Research Article
Noise Folding in Completely Perturbed Compressed Sensing

Limin Zhou, Xinxin Niu, and Jing Yuan

Information Security Center, Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Limin Zhou; zhoulimin.s@163.com

This paper first presents a new generally perturbed compressed sensing (CS) model $y = (A + E)(x + u) + e$, which incorporated a general nonzero perturbation E into sensing matrix A and a noise u into signal x simultaneously based on the standard CS model $y = Ax + e$ and is called noise folding in completely perturbed CS model. Our construction mainly will whiten the new proposed CS model and explore in restricted isometry property (RIP) and coherence of the new CS model under some conditions. Finally, we use OMP to give a numerical simulation which shows that our model is feasible although the recovered value of signal is not exact compared with original signal because of measurement noise e, signal noise u, and perturbation E involved.

1. Introduction

Compressed sensing (CS) model, which was proposed by Candes et al. [1] and Donoho [2], had become a hot topic and attracted a lot of researchers to study it over the past years because it can recover a signal as a technique. Thus, it had been widely applied in many areas such as radar systems [3], signal processing [4], and image processing [5]. These applications depended on the main function of CS model to recover the original signal with some related algorithms including convex relaxation [6, 7], greedy pursuit [7], and Bayesian algorithm [8, 9], which were utilized to estimate the best approximation value of the original signal.

The classic and basic CS model in an unperturbed scenario and can be formulated as

$$y = Ax. \tag{1}$$

Here, $y \in R^m$ is the measurement vector or observation value and $A \in R^{m \times n}$ is a full rank measurement matrix with $m \ll n$. Signal $x \in R^n$ is k-sparse if no more than k entries of signal x are nonzero. Thus, x is called a k-sparse signal.

To date, the basic model has mature theories and there are a lot of different algorithms [6, 7] such as match pursuit (BP) [1, 10], orthogonal match pursuit (OMP) [11–15], Compressive Sampling Matching Pursuit (CoSaMP) [16], and Bayesian algorithm [8, 9], which can be recovered exactly as signal value and utilized in many areas [17–20].

But in practice, the measurement vector y in (1) was often contaminated by a noise or an error. More concretely, a noise term $e \in R^m$, called *an additive noise*, was incorporated into $y = Ax$ to result in a *partially perturbed model* [21–23]:

$$y = Ax + e, \tag{2}$$

where a noise or an error e ($e \neq 0$) was uncorrelated with signal x. There were two methods to model noise e in [24]. Here, a noise e was randomly sampled from Gaussian distribution. This model was used in many areas [21–23] and naturally had mature theories in recent years. For example, a number of accuracy algorithms on (2) emerged, for example, BP [1, 21], OMP [21], CoSaMP [16], and Bayesian algorithm [8, 9].

In 2010, Herman and Strohmer [25] first incorporated a randomly nontrivial perturbation E into matrix A in (2) to generate a *general perturbation model* [25–27] as follows:

$$y = (A + E) x + e, \tag{3}$$

where $E \in R^{m \times n}$ was called *a general perturbation* or *a multiplicative noise*. They studied influence of E on signal x and indicated that considering this CS model was a must [25–27]. Intuitively, it was harder to analyze the multiplicative noise E compared to the additive noise e because E was related to signal x with Ex.

As for (3), there were two different scenarios from different points of view [25–27]. First, from user's point of view, the sensing process can be formulated as follows:

$$\widehat{y} = Ax + e,$$
$$\widehat{A} = A + E. \tag{4}$$

Its recovery process can be expressed by

$$\widehat{x} = \left(\widehat{y}, \widehat{A}, \ldots\right). \tag{N_1}$$

Thus, the useful measurement matrix was the perturbed matrix $\widehat{A}$ not the original measurement matrix A. The system was researched on the recovery signal with BP in [16, 25] and OMP in [26, 27].

The second model was from designer's perspective [25–27]. The sensing process was just written as

$$\widehat{y} = \widehat{A}x + e,$$
$$\widehat{A} = A + E \tag{5}$$

and the recovery process was written as

$$\widehat{x} = R\left(\widehat{y}, A, \ldots\right). \tag{N_1'}$$

The useful sensing matrix was A not $\widehat{A}$ and the observation value was $\widehat{y}$. To the best of our knowledge, no work focused on recovery signal in the context of general perturbation E except for [25–27].

In some practical scenarios, signal itself was often contaminated by noise and such case was applied in sub-Nyquist converter. Though introducing noise to signal was significant, no prolific paper studied such signal noise u except for [24] which first added an unknown random noise $u \in R^n$ to sinal x of $y = Ax + e$ to produce *noise folding CS model* [24]:

$$y = A(x + u) + e. \tag{6}$$

They analyzed the RIP and *coherence* of the equivalent system after whitening and showed that the difference of the RIP and *coherence* between original A and whitened matrix was small [24]. Based on [24–27], we propose a new CS model and study its related properties in Section 3.

2. Preliminaries

In this paper, we will restrict our attention to RIP and *coherence* of our new CS model. By convention, sensing matrix A and perturbation E are assumed to sample independently and identically distributed (i.i.d.) Gaussian random variables since such matrix satisfies RIP and coherence, and so forth [7, 24], with probability one.

Definition 1 (see [7]). A sensing matrix A satisfies the restricted isometry property (RIP) of order k if there exists $\delta_k \in (0, 1)$ s.t.:

$$(1 - \delta_k) \|x\|_2^2 \le \|Ax\|_2^2 \le (1 + \delta_k) \|x\|_2^2 \tag{7}$$

for any k-sparse vector with $k = 1, 2, \ldots$, where δ_k is the smallest nonnegative number called the restricted isometry constant (RIC).

Definition $1'$ (see [24]). For (1) and (2), there is another equivalent statement for RIP of A, denoted by RIP$'$, in some special cases. For any index set $\Lambda \subset \{1, \ldots, N\}$ of size k, let A_Λ denote the submatrix of A consisting of the column vectors indexed by Λ, and the matrix A possesses RIP$'$ with constants $0 < \alpha_k \le \beta_k$, if

$$\alpha_k \|h\|^2 \le \|A_\Lambda h\|^2 \le \beta_k \|h\| \quad \forall h \in R^k \tag{8}$$

for any index set $\Lambda \subset \{1, \ldots, N\}$ of size k, where N is a positive integer.

For (6), there existed another form of RIP for matrix A which was given by Lemma 2 [24] since matrix A was whitened.

Lemma 2 (see [24]). *As for folding noise model (6), RIP for whitened A can be formulated as*

$$\alpha_k (1 - \rho_1) \|h\|_2^2 \le \|B_\Lambda h\|_2^2 \le \beta_k (1 + \rho_1) \|h\|_2^2, \tag{9}$$

where $\rho_1 = \rho/(1 - \rho)$ with $0 < \rho < 1/2$ and matrix B was obtained after whitening sensing matrix A.

The perturbation E and sensing matrix A in (3) can be quantified below in [25–27]

$$\frac{\|E\|_2}{\|A\|_2} \le \varepsilon_A,$$
$$\frac{\|E\|_2^{(k)}}{\|A\|_2^{(k)}} \le \varepsilon_A^{(k)}, \tag{10}$$
$$\|A\|_2^{(k)} = \sigma_{\max}^{(k)}(A),$$

where the symbol $\|A\|_2$ denotes spectral norm of a matrix A, $\|A\|_2^{(k)}$ denotes the largest spectral norm taken over all k-column submatrices of matrix A, and $\sigma_{\max}^{(k)}(A)$ [25] denotes the largest nonzero singular value taken over all k-column submatrices of matrix A. It was appropriate to assume $0 < \varepsilon_A$, $\varepsilon_A^{(k)}$, and $\varepsilon_y \ll 1$.

Lemma 3 (RIP for $\widehat{A}$ [25]). *For $k = 1, 2, \ldots$, given the RIC associated with matrix A in (3) and the relative perturbation $\varepsilon_A^{(k)}$, fix the constant*

$$\widehat{\delta}_{k,\max} = (1 + \delta_k)\left(1 + \varepsilon_A^{(k)}\right)^2 - 1. \tag{11}$$

Assume that the RIC $\widehat{\delta}_k \le \widehat{\delta}_{k,\max}$ for matrix $\widehat{A} = A + E$ is the smallest nonnegative number, and the RIP for $\widehat{A}$ can be written as

$$(1 - \widehat{\delta}_k) \|x\|_2^2 \le \|\widehat{A}x\|_2^2 \le (1 + \widehat{\delta}_k) \|x\|_2^2 \tag{12}$$

for any k-sparse vector x.

From Lemma 3 with (6), there is one equivalent statement for the RIP for $\widehat{A}$ in some special cases given by Lemma 3' [24].

Lemma 3' (see [24]). *For any index set $\Lambda \subset \{1, \ldots, N\}$ of size k, let $\widehat{A}_\Lambda$ denote the submatrix of $\widehat{A}$ consisting of the column vectors indexed by Λ. A matrix $\widehat{A}$ possesses RIP with constants $0 < \widehat{\alpha}_k \le \widehat{\beta}_k$, if*

$$\widehat{\alpha}_k \|h\|^2 \le \left\|\widehat{A}_\Lambda h\right\|^2 \le \widehat{\beta}_k \|h\|, \quad \forall h \in R^k \qquad (13)$$

for any index set $\Lambda \subset \{1, \ldots, N\}$ of size k, where N is a positive integer.

Definition 4 (see [7]). *The coherence, $\mu(A)$, of a matrix A is the largest absolute inner product between any two columns $A_i, A_j, i \ne j$, of matrix A as follows:*

$$\mu(A) = \max_{1 \le i < j \le n} \frac{\left|A_i^T A_j\right|}{\|A_i\|_2 \|A_j\|_2}. \qquad (14)$$

3. Constructions

3.1. A New Completely Perturbed CS Model.
As mentioned above, for (2), (3), and (6), only one noise in (2) or two noises in (3) and (6) affected the CS model. Maybe a noise e, a noise u, and a perturbation E simultaneously affect the CS model although no paper studies this. In terms of the idea, [24] together with [25–27] motivate us to introduce a noise u to *general perturbation model* (3) to generate *noise folding in generally perturbed situation* or to incorporate a nontrivial perturbation E into (6) to produce *a complete perturbation CS model with folding noise*, which for the first time yields so called *noise folding in completely perturbed CS model*. We formulate the CS model as

$$y = (A + E)(x + u) + e, \qquad (15)$$

where $e \in R^m$ is a random noise vector with covariance $\sigma^2 I$ and $u \in R^n$ presents a random premeasurement noise vector whose covariance is $\sigma_0^2 I$ independent of e. Here e and u are regarded as *additive noise*. $E \in R^{m \times n}$ is a random perturbation matrix and more details on perturbation E can be seen in [25]. Here we call CS model (15) *noise folding in completely perturbed CS model*. Analogous to (3) in [25–27], (15) can also be considered in two different situations. Similarly, from user's point of view, an incorrect sensing matrix can be obtained via an unknown measurement model:

$$\begin{aligned} \widehat{y} &= A(x + u) + e, \\ \widehat{A} &= A + E \end{aligned} \qquad (16)$$

and the recovery process algorithm can be written as

$$\widehat{x} = R\left(\widehat{y}, \widehat{A}, u, \ldots\right). \qquad (N_2)$$

The only difference between (N_1) and (N_2) is noise u in (N_2). From the designer's view, sensing process can be formulated as

$$\begin{aligned} \widehat{y} &= \widehat{A}(x + u) + e, \\ \widehat{A} &= A + E \end{aligned} \qquad (17)$$

and its recovery process is as

$$\widehat{x} = R\left(\widehat{y}, A, u, \ldots\right). \qquad (N_2')$$

Similarly, compared to (N_1'), noise u belongs to (N_2'). In this paper, we only study simply its properties: RIP and *coherence* after whitening. Obviously, (15) can be extended to general multiperturbation CS model:

$$y = \left(A + \sum_{i=1}^{n} E_i\right)(x + u) + e, \quad i = 1, 2, \ldots, n, \qquad (18)$$

where E_i is perturbation. System (18) can be viewed as a generalization of our proposed CS (15), which implies that the general conclusion of (18) can be obtained from the special conclusion of (15). The concrete results can be seen in the next section. Simultaneously, other general CS systems can be conjectured naturally as follows:

$$\begin{aligned} y &= \left(\sum_{i=1}^{s} A_i + E\right)(x + u) + e, \\ y &= \left(\sum_{i=1}^{s} A_i + E\right)\left(x + \sum_{i=1}^{s} u_i\right) + e, \\ y &= \left(\sum_{i=1}^{s} A_i + \sum_{i=1}^{s} E_i\right)\left(x + \sum_{i=1}^{s} u_i\right) + e. \end{aligned} \qquad (19)$$

Although their properties seem to be many but we do not know how to exploit and analyze them, we leave them as open problems. Here we mainly study relative RIP and *coherence* on (15) and (18). In the next section, we give general results.

3.2. Problem Formulation.
For (15), our goal is to analyze the effect of the premeasurement noise u and E on its RIP and *coherence*.

Throughout this paper, assume that e is a random noise vector with covariance $\sigma^2 I$, and u is a random noise vector with covariance $\sigma_0^2 I$ independent of e. Under these assumptions, (15) will be proved to be equivalent to $y = \widehat{B}x + w$, where $\widehat{B}$ is a matrix whose *coherence* and RIP constants are very close to that of A, w is whitened noise with variance $(\sigma^2 + (n/m)\sigma_0^2)I$, and I is identity matrix.

3.3. Equivalent Formulation.
To set up our conclusion, (15) can be expressed as

$$\begin{aligned} y &= (A + E)x + w, \\ w &= (A + E)u + e. \end{aligned} \qquad (20)$$

By hypothesis of whitened noise, the covariance of effective vector w is $Q : Q = \sigma^2 I + \sigma_0^2 (A + E)(A + E)^T$. Obviously, noise w is not whitened where the recovery process analysis becomes complicated. If w still preserves whitening, one case $\widehat{A} = A + E$ must be proportional to identity matrix. For example, suppose that $A + E$ consists of $r = n/m$ orthogonal basis such as

$$A = [A_1 + E_1, A_2 + E_2, \ldots, A_r + E_r] \tag{21}$$

in which $A_i + E_i$, $i = 1, 2, \ldots, r$, are $m \times m$ orthogonal matrices. Therefore, we have

$$
\begin{aligned}
(A + E)(A + E)^T &= (A_1 + E_1)(A_1 + E_1)^T + \cdots \\
&\quad + (A_r + E_r)(A_r + E_r)^T = rI \\
&= \frac{n}{m} I,
\end{aligned}
\tag{22}
$$

where noise covariance of w is $Q = \gamma I$, $\gamma = \sigma^2 + (n/m)\sigma_0^2$. Under the special case, $y = (A + E)(x + u) + e$ (or $y = (A + E)x + w$) is equivalent to $y = Ax + e$. Compared with noise covariance of e, noise covariance of w has increased by γ/σ^2. If $\sigma_0^2 \approx \sigma^2$, the noise covariance of w is increased by n/m, which was called *noise folding* [24].

3.4. RIP and Coherence of Our CS Model. We will show that the conclusion holds generally. In other words, if $(A + E)(A + E)^T$ is not proportional to the identity matrix I, (15) and (20) are roughly equivalent really. Now we describe it in detail.

Note that if E is one random matrix, $\widehat{A} = A + E$ is a random matrix. To study RIP and *coherence* of $\widehat{A}$, we must whiten noise w by multiplying $Q_1^{-1/2}$ with $Q_1 = Q/\gamma$ and get the equivalent system:

$$
\begin{aligned}
y &= \widehat{B}x + v, \\
\text{s.t.} \quad \widehat{B} &= Q_1^{-1/2}(A + E), \\
v &= Q_1^{-1/2}w.
\end{aligned}
\tag{23}
$$

Note that noise vector v is whitened with covariance matrix γI exactly if $(A + E)(A + E)^T$ is proportional to identity matrix. But the biggest difference lies in measurement matrix changing from original matrix $\widehat{A}$ to $\widehat{B}$ by whitening. The changing range is measured with three important indexes: RIP constant, *coherence*, and *stable*. Our theory mainly depends on approximating $(A + E)(A + E)^T$ with $(n/m)I$ and even $\widehat{A}$ is an arbitrary matrix. Let

$$\eta = \left\| I - \frac{m}{n}(A + E)(A + E)^T \right\|_2 \tag{24}$$

measure accuracy of the approximating, in which $\|\cdot\|$ denotes the standard operator norm in R^n. For derivation convenient, assume that η is very small and show that the *coherence* and RIP constant of $\widehat{B}$ are very close to that of A. By convention, the entries of A are i.i.d. mean zero and variance $1/m$ random

variables with Gaussian distribution; thus, it is easy to justify that η is always small.

Another useful formula can be formulated:

$$\eta_0 = \left\| I - \frac{m}{n} AA^T \right\|_2 \tag{25}$$

which was introduced in [24]. The fact that η_0 was very small had been proved in [24] with restrictions on only matrix A. It is natural to think whether the difference between η and η_0 is very small. Theorem 5 confirms our conjecture and further inspires us to think whether the difference between $\widehat{B}$'s *coherence* and RIP and $\widehat{A}$'s and A's is very small, respectively. The later related theorems will give us the positive answers.

Theorem 5 shows the relation between η_0 and η under the context $\|E\|_2 / \|A\|_2 \leq \varepsilon_A$.

Theorem 5. *Assume that sensing matrix $A \in R^{m \times n}$, an unknown random matrix $E \in R^{m \times n}$, $m \ll n$, $\|E\|_2 / \|A\|_2 \leq \varepsilon_A$, $0 < \varepsilon_A \ll 1$,*

$$
\begin{aligned}
\eta_0 &= \left\| I - \frac{m}{n} AA^T \right\|_2, \\
\eta &= \left\| I - \frac{m}{n}(A + E)(A + E)^T \right\|_2
\end{aligned}
\tag{26}
$$

$0 < \eta_0 < 1/2$, $\|A\| = \sigma_1$, where σ_1 is the largest nonzero positive singular value of A; then

$$\eta_0 - \frac{m}{n}\left(2\varepsilon_A + \varepsilon_A^2\right)\sigma_1^2 \leq \eta \leq \eta_0 + \frac{m}{n}\left(2\varepsilon_A + \varepsilon_A^2\right)\sigma_1^2. \tag{27}$$

Proof. The detailed proof is postponed to the Appendix. □

Remark 6. For (27), by assumption $m \ll n$, $n \to \infty$, obtain

$$
\begin{aligned}
\frac{m}{n} &\longrightarrow 0, \\
\frac{m}{n}\left(2\varepsilon_A + \varepsilon_A^2\right)\sigma_1^2 &\longrightarrow 0
\end{aligned}
\tag{28}
$$

due to $0 < \varepsilon_A \ll 1$, a positive number σ_1. Thus, $\eta_0 \leq \eta_1 \leq \eta_0$; that is, $\eta_0 = \eta$, such that $m \ll n$, $n \to \infty$, $0 < \varepsilon_A \ll 1$ and there is a positive number σ_1. Theorem 5 shows the relation between η and η_0 which implies that $\eta = \eta_0$ under some special conditions. Therefore, we can let $\eta < 1/2$ like $\eta_0 < 1/2$ in [24].

Theorem 7 shows the RIP of $\widehat{B}$ in the case of $\eta < 1/2$ though $0 < \eta < 1$ is sufficient for the proof of the RIC for $\widehat{B}$.

Theorem 7. *Assume that sensing matrix $A \in R^{m \times n}$, an unknown random matrix $E \in R^{m \times n}$, $m \ll n$. Let $\|E\|_2 / \|A\|_2 \leq \varepsilon_A$, $0 < \varepsilon_A \ll 1$,*

$$
\begin{aligned}
\eta &= \left\| I - \frac{m}{n}(A + E)(A + E) \right\|_2, \\
\eta_0 &= \left\| I - \frac{m}{n} AA^T \right\|_2
\end{aligned}
\tag{29}
$$

$0 < \eta < 1/2$, $0 < \eta_0 < 1/2$, suppose that $\widehat{A}$ satisfies the RIP of order k with $0 < \widehat{\alpha}_\Lambda \le \widehat{\beta}_\Lambda$, and $\sigma_1 > 0$ is the largest singular value of matrix A; then $\widehat{B}$ satisfies the RIP of order k with different constants below:

$$\widehat{\alpha}_k \left(1 - \mu_1'\right) \|h\|_2^2 \le \left\|\widehat{B}_\Lambda h\right\|_2^2 \le \widehat{\beta}_k \left(1 + \mu_1'\right) \|h\|_2^2,$$
$$\mu_1' = \frac{\mu_1}{1 - \mu_1}$$

$$\widehat{\alpha}_k \left(1 - \eta_1'\right) \|h\|_2^2 \le \left\|\widehat{B}_\Lambda h\right\|_2^2 \le \widehat{\beta}_k \left(1 + \eta_1'\right) \|h\|_2^2,$$
$$\eta_1' = \frac{\eta_1}{1 - \eta_1} \tag{30}$$

$$\widehat{\alpha}_k \left(1 - \mu_2'\right) \|h\|_2^2 \le \left\|\widehat{B}_\Lambda h\right\|_2^2 \le \widehat{\beta}_k \left(1 + \mu_2'\right) \|h\|_2^2,$$
$$\mu_2' = \frac{\mu_2}{1 - \mu_2}.$$

Proof. The detailed proof is postponed to the Appendix. □

Remark 8. In Theorems 5 and 7, the condition $\|E\|/\|A\| \le \varepsilon_A$ with $0 < \varepsilon_A \ll 1$ can be taken in place of $E = \varepsilon A$ [25], in which E is a simple version of A, so that we can get another result. Due to paper volume, they are omitted here. But their proofs are very simple that researchers can prove them and yield perfect results.

Multiperturbation CS system (18) can be viewed as a generalization of the new proposed CS system (15) so that the general conclusion of the (18) can come from that of (15). Theorems 9 and 11 give us the results.

Theorem 9. *Assume that $A \in R^{m \times n}$ is sensing matrix, and $E_i \in R^{m \times n}$ is an unknown random matrix with $m \ll n$. Let*

$$\eta_0 = \left\|I - \frac{m}{n} A A^T\right\|_2,$$
$$\widetilde{\eta} = \left\|I - \frac{m}{n} \left(A + \sum_{i=1}^s E_i\right) \left(A + \sum_{i=1}^s E_i\right)^T\right\|_2. \tag{31}$$

$\|E_i\|_2/\|A\|_2 \le \varepsilon_A$, $0 < \varepsilon_A \ll 1$, *and let σ_1 be the largest singular value of matrix A, $\|A\| = \sigma_1$; and suppose that matrix A satisfies RIP, a number s is an integer, $s < n$; then the relation between $\widetilde{\eta}$ and η_0 can be formulated as*

$$\eta_0 - \frac{m}{n} \left(2s\varepsilon_A\sigma_1 + s^2\varepsilon_A^2\sigma_1\right) \le \widetilde{\eta}$$
$$\le \eta_0 + \frac{m}{n} \left(2s\varepsilon_A\sigma_1 + s^2\varepsilon_A^2\sigma_1\right). \tag{32}$$

Proof. The detailed proof is postponed to the Appendix. □

Remark 10. For (32), since $1 \le s \ll n$, $m \ll n$, and s, m, n are positive integers and σ_1 is a constant, $(m/n)(2s\varepsilon_A\sigma_1^2 + s^2\varepsilon_A^2\sigma_1^2) \to 0$ when $n \to \infty$ which implies $\widetilde{\eta} \to \eta_0$.

Theorem 11. *Assume that $A \in R^{m \times n}$ is sensing matrix, and $E_i \in R^{m \times n}$ is an unknown random matrix with $m \ll n$. Let*

$$\eta_0 = \left\|\frac{m}{n} A A^T - I\right\|_2,$$
$$\widetilde{\eta} = \left\|I - \frac{m}{n} \left(A + \sum_{i=1}^s E_i\right) \left(A + \sum_{i=1}^s E_i\right)^T\right\|_2, \tag{33}$$

$\|E_i\|_2/\|A\|_2 \le \varepsilon_A$, *and $0 < \varepsilon_A \ll 1$, and suppose that $\widetilde{\eta} < 1/2$, σ_1 is the largest singular value of matrix A, s and n are integers, $s < n$, and*

$$Q = \sigma^2 I + \sigma_0^2 \left(A + \sum_{i=1}^s E_i\right) \left(A + \sum_{i=1}^s E_i\right)^T,$$
$$\gamma = \sigma^2 + \frac{n}{m}\sigma_0^2, \tag{34}$$
$$Q_1 = \frac{Q}{\gamma};$$

then,

$$\widehat{\alpha}_k \left(1 - \mu_3'\right) \|h\|_2^2 \le \left\|\widehat{B}_\Lambda h\right\|_2^2 \le \widehat{\beta}_k \left(1 + \mu_3'\right) \|h\|_2^2,$$
$$\mu_3' = \frac{\mu_3}{1 - \mu_3}$$

$$\widehat{\alpha}_k \left(1 - \eta_3'\right) \|h\|_2^2 \le \left\|\widehat{B}_\Lambda h\right\|_2^2 \le \widehat{\beta}_k \left(1 + \eta_3'\right) \|h\|_2^2,$$
$$\eta_3' = \frac{\eta_3}{1 - \eta_3} \tag{35}$$

$$\widehat{\alpha}_k \left(1 - \mu_4'\right) \|h\|_2^2 \le \left\|\widehat{B}_\Lambda h\right\|_2^2 \le \widehat{\beta}_k \left(1 + \mu_4'\right) \|h\|_2^2,$$
$$\mu_4' = \frac{\mu_4}{1 - \mu_4}.$$

Proof. The detailed proof is postponed to the Appendix. □

Remark 12. Though $0 < \widetilde{\eta} < 1$ is sufficient for the proof, the RIC for $\widehat{B}$ is positive in the restriction of $\widetilde{\eta} < 1/2$.

Next, we compare the *coherence* of $\widehat{B}$ after whitening $\widehat{A}$ to that of $\widehat{A}$. A_i, $i = 1, \dots, n$, is used to denoted the ith column vector of a matrix A. Similar to the coherence of A, the coherence of $\widehat{A}$ is first given in Definition 13.

Definition 13. Assume that $A \in R^{m \times n}$ is a random matrix, $E \in R^{m \times n}$ is an unknown random matrix in CS, and $\widehat{A} = A + E$; then *coherence* of $\widehat{A}$, denoted by $\mu(\widehat{A})$, can be formulated as

$$\mu\left(\widehat{A}\right) = \max_{1 \le i < j \le n} \frac{\left|\widehat{A}_i^T \widehat{A}_j\right|}{\left\|\widehat{A}_i\right\|_2 \left\|\widehat{A}_j\right\|_2}. \tag{36}$$

In fact, $\mu(\widehat{A})$ is the largest absolute inner product between any columns $\widehat{A}_i$, $\widehat{A}_j$, $i \ne j$.

As mentioned above, $\|Q_1 - I\| < 1$, respectively, in some special contexts with $\|E\|_2/\|A\|_2 \leq \varepsilon_A$ in Theorem 7. We can take advantage of $\|Q_1 - I\| < 1$ to prove Theorem 14. For the lack of space, we only take $\|Q_1 - I\| < \mu_2$, $\|Q_1^{-1} - I\| < \mu_2'$ as an example with $\eta < 1/2$. The proofs of the rest of the cases, including $\|Q_1 - I\| < \mu_1, \eta_1$, and $\|Q_1^{-1} - I\| < \mu_1', \eta_1'$, are similar, and we leave them to readers. As for the general results $\mu(A + \sum_{i=1}^n E_i)$ of the general CS model $y = (A + \sum_{i=1}^n E_i)(x+u)+e$, we omit them too due to space constrains. The proof of general coherence of $Q_1^{-1/2}(A + \sum_{i=1}^n E_i)$ is similar too. Theorem 14 demonstrates the relation between coherence of $\widehat{B}$ and that of $\widehat{A}$.

Theorem 14. *Assume that $\mu_2 < 3/4$ in $\|Q_1 - I\| \leq \mu_2$, $\widehat{B} = Q_1^{-1/2}\widehat{A}$ with $\widehat{A} = A + E$; then,*

$$\mu\left(\widehat{B}\right) \leq \frac{(1 + \widehat{\mu}_2)}{(1 - \widehat{\mu}_2)^2}\mu\left(\widehat{A}\right), \tag{37}$$

where $\widehat{\mu}_2 = (1 - \mu_2)^{-1/2} - 1$, $\mu(\widehat{B}) = \max(|\widehat{B}_i^T\widehat{B}_j|/\|\widehat{B}_i\|_2\|\widehat{B}_j\|_2)$. $\widehat{B}_j$ denotes the jth column vector of whitening matrix $\widehat{B}$; $\widehat{A}_j$ denotes the jth column vector of $\widehat{A}$; that is, $\widehat{A}_j = A_j + E_j$.

Proof. The detailed proof is postponed to the Appendix. □

In [25], E is simply version of random matrix A such as $E = \varepsilon A$ with $0 < \varepsilon \ll 1$. The relation between $\mu(\widehat{A})$ and $\mu(A)$ will be seen from Theorem 15 below in the case of $E = \varepsilon A$, $0 < \varepsilon \ll 1$.

Theorem 15. *Let $\widehat{A} = A + E$, $E = \varepsilon A$, $0 < \varepsilon \ll 1$. The correlation of coherence between A and $\widehat{A}$ proceeds as*

$$\mu\left(\widehat{A}\right) \leq \frac{(1 + \varepsilon)^2}{(1 - \varepsilon)^2}\mu(A) \tag{38}$$

with $0 < \varepsilon \ll 1$.

Proof. The proof of Theorem 15 is similar to that of Theorem 14; here we omit it. □

4. Numerical Experiment

Vertical coordinate (sinusoidal) denotes the degree of recovery signal. Horizontal coordinate denotes time whose unit is seconds. Black line denotes original signal and red line denotes recovery signal.

Here we use OMP to give three numerical simulation results which demonstrate that our new proposed generally perturbed CS is feasible. To compare signal recovering with OMP from three figures, signal recovery from measurement noise model $y = Ax + e$ is almost exact because of only noise e in basic CS model $y = Ax + e$. There are a lot of differences between recovery signal and original signal in both $y = (A + E)(x + u) + e$ and $y = A(x + u) + e$ CS models because there are noises u, E, e in the two CS models. Comparing the change between recovered signal and original signal of Figure 1(b) with that of Figure 1(c), the change of
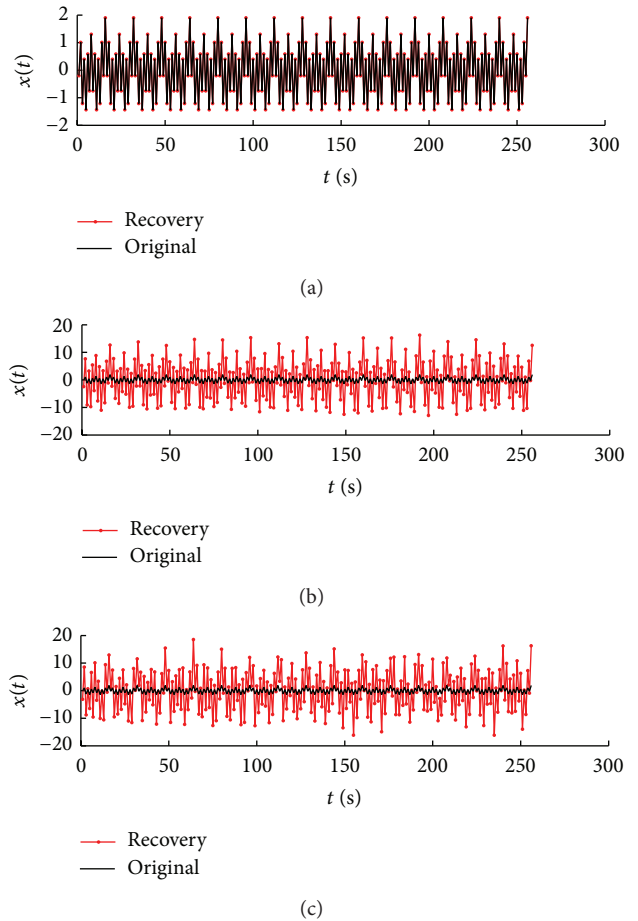

FIGURE 1: (a) Recover x from $y = Ax + e$. The error of the recovered signal value and original signal value is $6.4835e - 004$ by the Euclidean norm which indicates that the recovered signal and original signal are almost the same. (b) Recover x from $y = A(x+u)+e$. The error of the recovered signal value and the original signal value is 6.2258 by the Euclidean norm which means that the recovered signal and original signal are quite different. (c) Recover x from $y = (A+E)(x+u)+e$. The error of the recovered signal value and the original signal value is 5.8855 by the Euclidean norm which indicates that the recovered signal is much different from the original signal. Compared with the change (error) between the recovery signal and the original signal in (a), the changes (errors) in (b) and (c) differ little. What is more, comparing the change between the recovery signal and the original signal of (c) with that of (b), the change of (c) is quite a little bigger because E is involved in (c) but E is not involved in (b).

Figure 1(c) is a bit bigger than that of Figure 1(b) because perturbation E is involved in Figure 1(c) ($y = (A+E)(x+u)+e$) and E is not involved in Figure 1(b) ($y = A(x + u) + e$), which shows that the different noises E, u, e, have a different impact on signal recovering.

Compared with the change (error) between the recovery signal and the original signal in Figure 1(a), the changes (errors) in Figures 1(b) and 1(c) differ little. Namely, the differences between recovered signal and original signal from $y = (A + E)(x + u) + e$ are almost the same as the differences

between recovery signal and original signal from $y = A(x + u) + e$, which indicates that our proposed CS model is feasible.

Comparing the change between recovered signal and original signal in Figure 1(a), the changes in Figures 1(b) and 1(c) are quite different. The fact shows that OMP is not the best algorithm to recover x from $y = (A + E)(x + u) + e$ and $y = A(x + u) + e$ although OMP is used to recover exact original signal from $y = Ax + e$. Thus, it is important to search for a powerful algorithm or more algorithms to recover exactly for original sparse signal from $y = A(x + u) + e$ and $y = (A + E)(x + u) + e$ as open problems. And here leave these problems to the interested researchers to exploit them because the paper cannot focus on searching for optimal algorithms to recover exactly original signal from CS models $y = A(x + u) + e$ and $y = (A + E)(x + u) + e$.

5. Conclusion

We first propose a new CS system (15) by introducing a multiplicative noise E, a signal noise u, and an additive noise e into unperturbed CS model (1). We derive RIP and *coherence* for $\widehat{A} = A + E$ after whitening (15). As a matter of fact, this paper proves that our proposed completely perturbed CS model (15) equals to the classic CS model (2). The only difference is the changed measurement matrix by incorporating a nontrivial perturbation matrix E to measurement matrix and a nontrivial noise u to signal x. And thus this induces noise variance increased by a factor of n/m so that a tighter upper bound and lower bound of RIC is produced. As for *coherence* of deformed measurement matrix $\widehat{A} = A + E$ in CS model (15), the constant is nearly invariant essentially with $n/m \to 0$, $m, n \to \infty$. Finally, we use OMP to give three figures to recover signal from CS model $y = Ax + e$, $y = A(x + u) + e$, $y = (A + E)(x + u) + e$, respectively. Figures 1(b) and 1(c) in our experiment demonstrate that the change between recovered signal and original signal is much bigger than that in Figure 1(a) which indicates our proposed CS model is feasible and OMP is not fit for recovering signal from $y = A(x + u) + e$ and $y = (A + E)(x + u) + e$. Thus, we can try to search one optimal algorithm or more algorithms to recover signal exactly from the two CS model although OMP is the best algorithm to recover exactly original signal from $y = Ax + e$ now.

6. Future Work

Thanks to the features of our proposed CS model (15), there are many works to do. The change between recovered signal and original signal in Figures 1(a), 1(b), and 1(c) indicates that our proposed CS in this paper is feasible although the differences between the recovery signal and original signal in Figures 1(b) and 1(c) are much bigger with OMP than the differences in Figure 1(a). Thus, an obvious problem is to search one algorithm or more optimal algorithms suitable for $y = A(x + u) + e$ and $y = (A + E)(x + u) + e$ to recover signal exactly.

The related RIP of E in [24] further motivates us to think that E as a perturbation sensing matrix could form one perturbed CS model as $y_1 = E(x + u) + e$. Thus, (15) may consist of two similar systems $y_2 = A(x + u) + e$ and $y_1 = E(x + u) + e$. Similarly, our model may be divided into another two

models $y_3 = (A + E)x + e$, $y_4 = (A + E)u + e$, or three basic parts $y'_1 = Ax + e$, $y'_2 = Au + e$, and $y'_3 = E(x + u) + e$. If possible, what can we do to reduce or eliminate the influence of an error CS system $y'_3 = E(x + u) + e$? Can we recover signal x from the error system $y'_3 = E(x + u) + e$? And if can, how to do it? Maybe there exists CS models $y = A(x + u)$ and $y = (A + E)(x + u)$. In addition, maybe we can also consider the impulse noise and use $e + I$ instead of e where I is the impulse noise. If so, maybe it can generalize our model and we can get very good results in impulse noise model. Here we cannot study such impulse noise model and leave it as an open problem, too.

These open problems are worth considering and are to be waited for studying in future work. This paper only does some elementary researches on our proposed CS and we hope that the idea and simple study in this paper will be helpful to study its wide application in the future. We hope that higher level compressed sensing model will be put forward and more and more people explore this areas in the future.

Appendix

Proof of Theorem 5. On the one hand,

$$
\begin{aligned}
\eta &= \left\| I - \frac{m}{n} (A + E) (A + E)^T \right\|_2 \\
&= \left\| I - \frac{m}{n} AA^T - \frac{m}{n} \left(AE^T + EA^T + EE^T \right) \right\|_2 \\
&\geq \left\| I - \frac{m}{n} AA^T \right\|_2 - \frac{m}{n} \left\| AE^T \right\|_2 - \frac{m}{n} \left\| EA^T \right\|_2 \\
&\quad - \frac{m}{n} \left\| EE^T \right\|_2 \\
&\geq \eta_0 - \frac{m}{n} \|A\|_2 \left\| E^T \right\|_2 - \frac{m}{n} \|E\|_2 \left\| A^T \right\|_2 \\
&\quad - \frac{m}{n} \|E\|_2 \left\| E^T \right\|_2 \\
&\geq \eta_0 - \frac{m}{n} \|A\|_2 \, \varepsilon_A \left\| A^T \right\|_2 - \frac{m}{n} \varepsilon_A \|A\|_2 \left\| A^T \right\|_2 \\
&\quad - \frac{m}{n} \varepsilon_A \|A\|_2 \, \varepsilon_A \left\| A^T \right\|_2 \\
&= \eta_0 - \frac{m}{n} \left(2\varepsilon_A + \varepsilon_A^2 \right) \|A\|_2 \left\| A^T \right\|_2 \\
&= \eta_0 - \frac{m}{n} \left(2\varepsilon_A + \varepsilon_A^2 \right) \sigma_1^2.
\end{aligned}
\tag{A.1}
$$

Equation (A.1) holds because of $\|A\|_2 = \sigma_1$ and $\|A\|_2 = \|A^T\|_2$.

On the other hand,

$$
\begin{aligned}
\eta &= \left\| I - \frac{m}{n} (A + E) (A + E)^T \right\|_2 = \left\| I - \frac{m}{n} AA^T \right. \\
&\quad \left. - \frac{m}{n} \left(AE^T + EA^T + EE^T \right) \right\|_2 \leq \left\| I - \frac{m}{n} AA^T \right\|_2 \\
&\quad + \frac{m}{n} \left(\left\| AE^T \right\|_2 + \left\| EA^T \right\|_2 + \left\| EE^T \right\|_2 \right) \leq \eta_0 \\
&\quad + \frac{m}{n} \left(\|A\|_2 \left\| E^T \right\|_2 + \|E\|_2 \left\| A^T \right\|_2 + \|E\|_2 \left\| E^T \right\|_2 \right)
\end{aligned}
$$

$$\leq \eta_0 + \frac{m}{n} \left(\|A\|_2 \, \varepsilon_A \left\|A^T\right\|_2 + \varepsilon_A \|A\|_2 \left\|A^T\right\|_2 \right.$$

$$+ \varepsilon_A \|A\|_2 \, \varepsilon_A \left\|A^T\right\|_2 \right) \leq \eta_0 + \frac{m}{n} \left(2\varepsilon_A + \varepsilon_A^2 \right) \|A\|_2$$

$$\cdot \left\|A^T\right\|_2 = \eta_0 + \frac{m}{n} \left(2\varepsilon_A + \varepsilon_A^2 \right) \sigma_1^2.$$

$$\text{(A.2)}$$

The last equation holds because of $\|A\|_2 = \sigma_1$, $\|A\|_2 = \|A^T\|_2$. Combine (A.1) with (A.2) to obtain (32). □

Proof of Theorem 7. The three different inequalities come from different proving processes but in essence they are the same. Here we only prove the first inequality in more detail; the proofs of the second and third inequality are similar. For convenience, we denote them by Cases 1, 2, and 3 related to Cases $1'$, $2'$, and $3'$, respectively. The proofs depend on one fact that $Q_1 = Q/\gamma$ is close to I due to the definition of η. Suppose that

$$\gamma = \sigma^2 + \frac{n}{m}\sigma_0,$$

$$Q = \sigma^2 I + \sigma_0^2 (A + E)(A + E)^T.$$

$$\text{(A.3)}$$

Assume that (15) can be written as $y = \widehat{B}x + v$, where

$$\widehat{B} = Q_1^{-1/2}(A + E) = Q_1^{-1/2}\widehat{A},$$

$$v = Q_1^{-1/2}((A + E)u + e).$$

$$\text{(A.4)}$$

Case 1. Consider

$$\|Q_1 - I\|_2 = \left\|\frac{Q}{\gamma} - I\right\|_2 = \frac{1}{\gamma}\|Q - \gamma I\|_2 = \frac{1}{\gamma}\left\|\sigma^2 I\right.$$

$$+ \sigma_0^2 (A + E)(A + E)^T - \gamma I\Big\|_2 = \frac{1}{\gamma}\left\|\sigma^2 I + \sigma_0^2 AA^2\right.$$

$$+ \sigma_0^2 AE^T + \sigma_0^2 EA^T + \sigma_0^2 EE^T - \sigma^2 I - \frac{n}{m}\sigma_0^2 I\Big\|_2$$

$$= \frac{(n/m)\sigma_0^2}{\gamma}\left\|\frac{m}{n}AA^T + \frac{m}{n}AE^T + \frac{m}{n}EA^T + \frac{m}{n}\right.$$

$$\cdot EE^T - I\Big\|_2 \leq \frac{(n/m)\sigma_0^2}{\gamma}\left(\frac{m}{n}\left(\|A\|_2\left\|A^T\right\|_2\right.\right.$$

$$+ 2\varepsilon_A \|A\|_2 \cdot \left\|A^T\right\|_2 + \varepsilon_A \|A\|_2 \, \varepsilon_A \left\|A^T\right\|_2\right) + \|I\|_2\Big)$$

$$= \frac{(n/m)\sigma_0^2}{\sigma^2 + (n/m)\sigma_0}\left(\frac{m}{n}\left(\sigma_1^2 + 2\varepsilon_A\sigma_1^2 + \varepsilon_A^2\sigma_1^2\right) + 1\right)$$

$$\triangleq \mu_1.$$

$$\text{(A.5)}$$

The last equation holds because of $\|A\|_2 = \|A^T\|_2 = \sigma_1$, $\|I\|_2 = 1$.

For (A.5), since $0 < \varepsilon_A \ll 1$, $m \ll n$, σ_1 is positive; thus,

$$\frac{m}{n}\left(\sigma_1^2 + 2\varepsilon_A\sigma_1^2 + \varepsilon_A^2\sigma_1^2\right) + 1 \longrightarrow 1 \qquad \text{(A.6)}$$

holds when $n \to \infty$; therefore, (A.5) $\to$ $(n/m)\sigma_0^2/(\sigma^2 + (n/m)\sigma_0) < 1$.

That is, $\|Q_1 - I\|_2 < \mu_1 < 1$ when $n \to \infty$.

Case 2. Consider

$$\|Q_1 - I\|_2 = \left\|\frac{Q}{\gamma} - I\right\|_2 = \frac{1}{\gamma}\|Q - \gamma I\|_2$$

$$= \frac{1}{\gamma}\left\|\sigma^2 I + \sigma_0^2 (A + E)(A + E)^T - \sigma^2 I - \frac{n}{m}\sigma_0 I\right\|_2$$

$$= \frac{(n/m)\sigma_0^2}{\sigma^2 + (n/m)\sigma_0^2}\left\|\frac{m}{n}(A + E)(A + E)^T - I\right\|_2$$

$$\leq \frac{(n/m)\sigma_0^2}{\sigma^2 + (n/m)\sigma_0^2}\eta \triangleq \eta_1 < \eta < \frac{1}{2}.$$

$$\text{(A.7)}$$

That is, $\|Q_1 - I\|_2 \leq \eta_1 < \eta < 1/2$.

Case 3. Consider

$$\|Q_1 - I\|_2 = \left\|\frac{Q}{\gamma} - I\right\|_2 = \frac{1}{\gamma}\|Q - \gamma I\|_2 = \frac{1}{\gamma}\left\|\sigma^2 I\right.$$

$$+ \sigma_0^2 (A + E)(A + E)^T - \gamma I\Big\|_2 = \frac{1}{\gamma}\left\|\sigma^2 I + \sigma_0^2 AA^2\right.$$

$$+ \sigma_0^2 AE^T + \sigma_0^2 EA^T + \sigma_0^2 EE^2 - \sigma^2 I - \frac{n}{m}\sigma_0^2 I\Big\|_2$$

$$\leq \frac{(n/m)\sigma_0^2}{\gamma}\left(\left\|\frac{m}{n}AA^T - I\right\|_2\right.$$

$$+ \frac{m}{n}\left(\|A\|_2\left\|E^T\right\|_2 + \|E\|_2\left\|A^T\right\|_2 + \|E\|_2\left\|E^T\right\|_2\right)\Big)$$

$$= \frac{(n/m)\sigma_0^2}{\sigma^2 + (n/m)\sigma_0}\left(\eta_0 + \frac{m}{n}\left(2\varepsilon_A\sigma_1^2 + \varepsilon_A^2\sigma_1^2\right)\right)$$

$$\triangleq \mu_2.$$

$$\text{(A.8)}$$

The last equation holds due to $\|A\|_2 = \|A^T\|_2 = \sigma_1$.

From (A.8), since $0 < \varepsilon_A \ll 1$, $m \ll n$, and $\|A\|_2 = \sigma_1$ is positive,

$$\frac{m}{n}\left(2\varepsilon_A\sigma_1^2 + \varepsilon_A^2\sigma_1^2\right) \longrightarrow 0 \qquad \text{(A.9)}$$

with $n \to \infty$; thus,

$$\text{(A.8)} \longrightarrow \frac{(n/m)\sigma_0^2}{\sigma^2 + (n/m)\sigma_0}\eta_0. \qquad \text{(A.10)}$$

That is, $\|Q_1 - I\|_2 < \mu_2 < \eta_0 < 1/2$ with $n \to \infty$.

As mentioned above, $\|Q_1 - I\|_2 < 1$ holds under some conditions. Using Cases 1, 2, and 3, we can obtain three different results, denoted by Cases $1'$, $2'$, and $3'$, respectively.

Here $Q_1^{-1} - I$ can be expressed as follows:

$$Q_1^{-1} - I = \frac{I - Q_1}{Q_1} = \left(\frac{Q_1}{I - Q_1}\right)^{-1} = \frac{I - Q_1}{I - (1 - Q_1)}$$

$$= \sum_{k \geq 1} (I - Q_1)^k.$$

$$\text{(A.11)}$$

Case 1′. Note that $Q_1^{-1} - I$ converges due to (A.5) $\|I - Q_1\|_2 < \mu_1 < 1$, where $\|\cdot\|_2$ is an operator norm. Take such norm on both sides of the above equality and utilize the triangle inequality to get

$$
\begin{aligned}
\left\|Q_1^{-1} - I\right\|_2 &= \left\|\sum_{k\geq 1}(I - Q_1)^k\right\|_2 \leq \sum_{k\geq 1}\|I - Q_1\|_2^k \\
&< \sum_{k\geq 1}\mu_1^k = \frac{\mu_1}{1 - \mu_1} \triangleq \mu_1'.
\end{aligned}
\tag{A.12}
$$

Let Λ be an index set of size k, $\forall h \in R^k$;

$$
\left\|\widehat{B}_\Lambda h\right\|_2^2 - \left\|\widehat{A}_\Lambda h\right\|_2^2 = h^T \widehat{A}_\Lambda^T \left(Q_1^{-1} - I\right)\widehat{A}_\Lambda h
\tag{A.13}
$$

holds. Since

$$
\begin{aligned}
\left|h^T \widehat{A}_\Lambda^T \left(Q_1^{-1} - I\right)\widehat{A}_\Lambda h\right| &\leq \left\|Q_1^{-1} - I\right\|_2 \left\|\widehat{A}_\Lambda h\right\|_2^2 \\
&\leq \mu_1' \left\|\widehat{A}_\Lambda h\right\|_2^2,
\end{aligned}
\tag{A.14}
$$

we obtain

$$
\left|\left\|\widehat{B}_\Lambda h\right\|_2^2 - \left\|\widehat{A}_\Lambda h\right\|_2^2\right| \leq \mu_1' \left\|\widehat{A}_\Lambda h\right\|_2^2.
\tag{A.15}
$$

Remove the absolute value to get

$$
\left(1 - \mu_1'\right)\left\|\widehat{A}_\Lambda h\right\|_2^2 \leq \left\|\widehat{B}_\Lambda h\right\|_2^2 \leq \left(1 + \mu_1'\right)\left\|\widehat{A}_\Lambda h\right\|_2^2.
\tag{A.16}
$$

Due to

$$
\widehat{\alpha}_k \|h\|_2^2 \leq \left\|\widehat{A}h\right\|_2^2 \leq \widehat{\beta}_k \|h\|_2^2, \quad \forall h \in R^k,
\tag{A.17}
$$

if set $\mu_1' = \mu_1/(1 - \mu_1)$, we have

$$
\widehat{\alpha}_k \left(1 - \mu_1'\right)\|h\|_2^2 \leq \left\|\widehat{B}_\Lambda h\right\|_2^2 \leq \widehat{\beta}_k \left(1 + \mu_1'\right)\|h\|_2^2.
\tag{A.18}
$$

Case 2′. Note that $Q_1^{-1} - I$ converges due to (A.7), $\|I - Q_1\|_2 < \eta_1 < 1/2$. Take spectral norm on both sides of $Q_1^{-1} - I$ and utilize the triangle inequality to get

$$
\begin{aligned}
\left\|Q_1^{-1} - I\right\|_2 &= \left\|\sum_{k\geq 1}(I - Q_1)^k\right\|_2 \leq \sum_{k\geq 1}\|I - Q_1\|_2^k \\
&< \sum_{k\geq 1}\eta_1^k = \frac{\eta_1}{1 - \eta_1} \triangleq \eta_1'.
\end{aligned}
\tag{A.19}
$$

The remaining proof of Case 2′ is similar to that of Case 1′ except for η_1' instead of μ_1'. At last we have

$$
\begin{aligned}
&\widehat{\alpha}_k \left(1 - \eta_1'\right)\|h\|_2 \leq \left\|\widehat{B}_\Lambda h\right\|_2^2 \leq \widehat{\beta}_k \left(1 + \eta_1'\right)\|h\|_2^2, \\
&\qquad\qquad\qquad\qquad\qquad \eta_1' = \frac{\eta_1}{1 - \eta_1}.
\end{aligned}
\tag{A.20}
$$

Case 3′. Note that $Q_1^{-1} - I$ converges from (A.8) due to $\|I - Q_1\|_2 < \mu_2 < 1/2$. Take spectral norm on both sides of $Q_1^{-1} - I$ and utilize the triangle inequality to get

$$
\begin{aligned}
\left\|Q_1^{-1} - I\right\|_2 &= \left\|\sum_{k\geq 1}(I - Q_1)^k\right\|_2 \leq \sum_{k\geq 1}\|I - Q_1\|_2^k \\
&< \sum_{k\geq 1}\mu_2^k = \frac{\mu_2}{1 - \mu_2} \triangleq \mu_2'.
\end{aligned}
\tag{A.21}
$$

The remaining proof of Case 3′ is similar to that of Case 1′ except for μ_2' instead of μ_1'. At last, we obtain

$$
\begin{aligned}
&\widehat{\alpha}_k \left(1 - \mu_2'\right)\|h\|_2 \leq \left\|\widehat{B}_\Lambda h\right\|_2^2 \leq \widehat{\beta}_k \left(1 + \mu_2'\right)\|h\|_2^2, \\
&\qquad\qquad\qquad\qquad\qquad \mu_2' = \frac{\mu_2}{1 - \mu_2}.
\end{aligned}
\tag{A.22}
$$

$\square$

Proof of Theorem 9. On the one hand,

$$
\begin{aligned}
\widetilde{\eta} &= \left\|I - \frac{m}{n}\left(A + \sum_{i=1}^s E_i\right)\left(A + \sum_{i=1}^s E_i\right)^T\right\|_2 = \left\|\frac{m}{n}\right. \\
&\quad \cdot AA^T - I + \frac{m}{n}\left(A\sum_{i=1}^s E_i^T + A^T\sum_{i=1}^s E_i\right. \\
&\quad \left.\left. + \left(\sum_{i=1}^s E_i\right)\left(\sum_{i=1}^s E_i\right)^T\right)\right\|_2 \geq \left\|\frac{m}{n}AA^T - I\right\|_2 \\
&\quad - \frac{m}{n}\left(\left\|A\sum_{i=1}^s E_i^T + A^T\sum_{i=1}^s E_i\right.\right. \\
&\quad \left.\left. + \left(\sum_{i=1}^s E_i\right)\left(\sum_{i=1}^s E_i\right)^T\right\|_2\right) \geq \eta_0 - \frac{m}{n}\left(\|A\|_2\right. \\
&\quad \cdot \sum_{i=1}^s\left\|E_i^T\right\|_2 + \|A^T\|_2\sum_{i=1}^s\|E_i\|_2 + \left(\sum_{i=1}^s\|E_i\|_2\right) \\
&\quad \left.\cdot\left(\sum_{i=1}^s\left\|E_i^T\right\|_2\right)\right) \geq \eta_0 - \frac{m}{n}\left(\|A\|_2\,\varepsilon_A\sum_{i=1}^s\|A^T\|_2\right. \\
&\quad + \|A^T\|_2\sum_{i=1}^s\varepsilon_A\|A\|_2 + \left(\sum_{i=1}^s\varepsilon_A\|A\|_2\right) \\
&\quad \left.\cdot\left(\sum_{i=1}^s\varepsilon_A\|A^T\|_2\right)\right) = \eta_0 - \frac{m}{n}\left(2s\varepsilon_A\sigma_1^2\right. \\
&\quad \left. + s^2\varepsilon_A^2\sigma_1^2\right).
\end{aligned}
\tag{A.23}
$$

The last equation holds due to $\|A^T\|_2 = \|A\|_2 = \sigma_1$.

On the other hand,

$$\tilde{\eta} = \left\| I - \frac{m}{n}\left(A + \sum_{i=1}^{s} E_i\right)\left(A + \sum_{i=1}^{s} E_i\right)^T \right\|_2 = \left\| \frac{m}{n} \right.$$

$$\cdot AA^T - I + \frac{m}{n}\left(A\sum_{i=1}^{s} E_i^T + A^T\sum_{i=1}^{s} E_i\right.$$

$$\left.+ \left(\sum_{i=1}^{s} E_i\right)\left(\sum_{i=1}^{s} E_i\right)^T\right)\right\|_2 \leq \left\| \frac{m}{n}AA^T - I\right\|_2$$

$$+ \frac{m}{n}\left(\left\| A\sum_{i=1}^{s} E_i^T + A^T\sum_{i=1}^{s} E_i\right.\right.$$

$$\left.\left.+ \left(\sum_{i=1}^{s} E_i\right)\left(\sum_{i=1}^{s} E_i\right)^T\right\|_2\right) \leq \eta_0 + \frac{m}{n}\left(\|A\|_2\right. \qquad \text{(A.24)}$$

$$\cdot \sum_{i=1}^{s}\left\| E_i^T\right\|_2 + \left\| A^T\right\|_2 \sum_{i=1}^{s}\left\| E_i\right\|_2 + \left(\sum_{i=1}^{s}\left\| E_i\right\|_2\right)$$

$$\cdot \left(\sum_{i=1}^{s}\left\| E_i^T\right\|_2\right)\right) \leq \eta_0 + \frac{m}{n}\left(\|A\|_2\, \varepsilon_A \sum_{i=1}^{s}\left\| A^T\right\|_2\right.$$

$$+ \left\| A^T\right\|_2 \sum_{i=1}^{s}\varepsilon_A \|A\|_2 + \left(\sum_{i=1}^{s}\varepsilon_A\|A\|_2\right)$$

$$\cdot \left(\sum_{i=1}^{s}\varepsilon_A\left\| A^T\right\|_2\right)\right) = \eta_0 + \frac{m}{n}\left(2s\varepsilon_A\sigma_1^2\right.$$

$$\left.+ s^2\varepsilon_A^2\sigma_1^2\right).$$

The last equation holds because of $\|A^T\|_2 = \|A\|_2 = \sigma_1$. As mentioned above, combine (A.23) with (A.24) to get (32). $\square$

Proof of Theorem 11. There are three different results of whitening $A + \sum_{i=1}^{s} E_i$ due to the different proving process. The proof depends on one fact that Q_1 is close to I due to the definition of $\tilde{\eta}$. Assume that (A.1) can be written as $y = \widehat{B}x + w$, where $\widehat{B} = Q_1^{-1/2}(A + \sum_{i=1}^{s} E_i)$, $w = Q_1^{-1/2}((A + \sum_{i=1}^{s} E_i)u + e)$.

Case 1. Consider

$$\|Q_1 - I\|_2 = \left\| \frac{Q}{\gamma} - I\right\|_2 = \frac{1}{\gamma}\left\| \sigma^2 I + \sigma_0^2\left(A + \sum_{i=1}^{s} E_i\right)\right.$$

$$\cdot \left(A + \sum_{i=1}^{s} E_i\right)^T - \gamma I\right\|_2 = \frac{1}{\gamma}\left\| \sigma^2 I + \sigma_0^2\left(A\right.\right.$$

$$\left.+ \sum_{i=1}^{s} E_i\right)\left(A + \sum_{i=1}^{s} E_i\right)^T - \left(\sigma^2 + \frac{n}{m}\sigma_0^2\right)I\right\|_2$$

$$= \frac{(n/m)\,\sigma_0^2}{\gamma}\left\| \frac{m}{n}\left(A + \sum_{i=1}^{s} E_i\right)\left(A + \sum_{i=1}^{s} E_i\right)^T\right.$$

$$\left. - I\right\|_2 = \frac{(n/m)\,\sigma_0^2}{\gamma}\left\| \frac{m}{n}AA^T - I + \frac{m}{n}\left(A\sum_{i=1}^{s} E_i^T\right.\right.$$

$$\left.+ A^T\sum_{i=1}^{s} E_i + \left(\sum_{i=1}^{s} E_i\right)\left(\sum_{i=1}^{s} E_i\right)^T\right)\right\|_2$$

$$\leq \frac{(n/m)\,\sigma_0^2}{\gamma}\left(\left\| \frac{m}{n}AA^T - I\right\|_2\right.$$

$$+ \frac{m}{n}\left(\|A\|_2\sum_{i=1}^{s}\left\| E_i^T\right\|_2 + \left\| A^T\right\|_2\sum_{i=1}^{s}\left\| E_i\right\|_2\right.$$

$$\left.\left.+ \left(\sum_{i=1}^{s}\left\| E_i\right\|_2\right)\left(\sum_{i=1}^{s}\left\| E_i^T\right\|_2\right)\right)\right) \leq \frac{(n/m)\,\sigma_0^2}{\gamma}\left(\eta_0\right.$$

$$+ \frac{m}{n}\left(\|A\|_2\,\varepsilon_A\sum_{i=1}^{s}\left\| A^T\right\|_2 + \left\| A^T\right\|_2\sum_{i=1}^{s}\varepsilon_A\|A\|_2\right.$$

$$\left.\left.+ \varepsilon_A^2\left(\sum_{i=1}^{s}\|A\|_2\right)\left(\sum_{i=1}^{s}\left\| A^T\right\|_2\right)\right)\right)$$

$$= \frac{(n/m)\,\sigma_0^2}{\gamma}\left(\eta_0 + \frac{m}{n}\left(\varepsilon_A\sum_{i=1}^{s}\sigma_1^2 + \varepsilon_A\sum_{i=1}^{s}\sigma_1^2\right.\right.$$

$$\left.\left.+ s^2\varepsilon_A^2\sigma_1^2\right)\right) = \frac{(n/m)\,\sigma_0^2}{\sigma^2 + (n/m)\,\sigma_0^2}\left(\eta_0\right.$$

$$\left.+ \frac{m}{n}\left(2s\varepsilon_A\sigma_1^2 + s^2\varepsilon_A^2\sigma_1^2\right)\right) \triangleq \mu_3.$$

$$\text{(A.25)}$$

The last equation holds because of $\|A\|_2 = \sigma_1$ and $\|A\|_2 = \|A^T\|_2$.

From (A.25), since $0 < \varepsilon_A \ll 1$, $m \ll n$, and $\|A\|_2 = \sigma_1$ is a constant, we have

$$\frac{m}{n}\left(2s\varepsilon_A\sigma_1^2 + s^2\varepsilon_A^2\sigma_1^2\right) \longrightarrow 0 \qquad \text{(A.26)}$$

when $n \to \infty$; therefore, (A.25) $\to$ $((n/m)\sigma_0^2/(\sigma^2 + (n/m)\sigma_0^2))\eta_0 < \eta_0 < 1/2$. That is, $\|Q_1 - I\|_2 < \mu_3 < 1/2$ when $n \to \infty$.

Case 2. Consider

$$\|Q_1 - I\|_2 = \left\| \frac{Q}{\gamma} - I\right\|_2 = \frac{1}{\gamma}\left\| \sigma^2 I\right.$$

$$+ \sigma_0^2\left(A + \sum_{i=1}^{s} E_i\right)\left(A + \sum_{i=1}^{s} E_i\right)^T - \gamma I\right\|_2$$

$$= \frac{1}{\gamma}\left\| \sigma^2 I + \sigma_0^2\left(A + \sum_{i=1}^{s} E_i\right)\left(A + \sum_{i=1}^{s} E_i\right)^T\right.$$

$$\left. - \left(\sigma^2 + \frac{n}{m}\sigma_0^2\right)I\right\|_2$$

$$= \frac{(n/m)\,\sigma_0^2}{\gamma}\left\|\frac{m}{n}\left(A + \sum_{i=1}^{s} E_i\right)\left(A + \sum_{i=1}^{s} E_i\right)^T\right.$$

$$\left. - I\right\|_2 = \frac{(n/m)\,\sigma_0^2}{\sigma^2 + (n/m)\sigma_0^2}\tilde{\eta} \triangleq \eta_2 < \tilde{\eta} < \frac{1}{2}.$$

(A.27)

From (A.27), we get $\|Q_1 - I\|_2 < 1$.

Case 3. Consider

$$\|Q_1 - I\|_2 = \left\|\frac{Q}{\gamma} - I\right\|_2 = \frac{1}{\gamma}\left\|\sigma^2 I + \sigma_0^2\left(A + \sum_{i=1}^{s} E_i\right)\right.$$

$$\cdot \left(A + \sum_{i=1}^{s} E_i\right)^T - \gamma I\right\|_2 = \frac{1}{\gamma}\left\|\sigma^2 I + \sigma_0^2\left(A\right.\right.$$

$$\left.+ \sum_{i=1}^{s} E_i\right)\left(A + \sum_{i=1}^{s} E_i\right)^T - \left(\sigma^2 + \frac{n}{m}\sigma_0^2\right) I\right\|_2$$

$$= \frac{(n/m)\,\sigma_0^2}{\gamma}\left\|\frac{m}{n}\left(A + \sum_{i=1}^{s} E_i\right)\left(A + \sum_{i=1}^{s} E_i\right)^T\right.$$

$$\left. - I\right\|_2 \leq \frac{(n/m)\,\sigma_0^2}{\gamma}\left(\|-I\|_2 + \frac{m}{n}\left(\|AA^T\|_2\right.\right.$$

$$+ \|A\|_2 \sum_{i=1}^{s}\|E_i^T\|_2 + \|A^T\|_2 \sum_{i=1}^{s}\|E_i\|_2$$

$$\left.\left.+ \left(\sum_{i=1}^{s}\|E_i\|_2\right)\left(\sum_{i=1}^{s}\|E_i^T\|_2\right)\right)\right) \leq \frac{(n/m)\,\sigma_0^2}{\gamma}\left(1\right.$$

$$+ \frac{m}{n}\left(\|A\|_2\|A^T\|_2 + \|A\|_2\,\varepsilon_A \sum_{i=1}^{s}\|A^T\|_2\right.$$

$$+ \|A^T\|_2 \sum_{i=1}^{s}\varepsilon_A\|A\|_2$$

$$\left.\left.+ \left(\sum_{i=1}^{s}\varepsilon_A\|A\|_2\right)\left(\sum_{i=1}^{s}\varepsilon_A\|A^T\|_2\right)\right)\right)$$

$$= \frac{(n/m)\,\sigma_0^2}{\gamma}\left(1 + \frac{m}{n}\left(\sigma_1^2 + \varepsilon_A \sum_{i=1}^{s}\sigma_1^2 + \varepsilon_A \sum_{i=1}^{s}\sigma_1^2\right.\right.$$

$$\left.\left.+ s^2\varepsilon_A^2\sigma_1^2\right)\right) = \frac{\sigma_0^2}{\sigma^2 + (n/m)\,\sigma_0^2}\left(1 + \frac{m}{n}\left(\sigma_1^2\right.\right.$$

$$\left.\left.+ 2s\varepsilon_A\sigma_1^2 + s^2\varepsilon_A^2\sigma_1^2\right)\right) \triangleq \mu_4.$$

(A.28)

The last equation holds due to $\|A\|_2 = \sigma_1$ and $\|A\|_2 = \|A^T\|_2$.

From (A.28), since $0 < \varepsilon_A \ll 1$, $m \ll n$, and $\|A\|_2 = \sigma_1$ is a constant, then

$$\frac{m}{n}\left(\sigma_1^2 + 2s\varepsilon_A\sigma_1^2 + s^2\varepsilon_A^2\sigma_1^2\right) \longrightarrow 0$$

(A.29)

when $n \to \infty$; thus, (A.28) $\to (n/m)\sigma_0^2/(\sigma^2 + (n/m)\sigma_0^2) < 1$. That is, $\|Q_1 - I\|_2 < \mu_4 < 1$. As mentioned above, $\|Q_1 - I\|_2 < 1$ under some conditions. Using the above three cases (Cases 1, 2, and 3), we can obtain three different results, denoted by Cases $1'$, $2'$, and $3'$, respectively.

Here $Q_1^{-1} - I$ can be expressed as follows:

$$Q_1^{-1} - I = \frac{I - Q_1}{Q_1} = \left(\frac{Q_1}{I - Q_1}\right)^{-1} = \frac{I - Q_1}{I - (1 - Q_1)}$$

$$= \sum_{k \geq 1}(I - Q_1)^k.$$

(A.30)

Case $1'$. Note that (A.30) converges due to Case 1: $\|I - Q_1\|_2 < \mu_3 < 1$. Take spectral norm on both sides of the equality and utilize the triangle inequality to get

$$\|Q_1^{-1} - I\|_2 = \left\|\sum_{k \geq 1}(I - Q_1)^k\right\|_2 \leq \sum_{k \geq 1}\|I - Q_1\|_2^k$$

$$< \sum_{k \geq 1}\mu_3^k = \frac{\mu_3}{1 - \mu_3} \triangleq \mu_3'.$$

(A.31)

Let Λ be an index set of size k; we have

$$\left\|\widehat{B}_\Lambda h\right\|_2^2 - \left\|\widehat{A}_\Lambda h\right\|_2^2 = h^T\widehat{A}_\Lambda^T\left(Q_1^{-1} - I\right)\widehat{A}_\Lambda h,$$

$$\forall h \in R^k.$$

(A.32)

Since

$$\left|h^T\widehat{A}_\Lambda^T\left(Q_1^{-1} - I\right)\widehat{A}_\Lambda h\right| \leq \left\|Q_1^{-1} - I\right\|_2\left\|\widehat{A}_\Lambda h\right\|_2^2$$

$$\leq \mu_3'\left\|\widehat{A}_\Lambda h\right\|_2^2,$$

(A.33)

we obtain that

$$\left|\left\|\widehat{B}_\Lambda h\right\|_2^2 - \left\|\widehat{A}_\Lambda h\right\|_2^2\right| \leq \mu_3'\left\|\widehat{A}_\Lambda h\right\|_2^2.$$

(A.34)

Remove the absolute value to get

$$\left(1 - \mu_3'\right)\left\|\widehat{A}_\Lambda h\right\|_2^2 \leq \left\|\widehat{B}_\Lambda h\right\|_2^2 \leq \left(1 + \mu_3'\right)\left\|\widehat{A}_\Lambda h\right\|_2^2.$$

(A.35)

Due to

$$\widehat{\alpha}_k\|h\|_2^2 \leq \left\|\widehat{A}h\right\|_2^2 \leq \widehat{\beta}_k\|h\|_2^2, \quad \forall h \in R^n,$$

(A.36)

we have

$$\widehat{\alpha}_k\left(1 - \mu_3'\right)\|h\|_2^2 \leq \left\|\widehat{B}_\Lambda h\right\|_2^2 \leq \widehat{\beta}_k\left(1 + \mu_3'\right)\|h\|_2^2,$$

$$\mu_3' = \frac{\mu_3}{1 - \mu_3}.$$

(A.37)

Case 2′. Note that (A.30) converges due to Case 2: $\|I - Q_1\|_2 < \eta_2 < 1$. Take spectral norm on both sides of the equality (A.45) and utilize the triangle inequality to get

$$
\begin{aligned}
\left\| Q_1^{-1} - I \right\|_2 &= \left\| \sum_{k \geq 1} \left(I - Q_1 \right)^k \right\|_2 \leq \sum_{k \geq 1} \left\| I - Q_1 \right\|_2^k \\
&< \sum_{k \geq 1} \eta_2^k = \frac{\eta_2}{1 - \eta_2} \triangleq \eta_2'.
\end{aligned}
\tag{A.38}
$$

The remaining proving is similar to that of Case 1′ except for η_2' instead of μ_3'. At last we have

$$
\widehat{\alpha}_k \left(1 - \eta_2' \right) \|h\|_2^2 \leq \left\| \widehat{B}_\Lambda h \right\|_2^2 \leq \widehat{\beta}_k \left(1 + \eta_2' \right) \|h\|_2^2. \tag{A.39}
$$

Case 3′. Note that (A.45) converges due to Case 3: $\|I - Q_1\|_2 < \mu_4 < 1$. Take spectral norm on both sides of the equality (A.45) and utilize the triangle inequality to get

$$
\begin{aligned}
\left\| Q_1^{-1} - I \right\|_2 &= \left\| \sum_{k \geq 1} \left(I - Q_1 \right)^k \right\|_2 \leq \sum_{k \geq 1} \left\| I - Q_1 \right\|_2^k \\
&< \sum_{k \geq 1} \mu_4^k = \frac{\mu_4}{1 - \mu_4} \triangleq \mu_4'.
\end{aligned}
\tag{A.40}
$$

The remaining proving is similar to that of Case 1′ except for μ_4' instead of μ_3'. At last we have

$$
\begin{aligned}
\widehat{\alpha}_k \left(1 - \mu_4' \right) \|h\|_2^2 \leq \left\| \widehat{B}_\Lambda h \right\|_2^2 &\leq \widehat{\beta}_k \left(1 + \mu_4' \right) \|h\|_2^2, \\
\mu_4' &= \frac{\mu_4}{1 - \mu_4}.
\end{aligned}
\tag{A.41}
$$

$\square$

Proof of Theorem 14. To prove the theorem, we should find out an upper bound of the numerator $|\widehat{B}_i^T \widehat{B}_j|$ of $\mu(\widehat{B})$ and a lower bound of the denominator $\|\widehat{B}_i\|_2$. For $i \neq j$, by assumption, we obtain

$$
\begin{aligned}
\left| \widehat{B}_i^T \widehat{B}_j \right| &= \left| \widehat{A}_i^T Q_1^{-1} \widehat{A}_j \right| \leq \left| \widehat{A}_i^T \widehat{A}_j \right| + \left| \widehat{A}_i^T \left(Q_1^{-1} - I \right) \widehat{A}_j \right| \\
&\leq \left(1 + \mu_2' \right) \left| \widehat{A}_i^T \widehat{A}_j \right|.
\end{aligned}
\tag{A.42}
$$

Next, we estimate lower bound $\|\widehat{B}_i\|_2$ with restrictions on $\|A_i\|$ and μ_2. Similar to the proof of Theorem 7, $Q_1^{-1/2} - I$ can be expressed as a power series:

$$
Q_1^{-1/2} - I = \sum_{k \geq 1} c_k \left(I - Q_1 \right), \tag{A.43}
$$

where c_k is the coefficients in the Taylor expansion of $(1 - x)^{1/2}$. Both sides of the equality are taken norm obtaining

$$
\begin{aligned}
\left\| Q_1^{1/2} - I \right\|_2 &\leq \sum_{k \geq 1} c_k \left\| \left(Q_1 - I \right)^k \right\|_2^k \leq \sum_{k \geq 1} c_k \mu_2 \\
&= \left(1 - \mu_2 \right)^{-1/2} - 1 \triangleq \widehat{\mu}_2.
\end{aligned}
\tag{A.44}
$$

Thus,

$$
\begin{aligned}
\left\| \widehat{B}_i \right\|_2 &= \left\| Q_1^{-1/2} \widehat{A}_i \right\|_2 \geq \left\| \widehat{A}_i \right\|_2 - \left\| \left(Q_1^{-1/2} - I \right) \widehat{A}_i \right\|_2 \\
&\geq \left(1 - \widehat{\mu}_2 \right) \left\| \widehat{A}_i \right\|_2,
\end{aligned}
\tag{A.45}
$$

where $\widehat{\mu}_2 = (1 - \mu_2)^{-1/2} - 1$. Combine (A.42) with (A.45) to get the result in Theorem 14. $\square$

Competing Interests

The authors declare that there is no competing interests regarding the publication of this paper.

Acknowledgments

References

[1] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[2] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[3] R. Baraniuk and P. Steeghs, "Compressive radar imaging," in *Proceedings of the IEEE Radar Conference*, pp. 128–133, Boston, Mass, USA, April 2007.

[4] M. F. Duarte, M. A. Davenport, M. B. Wakin et al., "Sparse signal detection from incoherent projections," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '06)*, vol. 3, p. III, IEEE, Toulouse, France, May 2006.

[5] J. Romberg, "Imaging via compressive sampling: introduction to compressive sampling and recovery via convex programming," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 14–20, 2008.

[6] M. A. Davenport, M. F. Duarte, Y. C. Eldar et al., "Introduction to compressed sensing," in *Compressed Sensing: Theory and Applications*, vol. 93, pp. 1–64, Cambridge University Press, 2011.

[7] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Springer, Berlin, Germany, 2013.

[8] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2153–2164, 2004.

[9] H. Zayyani, M. Babaie-Zadeh, and C. Jutten, "An iterative Bayesian algorithm for sparse component analysis in presence of noise," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4378–4390, 2009.

[10] T. T. Cai, L. Wang, and G. Xu, "New bounds for restricted isometry constants," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4388–4394, 2010.

[11] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[12] M. A. Davenport and M. B. Wakin, "Analysis of orthogonal matching pursuit using the restricted isometry property," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4395–4401, 2010.

[13] S. Huang and J. Zhu, "Recovery of sparse signals using OMP and its variants:conver-gence analysis based on RIP," *Inverse Problems*, vol. 27, no. 3, 2011.

[14] A. K. Fletcher and S. Rangan, "Orthogonal matching pursuit: a Brownian motion analysis," *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 1010–1021, 2012.

[15] Q. Mo and Y. Shen, "A remark on the restricted isometry property in orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3654–3656, 2012.

[16] D. Needell and J. A. Tropp, "CoSaMP: iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[17] S. Li, F. Gao, G. Ge, and S. Zhang, "Deterministic construction of compressed sensing matrices via algebraic curves," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5035–5041, 2012.

[18] X. Wang, Z. Zhao, N. Zhao et al., "On the application of compressed sensing in communication networks," in *Proceedings of the 5th International ICST Conference on Communications and Networking in China (CHINACOM '10)*, pp. 1–7, IEEE, Beijing, China, August 2010.

[19] D. Malioutov and M. Malyutov, "Boolean compressed sensing: LP relaxation for group testing," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '12)*, pp. 3305–3308, Kyoto, Japan, March 2012.

[20] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," *IEEE Journal on Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 310–316, 2010.

[21] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 6–18, 2006.

[22] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[23] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[24] E. Arias-Castro and Y. C. Eldar, "Noise folding in compressed sensing," *IEEE Signal Processing Letters*, vol. 18, no. 8, pp. 478–481, 2011.

[25] M. A. Herman and T. Strohmer, "General deviants: an analysis of perturbations in compressed sensing," *IEEE Journal on Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 342–349, 2010.

[26] J. Ding, L. Chen, and Y. Gu, "Performance analysis of orthogonal matching pursuit under general perturbations," in *Proceedings of the International Conference on Computing, Networking and Communications (ICNC '12)*, pp. 892–896, Maui, Hawaii, USA, February 2012.

[27] J. Ding, L. Chen, and Y. Gu, "Perturbation analysis of orthogonal matching pursuit," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 398–410, 2013.