory. Consider a continuous-time linear time-invariant dynamical system

$$\dot{x} = Ax + Bu, \quad y = Cx,$$

with matrices $A \in \mathbb{R}^{n\times n}$, and B, C are matrices of appropriate dimensions. A stabilizing controller of order s for the above system is to specified as

$$\dot{z} = A_s z + B_s y, \quad u = C_s z + D_s y,$$

where $A_s \in \mathbb{R}^{s\times s}$. The question is to check, for a given s, whether such a controller (of fixed order) exists. Based on work by EI Ghaoui and Gahinet [8], Mesbahi and Papavassilopoulos [17] reformulated it as a rank minimization with linear matrix inequality constraint:

$$\min \text{rank}(X) \quad \text{s.t.} \quad \text{M(X)} + \text{Q} \succeq 0, \quad \text{X} \succeq 0,$$

where $M(X) = X - \sum_{i=1}^{k} M_i X M_i^T$ is a symmetry preserving linear map on $\mathbb{S}^n$ with matrices $M_i \in \mathbb{R}^{n\times n}$ for $i = 1, 2, \cdots, k$, Q is a symmetric matrix (of appropriate dimensions).

The Multidimensional Scaling Multidimensional scaling (MDS) is to discover some interesting and important information hidden in multidimensional data, see, e.g., [16]. Suppose a matrix $D = (d_{ij}) \in \mathbb{S}^n$ is given, where d_{ij} is the distance between points i and j. We aim to find n points $\{x_1, \cdots, x_n\}$ in a low-dimensional metric space such that the metric distance between

x_i and x_j matches the distance d_{ij}(or as close as possible in the noise case). Recently, from the monograph [21], we can reformulate MDS as the following rank minimization problem:

$$\min \operatorname{rank}(X) \quad \text{s.t.} \quad \mathcal{A}(\mathrm{X}) = \mathrm{D} \circ \mathrm{D}, \mathrm{X} \succeq 0,$$

where $\mathcal{A}(X) = diag(X)e^T + ediag(X)^T - 2X$ is a linear operator, "$\circ$" is the Hadamard product of matrices. Clearly, the above MDS model belongs to the problem (1).

The Phase Retrieval Problem Phase retrieval is the problem of finding the phase that satisfies a set of constraints for a measured amplitude, which has many important applications in X-ray crystallography, transmission electron microscopy and coherent diffractive imaging, etc. Recently, Candès et al [2, 3] formulated the phase retrieval problem as a problem (1). Let $x \in \mathcal{C}^n$ be a discrete signal and the given observation b_i be the squared modulus of the inner product of the signal x and some known vectors z_i, i.e., $b_i = |\langle z_i, x\rangle|^2, i = 1, 2, \cdots, m$. In other words, we can know the magnitude of $\langle z_i, x\rangle$and the phase information is lost. However, we want to record both phase and magnitude information of x from the known observation. Letting $\mathcal{A} : \mathbb{H}^{n\times n} \to \mathbb{R}^m$ be the linear transformation, where $\mathbb{H}^{n\times n}$ denotes the space of Hermitian matrices. By rewriting the data collection $|\langle z_i, x\rangle|^2 = b_i, i = 1, 2, \cdots, m$ as $\mathcal{A}(xx^*) = b$. Then, letting $X = xx^*$, we can represent the phase retrieval problem as a problem (1) (see [2, 3] for more details):

$$\min \operatorname{rank}(X) \quad \text{s.t.} \quad \mathcal{A}(\mathrm{X}) = \mathrm{b}, \quad \mathrm{X} \succeq 0.$$

Recently, there has been some increasing effort and activities on the problem (1), see, e.g., [2, 3, 11, 20]. Candès et al [2, 3] proposed the so-called PhaseLift method to solve problem (1) via SDP relaxation since it lifts up the problem of vector recovery from quadratic constraints into that of recovering a rank-one matrix from affine constraints. Wang, Xu and Tang [20] gave a necessary and sufficient condition under which the feasible set of problem (1) is a singleton and it is just its solution. Gross, Liu, Flammia, Becker, and Eisert [11] successfully used SDP relaxation to solve Quantum State Tomography. However, they did not give the conditions which guarantee the exact low-rank positive semidefinite solution to problem (1) via SDP relaxation. Besides the matrix case, there are much more attention to SNR, see, e.g., [7, 14, 15, 20, 22] to name a few.

The paper deals with recovery conditions for the low-rank semidefinite matrix recovery via SDP relaxation. In Section 2, we give an important lemma by employing the decomposition technique. Based on it, we prove the unitary property of the linear transformation for problem (1). In Section 3, we introduce s-semigoodness and semiNSP for a linear transformations in low-rank semidefinite matrix recovery, and show the equivalence between s-semigoodness and semiNSP for problem (1). This develops the s-semigoodness results coined by Juditsky and Nemirovski [13, 14] from the nonnegative cone to the nonpolytope cone of positive semidefinite symmetric matrices. In Section 4, we establish the exact and stable recovery results for problem (1) via SDP. We end this paper with some remarks.

2 The Unitary Property

We will present a unified technique to establish a bridge from the sparse nonnegative recovery to low-rank semidefinite matrix recovery based on a useful lemma related to eigenvalue inequality for positive semidefinite matrices. We start with defining an *unitary* property of a linear transformation $\mathcal{A}$, while Oymak, Mohan, Fazel and Hassibi [18] named it *extension* property.

For a linear transformation $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ and a unitary U, $\mathcal{A}_U$ is called the *restriction* of $\mathcal{A}$ to unitary U if we define

$$\mathcal{A}_U(x) := \mathcal{A}(UDiag(x)U^T)$$

for all $x \in \mathbb{R}^n$. In particular, $\mathcal{A}_U$ can be represented by a matrix $A_U \in \mathbb{R}^{m\times n}$.

Definition 2.1 *Let* $\boldsymbol{P}$ *be a property defined for matrices* $A \in \mathbb{R}^{m\times n}$*. We say that a linear transformation* $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ *satisfies the* unitary property, $\boldsymbol{P}^{\mathbf{u}}$, *if its restriction* A_U *has property* $\boldsymbol{P}$ *for any unitary* U*.*

For simplicity, let $\|\cdot\|_v$ be an arbitrary norm on $\mathbb{R}^n$ and let $\|\cdot\|_m$ be the corresponding unitarily invariant matrix norm such that $\|X\|_m = \|\lambda(X)\|_v$ where $\lambda(X)$ is the vector of the eigenvalues of X. For convenient statements regarding recovery of sparse nonnegative vectors, we define the notations

- v_1: The matrix $A : \mathbb{R}^n \to \mathbb{R}^m$ satisfies a property $\mathbf{P}$.
- v_2: In problem (3), for any w, $b = Aw + \varsigma$ with $\|\varsigma\| \leq \varepsilon$ and any x as good as w with respect to b (i.e., $\|Ax - b\| \leq \varepsilon$ and $\|x\|_1 \leq \|w\|_1$), we have

  $$\|x - w\|_v \leq h(w, \varepsilon)$$

  for some real-valued function $h : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$.
- v_3: For any nonzero x with $Ax = 0$, x satisfies a property $\mathbf{Q}$.

Similarly, for statements regarding recovery of low-rank positive semidefinite matrices, we define the corresponding notations in the matrix setting

- M_1: The linear transformation $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ satisfies the extension property $\mathbf{P}^{\mathbf{u}}$.
- M_2: In problem (2), for any W, $b = \mathcal{A}W + \varsigma$ with $\|\varsigma\| \leq \varepsilon$ and any X as good as W with respect to b (i.e., $\|\mathcal{A}X - b\| \leq \varepsilon$ and $\|X\|_* \leq \|W\|_*$), we have

  $$\|X - W\|_m \leq h(\lambda(W), \varepsilon).$$
- M_3: For any nonzero X with $\mathcal{A}X = 0$, $\lambda(X)$ satisfies a property $\mathbf{Q}$.

In order to study the unitary property of $\mathcal{A}$, we need the following very useful lemma related to eigenvalue inequality for positive semidefinite matrices. This lemma is different from the useful Lemma 2 of [18] in low-rank matrix recovery and the key singular value inequality in [12]. Our proof is mainly based on the spectral decomposition of a symmetric matrix, space orthogonal decomposition of the set of all symmetric matrices, and the fact $\|X\|_* = \langle I, X\rangle$ for $X \succeq 0$.

Lemma 2.2 *For a given matrix* $Z \in \mathbb{S}^n$ *with spectral decomposition* $Z = UDiag(\lambda(Z))U^T$*, suppose that there is a positive semidefinite matrix* $W \succeq 0$ *satisfying* $W+Z \succeq 0$ *and* $\|W+Z\|_* \leq \|W\|_*$*. Then for* $X = UDiag(d)U^T$ *with* $d = \lambda(W)$*, one has* $X + Z \succeq 0$ *and* $\|X + Z\|_* \leq \|X\|_*$*.*

Proof From the assumption that $Z = UDiag(\lambda(Z))U^T$. Without loss of generality, letting $U = [q_1, q_2, \cdots, q_n]$ be the corresponding orthogonal matrix, we then have $Z = \sum_{i=1}^n \lambda_i(Z) q_i q_i^T$. Fix the rank-one matrices $\{q_1 q_1^T, q_2 q_2^T, \cdots, q_n q_n^T\}$. Let

$$\Omega := span\{q_1 q_1^T, q_2 q_2^T, \cdots, q_n q_n^T\}.$$

Clearly, $\Omega = \{X : X = UDiag(d)U^T, d \in \mathbb{R}^n\}$ is a subspace in $\mathbb{S}^n$ and $I \in \Omega$. It holds from space decomposition theorem, $\mathbb{S}^n = \Omega \bigoplus \Omega^\perp$ with $\Omega^\perp$ is the orthogonal subspace of Ω. Thus, for W, we have

$$W = \sum_{i=1}^n w_i q_i q_i^T + M$$

with $M \in \Omega^\perp$. Then $\langle q_i q_i^T, M\rangle = 0$ and $\langle I, W\rangle = 0$. Since $W \succeq 0$, it holds by self-duality properties of the the cone of positive semidefinite matrices,

$$\langle q_i q_i^T, W\rangle = w_i \geq 0 \quad \text{for} \quad i \in \{1, 2, \cdots, n\}.$$

Similarly, by $W + Z \succeq 0$, $\langle q_i q_i^T, W + Z\rangle = w_i + \lambda_i(Z) \geq 0$ for $i \in \{1, 2, \cdots, n\}$. Take

$$X = \sum_{i=1}^n w_i q_i q_i^T = UDiag(d)U^T,$$

where $d = (w_1, w_2, \cdots, w_n)^T$ and $d \geq 0$. It is easy to see that $X \succeq 0$ and

$$\|X\|_* = \sum_{i=1}^n w_i = \langle X, I\rangle = \langle X + M, I\rangle = \|W\|_*.$$

Moreover, from $X + Z = \sum_{i=1}^n (w_i + \lambda_i(Z)) q_i q_i^T$, we immediately obtain that $X + Z \succeq 0$ and

$$\|X + Z\|_* = \|W + Z\|_*.$$

Therefore, the conclusion $\|X + Z\|_* \leq \|X\|_*$ follows directly, and we complete the proof. □

We are ready to state our main theorem of the unitary property of $\mathcal{A}$.

Theorem 2.3 *For a given property* $\boldsymbol{P}$*, the following implications hold:*

$$(v_1 \Rightarrow v_2) \quad \Rightarrow \quad (M_1 \Rightarrow M_2), \tag{4}$$

$$(v_1 \Rightarrow v_3) \quad \Rightarrow \quad (M_1 \Rightarrow M_3). \tag{5}$$

Proof For a given property $\mathbf{P}$, we first prove $(v_1 \Rightarrow v_2) \Rightarrow (M_1 \Rightarrow M_2)$. Let $v_1 \Rightarrow v_2$ and that M_1 holds. For problem (2), we set the measurements $b_0 = \mathcal{A}W + \varsigma_0$ with $\|\varsigma_0\| \leq \varepsilon$. Below we show that M_2 holds. That is, we need show that for any positive semidefinite matrix X which is as good as W with respect to b_0, it holds

$$\|X - W\|_m \leq h(\lambda(W), \varepsilon).$$

Consider any such X and let $Z = X - W$. This implies that $W + Z \succeq 0$, $\|W + Z\|_* \leq \|W\|_*$ and

$$\|\mathcal{A}_1(W + Z) - b_0\| \leq \varepsilon,$$

since $\|\mathcal{A}_1 X - b_0\| \leq \varepsilon$ and $\|X\|_* \leq \|W\|_*$. Let Z have the spectral decomposition $Z = UDiag(\lambda(Z))U^T$. Then, from Lemma 4.1, there exists a positive semidefinite matrix $X = UDiag(d)U^T$ with $d \geq 0$ such that $W + Z \succeq 0$, $\|X + Z\|_* \leq \|X\|_*$. Taking $b_1 = \mathcal{A}X + \varsigma_0$, we easily obtain

$$\|\mathcal{A}(X + Z) - b_1\| = \|\mathcal{A}(W + Z) - b_0\| \leq \varepsilon.$$

Therefore, $X + Z$ is as good as X with respect to b_1. Below we deal with problem (3) where A_U is the given measurement matrix, b_1 is the measurement, $x = d$ is unknown vector and $z = \lambda(Z)$ is the perturbation. Thus, we obtain that $x + z$ is as good as x with respect to b_1 since $X + Z$ is as good as X with respect to b_1. From the fact that $\mathcal{A}$ has $\mathbf{P^u}$ implies that A_U has $\mathbf{P}$, we obtain that v_1 holds for A_U. Thus, noting $v_1 \Rightarrow v_2$, we claim

$$\|X - W\|_m = \|Z\|_m = \|z\|_v \leq h(d, \varepsilon) = h(\lambda(W), \varepsilon).$$

Therefore, it holds that $M_1 \Rightarrow M_2$.

We now prove $(v_1 \Rightarrow v_3) \Rightarrow (M_1 \Rightarrow M_3)$. Similarly, let $v_1 \Rightarrow v_3$ and that M_1 holds. For any nonzero X with spectral decomposition $X = UDiag(\lambda(X))U^T$ such that $\mathcal{A}X = 0$. Since A_U satisfies $\mathbf{P}$ and hence v_1 holds. From $v_1 \Rightarrow v_3$, we immediately obtain $\lambda(X)$ satisfies the property $\mathbf{Q}$. Then we complete the proof. □

3 S-semigoodness

We will introduce several basic notions related to recovery conditions for problem (1) via SDP. Then we give the characterization and intimate links between them. We begin with the definitions of s-semigoodness and semiNSP, which are generalizations of NSP and s-goodness from SNR to problem (1). As we mentioned in Introduction, in view of the optimality condition of (3), Juditsky, Karzan and Nemirovski [14] gave a necessary and sufficient condition for SNR, s-semigoodness. Here we extend it to the matrix setting.

Definition 3.1 *(s-**semigoodness**) We say that $\mathcal{A}$ is s-semigood, if for every matrix $W \in \mathbb{S}^n$ with $W \succeq 0$ being a s-rank matrix, W is the unique solution to the following semidefinite program*

$$\min_{X \in \mathbb{S}^n} \{\langle I, X\rangle : \mathcal{A}X = \mathcal{A}W, \quad X \succeq 0\}. \tag{6}$$

In order to characterize the above s-semigoodness of a linear transformation $\mathcal{A}$, we introduce the following useful *s-semigoodness conditions* related to some parameters which will be useful to establish the stable recovery result for problem (1) via SDP.

Definition 3.2 *Let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ be a linear transformation and the sparsity s be given.*

i) $\mathcal{A}$ satisfies the condition $SG_s(\zeta, \theta)$ with parameters ζ and θ*: for every index set $J \subseteq \{1, 2, \cdots, n\}$ with $|J| \leq s$ and every nonzero matrix $W \in \mathbb{S}^n$ with the spectral decomposition $W = UDiag(\lambda(W))U^T$ and $\lambda(W) \in \mathbb{R}^n$ such that $\mathcal{A}W = 0$, one has*

$$\sum_{i \in J} \lambda_i(W) \leq \zeta \left(\sum_{i \in \bar{J}} \psi(\lambda_i(W)) \right), \quad \psi(t) = \max[-t, \theta t],$$

or, equivalently: for all $W \in \mathbb{S}^n$ *with the spectral decomposition* $W = UDiag(\lambda(W))U^T$ *and* $\lambda(W) \in \mathbb{R}^n$ *such that* $\mathcal{A}W = 0$, $\Theta(\lambda(W)) \leq \zeta\Psi(\lambda(W))$ *where*

$$\begin{aligned}\Theta(\lambda(W)) &:= \max_{|J|\leq s}\left[\sum_{i\in J}\max[(1-\zeta)\lambda_i(W),(1+\theta\zeta)\lambda_i(W)]\right],\\ \Psi(\lambda(W)) &:= \sum_{i=1}^{n}\max[-\lambda_i(W),\theta\lambda_i(W)].\end{aligned}$$

ii) $\mathcal{A}$ satisfies the condition $SG_{s,\beta}(\zeta,\theta)$ with parameters ζ, θ and β*: for every index set* $J \subseteq \{1,2,\cdots,n\}$ *with* $|J| \leq s$ *and any* $W \in \mathbb{S}^n$ *with the spectral decomposition* $W = UDiag(\lambda(W))U^T$ *and* $\lambda(W) \in \mathbb{R}^n$*, one has*

$$\sum_{i\in J}\lambda_i(W) \leq \beta\|\mathcal{A}W\| + \zeta\sum_{i\in \bar{J}}\psi(\lambda_i(W)), \quad \psi(t) = \max[-t,\theta t].$$

iii) $\mathcal{A}$ satisfies the condition $SG_{s,\beta}(\zeta)$ with parameters ζ and β*: for every index set* $J \subseteq \{1,2,\cdots,n\}$ *with* $|J| \leq s$ *and any* $W \in \mathbb{S}^n$ *with the spectral decomposition* $W = UDiag(\lambda(W))U^T$ *and* $\lambda(W) \in \mathbb{R}^n$ *with* $\lambda_i(W) \leq 0$ *for* $i \in \bar{J}$*, one has*

$$\sum_{i\in J}\lambda_i(W) \leq \beta\|\mathcal{A}W\| + \zeta\sum_{i\in \bar{J}}|\lambda_i(W)|.$$

We next give the following important equivalent results related to s-semigoodness conditions.

Theorem 3.3 *Let* $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ *be a linear transformation and the sparsity* s *be given. Then the following statements hold equivalently:*

a) $\mathcal{A}$ *is* s*-semigood.*

b) There exist $\zeta \in (0,1)$ *and* $\theta \in [1,\infty)$ *such that* $\mathcal{A}$ *satisfies the condition* $SG_s(\zeta,\theta)$*.*

c) There exist $\zeta \in (0,1)$*,* $\theta \in [1,\infty)$ *and* $\beta \in [0,\infty)$ *such that* $\mathcal{A}$ *satisfies the condition* $SG_{s,\beta}(\zeta,\theta)$*.*

e) There exist $\zeta \in (0,1)$ *and* $\beta \in [0,\infty)$ *such that* $\mathcal{A}$ *satisfies the condition* $SG_{s,\beta}(\zeta)$*.*

Proof Let $A \in \mathbb{R}^{m\times n}$ be a sensing matrix and the sparsity s be given. As in the proof of Theorem 3.3, by specifying Proposition 1 in [14] to the nonnegative vector case of all sign restrictions, we easily obtain the following equivalence for SNR, for details see [14] and the reference therein. That is, the following statements are equivalent:

1) A is s-semigood.

2) There exist $\zeta \in (0,1)$ and $\theta \in [1,\infty)$ such that A satisfies the condition $SG_s(\zeta,\theta)$ as follows: for every index set $J \subseteq \{1,2,\cdots,n\}$ with $|J| \leq s$ and any $w \in \mathbb{R}^n$ such that $Aw = 0$ one has

$$\sum_{i\in J} w_i \leq \zeta\left(\sum_{i\in\bar{J}}\psi(w_i)\right), \quad \psi(t) = \max[-t,\theta t],$$

or, equivalently: for all $w \in \mathbb{R}^n$ such that $Aw = 0$, $\Theta(w) \leq \zeta\Psi(w)$ where

$$\Theta(w) := \max_{|J|\leq s}\left[\sum_{i\in J}\max[(1-\zeta)w_i,(1+\theta\zeta)w_i]\right], \Psi(w) := \sum_{i=1}^{n}\max[-w_i,\theta w_i].$$

3) There exist $\zeta \in (0,1)$, $\theta \in [1,\infty)$ and $\beta \in [0,\infty)$ such that A satisfies the condition $SG_{s,\beta}(\zeta,\theta)$ as follows: for every index set $J \subseteq \{1,2,\cdots,n\}$ with $|J| \leq s$ and any $w \in \mathbb{R}^n$, one has

$$\sum_{i\in J} w_i \leq \beta\|Aw\| + \zeta \sum_{i\in \bar{J}} \psi(w_i), \quad \psi(t) = \max[-t, \theta t].$$

4) There exist $\zeta \in (0,1)$ and $\beta \in [0,\infty)$ such that A satisfies the condition $SG_{s,\beta}(\zeta)$ as follows: for every index set $J \subseteq \{1,2,\cdots,n\}$ with $|J| \leq s$ and any $w \in \mathbb{R}^n$ with $w_i \leq 0$ for $i \in \bar{J}$, one has

$$\sum_{i\in J} w_i \leq \beta\|Aw\| + \zeta \sum_{i\in \bar{J}} |w_i|.$$

Together with the above arguments, the desired conclusion holds by Theorem 2.3 and the definitions of s-semigoodness conditions. □

As one of important applications, we will utilize our main Theorem 2.3 to transfer the connections between s-semigoodness and semiNSP. We state the definition of semiNSP, which is a generalization of null space property and s-goodness from SNR to PSLMR. In the vector case, it reduces to the well-known *(nonnegative) NSP* in [7, 15, 20, 22], which provides a necessary and sufficient condition for exactly recovering sparse nonnegative vectors via linear program (3).

Definition 3.4 *(**semiNSP**) We say that $\mathcal{A}$ satisfies semiNSP of order s, if for any index set $J \subseteq \{1,2,\cdots,n\}$ with $|J| = s$ and every nonzero matrix $W \in \mathbb{S}^n$ with the spectral decomposition $W = U\mathrm{Diag}(w)U^T$ with $w \in \mathbb{R}^n$ such that $\mathcal{A}W = 0$ and $w_i \leq 0$ for $i \in \bar{J}$, then we have*

$$\sum_{i\in J} w_i < \sum_{i\in \bar{J}} |w_i|.$$

We are ready to provide the equivalence between semiNSP and s-semigoodness conditions.

Theorem 3.5 *Let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ be a linear transformation and the sparsity s be given. Then, $\mathcal{A}$ is s-semigood if and only if $\mathcal{A}$ satisfies semiNSP.*

Proof In the vector case, let $A \in \mathbb{R}^{m\times n}$ be a sensing matrix and the sparsity s be given. Proposition 1 in [14] implies that A is s-semigood if and only if A satisfies semiNSP. Then, the desired conclusion holds by Theorem 2.3 and the definitions of semiNSP, s-semigoodness. □

4 Exact and Stable Recovery

In many applications, the problem (1) is not necessary to be solved exactly because of noise measurement and the matrix variable W being not s-rank but rather compressible. In this sense, we need approximate solutions which are robust to noise and can approximately recover compressible matrices. Below, we consider approximate solutions to the problem

$$Opt(b_*) = \min_{X\in S^n} \{\|X\|_* : \|\mathcal{A}X - b\| \leq \varepsilon, X \succeq 0\} \tag{7}$$

where $\varepsilon \geq 0$ and $b = \mathcal{A}X + \varsigma,\ \ \varsigma \in \mathbb{R}^m$ with $\|\varsigma\| \leq \varepsilon$. In what follows, let W^s stands for the *best s-rank approximation* of W, i.e., the matrix obtained from W by replacing all but the s largest in magnitude eigenvalues in W with zeros.

Theorem 4.1 *Let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ be a linear transformation. Let $\varepsilon \geq 0$ and let $W \succeq 0$ and b in (7) be such that $\|\mathcal{A}W - b\| \leq \varepsilon$. Let X be a (ϑ, υ)-optimal solution to the problem (7), meaning that $\|\mathcal{A}X - b\| \leq \vartheta$ and $\|X\|_* \leq Opt(b) + \upsilon$.*

1) If $\mathcal{A}$ satisfies the condition $SG_{s,\beta}(\zeta, \theta)$ with some $\zeta \in (0,1)$, $\theta \in [1, \infty)$ and $\beta \in [0, \infty)$, then

$$\|X - W\|_* \leq \frac{2\beta(\vartheta + \varepsilon) + 2(1 + \zeta\theta)\|W - W^s\|_* + (1 + \zeta)\upsilon}{1 - \zeta}.$$

2) If A satisfies the condition $SG_{s,\beta}(\zeta)$ with some $\zeta \in (0,1)$ and $\beta \in [0, \infty)$, then

$$\|X - W\|_* \leq \frac{2\beta(\vartheta + \varepsilon) + 2(1 + \beta\alpha)\|W - W^s\|_* + (1 + \zeta)\upsilon}{1 - \zeta},$$

where $\alpha := \max_{X \in \mathbb{S}^n} \frac{\|\mathcal{A}X\|}{\|X\|_*}$.

Proof As in the proof of Proposition 3.5 , for the nonnegative vector case, we have the following stable result from Proposition 2 in [14]. Let $\varepsilon \geq 0$ and let w and b in the problem

$$Opt(b_1) = \min_{x \in \mathbb{R}^n} \{\|x\|_1 : \|Ax - b\| \leq \varepsilon, x \geq 0\} \tag{8}$$

where $\varepsilon \geq 0$ and $b = Aw + \varsigma$, $\varsigma \in \mathbb{R}^m$ with $\|\varsigma\| \leq \varepsilon$. Let w^s be the *best s-sparse approximation* of w, i.e., the vector obtained from w by replacing all but the s largest in magnitude entries in w with zeros. Let x be a (ϑ, υ)-optimal solution to the problem (8), meaning that $\|Ax - b\| \leq \vartheta$ and $\|x\|_1 \leq Opt(b) + \upsilon$.

1) If A satisfies the condition $SG_{s,\beta}(\zeta, \theta)$ with some $\zeta \in (0,1)$, $\theta \in [1, \infty)$ and $\beta \in [0, \infty)$, then

$$\|x - w\|_1 \leq \frac{2\beta(\vartheta + \varepsilon) + 2(1 + \zeta\theta)\|w - w^s\|_1 + (1 + \zeta)\upsilon}{1 - \zeta}.$$

2) If A satisfies the condition $SG_{s,\beta}(\zeta)$ with some $\zeta \in (0,1)$ and $\beta \in [0, \infty)$, then

$$\|x - w\|_1 \leq \frac{2\beta(\vartheta + \varepsilon) + 2(1 + \beta\alpha)\|w - w^s\|_1 + (1 + \zeta)\upsilon}{1 - \zeta},$$

where α stands for the maximum of $\|\cdot\|$-norms of the columns in A.

Note that the above maximum of $\|\cdot\|$-norms of the columns in A is $\max_{x \in \mathbb{R}^n} \frac{\|Ax\|}{\|x\|_1}$. In the matrix case, it generalizes to $\max_{X \in \mathbb{S}^n} \frac{\|\mathcal{A}X\|}{\|X\|_*}$. Then we obtain the desired stable recovery result by applying Theorem 2.3. $\square$

The above result is a generalization of the corresponding one for SNR given by Juditsky, Karzan and Nemirovski [14]. We showed that in the "non-ideal case", when W is "nearly s-rank" and (7) is solved to near-optimality, the error (via nuclear norm) of the problem (1) can be bounded in terms of $SG_{s,\beta}(\zeta, \theta)$ or $SG_{s,\beta}(\zeta)$, measurement error ε, "s-tail" $\|W - W^s\|_*$ and the accuracy (ϑ, υ).

When we set $\vartheta = 0$, $\upsilon = 0$, and let $W \succeq 0$ be a s-rank matrix such that $\mathcal{A}W = b$, from above theorem, we immediately have the exact recovery result for problem (1) via SDP.

Theorem 4.2 *Let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ be a linear transformation. Let W be a s-rank matrix such that $\mathcal{A}W = b$ and $W \succeq 0$. Let X be a solution to the SDP problem (2). Then $X = W$ if and only if $\mathcal{A}$ is s-semigood.*

Proof If $\mathcal{A}$ is s-semigood, it is immediate to show $X = W$ by Theorem 4.1. On the contrary, if $X = W$ is the unique solution to the SDP problem (2), we then obtain that $\mathcal{A}$ is semiNSP by Theorem 2 in [22]. This says that $\mathcal{A}$ is s-semigood from Theorem 3.5. □

5 Conclusion

In this paper, we characterized the s-semigoodness for the linear transformations in low-rank semidefinite matrix recovery. We showed that both s-semigoodness and semiNSP are necessary and sufficient conditions for exact s-rank semidefinite matrix recovery via SDP. Applying the s-semigoodness characteristic, we gave the exact and stable low-rank semidefinite matrix recovery results via SDP.

Acknowledgments The work was supported in part by the National Basic Research Program of China (2010CB732501), the National Natural Science Foundation of China (11171018,10831006).

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, 51(1), pp. 34-81, (2009)

[2] E. J. Candès, Y. Eldar, T. Strohmer and V. Voroninski. Phase retrieval via matrix completion. 2011

[3] E. J. Candès, T. Strohmer and V. Voroninski. PhaseLift: exact and stable signal recovery from magnitude measurements via convex programming. To appear in Communications on Pure and Applied Mathematics.

[4] E. J. Candès, J. Romberg, and T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inform. Theory, 52(2), pp. 489-509 (2006)

[5] E. J. Candès and T. Tao, Decoding by linear programming, IEEE Trans. Inform. Theory, 51(12), pp. 4203-4215 (2005)

[6] D. L. Donoho, Compressed sensing. IEEE Trans. Inform. Theory, 52(4), pp.1289-1306 (2006)

[7] D. L. Donoho, and J. Tanner, Sparse nonnegative solutions of underdetermined linear equations by linear programming *Proc. Natl. ACAD. Sci* 102 9446-9451 (2005)

[8] L. El Ghaoui and P. Gahinet, Rank minimization under LMI constraints: A framework for output feedback problems, in Proc. European Contr. Conf., July (1993)

[9] Y.C. Eldar and G. Kutyniok, Compressed Sensing: Theory and Applications. Cambridge University Press, 2012.

[10] M. Fazel, H. Hindi, and S. Boyd, A rank minimization heuristic with application to minimum order system approximation. In Proceedings American Control Conference, 2001.

[11] D. Gross, Y.-K. Liu, S.T. Flammia, S. Becker, and J. Eisert, Quantum State Tomography via Compressed Sensing, PRL 105, 150401 (2010)

[12] R.A. Horn and C.R. Johnson, Topics in Matrix Analysis, (1990)

[13] A. Juditsky, and A. S. Nemirovski, On verifiable sufficient conditions for sparse signal recovery via ℓ_1 minimization, Math. Program., 127(1), pp. 57-88 (2011)

[14] A. Juditsky, F. Karzan and A. S. Nemirovski, Verifiable conditions of ℓ_1-recovery of sparse signals with sign restrictions, Math. Program., 127(1), pp. 89-122 (2011)

[15] M. A. Khajehnejad, A. G. Dimakis, W. Xu, and B. Hassibi, Sparse recovery of nonnegative signals with minimal expansion *IEEE Trans. Signal Processing* 59(1) 196-208 (2011)

[16] J. de Leeuw and W. Heiser, Theory of multidimensional scaling, in Handbook of Statistics, Vol. 2, P.R. Krishnaiah and L.N. Kanal, eds., Amsterdam: North-Holland, 1982, pp. 285-316.

[17] M. Mesbahi and G. P. Papavassilopoulos, On the rank minimization problem over a positive semidefinite linear matrix inequality, IEEE Transactions on Automatic Control, 42(2), pp. 239-243 (1997)

[18] S. Oymak, K. Mohan, M. Fazel and B. Hassibi, A simplified approach to recovery conditions for low rank matrices, 2011.

[19] B. Recht, M. Fazel, and P. Parrilo, Guaranteed minimum rank solutions of matrix equations via nuclear norm minimization, SIAM Review, 52(3), pp. 471-501 (2010)

[20] M. Wang, W. Xu and A. Tang, A unique “nonnegative” solution to an underdetermined systems: from vector to matrices *IEEE Trans. Signal Processing* 59(3) 1007-1016 (2011)

[21] H. Wolkowicz, R. Saigal, L. Vandenberghe, Handbook of Semidefinite Programming: Theory, Algorithms, and Applications, Kluwer Academic Publishers, Boston, 2000.

[22] Y. Zhang , A simple proof for the recoverability of l_1 -minimization (II): the nonnegative case, *Technical report TR05-10, Department of Computational and Applied Mathematical, Rice University, Houston, TX* (2005)