A NULLSPACE ANALYSIS OF THE NUCLEAR NORM HEURISTIC FOR RANK MINIMIZATION

Krishnamurthy Dvijotham

University of Washington
Department of Computer Science and Engineering

Maryam Fazel[*]

University of Washington
Department of Electrical Engineering

ABSTRACT

The problem of minimizing the rank of a matrix subject to linear equality constraints arises in applications in machine learning, dimensionality reduction, and control theory, and is known to be NP-hard. A popular heuristic minimizes the nuclear norm (sum of the singular values) of the matrix instead of the rank, and was recently shown to give an exact solution in several scenarios. In this paper, we present a new analysis for this heuristic based on a property of the nullspace of the operator defining the constraints, called the *spherical section* property. We give conditions for the exact recovery of all matrices up to a certain rank, and show that these conditions hold with high probability for operators generated from random Gaussian ensembles. Our analysis provides simpler proofs than existing isometry-based methods, as well as robust recovery results when the matrix is not exactly low-rank.

Index Terms— Matrix rank minimization, compressed sensing, convex optimization

1. INTRODUCTION

The affine rank minimization problem consists of finding a matrix of minimum rank that satisfies a given system of linear equality constraints, i.e.,

$$\begin{array}{ll} \text{minimize} & \text{rank}(X) \\ \text{subject to} & \mathcal{A}(X) = \mathcal{A}(X_0), \end{array} \quad (1)$$

where $X \in \mathbf{R}^{m \times n}$ is the optimization variable, $\mathcal{A} : \mathbf{R}^{m \times n} \mapsto \mathbf{R}^p$ is a linear operator, and X_0 is a minimum rank solution. This problem arises in a diverse set of fields, where notions of order or dimension are expressed by means of the rank of a matrix. Applications include dynamical system identification, collaborative filtering in machine learning, and Euclidean embedding and dimensionality reduction (see [1] and references therein). Problem (1) is in general NP-hard. A popular convex heuristic [2] replaces rank with the *nuclear norm* (also known as the Schatten 1-norm or trace norm) of the matrix, denoted by $\|X\|_* = \sum_i^r \sigma_i(X)$ where $\sigma_i(X)$ are the singular values and $r = \text{rank}(X)$. The heuristic solves the convex problem

$$\begin{array}{ll} \text{minimize} & \|X\|_* \\ \text{subject to} & \mathcal{A}(X) = \mathcal{A}(X_0), \end{array} \quad (2)$$

which can be cast as a semidefinite program and solved efficiently. Note that if the matrix X is diagonal, (1) reduces to finding the sparsest vector in the feasible set, and (2) reduces to its well-known ℓ_1 relaxation.

Until recently, despite its practical success, no theoretical guarantees for this heuristic were available. The paper [1] gave the

first theoretical result by extending analysis tools from the area of Compressed Sensing. It was shown that the heuristic succeeds in finding the minimum rank matrix if the linear operator defining the constraints satisfies a certain Restricted Isometry Property (RIP). Several classes of random operators satisfy this property with high probability, if the number of constraints is large enough. Under the RIP, the minimum rank solution is unique, which means it is possible to "recover" an unknown low-rank matrix X_0 given a set of linear constraints (or measurements) on the matrix. The analysis in [1] guarantees recovery of matrices up to rank r from a number of measurements on the order of $r(m + n)\log(mn)$.

Very recent work (that came out after the submission of this paper) by Candes et al [3] shows that for Gaussian ensembles, the number of measurements can scale as $\max(m, n)r$ which is shown to be nearly minimax optimal. In light of this recent work, it seems that the RIP is a powerful tool that can be used to prove optimal bounds. However, it still has certain drawbacks (also pointed out in [4] for the vector case). First, the constraints can equivalently be expressed as $\mathcal{G}\mathcal{A}(X) = \mathcal{G}\mathcal{A}(X_0)$ for any invertible map $\mathcal{G}$ so the recovery condition should be invariant to this operation; whereas the RIP is not. In fact, the ability to recover X_0 should only depend on the *nullspace* of $\mathcal{A}$, denoted by $\text{Null}(\mathcal{A})$. Further, it is not easy to extend the RIP analysis to incorporate additional constraints on the matrix being recovered (for example positive-definiteness). Another related work that also uses a nullspace-based approach is the paper [5] which gives a necessary and sufficient recovery condition. However, this condition requires a much more involved analysis, and more importantly, it has not been shown to have robustness properties with respect to approximately low rank matrices.

In this paper, we present a new analysis for the nuclear norm heuristic based on the *spherical section property* of the nullspace. In section 2.1 we derive deterministic (sufficient) conditions under which exact low rank matrix recovery is possible: if the nullspace of $\mathcal{A}$ satisfies this condition, *any* matrix X_0 up to a certain rank can be recovered. Section 2.2 shows the recovery is robust by giving an *error bound* for the case where X_0 is only approximately low rank, which is a more realistic assumption in practice. The error is only a constant factor times the best possible low-rank approximation error. Section 3 presents a probabilistic analysis for exact recovery. We obtain conditions, as a function of the matrix dimensions and rank and the number of constraints, such that our conditions for success hold with high probability for operators generated from random Gaussian ensembles.

The spherical section parameter Δ, defined below, is (inversely) related to a well-known distortion measure for subspaces of $\mathbf{R}^n$ appearing in dimensionality reduction, approximation theory, and compressed sensing (see [6, 4] and references therein). Our definition

extends that idea to subspaces of matrices. As we show in the next section, this property essentially ensures that the subspace $\mathrm{Null}(\mathcal{A})$ does not include matrices with small rank.

Definition We say that $\mathrm{Null}(\mathcal{A})$ satisfies the Δ spherical section property if $\frac{\|Z\|_*}{\|Z\|_F} \geq \sqrt{\Delta}$ for all $Z \in \mathrm{Null}(\mathcal{A})$, $Z \neq 0$.

2. EXACT RECOVERY UNDER THE SPHERICAL SECTION ASSUMPTION

In this section, we discuss conditions on $\mathrm{Null}(\mathcal{A})$ that ensure success of the nuclear norm heuristic for any matrix up to a certain rank. The following theorem has two parts: the first describes conditions under which there is a unique low-rank matrix satisfying the constraints and the second describes conditions under which this low-rank matrix can be recovered by minimizing the nuclear norm.

Theorem 2.1 *Suppose* $\mathrm{Null}(\mathcal{A})$ *has the* Δ*-spherical section property. Let* $X_0 \in \mathbf{R}^{m \times n}$, $m \leq n$, $X_0 \neq 0$. *Then*

   (a) X_0 *is the unique solution to* (1) *if* $\mathrm{rank}(X_0) < \frac{\Delta}{2}$.

   (b) *If* $\mathrm{rank}(X_0) < \min\left(\frac{3m}{4} - \sqrt{\frac{9m^2}{16} - \frac{m\Delta}{4}}, m/2\right)$, X_0 *is the unique solution to* (2). *A simpler sufficient condition for the above is* $\mathrm{rank}(X_0) < \frac{\Delta}{6}$.

Proof (a) Suppose there is a solution $Y \neq X_0$ to $\mathcal{A}(Y) = \mathcal{A}(X_0)$ with $\mathrm{rank}(Y) \leq \mathrm{rank}(X_0)$. Let $Z = X_0 - Y$ and note that $Z \in \mathrm{Null}(\mathcal{A})$. Using the spherical section property and standard norm inequalities we have

$$\|Z\|_* \geq \sqrt{\Delta}\|Z\|_F \geq \sqrt{\Delta}/\sqrt{\mathrm{rank}(Z)}\,\|Z\|_*,$$

thus $\mathrm{rank}(Z) \geq \Delta$. However, $\mathrm{rank}(Z) \leq \mathrm{rank}(X_0) + \mathrm{rank}(Y) \leq 2\,\mathrm{rank}(X_0)$. Since $\mathrm{rank}(X_0) < \frac{\Delta}{2}$, we have a contradiction, hence X_0 is the unique solution. ∎

   (b) Suppose there is a $Y \neq X_0$ such that $\|Y\|_* \leq \|X_0\|_*$ and $\mathcal{A}(Y) = \mathcal{A}(X_0)$. Let $X_0 = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^T$ be the SVD of X_0 where $\Sigma \in \mathbf{R}^{r \times r}$ is diagonal. Let $Z = X_0 - Y$, and $\bar{Z} = U^T Z V = \begin{bmatrix} \bar{Z}_{11} & \bar{Z}_{12} \\ \bar{Z}_{21} & \bar{Z}_{22} \end{bmatrix}$. Then

$$\begin{aligned} \|X_0\|_* &\geq \|X_0 + Z\|_* \\ &= \left\| \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \bar{Z}_{11} & \bar{Z}_{12} \\ \bar{Z}_{21} & \bar{Z}_{22} \end{bmatrix} \right\|_* \\ &\geq \left\|\Sigma + \bar{Z}_{11}\right\|_* + \left\|\bar{Z}_{22}\right\|_* \quad \text{([5] Lemma 3.2)} \end{aligned}$$
$$\|\Sigma\|_* \geq \|\Sigma\|_* - \left\|\bar{Z}_{11}\right\|_* + \left\|\bar{Z}_{22}\right\|_*$$

which yields $\left\|\bar{Z}_{22}\right\|_* \leq \left\|\bar{Z}_{11}\right\|_*$. Now consider the problem

$$\begin{aligned} \text{maximize} \quad & \frac{\|Z\|_*}{\|Z\|_F} \\ \text{subject to} \quad & \|Z_{22}\|_* \leq \|Z_{11}\|_*, \end{aligned}$$

where $Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}$, $Z_{11} \in \mathbf{R}^{r \times r}$, $Z_{22} \in \mathbf{R}^{(m-r) \times (n-r)}$, $Z_{12} \in \mathbf{R}^{(m-r) \times r}$, and $Z_{21} \in \mathbf{R}^{r \times (n-r)}$. It can be shown [7] that the optimal value of this program is bounded above by $\sqrt{2r(3m-2r)}/\sqrt{m}$ if $r \leq m/2$. By the spherical section property of $\mathrm{Null}(\mathcal{A})$, we have

$$\Delta \leq \frac{2r(3m-2r)}{m} \implies r \geq \frac{3m}{4} - \sqrt{\frac{9m^2}{16} - \frac{m\Delta}{4}}.$$

Thus, if $r < \frac{3m}{4} - \sqrt{\frac{9m^2}{16} - \frac{m\Delta}{4}}$, we have a contradiction, hence X_0 is the unique optimum of the problem (2). ∎

   The next theorem extends the above result to the case where X_0 is only approximately low-rank. It shows that in this case, the heuristic finds a matrix that is close to X_0 within a constant factor of the error achieved by the *best rank-k approximation* to X_0.

Theorem 2.2 *Suppose* $\mathrm{Null}(\mathcal{A})$ *has the* Δ*-spherical section property. Let* $X_0 \in \mathbf{R}^{m \times n}$, $m \leq n$, $X_0 \neq 0$, *and let* X_0^k *be the best rank-k approximation of* X_0. *If* X^* *is the solution to* (2), *then*

$$\|X^* - X_0\|_* \leq c \left\|X_0 - X_0^k\right\|_*$$

provided $c > 2$ *and* $k < \min\left(\frac{m(c-2)}{2(c+2)}, \frac{\Delta(c-2)^2}{6c^2}\right)$.

Proof Let $X_0 = U \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma' \end{bmatrix} V^T$ be the SVD of X_0 where Σ consists of the largest k singular values of X_0. Let $Z = -X_0 + X^*$ and $\bar{Z} = U^T Z V$, and note that $\|X_0\|_* \geq \|X^*\|_*$. Since norms are preserved by orthogonal transformations,

$$\left\| \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma' \end{bmatrix} \right\|_* \geq \left\| \begin{bmatrix} \Sigma + \bar{Z}_{11} & \bar{Z}_{12} \\ \bar{Z}_{21} & \Sigma' + \bar{Z}_{22} \end{bmatrix} \right\|_*$$
$$\|\Sigma\|_* + \|\Sigma'\|_* \geq \left\|\Sigma + \bar{Z}_{11}\right\|_* + \left\|\Sigma' + \bar{Z}_{22}\right\|_*$$
$$\geq \|\Sigma\|_* - \left\|\bar{Z}_{11}\right\|_* + \left\|\bar{Z}_{22}\right\|_* - \|\Sigma'\|_*$$

Thus we get $2\|\Sigma'\|_* + \left\|\bar{Z}_{11}\right\|_* \geq \left\|\bar{Z}_{22}\right\|_*$. Now, $\|\Sigma'\|_* = \left\|X_0 - X_0^k\right\|_*$ by definition. Suppose that $c\|\Sigma'\|_* < \|Z\|_*$, otherwise we are done. Then by the trivial upper bound,

$$c\|\Sigma'\|_* < \left\|\bar{Z}_{11}\right\|_* + \left\|\bar{Z}_{22}\right\|_* + \left\|\bar{Z}_{12}\right\|_* + \left\|\bar{Z}_{21}\right\|_*,$$

so that

$$(\eta + 1)\left\|\bar{Z}_{11}\right\|_* + \eta(\left\|\bar{Z}_{12}\right\|_* + \left\|\bar{Z}_{21}\right\|_*) > (1 - \eta)\left\|\bar{Z}_{22}\right\|_*,$$

where $\eta = 2/c$. Now consider the problem

$$\begin{aligned} \text{maximize} \quad & \frac{\|Z\|_*}{\|Z\|_F} \\ \text{subject to} \quad & \frac{1+\eta}{1-\eta}\|Z_{11}\|_* + \frac{\eta}{1-\eta}(\|Z_{12}\|_* + \|Z_{21}\|_*) > \|Z_{22}\|_*. \end{aligned}$$

We can show [7] that the optimal value is bounded above by

$$\frac{\sqrt{k}}{\sqrt{-\frac{\eta^2 + 2\eta}{2} + \frac{(1+2\eta)^2}{4(1.5 - \frac{k}{m})}}},$$

provided that $k < \frac{m(1-\eta)}{2(1+\eta)}$. If $\sqrt{\Delta}$ is greater than this value, we have a contradiction, hence $\|X_0 - X^*\|_* \leq c\|X_0 - X_0^k\|_*$. However, this condition is hard to interpret, and it can be easily shown that the simpler condition $k < \min\left(\frac{(1-\eta)}{2(1+\eta)}m, \frac{\Delta(1-\eta)^2}{6}\right)$ suffices to ensure the one above. ∎

Letting $c \to \infty$ when $X_0 = X_0^k$, we recover the result in theorem 2.1.

3. PROBABILISTIC ANALYSIS OF THE NULLSPACE CONDITION

In this section, we show that if the matrix representation of $\mathcal{A}$ has iid zero-mean, unit-variance Gaussian (i.e., $\mathcal{N}(0,1)$) entries, the nullspace condition described above holds with high probability. In practice, in many rank minimization problems, it is not possible to choose $\mathcal{A}$ to have completely random entries, since the application dictate that $\mathcal{A}$ should have certain structure. For example, in the matrix completion problem [8], $\mathcal{A}$ is an operator that just picks entries of the matrix and cannot take arbitrary linear combinations of the entries. However, there are some applications (such as quantum tomography [9]) where this is still a reasonable assumption. Also, as has happened in the vector case, we hope that analyzing the simpler case of the Gaussian ensemble will pave the way for more sophisticated analyses of structured $\mathcal{A}$.

Further, recent results based on RIP [10] suggest that some results for more structured $\mathcal{A}$ can be derived from general results. They show that although RIP does not holds for all low-rank matrices, it holds for a large class of low-rank matrices satisfying certain incoherence properties and that matrices generated from certain random models have these incoherence properties with high probability. This seems like a promising way to extend general conditions to handle more specific and structured cases.

Assume without loss of generality that $\gamma \leq 1$. The nullspace of random operator $\mathcal{A}$ can be characterized [5] as consisting of matrices

$$Z = \sum_{k=1}^{M} v_k G^k,$$

where $G^k \in \mathbf{R}^{\gamma n \times n}$, $M = \gamma n^2 - p$ have iid $\mathcal{N}(0,1)$ entries. Thus, our aim reduces to showing that the probability

$$P\left(\inf_{v \neq 0} \frac{\left\|\sum_i v_i G^i\right\|_*}{\left\|\sum_i v_i G^i\right\|_F} \leq \beta\right) =$$

$$P\left(\inf_{v:\|v\|_2=1} \left\|\sum_i v_i G^i\right\|_* - \beta \left\|\sum_i v_i G^i\right\|_F \leq 0\right)$$

is very small, where the equality above follows because the ratio of norms is invariant to scaling of v, so that we can fix $\|v\|_2 = 1$. Let $\sigma = \max_{M \in \mathbf{R}^{\gamma n \times n}, \|M\|_2 = 1} \|M\|_F = \sqrt{\gamma n}$, and

$$\Psi(\gamma) = \int_{(1-\sqrt{\gamma})^2}^{(1+\sqrt{\gamma})^2} \sqrt{\frac{(z-s_1)(s_2-z)}{z}} \, dz.$$

Lemma 3.1 *Let $G^i \in \mathbf{R}^{\gamma n \times n}$, $i = 1, 2, \ldots, M$, have iid $\mathcal{N}(0,1)$ entries. Then there exists a constant $c_* > 0$ such that*

$$E\left[\inf_{\|v\|_2=1} \left\|\sum_{i=1}^{M} v_i G^i\right\|_* - \beta \left\|\sum_{i=1}^{M} v_i G^i\right\|_F\right] \geq$$
$$c_* \Psi(\gamma) n^{3/2} - n\sqrt{\gamma}\beta - (\sigma + \beta)\sqrt{\gamma n^2 - p} \quad (3)$$

Proof Consider the space of $\mathbf{R}^{\gamma n \times n}$ matrices with the inner product $\langle A, B \rangle = \text{tr}(A^T B)$. We first rewrite the expression variationally using the dual norm:

$$\inf_{\|v\|_2=1} \sup_{\|Z\|_2=1} \inf_{\|Y\|_F=1} \left\langle \sum_{i=1}^{M} v_i G^i, Z - \beta Y \right\rangle$$

We use corollary 3.13 from [11] by constructing an auxiliary Gaussian process to help bound the desired probability. Let $n \in \mathbf{R}, n \sim \mathcal{N}(0,1), g \in \mathbf{R}^M, g_i \sim \mathcal{N}(0,1)$. Define

$$P_L(v, Z, Y) = \left\langle \sum_{i=1}^{M} v_i G^i, Z - \beta Y \right\rangle + n(\sigma + \beta)$$

$$P_R(v, Z, Y) = \langle G_0, Z - \beta Y \rangle + (\sigma + \beta)\langle g, v \rangle$$

P_L and P_R are standard Gaussian processes. By simple algebraic manipulations and using the Cauchy-Schwarz and norm inequalities, we can show that $E[P_L(v, Z, Y) P_L(v', Y', Z')] \geq E[P_R(v, Z, Y) P_R(v', Y', Z')]$ with equality if $v = v', Y = Y'$. Thus, by corollary 3.13 from [11] and the compactness argument from appendix A.3 in [5], we have

$$E\left[\min_{\|v\|_2=1, \|Y\|_F=1} \max_{\|Z\|_2=1} P_L(v, Y, Z)\right] \geq$$
$$E\left[\min_{\|v\|_2=1, \|Y\|_F=1} \max_{\|Z\|_2=1} P_R(v, Y, Z)\right] =$$
$$E[\|G_0\|_*] - \beta E[\|G_0\|_F] - (\sigma + \beta)E[\|g\|_2] =$$

From standard results on the limiting spectral distributions of random Gaussian matrices [12], we know that the expected value of the nuclear norm of a $\gamma n \times n$ random Gaussian matrix is given asymptotically by

$$\Psi(\gamma)n^{3/2} + o(n^{3/2})$$

Thus, there exists a constant $c_* > 0$ such that $E[\|G_0\|]_* \geq c_* \Psi(\gamma) n^{3/2}$. Also $E[\|G_0\|_F] \leq \sqrt{E[\|G_0\|_F^2]} = n\sqrt{\gamma}$. Thus, we get the following lower bound on the expectation:

$$c_* \Psi(\gamma) n^{3/2} - n\beta\sqrt{\gamma} - (\sigma + \beta)\sqrt{\gamma n^2 - p}.$$

∎

We also use the following lemma, a fairly elementary proof can be found in [7].

Lemma 3.2 *The function*

$$f(G^1, G^\cdot \ldots, G^M) = \inf_{\|v\|_2=1} \left\|\sum_i v_i G^i\right\|_* - \beta \left\|\sum_i v_i G^i\right\|_F$$

is Lipschitz with respect to the l_2 norm

$$\left\|\left(G^1, G^2, \ldots, G^M\right)\right\| = (\sum_{i=1}^{M} \left\|G^i\right\|_F^2)^{1/2}$$

with constant $\sigma + \beta$.

3.1. Recovery results

Theorem 3.3 *If $\mathcal{A} : \mathbf{R}^{\gamma n \times n} \mapsto \mathbf{R}^p$ is a random linear operator whose matrix representation has iid $\mathcal{N}(0,1)$ entries then if $p = \mu n^2$, X_0 is the unique optimal solution to (2) for all X_0 of rank up to $\frac{\alpha}{6}n$ with probability at least*

$$1 - \exp\left(-\frac{\kappa^2 n^2}{2(1 + \sqrt{\alpha})^2}\right)$$

provided that $c_* \Psi(\gamma) - \sqrt{\alpha\gamma} - \kappa > 0$ and

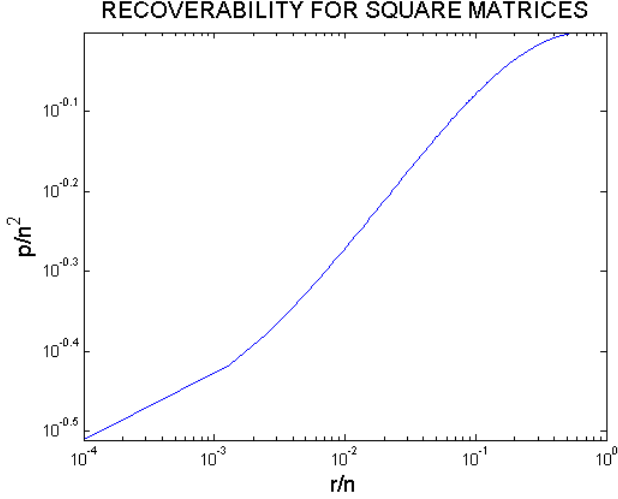$$\mu > \gamma - \left(\frac{c_* \Psi(\gamma) - \sqrt{\alpha\gamma} - \kappa}{1 + \sqrt{\alpha}}\right)^2.$$

RECOVERABILITY FOR SQUARE MATRICES

Fig. 1.

Proof An L-Lipschitz function of Gaussian random variables satisfies the following concentration bound [13]:

$$P[f(X) - E[f(X)] \leq -t] \leq \exp\left(-\frac{t^2}{2L^2}\right)$$

for any $t > 0$. Choose $t = c_*\Psi(\gamma)n^{3/2} + o(n^{3/2}) - n\sqrt{\gamma}\beta - (\sigma + \beta)\sqrt{\gamma n^2 - p}$. Let $\beta = \sqrt{\alpha n}$, $p = \mu n^2$. Then we have $t = (c_*\Psi(\gamma) - \sqrt{\alpha\gamma} - (1 + \sqrt{\alpha})\sqrt{\gamma - \mu})n^{3/2}$. To ensure that $t > 0$, we must have $c_*\Psi(\gamma) - \sqrt{\alpha\gamma} - \kappa > 0$,

$$\mu > \gamma - \left(\frac{c_*\Psi(\gamma) - \sqrt{\alpha\gamma} - \kappa}{1 + \sqrt{\alpha}}\right)^2,$$

and using (3.1), $P_{Z:\mathcal{A}(Z)=0}[f(Z) \leq 0] \leq e^{\left(-\frac{\kappa^2 n^2}{2(1+\sqrt{\alpha})^2}\right)}$.   ∎

Note that it has been observed that even for moderately sized square matrices ($n \geq 40$), the asymptotic estimates are fairly accurate [5], so that we can take $c_* \approx 1$. Thus, plugging in values, we can recover ranks up to $0.0001n$ with approximately $0.28n^2$ measurements. We have plotted results of number of measurements versus maximum rank recoverable in Figure 1 for a fixed probability of failure of 0.001.

4. CONCLUSIONS AND FUTURE WORK

We presented an analysis of the nuclear norm heuristic based on the spherical section property of the nullspace of $\mathcal{A}$. If $\mathcal{A}$ is picked from the random Gaussian ensemble, we showed that it satisfies the desired property with high probability provided there are enough measurements. We obtained bounds that demonstrate explicit tradeoffs between the rank of X_0, its aspect ratio γ, the measurements ratio μ, and the recovery probability. These kind of bounds are referred to as *Strong Bounds* in [5], as they guarantee that the heuristic succeeds in recovering matrices up to some rank, no matter what the optimal solution is (in fact, the mere existence of such bounds is very interesting from a theoretical viewpoint).

Future research will aim at obtaining better probabilistic bounds, extending to non-Gaussian operators, and analyzing other scaling regimes that may be useful in practice. In addition, we will pursue tighter results for the special case of positive semidefinite low-rank recovery, where numerical examples suggest recovery is possible for a far larger region than general matrices. Another open issue is whether efficient relaxations can be developed for testing the spherical section property given a specific $\mathcal{A}$, along the lines of the results in [14].

It is also interesting to analyze operators $\mathcal{A}$ that have both random and structured features, arising naturally in applications such as Hankel system identification, low-dimensional Euclidean embedding, and low-rank matrix completion [1].

5. REFERENCES

[1] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum rank solutions to linear matrix equations via nuclear norm minimization," Accepted for publication, *SIAM Review*.

[2] M. Fazel, H. Hindi, and S. Boyd, "A rank minimization heuristic with application to minimum order system approximation," in *Proceedings of the American Control Conference*, 2001.

[3] Emmanuel J. Candes and Yaniv Plan, "Tight oracle bounds for low-rank matrix recovery from a minimal number of random measurements," 2010.

[4] Y. Zhang, "On the theory of compressed sensing via ℓ_1 minimization: simple derivations and extensions," Available at http://www.caam.rice.edu/ zhang/reports/tr0811.pdf, 2008.

[5] B. Recht, W. Xu, and B. Hassibi, "Necessary and sufficient condtions for success of the nuclear norm heuristic for rank minimization," Tech. Rep., Caltech, 2008, Available at http://www.ist.caltech.edu/ brecht/papers/08.RecXuHas.TR.pdf.

[6] B. S. Kashin and V. N. Temlyakov, "A remark on compressed sensing," 2007, Available at http://www.dsp.ece.rice.edu/cs/KT2007.pdf.

[7] K. Dvijotham and M. Fazel, "Nullspace conditions for the nuclear norm heuristic," In preparation.

[8] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. of Comput. Math.*, , no. 9, 2008.

[9] David Gross, Yi-Kai Liu, Steven T. Flammia, Stephen Becker, and Jens Eisert, "Quantum state tomography via compressed sensing," 2009.

[10] Raghu Meka, Prateek Jain, and Inderjit S. Dhillon, "Guaranteed rank minimization via singular value projection," 2009.

[11] M. Ledoux and M. Talagrand, *Probability in Banach Spaces*, Springer-Verlag, 1991.

[12] Z. Bai, "Methodologies in the spectral analysis of large dimensional random matrices, a review," *Statistica Sinica*, 1999.

[13] P. Massart, *Concentration Inequalities and Model Selection*, vol. Lecture Notes in Mathematics, Springer-Verlag, 2007.

[14] A. d'Aspremont and L. El Ghaoui, "Testing the nullspace property using semidefinite programming," To appear in *Mathematical Programming*.