# MATCHED SUBSPACE DETECTION USING COMPRESSIVELY SAMPLED DATA

Dejiao Zhang Laura Balzano

Department of Electrical Engineering and Computer Science, University of Michigan dejiao, girasole@umich.edu

# ABSTRACT

We consider the problem of detecting whether a high dimensional signal lies in a given low dimensional subspace using only a few compressive measurements of it. By leveraging modern random matrix theory, we show that, even when we are short on information, a reliable detector can be constructed via a properly defined measure of energy of the signal outside the subspace. Our results extend those in [1] to a more general sampling framework. Moreover, the test statistic we define is much simpler than that required by [1], and it results in more efficient computation, which is crucial for highdimensional data processing.

*Index Terms* – High dimension, compressive measurements, random matrix theory.

# 1. INTRODUCTION

We focus on testing whether a signal vector  $v \in \mathbb{R}^n$ lies in a known *d*-dimensional subspace  $S \subset \mathbb{R}^n (d \ll n)$ , given few compressive measurements of the vector. This problem arises in a wide range of applications including medical imaging [2], hyperspectral target detection [3], anomaly detection [4], radar signal processing [5, 6, 7] and inference estimation [8]. Among these applications they either aim at finding the signal within some low-dimensional subspace or leverage the subspace as a model to detect signals of interest outside the subspace.

This problem can be modeled as a binary hypothesis test with hypotheses  $\mathcal{H}_0 : v \in S$  and  $\mathcal{H}_1 : v \notin S$ . When full observation is available, let  $v_{\perp}$  denote the projection residual of v onto S; then the above test can be constructed as

$$\mathcal{H}_0: \|v_{\perp}\|_2 = 0 \quad \text{vs.} \quad \mathcal{H}_1: \|v_{\perp}\|_2 > 0 \qquad (1)$$

However, with only compressive measurements of v, we cannot compute  $v_{\perp}$  directly. Although in recent years exciting results emerging on successfully recovering the underlying signal from compressive measurements under mild assumptions, the reconstruction can result in extra expensive computation and it can also be arbitrarily poor when the underlying signal does not belong to

the given subspace. Therefore, in this paper, we seek to detect whether the compressively sampled vector lies in the given subspace without reconstructing it. As we will show in Section 3, under mild assumptions on the sampling matrix, a reliable detector can be obtained for both noise-free and noisy data.

In [1], the authors also consider the matched subspace detection problem using compressive measurements. However, they require the sampling matrix and noise to be Gaussian distributed with independent and identically distributed entries, while we only require sub-gaussian sampling matrix with independent rows. Moreover, the test statistic defined in [1] is complicated, requiring extra expensive computations that can be prohibitive in the high dimensional setting. [9] is also closely related to our work, where the authors study the missing data case, where only a few of coordinates are drawn uniformly at random for each underlying vector. However, this case can be modeled such that each entry of the sampling matrix is a Bernoulli random variable that equals 1 with probability  $\frac{m}{n}$  and 0 with probability  $1 - \frac{m}{n}$  (where m is the number of observed entries). This is also an instance of sub-gaussian random matrix. Therefore, our results also generalize that of [9].

#### 2. TEST STATISTIC

Let S be a known d-dimensional subspace in  $\mathbb{R}^n$ , described by an  $n \times d$  matrix U whose orthonormal columns span S. We seek to detect whether the unknown vector  $v \in \mathbb{R}^n$  lies in S given only a small number of compressive measurements of the form:

$$x = Av + \xi \tag{2}$$

where  $x \in \mathbb{R}^m$  is the observed vector,  $A \in \mathbb{R}^{m \times n} (m \ll n)$  is the sampling matrix with independent sub-gaussian rows, and  $\xi \in \mathbb{R}^m$  is additive noise with entries being independent and identically distributed sub-gaussian random variables.

Analogously to [9], we define the test statistic as

$$T = \left\| \left( \mathbb{I}_m - \mathcal{P}_{AU} \right) x \right\|_2^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \eta_\sigma \tag{3}$$

where  $\mathcal{P}_{AU}$  denotes the orthogonal projection operator onto the column space of AU. Throughout this paper, we assume AU has full column rank. Note that for the fully sampled case where the sampling matrix is the identity matrix, the test statistic defined in (3) is the projection residual of v onto S. Comparing with the test statistic  $T = ||\mathbb{I}_m - \mathcal{P}_{BU}(A^T A)^{-\frac{1}{2}}x||_2^2, B =$  $(AA^T)^{-\frac{1}{2}}A$  defined in [1], we do not need the extra computation of  $(AA^T)^{-\frac{1}{2}}$  which can be expensive in the high dimensional setting.

Decompose the underlying signal vector as  $v = v_{\parallel} + v_{\perp}$  where  $v_{\parallel} \in S$  and  $v_{\perp} \in S^{\perp}$ . It then follows that  $\left\| \left( \mathbb{I}_m - \mathcal{P}_{AU} \right) A v_{\parallel} \right\|^2 = 0$ . Therefore, (3) is equivalent to

$$T = \left\| \left( \mathbb{I}_m - \mathcal{P}_{AU} \right) \left( Av_\perp + \xi \right) \right\|_2^2 \underset{\mathcal{H}_0}{\overset{2}{\approx}} \eta_\sigma \qquad (4)$$

In Section 3.1, we prove that, when there is no noise contained in the observation, the test statistic defined in (4) concentrates around  $||v_{\perp}||_2^2$  scaled by the sampling density, m/n. Therefore, it's natural for us to set  $\eta_{\sigma} = 0$  for noise-free data. When the observation contains noise, we set  $\eta_{\sigma}$  to be some positive value that depends on the noise level. As we present in Section 3.2, with subgaussian distributed noise, a reliable detector can be obtained as long as the energy of the vector outside the given subspace scales with n.

### 3. MAIN RESULTS

Sub-gaussian random variables form a quite wide class whose distributions can be dominated by the distribution of a centered Gaussian random variable. That is, X is a sub-gaussian random variable if

$$\mathbb{P}\left(|X| \ge t\right) \le 2\exp(-ct^2) \tag{5}$$

where c > 0 is a constant depending on the sub-gaussian norm of X defined as

$$||X||_{\Psi_2} = \inf \{K > 0 : \mathbb{E} \exp (X^2/K^2) \le 2\}$$
 (6)

Classical examples of sub-gaussian random variables include Gaussian, Bernoulli, and all bounded random variables. Given (6), the  $\Psi_2$  norm of a sub-gaussian random vector <sup>1</sup>  $Y \in \mathbb{R}^n$  is defined as

$$||Y||_{\Psi_2} = \sup_{||y||_2 = 1} ||\langle Y, y \rangle||_{\Psi_2}$$
(7)

Now we call out our main assumptions on the sampling matrix and noise. In this paper, we use  $M_i$  to denote the  $i^{th}$  row of the matrix M. **Condition 1.** The sampling matrix can be generated as  $A = \frac{1}{\sqrt{n}}B$  with the rows of B being independent sub-gaussian random vectors with mean zero. Let  $K = \max_i ||B_i||_{\Psi_2}$  denote the maximal sub-gaussian norm of the rows of B. We also assume that each row of B is isotropic, i.e.,  $\mathbb{E}[B_iB_i^T] = \mathbb{I}_n, i = 1, ..., m$ . This is equivalent to

$$\mathbb{E}\langle B_i, z \rangle^2 = \|z\|^2 \qquad \forall z \in \mathbb{R}^n$$

**Condition 2.** Suppose the entries of the noise vector are *i.i.d* sub-gaussian random variables with  $\mathbb{E}[\xi_i] = 0$ ,  $Cov(\xi_i) = \sigma^2$  and  $\|\xi_i\|_{\Psi_2} \leq K_1$ .

We extend [1] to more general and simpler sampling framework that is more applicable in practice. For example, in applications, sampling matrices are often sparse 0-1 matrices or with bounded entries, both of which are sub-gaussian random matrices. As we prove in the following section, this generality comes without sacrificing on the performance of the defined test statistic.

#### 3.1. Noiseless Data

We first consider noise-free data. As we present in Theorem 1, the test statistic  $\|(\mathbb{I}_m - \mathcal{P}_{AU}) Av_{\perp}\|_2^2$  concentrates around  $\frac{m}{n} \|v_{\perp}\|^2$  with high probability.

**Theorem 1.** Let  $\Gamma(a) = \min \{a^2/K^4, a/K^2\}$ , then with probability at least  $1 - \exp [-mC_1\Gamma(\beta_1 - 1)] - \exp [-mC_1\Gamma(1 - \alpha_1)] - \exp [-dC_2\Gamma(\beta_0 - 1)]$  we have

$$\left(\alpha_1 \frac{m}{n} - \beta_0 \frac{d}{n}\right) \|v_{\perp}\|^2 \le \|(\mathbb{I}_m - \mathcal{P}_{AU}) A v_{\perp}\|^2 \le \beta_1 \frac{m}{n} \|v_{\perp}\|^2$$
where  $0 < \alpha_1 < 1$  by  $\beta_0 > 1$  and  $C_1 C_2 > 0$  are

where  $0 < \alpha_1 < 1, \beta_1, \beta_0 > 1$ , and  $C_1, C_2 > 0$  are absolute constants.

This implies the following corollary by setting  $\alpha_1 = 2/\log n, \beta_0 = \frac{m}{2ed}$  and  $\beta_1 = e$ .

**Corollary 1.** If m > 2ed, then with probability at least  $1 - 3 \exp[-\tau_2 m]$  we obtain

$$\frac{m}{2en} \|v_{\perp}\|^2 \le \|(\mathbb{I}_m - \mathcal{P}_{AU}) A v_{\perp}\|^2 \le e\frac{m}{n} \|v_{\perp}\|^2$$
  
here  $\tau_2 = \min\left\{C_1 \Gamma(1 - \frac{1}{e}), \frac{C_2(\frac{1}{2e} - \frac{d}{m})}{K^2}, \frac{C_2 \frac{m}{d}(\frac{1}{2e} - \frac{d}{m})^2}{K^4}\right\}.$ 

We need the following results for the proof of Theorem 1.

Lemma 1. With the same notation as Theorem 1,

$$\mathbb{P}\left(\|Av_{\perp}\|^{2} \leq \alpha_{1}\frac{m}{n}\|v_{\perp}\|^{2}\right) \leq \exp\left[-mC_{1}\Gamma(1-\alpha_{1})\right]$$
$$\mathbb{P}\left(\|Av_{\perp}\|^{2} \geq \beta_{1}\frac{m}{n}\|v_{\perp}\|^{2}\right) \leq \exp\left[-mC_{1}\Gamma(\beta_{1}-1)\right]$$

w

<sup>&</sup>lt;sup>1</sup>Similarly to the definition of Gaussian random vectors, a random vector  $Y \in \mathbb{R}^n$  is called sub-gaussian if the one-dimensional marginals  $\langle Y, y \rangle$  are sub-gaussian random variables for all  $y \in \mathbb{R}^n$ .

**Lemma 2.** With the same notation as Theorem 1,  $\|\mathcal{P}_{AU}(Av_{\perp})\|^2 \leq \beta_0 d \|v_{\perp}\|^2/n$  holds with probability at least  $1 - \exp\left[-dC_2\Gamma(\beta_0 - 1)\right]$ .

Proof of Theorem 1. Apply the above Lemmas to  $\|(\mathbb{I}_m - \mathcal{P}_{AU}) Av_{\perp}\|^2 = \|Av_{\perp}\|^2 - \|\mathcal{P}_{AU}(Av_{\perp})\|^2 \le \|Av_{\perp}\|^2$ . Then taking the union bounds of Lemma 1 and Lemma 2 completes the proof.  $\Box$ 

#### 3.2. Noisy Data

When the observation contains noise, we set the test threshold  $\eta_{\sigma}$  to be some properly chosen positive value. As we present in Theorem 2, a reliable detector can be obtained as long as  $||v_{\perp}||^2$  scales with n under  $\mathcal{H}_1$ .

**Theorem 2.** Let  $\Gamma_1(a) = \min \left\{ a^2 \sigma^4 / K_1^4, a \sigma^2 / K_1^2 \right\}$ and  $\eta_{\sigma} = e(m-d)\sigma^2$ . Then

$$\mathbb{P}\left(T \ge \eta_{\sigma} | \mathcal{H}_0\right) \le \exp\left[-(m-d)C_3\Gamma_1(e-1)\right].$$

Additionally suppose m > 2ed and

$$||v_{\perp}||^2 \ge 4e(e+2)(1-d/m)n\sigma^2$$
 (8)

holds for any  $v \notin S$ , then

$$\mathbb{P}\left(T \leq \eta_{\sigma} \left| \mathcal{H}_{1}\right) \leq 2 \exp\left[-\tau_{3}m\right] + 2 \exp\left[-\tau_{4}(m-d)\right],$$
where  $\tau_{3} = \min\left\{C_{1}\Gamma\left(1-\frac{1}{e}\right), \frac{C_{2}\left(\frac{1}{2e}-\frac{d}{m}\right)}{K^{2}}, \frac{C_{2}\frac{m}{d}\left(\frac{1}{2e}-\frac{d}{m}\right)^{2}}{K^{2}}\right\}$ 
and  $\tau_{4} = \left\{-C_{4}\sigma^{2}/K_{1}^{2}, C_{3}\Gamma_{1}(1-1/e)\right\}$  with  $C_{3}, C_{4}$ 
be absolute constants and  $C_{1}, C_{2}, \Gamma$  be the same as Theorem 1.

Theorem 2 states that, with sub-gaussian noise, the probability of false alarm decays exponentially in terms of m - d. Given  $v \notin S$ , the probability Type II error also decays exponentially as long as the energy of  $v_{\perp}$  scales with n. This is caused by the fact that  $||Av_{\perp}||^2 \approx \frac{m}{n} ||v_{\perp}||^2$  while as we present in the following  $||\xi||^2 \approx m\sigma^2$ .

Now we call out the following lemmas for the proof of Theorem 2, which quantify the perturbation terms induced by the additive noise.

**Lemma 3.** With the same notation as that of Theorem 2, let  $Y = \|(\mathbb{I}_m - \mathcal{P}_{AU})\xi\|^2$ , then

$$\mathbb{P}\left[Y \le \alpha_2(m-d)\sigma^2\right] \le \exp\left[-(m-d)C_3\Gamma_1(1-\alpha_2)\right]$$
$$\mathbb{P}\left[Y \ge \alpha_2(m-d)\sigma^2\right] \le \exp\left[-(m-d)C_3\Gamma_1(1-\alpha_2)\right]$$

$$\mathbb{P}\left[Y \ge \beta_2(m-d)\sigma^2\right] \le \exp\left[-(m-d)C_3\Gamma_1(\beta_2-1)\right]$$

where  $0 < \alpha_2 < 1$ ,  $\beta_2 > 1$  and  $C_3$  is an absolute constant.

**Lemma 4.** Let  $Z = (\mathbb{I}_m - \mathcal{P}_{AU}) Av_{\perp}$ . Then for any  $\gamma > 0$  we have

$$\mathbb{P}\left[\xi^T Z \le -\gamma \sqrt{m-d}\sigma \|Z\|\right] \le \exp\left[-C_4 \frac{\gamma^2 \sigma^2 (m-d)}{K_1^2}\right]$$

where  $C_4 > 0$  is an absolute constant.

Proof of Theorem 2. Given  $v \in S$ ,  $T = \|(\mathbb{I}_m - \mathcal{P}_{AU})\xi\|^2$ . Hence, Lemma 3 directly yields the first part.

Let  $Z = (\mathbb{I}_m - \mathcal{P}_{AU}) Av_{\perp}$ , and set  $\alpha_2 = 1/e$  and  $\gamma = 1$  in Lemma 3 and 4 correspondingly, then with probability at least  $1 - \exp\left[-(m-d)C_3\Gamma_1(1-\frac{1}{e})\right] - \exp\left[-C_4(m-d)\sigma^2/K_1^2\right]$  we have

$$T \ge \|Z\|^2 - 2\sqrt{m-d}\sigma\|Z\| + (m-d)\sigma^2/e$$
 (9)

Therefore,  $T \ge e(m-d)\sigma^2$  is equivalent to

$$||Z||^2 \ge 2(e+2)(m-d)\sigma^2$$

Set  $\alpha_1 = 1/e$  and  $\beta_0 = \frac{m}{2ed}$  for the lower bound in Theorem 1, then together with (8) we have

$$||Z||^2 \ge \frac{m}{2en} ||v_{\perp}||^2 \ge 2(e+2)(m-d)\sigma^2 \qquad (10)$$

holds with probability at least  $1 - \exp\left[-C_1\Gamma(1-\frac{1}{e})\right] - \exp\left[-dC_2\Gamma\left(\frac{m}{2ed}-1\right)\right]$ . The probability bound is obtained by 1 minus the union bounds of those yielding (9) and (10).

# 4. NUMERICAL RESULTS



Figure 1: Illustration of Theorem 1 (left) and Theorem 2 (right) with n = 5000 and d = 50. The sampling matrices for both (a) and (b) are generated as a sparse matrix such that  $A_{ij} = \pm \sqrt{3/n}$ , w.p. 1/6 and  $A_{ij} = 0$ , w.p. 2/3. The entries of noise vectors (for (b)) are generated as i.i.d uniform random variables with mean zero and unit covariance.

In this section, we examine our main results with synthetic data. In order to demonstrate that incoherence assumptions are not required when sub-gaussian random matrices are used, both v and U are generated as sparse ensembles with sparsity on the order of  $\log(n)/n$ .

Fig (1a) examines Theorem 1 over 100 simulations, each with a fixed subspace S and different sampling densities m/n. For each sampling density, we sample 50 instances of  $Av_{\perp}$ , and then compute the mean, maximum and minimum of our defined test statistic. As we can see the values of our test statistic concentrate well around  $\frac{m-d}{n} ||v_{\perp}||^2$  as long as m > d. In Fig (1b), Type II error is averaged over 100 simulations with different sampling densities and different instances of  $v_{\perp}$  that scales variational with n. As that predicted in Theorem 2, the probability of Type II error decays exponentially as long as  $||v_{\perp}||^2/n$  is bounded from below by some constant.

### 5. CONCLUSION

We have shown that it is possible to test whether a high dimensional vector lies in a known subspace with only few compressive measurements. By leveraging modern random matrix theory, we extend the results in [1] to a more general sampling framework with a much simpler and more efficient test statistic. We prove that when the sampling matrix has independent sub-gaussian rows, the energy outside the subspace is preserved. For noisy data, we show that our defined test statistic is reliable as long as, under  $\mathcal{H}_1$ , the energy of the underlying signal vector outside the given subspace scales with n.

# **Proof of Main Results**

We need the following for the proof of our main results.

**Lemma 5.** [10] A random variable X is sub-gaussian if and only if  $X^2$  is sub-exponential. Moreover,  $\|X^2\|_{\Psi_1} = \|X\|_{\Psi_2}^2$  and  $\|X^2 - \mathbb{E}X^2\|_{\Psi_1} \le C\|X^2\|_{\Psi_1}^2$ .

**Lemma 6.** (Hoeffding-type inequality, [10]). Let  $X_1, \ldots, X_N$  be independent centered sub-gaussian random variables, and let  $K = \max_i ||X_i||_{\Psi_2}$ . For every  $a = (a_1, \cdots, a_N)$ , let  $Y = \sum_{i=1}^N a_i X_i$ , then

$$\mathbb{P}\left(|Y| \ge t\right) \le 2\exp\left(-ct^2/K^2 \|a\|_2^2\right) \quad \forall t > 0$$

where c > 0 is an absolute constant.

**Lemma 7.** (Bernstein's inequality, [10]). Let  $X_1, \ldots, X_N$  be independent, mean zero, subexponential random variables. Let  $Y = \sum_{i=1}^N X_i$ , then  $\forall t \ge 0$  we have

$$\mathbb{P}\left\{|Y| \ge t\right\} \le 2 \exp\left[-c_1 \min\left(\frac{t^2}{\Sigma_1(X)}, \frac{t}{\Sigma_2(X)}\right)\right]$$
  
with  $\Sigma_1(X) = \sum_{i=1}^N \|X_i\|_{\Psi_1}^2$  and  $\Sigma_2(X) = \max_i \|X_i\|_{\Psi_1}$ 

**Lemma 8.** (Hanson-Wright inequality, [11]). Let  $X = (X_1, \ldots, X_n) \in \mathbb{R}^n$  be a random vector with independent, mean zero, sub-gaussian coordinates which satisfy  $\max_i ||X_i||_{\Psi_2} \leq K$ . Let M be an  $n \times n$  matrix and  $Y = X^T M X - \mathbb{E} X^T M X$ , then  $\forall t \geq 0$  we have

$$\mathbb{P}\left\{|Y| \ge t\right\} \le \exp\left[-c_2 \min\left(\frac{t^2}{K^4 \|M\|_F^2}, \frac{t}{K^2 \|M\|}\right)\right]$$

Proof of Lemma 1. By assumption  $A_i = \frac{1}{\sqrt{n}}B_i$  are independent sub-gaussian random vectors with  $||B_i||_{\Psi_2} \leq K$ . Thus  $X_i = \langle B_i, v_\perp \rangle$  are independent sub-gaussian random variables with  $\mathbb{E}X_i^2 = ||v_\perp||^2$  and  $||X_i||_{\Psi_2} = ||\langle B_i, \frac{v_\perp}{||v_\perp||_2}\rangle||_{\Psi_2} ||v_\perp||_2 \leq ||B_i||_{\Psi_2} ||v_\perp|| \leq K ||v_\perp||$ . Let  $Y_i = X_i^2 - \mathbb{E}X_i^2 = X_i^2 - ||v_\perp||^2$ , then Lemma 5 implies  $Y_i$  are independent, mean zero sub-exponential random variables with  $||Y_i||_{\Psi_1} \leq CK^2 ||v_\perp||^2$ .

Note that  $||Av_{\perp}||^2 - \frac{m}{n} ||v_{\perp}||^2 = \frac{1}{n} \sum_{i=1}^{m} Y_i$ , hence

$$\mathbb{P}\left[\|Av_{\perp}\|^{2} \leq \alpha \frac{m}{n} \|v_{\perp}\|^{2}\right] = \mathbb{P}\left[\sum_{i=1}^{m} Y_{i} \leq (1-\alpha)m\|v_{\perp}\|^{2}\right]$$

$$\stackrel{\text{Lemma 7}}{\leq} \exp\left\{-C_{1}\min\left(\frac{m(1-\alpha)^{2}}{C^{2}K^{4}}, \frac{(1-\alpha)m}{CK^{2}}\right)\right\}$$

Following similar argument for the second part completes the proof.  $\Box$ 

Proof of Lemma 2. As we argued previously  $X_i = \langle B_i, v_\perp \rangle$ are independent sub-gaussian random variables with mean zero and  $\mathbb{E}X_i^2 = \|v_\perp\|^2$ ,  $\|X_i\|_{\Psi_2} \leq K\|v_\perp\|$ . Note that  $\mathcal{P}_{AU}$  is the orthogonal projection operator with rank d, therefore  $M = \mathcal{P}_{AU} = QQ^T$  where  $\mathcal{P}_{AU} = Q\Sigma Q^T$ is thin singular value decomposition of  $\mathcal{P}_{AU}$ . Let  $X = (X_1, \ldots, X_m)$ , then  $\mathbb{E}X^TMX = \sum_{i=1}^m \mathbb{E}\langle X, Q_i \rangle = d\|v_\perp\|^2$ . Let  $Z_1 = \mathcal{P}_{AU}(Av_\perp)$  and  $Z_2 = X^TMX - \mathbb{E}X^TMX$ , it then follows that

$$\mathbb{P}\left[\|Z_1\|_2^2 \ge \beta_2 \frac{d}{n} \|v_\perp\|^2\right] = \mathbb{P}\left[\|Z_2\|^2 \ge (\beta_2 - 1)d\|v_\perp\|^2\right]$$

$$\stackrel{\text{Lemma 8}}{\le} \exp\left[-C_1 \Gamma(\beta_2 - 1)d\right]$$
where  $\Gamma(a) = \min\left\{a^2/K^4, a/K^2\right\}.$ 

Proof of Lemma 3. Note that  $\mathbb{I}_m - \mathcal{P}_{AU}$  is an projection matrix with rank m - d, which is independent of  $\xi$ . Therefore  $\mathbb{E} \| (\mathbb{I}_m - \mathcal{P}_{AU}) \xi \|^2 = \mathbb{E} \left[ \text{tr} \left\{ (\mathbb{I}_m - \mathcal{P}_{AU}) \mathbb{E} \xi \xi^T \right\} \right] = (m - d)\sigma^2$ . Then applying Lemma 8 with  $\delta = 1 - \alpha_2$  and  $\delta = \beta_2 - 1$  separately yields the results.  $\Box$ 

Proof of Lemma 4. Let  $Y = (\mathbb{I}_m - \mathcal{P}_{AU}) Av_{\perp}$ , then  $\mathbb{E} [\xi^T Y] = \mathbb{E}_Y [\mathbb{E}(\xi^T Y | Y)] = 0$ . Hence, Lemma 6 yields

$$\mathbb{P}\left(\xi^{T}Y \leq -\gamma\sigma\sqrt{m-d}\|Y\|\right) \leq \exp\left\{-C_{4}\frac{\gamma^{2}\sigma^{2}(m-d)}{K_{1}^{2}}\right\}$$

#### 6. REFERENCES

 M. Azizyan and A. Singh, "Subspace detection of high-dimensional vectors using compressive sampling," in 2012 IEEE Statistical Signal Processing Workshop (SSP). IEEE, 2012, pp. 724–727.

<sup>&</sup>lt;sup>2</sup>A random variable X with at least an exponential tail decay is called sub-exponential random variable. The sub-exponential norm of X is denoted as  $||X||_{\Psi_1}$ . For more information, please refer to [10].

- [2] B. A. Ardekani, J. Kershaw, K. Kashikura, and I. Kanno, "Activation detection in functional mri using subspace modeling and maximum likelihood estimation," *IEEE Transactions on Medical Imaging*, vol. 18, no. 2, pp. 101–114, 1999.
- [3] H. Kwon and N. M. Nasrabadi, "Kernel matched subspace detectors for hyperspectral target detection," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 2, pp. 178–194, 2006.
- [4] D. W. Stein, S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly detection from hyperspectral imagery," *IEEE signal processing magazine*, vol. 19, no. 1, pp. 58–69, 2002.
- [5] M. Rangaswamy, F. C. Lin, and K. R. Gerlach, "Robust adaptive signal processing methods for heterogeneous radar clutter scenarios," *Signal Processing*, vol. 84, no. 9, pp. 1653–1665, 2004.
- [6] K. Ahmed, S. Kothuri, M. Patwary, and M. Abdel-Maguid, "Subspace compressive glrt detector for airborne mimo radar," in 2010 16th Asia-Pacific Conference on Communications (APCC). IEEE, 2010, pp. 302–306.
- [7] Y. Yu, A. P. Petropulu, and H. V. Poor, "Measurement matrix design for compressive sensing-based mimo radar," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5338–5352, 2011.
- [8] M. L. McCloud and L. L. Scharf, "Interference estimation with applications to blind multipleaccess communication over fading channels," *IEEE Transactions on Information Theory*, vol. 46, no. 3, pp. 947–961, 2000.
- [9] L. Balzano, B. Recht, and R. Nowak, "Highdimensional matched subspace detection when data are missing," in 2010 IEEE International Symposium on Information Theory. IEEE, 2010, pp. 1638–1642.
- [10] R. Vershynin, "Introduction to the non-asymptotic analysis of random matrices," *arXiv preprint arXiv:1011.3027*, 2010.
- [11] M. Rudelson, R. Vershynin *et al.*, "Hanson-wright inequality and sub-gaussian concentration," *Electron. Commun. Probab*, vol. 18, no. 82, pp. 1–9, 2013.