Performance Estimation of Sparse Signal Recovery Under Bernoulli Random Projection with Oracle Information

Ruiyang Song, Laming Chen, and Yuantao Gu
State Key Laboratory on Microwave and Digital Communications
Tsinghua National Laboratory for Information Science and Technology
Department of Electronic Engineering, Tsinghua University, Beijing 100084, CHINA
{songry12, chen-lm06}@mails.tsinghua.edu.cn, gyt@tsinghua.edu.cn

***Abstract*—This article discusses the performance of the oracle receiver in recovering high dimensional sparse signals, which possesses the knowledge of the signals' support set. We consider a general framework, in which the sensing matrix and the measurements are disturbed simultaneously. The entries of the sensing matrix are i.i.d. Bernoulli random variables. We introduce the lower and upper bounds of the normalized mean square error of the reconstruction, which are proved to hold with high probability and verified by numerical simulations. The result is then compared with previous works on Gaussian sensing matrices. The average recovery error is derived as a generalization of the conclusion in [12] for the Gaussian ensemble and measurement noise only case.**

***Index Terms*—Sparse signal reconstruction, oracle receiver, Bernoulli random matrices, mean square error analysis**

I. Introduction

One of the fundamental issues in compressed sensing [5], [8], [15] is sparse signal recovery, which aims to recover a sparse or compressible vector from an underdetermined linear system. Unfortunately, the noises and perturbations are inevitable in this real world. Considering additive noise to the measurements, the model

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \tag{1}$$

has been well studied from several perspectives. In [9] it is shown that the convex programming problem returns a good recovery since the ℓ_2 distance between the solution and the sparse signal is bounded by a constant times the ℓ_2 norm of $\mathbf{n}$ [10]. A Dantzig selector on the noise model has been introduced in [7] and analyzed in [4]. This setup is also discussed in [3] where the Cramer-Rao bound is derived.

Models with more general noises and perturbations have also been introduced in literatures. Considering the additive noise to the sparse signal, the model $\mathbf{y} = \mathbf{A}(\mathbf{x} + \mathbf{e}) + \mathbf{n}$ has been discussed in [1] and [13]. Note that the term $\mathbf{Ae}$ involves the noise folding effect, and it can be equivalently regarded as the measurement noise after carefully calculating the equivalent signal-to-noise ratio [1]. Another interesting model considers the perturbation of the sensing matrix, which can be elaborated as $\mathbf{y} = (\mathbf{A} + \mathbf{E})\mathbf{x} + \mathbf{n}$ [11], [14]. This setup is also discussed in [19] and two lower bounds of the mse are introduced, including the constrained Cramer-Rao bound (CCRB) and the Hammersley Chapman-Robbins bound (HCRB). Note that in this model $\mathbf{Ex}$ contains information of $\mathbf{x}$ and cannot be directly regarded as the measurement noise [19].

Based on the mathematical results of Wishart distribution [17], [23], Coluccia et al. [12] analyzed the oracle receiver which possesses the exact information of the support set with Gaussian sensing matrix $\mathbf{A}$ in model (1). Besides Gaussian sensing ensembles, Bernoulli ones are another popular choice due to its low computational complexity involved. The Restricted Isometry Properties of these matrices are studied in [18]. The authors of [6] analyzed the Gaussian case in depth and presented the probability of exact reconstruction, and generalized the results to sub-Gaussian case. In [21], the authors also considered the noisy setup and particularly mentioned the Gaussian and Bernoulli ensembles. Concrete discussions of the Bernoulli case can also be found in [24]. Similar to [12], in this paper we aim to study the performance of sparse signal recovery with oracle information but using Bernoulli sensing matrices and considering the general noise model instead.

A. Notations

Denote the support set of vector $\mathbf{v}$ by $\mathrm{supp}(\mathbf{v})$. For a finite subset $\mathcal{S}$ of positive integers and any matrix $\mathbf{B}$, denote $\mathbf{B}_\mathcal{S}$ as the submatrix of $\mathbf{B}$ generated by picking the columns indexed in $\mathcal{S}$. Similarly, for any vector $\mathbf{v}$, denote by $\mathbf{v}_\mathcal{S}$ the vector derived by taking the entries of $\mathbf{v}$ indexed in $\mathcal{S}$. Let $s_{\max}(\mathbf{B})$ and $s_{\min}(\mathbf{B})$ be the largest and smallest singular values, respectively. For a square matrix $\mathbf{D}$, denote $\lambda(\mathbf{D})$ the set of its eigenvalues. The notations $\|\cdot\|_0$, $\|\cdot\|_2$, and $\|\cdot\|_F$ represent the ℓ_0, ℓ_2, and Frobenius norm respectively.

B. Problem setup

In this article, we consider the general framework formulated as

$$\mathbf{y} = (\mathbf{A} + \mathbf{E})\mathbf{x} + \mathbf{n}. \tag{2}$$

Suppose the underlying sparse signal $\mathbf{x}$ can be decomposed as $\mathbf{x} = \mathbf{\Psi}\boldsymbol{\theta}$, where $\mathbf{\Psi} \in \mathbb{R}^{N\times N}$ represents an orthonormal matrix, and $\boldsymbol{\theta}$ is a vector in $\mathbb{R}^N$ with $\mathrm{supp}(\boldsymbol{\theta}) = \Omega$ and $|\Omega| = K \ll N$. The sensing matrix $\mathbf{A} \in \mathbb{R}^{M\times N} (K < M < N)$ comprises of i.i.d. Bernoulli entries with sample space $\{-\sigma_A, \sigma_A\}$ and probability parameter $p = 1/2$, where $\sigma_A > 0$ is a constant. $\mathbf{E} \in \mathbb{R}^{M\times N}$ is a perturbation matrix with i.i.d. Gaussian distributed entries with zero mean and variance σ_E^2. The measurement noise $\mathbf{n} \sim \mathcal{N}(0, \mathbf{\Sigma}_n)$, where $\mathbf{\Sigma}_n = \sigma_n^2\mathbf{I}$.

The performance of the receiver can be evaluated by the normalized mean square error (mse_0) of the reconstructed signal $\widehat{\mathbf{x}}$ defined as follows:

$$\mathrm{mse}_0 = \mathbb{E}\|\mathbf{x} - \widehat{\mathbf{x}}\|_2^2/\mathbb{E}\|\mathbf{x}\|_2^2. \tag{3}$$

Consider the oracle estimator which possesses the exact knowledge of the support set Ω. Let $\mathbf{U} = \mathbf{A}\mathbf{\Psi}$. The estimator can therefore reconstruct $\widehat{\boldsymbol{\theta}}$ as $\widehat{\boldsymbol{\theta}}_{\mathbf{\Omega}} = \mathbf{U}_{\mathbf{\Omega}}^\dagger\mathbf{y}$, $\widehat{\boldsymbol{\theta}}_{\mathbf{\Omega}^c} = \mathbf{0}$, and therefore recover the signal $\mathbf{x}$. This paper develops an approach based on theories in random matrices to achieve concrete and explicit results of the oracle estimator for the general noise model defined in (2). We primarily dedicate in studying the performance of Bernoulli random matrices, while we also generalize the result on Gaussian sensing matrix [12] to (2) in Section V.

II. Main Result

We introduce some notations for better demonstration of our results. Denote by

$$\gamma = M/K \tag{4}$$

the redundancy coefficient. Let

$$p_x = \mathbb{E}(\mathbf{x}^\mathsf{T}\mathbf{x}) \tag{5}$$

be the average intensity of signal. As in [13], we introduce MSNR which is the SNR concerning the measurement noise $\mathbf{n}$:

$$\mathrm{MSNR} = M\sigma_A^2 p_x/\mathrm{tr}(\mathbf{\Sigma}_n). \tag{6}$$

Besides, we define the average relative power of measurement to perturbation as

$$\eta = \sigma_A^2/\sigma_E^2. \tag{7}$$

The following proposition summarizes the main result of this article, showing the probabilistic lower and upper bounds for the normalized mean square error. A sketch of the proof is given in Section III-D.

Proposition 1 (Lower and upper bounds of mse_0)**.** *Assume $\mathbf{\Psi} = \mathbf{I}$. The entries of the sensing matrix $\mathbf{A} \in \mathbb{R}^{M\times N}$ are i.i.d. randomly selected from $\{-\sigma_A, \sigma_A\}$ with equal probability, the perturbation matrix $\mathbf{E} \in \mathbb{R}^{M\times N}$ has i.i.d. entries that follow $\mathcal{N}(0, \sigma_E^2)$, and the measurement noise $\mathbf{n}$ is white Gaussian and follows $\mathcal{N}(0, \sigma_n^2\mathbf{I})$. Further assume $\gamma^2 \gg 1$ and $K \gg 1$. For some $t > 0$ and constants $C, c > 0$ that depend only on $\sqrt{M}$ and σ_A, let*

$$\theta_0 = \left(\frac{\sqrt{M} - C\sqrt{K} - t}{\sqrt{M}}\right)^2, \tag{8}$$

$$f_{M,K,\theta_0} \triangleq \frac{1}{\theta_0}\left(1 - \frac{(1-\theta_0)^2}{1-\theta_0+\frac{K+\gamma-1}{\gamma K}}\right), \tag{9}$$

and

$$\gamma_f = \gamma/f_{M,K,\theta_0}. \tag{10}$$

Suppose the sparse signal $\mathbf{x}$ satisfies $\|\mathbf{x}\|_0 = K$ and the support set Ω is provided to the receiver by an oracle, then the normalized mean square error is lower bounded by

$$\mathrm{mse}_0 \geq \frac{1}{\gamma\eta} + \frac{1}{\gamma}\frac{1}{\mathrm{MSNR}} \tag{11}$$

with probability $1 - p_S$, and upper bounded by

$$\mathrm{mse}_0 \leq \frac{1}{\gamma_f\eta} + \frac{1}{\gamma_f}\frac{1}{\mathrm{MSNR}} \tag{12}$$

with probability exceeding $1 - 1/\gamma^2 - 2\exp(-ct^2) - p_S$, where

$$p_S = (1+o(1))K(K-1)/2^M = \mathcal{O}(K^2/2^M). \tag{13}$$

Remark 1. *Proposition 1 presents a lower bound of the normalized average recovery error of the oracle estimator that holds with overwhelmingly high probability. The bounds in (11) and (12) are both linear in either one of $1/\mathrm{MSNR}$ and $1/\eta$ when the other is given.*

Remark 2. *This proposition establishes a probablistic upper bound of mse_0 in the Bernoulli case. By taking γ^2 and parameter t sufficiently large, we can derive an upper bound that holds with high probability.*

Recall that in Proposition 1 we assume $\gamma^2 \gg 1, K \gg 1$. Furthermore, if we have $\gamma \gg 1$, then $\theta_0 \approx 1 - 2C/\sqrt{\gamma}$. Hence, the following corollary can be naturally derived:

Corollary 1. *When $\gamma \gg 1$, let $\widehat{\gamma} = \gamma^2/(\gamma+1)$. We have $\gamma_f \geq \widehat{\gamma}$. Thus, we can introduce the following lower and upper bounds*

$$\frac{1}{\gamma\eta} + \frac{1}{\gamma}\frac{1}{\mathrm{MSNR}} \leq \mathrm{mse}_0 \leq \frac{1}{\widehat{\gamma}\eta} + \frac{1}{\widehat{\gamma}}\frac{1}{\mathrm{MSNR}}, \tag{14}$$

which hold with high probability.

Remark 3. *When the matrix perturbation $\mathbf{E} = \mathbf{0}$, we have $1/\eta = 0$. Hence, the inequalities can be simplified as*

$$\frac{1}{\gamma}\frac{1}{\mathrm{MSNR}} \leq \mathrm{mse}_0 \leq \frac{1}{\widehat{\gamma}}\frac{1}{\mathrm{MSNR}}. \tag{15}$$

Thus, the lower and upper bounds of the normalized mean square error are proportional to $1/\mathrm{MSNR}$.

Remark 4. *When the measurement noise is zero, i.e. $\mathbf{n} = \mathbf{0}$, we have $\mathrm{MSNR} = \infty$. In this case,*

$$\frac{1}{\gamma\eta} \leq \mathrm{mse}_0 \leq \frac{1}{\widehat{\gamma}\eta}. \tag{16}$$

Hence, the lower and upper bounds of mse_0 are proportional to $1/\eta$.

III. Theoretical Analysis

A. Preparation

A tall Bernoulli matrix is full-rank with high probability [16], [20]. The following lemma will be used to prove the main result.

Lemma 1 (Column independence [16])**.** *Suppose $\mathbf{A}_\Omega$ is a ± 1 random matrix in $\mathbb{R}^{M\times K}$. The probability that the columns of $\mathbf{A}_\Omega$ are linearly dependent is p_S as defined in (13).*

Given the condition that $N > M > K$, the columns of $\mathbf{A}_\Omega$ are linearly independent with high probability.

In later sections, we need bounds for the trace of the inverse of a positive definite matrix. The following lemma from [2] comes in handy.

Lemma 2 (Bounds of trace [2])**.** *Let $\mathbf{A}$ be an n-by-n symmetric positive definite matrix. Let $\mu_1 = \mathrm{tr}(\mathbf{A})$, $\mu_2 = \|\mathbf{A}\|_F^2$, and the set of eigenvalues $\lambda(\mathbf{A}) \subset [\alpha, \beta]$ with $0 < \alpha \leq \beta$. Then we have*

$$\begin{bmatrix}\mu_1 & n\end{bmatrix}\mathbf{J}_A^{-1}(\beta)\begin{bmatrix}n\\1\end{bmatrix} \leq \mathrm{tr}(\mathbf{A}^{-1}) \leq \begin{bmatrix}\mu_1 & n\end{bmatrix}\mathbf{J}_A^{-1}(\alpha)\begin{bmatrix}n\\1\end{bmatrix}, \tag{17}$$

where

$$\mathbf{J}_A(\xi) = \begin{bmatrix}\|\mathbf{A}\|_F^2 & \mathrm{tr}(\mathbf{A})\\ \xi^2 & \xi\end{bmatrix}. \tag{18}$$

Lemma 2 shows an upper bound of $\mathrm{tr}(\mathbf{A}^{-1})$, which is directly related to $\mathbf{A}$'s smallest eigenvalue. In [22], a probabilistic result is shown concerning the lower bound of the least singular value of random matrices. Before introducing the lemma, it is necessary to specify the definition of an isotropic random vector.

Definition 1 (isotropic vector [22])**.** *A random vector $\mathbf{X} \in \mathbb{R}^n$ is called isotropic if $\mathbb{E}((\mathbf{X}^\mathsf{T}\mathbf{x})^2) = \|\mathbf{x}\|_2^2$ for all $\mathbf{x} \in \mathbb{R}^n$.*

Obviously, a ± 1 Bernoulli random vector is isotropic, which is a simple result also shown in [22]. The following lemma presents a probabilistic result on tall matrices with independent sub-Gaussian isotropic columns.

Lemma 3 (Bounds of singular values [22])**.** *Let $\mathbf{H}$ be an N-by-n matrix ($n \leq N$) whose columns are independent sub-Gaussian isotropic random vectors in $\mathbb{R}^N$ with $\|\mathbf{H}_j\|_2 = \sqrt{N}$. Then for every $t \geq 0$, the inequality $\sqrt{N} - C\sqrt{n} - t \leq s_{\min} \leq s_{\max} \leq \sqrt{N} + C\sqrt{n} + t$ holds with probability at least $1 - 2\exp(-ct^2)$, where $C, c > 0$ depend only on the sub-Gaussian norm $\max_j \|\mathbf{H}_j\|_2$ of the columns.*

Recall that for ± 1 Bernoulli matrices, $\|\mathbf{H}_j\|_2 = \sqrt{M}$. Therefore, for the sensing matrix in Proposition 1, C, c are constants depending only on $\sqrt{M}$ and σ_A.

The framework in (2) can be reformulated as:

$$\mathbf{y} = \mathbf{A}\mathbf{\Psi}\boldsymbol{\theta} + (\mathbf{E}\mathbf{\Psi}\boldsymbol{\theta} + \mathbf{n}) = \mathbf{U}\boldsymbol{\theta} + \mathbf{z},$$

where $\mathbf{U} = \mathbf{A}\mathbf{\Psi}$ and $\mathbf{z} = \mathbf{E}\mathbf{\Psi}\boldsymbol{\theta} + \mathbf{n}$. Note that $\mathbf{A}$, $\boldsymbol{\theta}$ and $\mathbf{n}$ are independent, $\mathbf{U}_\Omega$ and $\mathbf{z}$ are independent. When the columns of $\mathbf{U}_\Omega$ are independent,

$$\text{mse} = \mathbb{E}\|\boldsymbol{\theta}_\Omega - \widehat{\boldsymbol{\theta}}_\Omega\|_2^2 = \mathbb{E}\|\mathbf{U}_\Omega^\dagger \mathbf{z}\|_2^2 = \mathbb{E}[\mathbf{z}^\text{T}(\mathbf{U}_\Omega\mathbf{U}_\Omega^\text{T})^\dagger \mathbf{z}]. \quad (19)$$

The expectation of $\mathbf{z}^\text{T}\mathbf{z}$ can be decomposed into

$$\mathbb{E}(\mathbf{z}^\text{T}\mathbf{z}) = \text{tr}(\mathbf{\Sigma}_n) + M\sigma_E^2 p_x. \quad (20)$$

B. Lower Bound

In the following analysis we assume $\mathbf{\Psi} = \mathbf{I}$. Hence, we have $\mathbf{U} = \mathbf{A}\mathbf{\Psi} = \mathbf{A}$, and the measurement

$$\mathbf{y} = \mathbf{A}\boldsymbol{\theta} + (\mathbf{E}\boldsymbol{\theta} + \mathbf{n}).$$

Together with (19),

$$\text{mse} = \mathbb{E}[\mathbf{z}^\text{T}\mathbb{E}[(\mathbf{A}_\Omega\mathbf{A}_\Omega^\text{T})^\dagger]\mathbf{z}].$$

Let $\mathbf{B} = \mathbf{A}_\Omega\mathbf{A}_\Omega^\text{T}$, and $\mathbf{C} = \mathbf{A}_\Omega^\text{T}\mathbf{A}_\Omega$. With Lemma 1, the columns of $\mathbf{A}_\Omega$ are independent with high probability. When $\mathbf{A}_\Omega$ is full-rank, $\mathbf{B}$ has K positive eigenvalues $\lambda_1, \ldots, \lambda_K$. Suppose $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \cdots, \lambda_K\}$. Hence, the set of nonzero eigenvalues of $\mathbf{B}^\dagger$ are the diagonal terms of $\mathbf{\Lambda}^\dagger = \text{diag}\{1/\lambda_1, \cdots, 1/\lambda_K\}$. Also, when $\text{rank}(\mathbf{A}_\Omega) = K$, $\mathbf{C}$ is positive definite and $\text{tr}(\mathbf{B}^\dagger)$=$\text{tr}(\mathbf{C}^{-1})$.

The following lemma shows a lower bound of the trace of $\mathbf{B}^\dagger$. This lemma can be easily derived by applying Cauchy's inequality.

Lemma 4 (lower bound)**.** *The trace of $\mathbf{B}^\dagger$ is lower bounded as follows:*

$$\text{tr}(\mathbf{B}^\dagger) \geq \sigma_A^{-2}K/M.$$

C. Upper Bound

In the following discussion, suppose $\gamma^2 \gg 1$, where $\gamma = M/K$ is defined in (4). Thus we have $N > M = \gamma K > K \gg 1$. In the following analysis, let $\mu_1 = \text{tr}(\mathbf{C})$, and $\mu_2 = \|\mathbf{C}\|_F^2$. When $\mathbf{C}$ is positive definite, Lemma 2 leads to

$$\text{tr}(\mathbf{B}^\dagger) = \text{tr}(\mathbf{C}^{-1}) \leq \begin{bmatrix}\text{tr}(\mathbf{C}) & K\end{bmatrix} \mathbf{J}_C^{-1}(\alpha) \begin{bmatrix}K \\ 1\end{bmatrix} \triangleq U_\alpha, \quad (21)$$

where the matrix $\mathbf{J}$ is defined in (18). α is the smallest eigenvalue of $\mathbf{C}$ and also the square of the least singular value of $\mathbf{A}_\Omega$. The following lemma provides a concentration analysis of μ_2, which can be easily proved by applying Chebyshev's inequality.

Lemma 5 (Concentration analysis)**.** *If $\gamma^2 \gg 1$, then with high probability, μ_2 is close to its mean $\mathbb{E}(\mu_2)$. More rigorously, we have*

$$P(|\mu_2 - \mathbb{E}(\mu_2)| \geq M^2\sigma_A^4) \ll 1.$$

Note that the upper bound U_α as defined in (21) is a funciton of α. In the following lemma we find the probability that U_α is monotonically decreasing. This can be proved by applying Lemma 5 and Chebyshev's inequality.

Lemma 6 (Monotonicity)**.** *When $\mathbf{C}$ is positive definite, if $K \gg 1$, the upper bound of $\text{tr}(\mathbf{C}^{-1})$, i.e. U_α, decreases as α increases with high probability:*

$$\mathbb{P}(\frac{\partial U_\alpha}{\partial \alpha} < 0) \approx 1.$$

The following lemma shows an upper bound for the expectation of $\text{tr}(B^\dagger)$.

Lemma 7 (Upper bound)**.** *If the columns of $\mathbf{A}_\Omega$ are independent, the following inequality*

$$\mathbb{E}(\text{tr}(\mathbf{B}^\dagger)) = \mathbb{E}(\text{tr}(\mathbf{C}^{-1})) \leq \frac{K}{M\sigma_A^2} f_{M,K,\theta_0}$$

holds with probability at least $1 - 1/\gamma^2 - \exp(-ct^2)$, where θ_0, f_{M,K,θ_0}, C, and $c > 0$ are defined in Proposition 1.

This lemma can be proved by applying Lemma 3, Lemma 5, and Lemma 6.

D. A sketch of the proof of Proposition 1

In this section we provide a brief proof of the main result. Given the symmetry of $\mathbf{B}$, let $\mathbb{E}(\mathbf{B}^\dagger) = \mathbf{Q}\mathbf{\Lambda}_0^\dagger\mathbf{Q}^\text{T}$, where $\mathbf{Q}$ is an orthonormal matrix and $\mathbf{\Lambda}_0^\dagger = \text{diag}\{1/\lambda_{1,0}, \cdots, 1/\lambda_{K,0}\}$. Lemma 1 and (19) lead to that with probability $1 - p_S$, the mse can be elaborated as $\text{mse} = \mathbb{E}[\mathbf{z}^\text{T}\mathbf{Q}^\text{T}\mathbf{\Lambda}_0^\dagger\mathbf{Q}\mathbf{z}]$. Let $\mathbf{w} = \mathbf{Q}\mathbf{z}$. Note that $\mathbf{\Psi} = \mathbf{I}$, we have

$$\text{mse} = \frac{\mathbb{E}(\mathbf{w}_1^2)}{\lambda_{1,0}} + \cdots + \frac{\mathbb{E}(\mathbf{w}_K^2)}{\lambda_{K,0}},$$

and

$$\mathbb{E}(\mathbf{w}_i^2) = \sigma_E^2 p_x + \sigma_n^2.$$

1) Proof of Lower Bound: With Lemma 4, $\text{tr}(\mathbf{B}^\dagger) \geq \sigma_A^{-2}K/M$, then $\text{tr}(\mathbb{E}[\mathbf{B}^\dagger]) \geq \sigma_A^{-2}K/M$. Therefore,

$$\text{mse} \geq \frac{K}{M}\frac{(\sigma_E^2 p_x + \sigma_n^2)}{\sigma_A^2} = \frac{1}{\gamma}(\frac{1}{\eta} + \frac{1}{\text{MSNR}})p_x$$

with probability $1 - p_S$.

2) Proof of Upper Bound: When the columns of $\mathbf{A}_\mathbf{\Omega}$ are independent, with Lemma 7,

$$\mathbb{E}(\text{tr}(\mathbf{B}^\dagger)) \leq \frac{K}{M\sigma_A^2} f_{M,K,\theta_0}$$

with probability at least $1 - 1/\gamma^2 - 2\exp(-ct^2)$. Hence, similarly, we can show that the upper bound in (12) holds with probability exceeding $1 - 1/\gamma^2 - 2\exp(-ct^2) - p_S$. C and $c > 0$ are constants depending only on $\sqrt{M}$ and σ_A.

IV. Numerical Results

In this section we present numerical results of the normalized mse comparing with theoretical bounds. In the simulations, set $N = 625$, $K = 20$, $t = 1$, $\sigma_A = 0.1$, and assume $C = 1$. The non-zero components of the sparse signal $\boldsymbol{\theta}_\Omega$ are generated i.i.d. by a normal distributed random variable with zero mean and covariance $\mathbf{\Sigma}_\theta = \mathbf{I}$.

In Fig. 1 and Fig. 2, we set $M = 100$. In Fig. 1, the result shows mse_0 versus the reciprocal of MSNR. In this setting, $\sigma_E = 0.1$, and $1/\text{MSNR}$ takes values in $\{0.1, 0.6, 1.1, 1.6, 2.1\}$. In Fig. 2, we show the relation between mse_0 and the reciprocal of η. Here $\sigma_n = 0.1$, and $1/\eta$ takes values from $\{0.02, 0.22, 0.42, 0.62, 0.82\}$. In Fig. 3, we vary the value of M and hence γ takes values 2.5, 5, 7.5, 10, and 12.5. The figure is plotted for mse_0 when $\sigma_n = \sigma_E = 0.1$.

All the results are averaged in 200 trials. The numerical experiments verified the lower and upper bounds of the normalized mean square error introduced in Proposition 1.

V. Related Work

A closely related work is shown in [12]. The author studied the performance analysis of Gaussian sensing matrix when the information of the support set is available at the receiver. We introduce the following proposition as a generalization of the result in [12] considering $\mathbf{n}$ and $\mathbf{E}$ simultaneously.

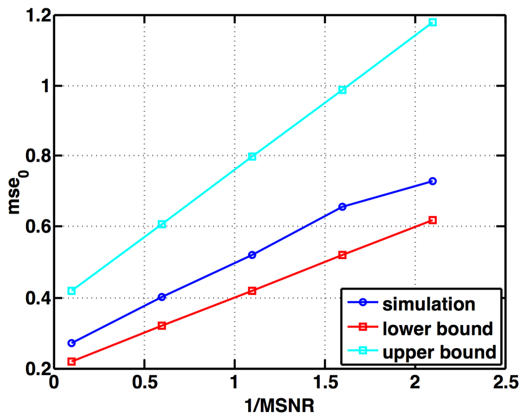

Fig. 1. The lower and upper bounds of mse_0 versus MSNR.

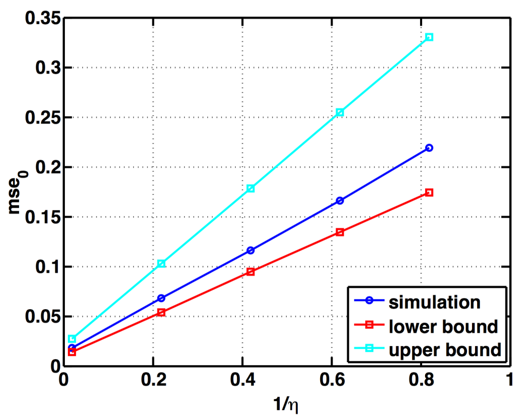

Fig. 2. The lower and upper bounds of mse_0 versus η.

Proposition 2 (Generalization of [12])**.** *For the general framework in (2), suppose the entries of the sensing matrix* $\mathbf{A} \in \mathbb{R}^{M\times N}$*, and the perturbation matrix* $\mathbf{E} \in \mathbb{R}^{M\times N}$*,* $(M < N)$*, are i.i.d. Gaussian random variables that follow* $\mathcal{N}(0, \sigma_A^2)$ *and* $\mathcal{N}(0, \sigma_E^2)$*, respectively.* $\mathbf{x} = \mathbf{\Psi}\boldsymbol{\theta}$*, where* $\mathbf{\Psi}$ *is an orthonormal matrix and* $\mathrm{supp}(\boldsymbol{\theta}) = \Omega$ *with* $K = |\Omega| < M - 3$*. Let* $\gamma_G \triangleq \gamma - 1 - 1/K$*. When the columns of* $\mathbf{U}_\Omega$ *are linearly independent, the average reconstruction error*

$$\mathrm{mse}_{0,G} = \frac{1}{\gamma_G \eta} + \frac{1}{\gamma_G}\frac{1}{\mathrm{MSNR}}. \tag{22}$$

Remark 5. *Note that in Proposition 2, when the perturbation matrix* $\mathbf{E} = \mathbf{0}$*, the equation reduces to the result in [12].*

Remark 6. *Note that this result has a very similar form as the bounds introduced for the Bernoulli case. Since* $\gamma_G < \gamma$*, it is interesting to see that the lower bound in (11) also holds for the Gaussian case.*

VI. Conclusion

In this work, we analyzed the performance of the oracle receiver in recovering high-dimensional sparse signal in a general framework with measurement noise and sensing matrix perturbation. The criterion of evaluating performance is the normalized mean square error of the reconstruction. We considered the case of Bernoulli sensing matrix and introduced the lower and upper bounds of the average recovery error. The impacts of perturbation matrix and measurement noise are discussed theoretically and also illustrated with numerical

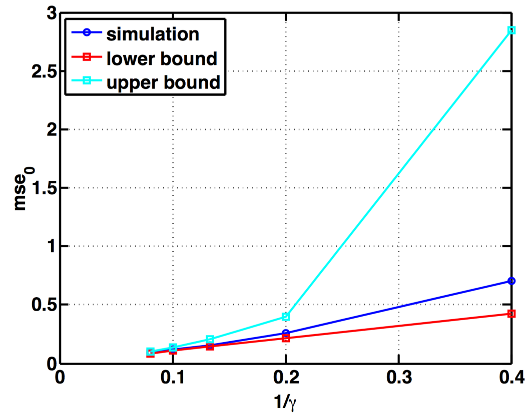

Fig. 3. The lower and upper bounds of mse_0 versus γ.

results. We also compare this work with previous results concerning Gaussian sensing matrices. Future work could be focused on finding the closed form solution of the mse in the Bernoulli case.

Acknowledgement

This work was partially supported by National Natural Science Foundation of China (NSFC 61371137, 61571263, 61531166005, 51459003) and Tsinghua University Initiative Scientific Research Program (Grant 2014Z01005).

References

[1] E. Arias-Castro and Y. C. Eldar. Noise folding in compressed sensing. *Signal Processing Letters, IEEE*, 18(8):478–481, 2011.

[2] Z. Bai and G. H. Golub. Bounds for the trace of the inverse and the determinant of symmetric positive definite matrices. *Annals of Numerical Mathematics*, 4:29–38, 1996.

[3] Z. Ben-Haim and Y. C. Eldar. The cramér-rao bound for estimating a sparse parameter vector. *Signal Processing, IEEE Transactions on*, 58(6):3384–3389, 2010.

[4] P. J. Bickel, Y. Ritov, and A. B. Tsybakov. Simultaneous analysis of lasso and dantzig selector. *The Annals of Statistics*, pages 1705–1732, 2009.

[5] E. Candès. Mathematics of sparsity (and a few other things). In *Proceedings of the International Congress of Mathematicians, Seoul, South Korea*, 2014.

[6] E. Candès and B. Recht. Simple bounds for recovering low-complexity models. *Mathematical Programming*, 141(1-2):577–589, 2013.

[7] E. Candes and T. Tao. The dantzig selector: statistical estimation when p is much larger than n. *The Annals of Statistics*, pages 2313–2351, 2007.

[8] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *Information Theory, IEEE Transactions on*, 52(2):489–509, 2006.

[9] E. J. Candes, J. K. Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on pure and applied mathematics*, 59(8):1207–1223, 2006.

[10] Emmanuel J Candes and Terence Tao. Decoding by linear programming. *Information Theory, IEEE Transactions on*, 51(12):4203–4215, 2005.

[11] L. Chen and Y. Gu. Oracle-order recovery performance of greedy pursuits with replacement against general perturbations. *Signal Processing, IEEE Transactions on*, 61(18):4625–4636, 2013.

[12] G. Coluccia, A. Roumy Aline, and E. Magli. Exact performance analysis of the oracle receiver for compressed sensing reconstruction. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pages 1005–1009. IEEE, 2014.

[13] M. Davenport, J. N. Laska, J. R. Treichler, and R. G. Baraniuk. The pros and cons of compressive sensing for wideband signal acquisition: Noise folding versus dynamic range. *Signal Processing, IEEE Transactions on*, 60(9):4628–4642, 2012.

[14] J. Ding, L. Chen, and Y. Gu. Robustness of orthogonal matching pursuit for multiple measurement vectors in noisy scenario. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 3813–3816. IEEE, 2012.

[15] D. L. Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, 52(4):1289–1306, 2006.

[16] J. Kahn, J. Komlós, and E. Szemerédi. On the probability that a random±1-matrix is singular. *Journal of the American Mathematical Society*, 8(1):223–240, 1995.

[17] D. Maiwald and D. Kraus. Calculation of moments of complex wishart and complex inverse wishart distributed matrices. *IEE Proceedings-Radar, Sonar and Navigation*, 147(4):162–168, 2000.

[18] H. Rauhut. Compressive sensing and structured random matrices. *Theoretical foundations and numerical methods for sparse recovery*, 9:1–92, 2010.

[19] Y. Tang, L. Chen, and Y. Gu. On the performance bound of sparse estimation with sensing matrix perturbation. *Signal Processing, IEEE Transactions on*, 61(17):4372–4386, 2013.

[20] T. Tao and V. Vu. On random±1 matrices: singularity and determinant. *Random Structures & Algorithms*, 28(1):1–23, 2006.

[21] Y. Tsaig and D. L. Donoho. Extensions of compressed sensing. *Signal processing*, 86(3):549–571, 2006.

[22] R. Vershynin. Introduction to the non-asymptotic analysis of random matrices. *arXiv preprint arXiv:1011.3027*, 2010.

[23] D. von Rosen. Moments for the inverted wishart distribution. *Scandinavian Journal of Statistics*, pages 97–109, 1988.

[24] G. Zhang, S. Jiao, X. Xu, and L. Wang. Compressed sensing and reconstruction with bernoulli matrices. In *Information and Automation (ICIA), 2010 IEEE International Conference on*, pages 455–460. IEEE, 2010.