IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 8 March 2024, accepted 4 April 2024, date of publication 8 April 2024, date of current version 15 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3386223

RESEARCH ARTICLE

An Improved Reweighted Method for Optimizing the Sensing Matrix of Compressed Sensing

LEI SHI [1,2] AND GANGRONG QU [2]

[1] Research Institute of Knowledge and Reason, Luster LightTech Company Ltd., Beijing 100094, China
[2] School of Mathematics and Statistics, Beijing Jiaotong University, Beijing 100044, China

Corresponding author: Lei Shi (17118429@bjtu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61931003.

ABSTRACT In compressed sensing (CS), coherence is a simple and practical measure of the quality of a sensing matrix. The smaller the coherence of the sensing matrix, the better the reconstruction result. Most strategies for optimizing the coherence of a sensing matrix are only applicable to gray-value matrices and are not to binary matrices with fast computation speed and smaller storage space. Under a certain condition, the coherence can be improved by weighting the sensing matrix to make its condition number equal to 1, i.e., all non-zero singular values become the same. In this paper, we propose a method for reweighting CS models to reduce the condition number of the sensing matrix so as to improve its coherence, prove that the condition number decreases monotonically to 1 as the weighting times approaches infinity, and then obtain a coherence improvement model equivalent to the original CS model. Finally, we give the RwOMP recovery algorithm based on the proposed reweighted method and verify its superiority by using different binary sensing matrices for CS experiments.

INDEX TERMS Compressed sensing, coherence, sensing matrix.

I. INTRODUCTION

Compressed sensing (CS) is a technique for recovering a sparse or compressible signal $x \in \mathbb{R}^N$ from measured data $y \in \mathbb{R}^m$. When the data acquisition process is linear, the reconstruction model is mathematically described as an underdetermined linear system [1]

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

where the sensing matrix $A \in \mathbb{R}^{m\times N}(m \ll N)$ models the linear measurement progress. The signal x is s-sparse if $\|x\|_0 \leq s \ll N$. The CS theory mainly consists of three core issues: the sparse representation of signals, the design of sensing matrices and the research of recovery algorithms.

The performance of CS recovery algorithms is closely related to the sensing matrix. Efficient recovery necessitates that the sensing matrix must meet certain properties such as restricted isometry, null space and the coherence.

Definition 1 ([2]): The restricted isometry constant (RIC) $\delta_s = \delta_s(A)$ of the matrix A with restricted isometry property (RIP) of order s is defined to be a smallest positive number such that

$$C_s(1-\delta_s)||x||_2^2 \leq ||Ax||_2^2 \leq C_s(1+\delta_s)||x||_2^2 \tag{2}$$

for some positive numbers C_s and all s sparse vectors $x \in \mathbb{R}^N$. The smaller the δ_s, the more accurate the reconstructions. Among the properties, RIP is most suited to assess the quality of a sensing matrix, and many researches have focused on the requirements of different recovery algorithms on the $\delta_s(A)$ to guarantee the accurate recovery of s sparse signals [3], [4], [5], [6], [7], [8]. The calculation of $\delta_s(A)$ is NP-hard [9], [10]. By comparison, coherence is simpler and easier to calculate [11]. The sensing matrix A with coherence μ satisfies the RIP of order s, i.e., $\delta_s = (s-1)\mu$ [12].

Coherence is usually used to guide the design and optimization of a sensing matrix. Currently, some random and deterministic sensing matrices have been presented [13], [14], [15], [16]. Over the years, people have studied many optimization methods for reducing the coherence of sensing matrices. Abolghasemi et al. [17], [18] utilized the idea of gradient descent to optimize the mutual coherence of the Gaussian random matrix and improved the performance of sparse recovery. Pan and Qiu [19] gave an orthogonal optimization method for reducing the mutual coherence of

The associate editor coordinating the review of this manuscript and approving it for publication was Laura Celentano.

the random measurement matrix based on QR factorization. Li et al. [20] designed a projection matrix using the estimated sparse representation to decrease the local cumulative coherence of the measurement dictionary. Jin et al. [21] gave an alternating projection strategy and an improved shrinkage method for reducing the average mutual coherence and the mutual coherence of the Gaussian random measurement matrix. These optimization methods are mainly used for the sensing matrix with continuous elements.

In addition to the gray-value matrices mentioned above, there are also some binary sensing matrices designed and applied in compressive imaging to save storage space of sensing matrices and accelerate the speed of data sampling [22], [23]. In compressive imaging, the sensing process is sometimes unfixed and determined by the designed mask for sampling data. Due to sampling speed limitations, the designed masks are generally binary, the above optimization methods are not applicable because the binarization of optimized sensing matrices generally eliminates the gains generated by optimization. Therefore, people are devoted to designing a binary sensing matrix with small coherence, and various design methods of binary sensing matrices have been presented [24], [25], [26], [27], [28]. Despite these methods, designing a satisfactory binary sensing matrix is still difficult, and these methods are applicable to the CS problem with an unfixed sensing process determined by the design sensing matrices. This affects the promotion and application of CS technology. Moreover, in some fields, the data sampling process is fixed and the sensing matrix determined by the fixed sampling process is immutable, once the sensing matrix is with a large coherence, it is very hard to recover successfully sparse signals. Thus, an optimization method for reducing the coherence of matrix A without changing the original sampling process is particularly important.

For the general CS problem, in [29] and [30], authors optimized the CS model (1) by weighting the sensing matrix to equalize all singular values. The weighted method of [29] requires the singular value decomposition (SVD) of A with a lot of numerical computations and the reciprocals of all singular values, which may cause large round-off errors for those appropriately small singular values. In [30], we gave a method of reweighting the sensing matrix in (1) by left-multiplying the normal equation of (1) by multiple symmetric positive-definite matrices related to $A^TA \in \mathbb{R}^{N\times N}$. In CS, the signal length is usually much larger than the number of measurements approximated to the signal sparsity, i.e., $m \approx s \ll N$. This reweighted method has a high computational complexity for a large-sized sparse signal x. To this end, in this paper, we propose an improved reweighted method to optimize the coherence of the sensing matrix for the general CS problems by left-multiplying the original CS model (1) a symmetric positive-definite matrix related to $AA^T \in \mathbb{R}^{m\times m}$, obtain a new CS model (12), $A_nx = y_n$, equivalent to the original model (1), and prove that the condition number of the new sensing matrix A_n decreases monotonically to 1 as the weighting times $n \to +\infty$ in Theorem 1 and the coherence of the new CS model (12) is smaller than that of the model (1) in Theorem 2. Besides, by combining the proposed reweighted method with the idea of the orthogonal matching pursuit (OMP) algorithm, we also establish the reweighted OMP (RwOMP) recovery algorithm, and carry out the CS experiments on one-dimensional sparse signals and two-dimensional non-sparse images using different random binary matrices and different algorithms to verify that the recovery using proposed RwOMP algorithm outperforms the direct OMP recovery and the recovery using the SVD-based weighted OMP algorithms [29]. Compared with the weighted method of [29], our reweighted method only requires the maximum singular value, which is easily estimated. Meanwhile, due to $m \ll N$ in CS problems, the order of the weighting matrix in our reweighted method is much smaller than that of [30]. So our method has lower computational complexity and better stability.

The organization of the paper is as follows. Section II introduces the notations and several preliminary results. Section III gives a reweighted method for the CS model and its theoretical analysis of improving coherence. Section IV presents the reweighted OMP recovery algorithms and implements CS experiments. Section V concludes the paper.

II. PRELIMINARIES

In this section, we give several important definitions and preliminary results. Let $\|\cdot\|_F$ and $\|\cdot\|_2$ be the Frobenius norm and 2-norm of a matrix, respectively. We use ^{T} to denote the transpose of a matrix or a vector. Let $\{\sigma_k\}_{k=1}^{\bar{m}}$ with the algebraic multiplicity p_k be the nonzero singular values of A such that $\sigma_1 > \sigma_2 > \ldots > \sigma_{\bar{m}}$ and $m = \sum_{k=1}^{\bar{m}} p_k$.

Definition 2: For a matrix $A \in \mathbb{R}^{m\times N}(m \ll N)$, denote its largest singular value and the smallest nonzero singular value by $\sigma_{\max}$ and $\sigma_{\min}$, respectively. The condition number of A is defined as

$$\mathcal{K}(A) := \frac{\sigma_{\max}}{\sigma_{\min}}. \tag{3}$$

For greedy algorithms such as the OMP, the highly correlated columns of the sensing matrix A can cause the wrong indices and thus cannot recover correctly sparse signals.

Definition 3 ([11]): For a matrix $A \in \mathbb{R}^{m\times N}$, its mutual coherence is

$$\mu(A) := \max_{1\le i\ne j\le N} \frac{|A_i^TA_j|}{\|A_i\|_2\|A_j\|_2}, \tag{4}$$

where A_i is the i-th column vector of A. The mutual coherence $\mu(A)$ measures the maximum absolute correlation between any two different normalized columns of A. The smaller the $\mu(A)$, the more accurate the reconstructions. It is known that $\mu(A) \ge \sqrt{\frac{N-m}{m(N-1)}}$, where $\sqrt{\frac{N-m}{m(N-1)}}$ is called as the Welch bound [31]. The equality holds if and only if the column-normalized vectors of A form an equiangular tight frame (ETF) [32]. However, the ETF usually does not exist in CS problems since $m \ll N$.

Assume that A is a matrix with ℓ_2-normalized columns. Let the Gram matrix $G(A) = [g_{i,j}]_{1\leq i,j\leq N}$. Then (4) can be rewritten into that $\mu(A) = \max_{1\leq i\neq j\leq N} |A_i^T A_j| = \max_{1\leq i\neq j\leq N} |g_{i,j}|$. The absolute off-diagonal entries, $|g_{i,j}|$ for $1 \leq i \neq j \leq N$, quantify the correlations between any two different columns of A. The coherence $\mu(A)$ only accounts for the largest absolute off-diagonal entry and reflects the worst similarity between different columns of A. Compared to the mutual coherence $\mu(A)$, the average mutual coherence of A can better reflect the overall orthogonality level of the submatrices of A [33], [34].

Definition 4 ([35]): For the sensing matrix A, its average mutual coherence

$$\bar{\mu}(A) := \frac{\sum_{1\leq i\neq j\leq N} {g_{i,j}}^2}{N(N-1)}, \tag{5}$$

where $g_{i,j}$ is the element in the i-th row and j-th column of $G(A)$. The nonzero eigenvalues of $G(A)$ are $\{{\sigma_k}^2\}_{k=1}^{\bar{m}}$ with the algebraic multiplicity p_k. From the properties of the Frobenius norm and it follows that $\|G(A)\|_F = \sum_{i,j=1}^{N} {g_{i,j}}^2 = \sum_{k=1}^{\bar{m}} p_k {\sigma_k}^4$. Since $g_{i,i} = 1$ for $i = 1, 2, \cdots, N$, $\sum_{1\leq i\leq N} {g_{i,i}}^2 = N$. By (5), we have that

$$\bar{\mu}(A) = \frac{\|G(A)\|_F - N}{N(N-1)}. \tag{6}$$

Lemma 1: Let $A \in \mathbb{R}^{m\times N}(m \ll N)$ be a column-normalized matrix of rank m, $\{\sigma_k\}_{k=1}^{\bar{m}}$ be its nonzero singular values with the algebraic multiplicity p_k such that $m = \sum_{k=1}^{\bar{m}} p_k$ and $G(A)$ be its Gram matrix. Let $\beta(\sigma) = \sum_{k=1}^{\bar{m}} p_k({\sigma_k}^2 - \frac{N}{m})^2$, where $\sigma = (\sigma_1, \sigma_2, \cdots, \sigma_{\bar{m}})$. Then we have that

- the smaller $\beta(\sigma)$, the smaller $\|G(A)\|_F$.
- $\|G(A)\|_F \geq \frac{N^2}{m}$ and equality holds if and only if ${\sigma_1}^2 = {\sigma_2}^2 \cdots = {\sigma_{\bar{m}}}^2 = \frac{N}{m}$.

The proof is provided in Appendix A.

By the equality (6) and the Lemma 1, it is illustrated that the smaller the value of $\beta(\sigma)$ where $\frac{N}{m}$ is the average of eigenvalues, the smaller the average mutual coherence $\bar{\mu}(A)$, and $\bar{\mu}(A)$ reaches to its minimum, $\frac{N-m}{m(N-1)}$, if and only if ${\sigma_1}^2 = \cdots = {\sigma_{\bar{m}}}^2 = \frac{N}{m}$ which means the condition number $\mathcal{K}(A) = 1$. Moreover, the smaller the condition number $\mathcal{K}(A)$, the more eigenvalues near its average $\frac{N}{m}$, which means a smaller value of $\beta(\sigma)$. Thus, a method of improving condition number $\mathcal{K}(A)$ can reduce the value of $\beta(\sigma)$ by improving the distribution of eigenvalues, and then improve the average mutual coherence $\bar{\mu}(A)$.

III. A REWEIGHTED METHOD AND THE IMPROVEMENT OF COHERENCE

In this section, we propose a reweighted method and prove that it improves the coherence of the sensing matrix under a certain condition.

Let r be the rank of matrix $A \in \mathbb{R}^{m\times N}$ such that $r \leq m \ll N$. Denote $\{\sigma_k\}_{k=1}^{\bar{m}}$ with the algebraic multiplicity p_k as the nonzero singular values of A such that $\sigma_1 > \sigma_2 > \ldots > \sigma_{\bar{m}}$ and $r = \sum_{k=1}^{\bar{m}} p_k$. For $j = 1, 2, \cdots p_k$ and $k = 1, \ldots, \bar{m}$, we use $\{u_{j,k}\}_{k=1}^{\bar{m}}$ and $\{v_{j,k}\}_{k=1}^{\bar{m}}$ to denote the left and right singular vectors corresponding to the singular value σ_k, respectively. Denote w_j $(1 \leq j \leq q)$ as the eigenvectors of the zero singular values of A with an algebraic multiplicity $q = N - r$. Let $U = (U_1, U_2, \cdots, U_{\bar{m}}) \in \mathbb{R}^{m\times m}$ and $V = (V_1, V_2, \cdots, V_{\bar{m}}, W) \in \mathbb{R}^{N\times N}$, where $W = (w_1, w_2, \cdots, w_q)$, $U_k = (u_{1,k}, u_{2,k}, \cdots, u_{p_k,k})$ and $V_k = (v_{1,k}, v_{2,k}, \cdots, v_{p_k,k})$ for $k = 1, 2, \cdots \bar{m}$. By the SVD of matrix A, then

$$U^T AV = \left(S, O_{m\times(N-m)}\right), \tag{7}$$

where $S = diag\left(\sigma_1 I_{p_1}, \cdots, \sigma_{\bar{m}} I_{p_{\bar{m}}}, O_{m-r}\right)$, I is a unit matrix, and O is a null matrix. Obviously, the matrix $AA^T \in \mathbb{R}^{m\times m}$ is symmetric semi-positive definite, its eigenvalues $\{{\sigma_k}^2\}_{k=1}^{\bar{m}}$ satisfy ${\sigma_1}^2 > {\sigma_2}^2 > \ldots > {\sigma_{\bar{m}}}^2 > 0$. Given a constant $a_0 > 1$, with the same U in (7), the matrix $a_0{\sigma_1}^2 I_m - AA^T$ can be diagonalized into

$$\begin{aligned} &U^T(a_0{\sigma_1}^2 I_m - AA^T)U \\ &= diag\big((a_0{\sigma_1}^2 - {\sigma_1}^2)I_{p_1}, \\ &\quad \cdots, (a_0{\sigma_1}^2 - {\sigma_{\bar{m}}}^2)I_{p_{\bar{m}}}, a_0{\sigma_1}^2 I_{m-r}\big). \end{aligned} \tag{8}$$

From ${\sigma_1}^2 > {\sigma_2}^2 > \ldots > {\sigma_{\bar{m}}}^2 > 0$ and it follows that $0 < a_0{\sigma_1}^2 - {\sigma_1}^2 < a_0{\sigma_1}^2 - {\sigma_2}^2 < \ldots < a_0{\sigma_1}^2 - {\sigma_{\bar{m}}}^2 < a_0{\sigma_1}^2$. So the matrix $a_0{\sigma_1}^2 I_m - AA^T$ is symmetric positive-definite. To improve the condition number of the CS model (1), we give the following reweighted method,

$$A_0 = A, \quad y_0 = y, * \tag{9a}$$

$$V_n^{-1} = a_{n-1}\sigma_{n-1,1}^2 I_m - A_{n-1}A_{n-1}^T, \tag{9b}$$

$$A_n = V_n^{-1} A_{n-1}, \quad y_n = V_n^{-1} y_{n-1}, \tag{9c}$$

where constants $a_{n-1} > 1$ are weighted parameters, $\sigma_{n-1,1}$ are the largest singular values of A_{n-1}, and the weighting times $n = 1, 2, \ldots$.

Theorem 1: For $n = 0, 1, \ldots$, assume that $a_n \geq 1 + 1/\mathcal{K}(A_n) + 1/\mathcal{K}^2(A_n)$, $\{a_n\}$ is a convergent sequence with $\lim_{n\to\infty} a_n = \bar{a} \geq 3$ and let $\{A_n\}$ be a matrix sequence defined by (9c). Then, we have

$$\mathcal{K}(A_{n+1}) < \mathcal{K}(A_n), \tag{10}$$

and

$$\lim_{n\to\infty} \mathcal{K}(A_n) = 1. \tag{11}$$

The proof is provided in Appendix B.

With the reweighted method (9a), the sensing process (1) can be written equivalently as

$$A_n x = y_n, \tag{12}$$

where A_n can be regarded as a new sensing matrix equivalent to the original sensing matrix A. We use the CS model (12) to recover x.

Theorem 2: Given a positive integer $n_1 \geq 1$ and two fixed positive numbers d_0 and d_1. Let $\bar{A}_{n_1}$ and $\bar{A}$ be the column-normalized matrices of A_{n_1} and A, respectively.

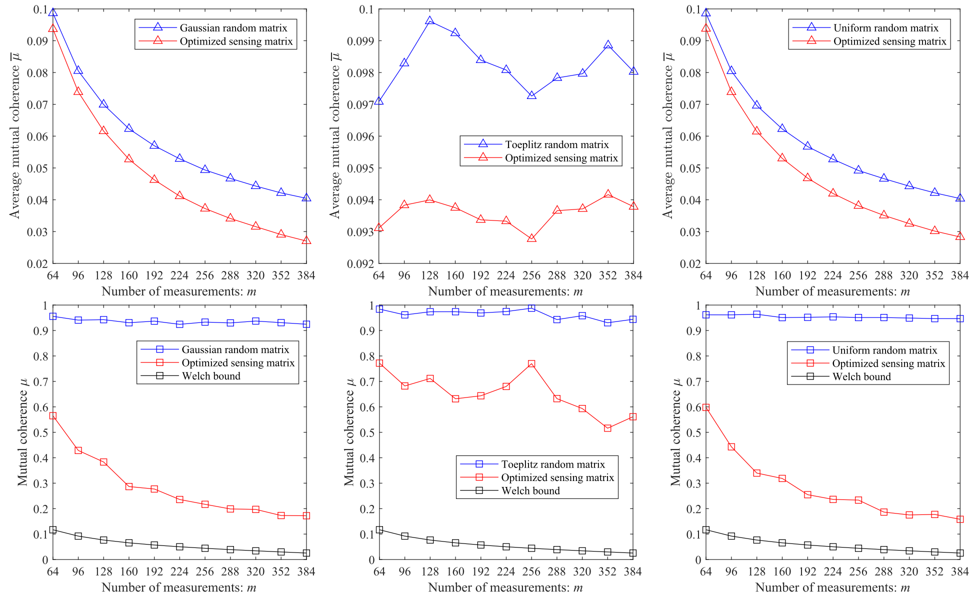

FIGURE 1. **Average mutual coherence (1st row) and mutual coherence (2nd row) as the functions of the number of measurements m, where the sensing matrix is Gaussian random matrix (left), Toeplitz random matrix (middle) and uniform random matrix (right), respectively. The weighting times $n = 5$ and the weighted parameters $a_k = 1.9$ for $k = 1, 2, \cdots, 5$.**

Assume that $|\mathcal{K}(\bar{A}) - \mathcal{K}(A)| \leq d_0$ and $|\mathcal{K}(\bar{A}_{n_1}) - \mathcal{K}(A_{n_1})| \leq d_{n_1}$. Then if $\mathcal{K}(A) - \mathcal{K}(A_{n_1}) > d_0 + d_{n_1}$, we have

$$\bar{\mu}(A_{n_1}) < \bar{\mu}(A). \tag{13}$$

The proof is given in Appendix C.

IV. EXPERIMENT

In this section, we implement several experiments consisted of two parts, i.e., the one-dimensional sparse signal recovery in subsection IV-A and two-dimensional compressive coded aperture imaging in subsection IV-B, respectively. By combining our reweighted method with the OMP idea, we give a new algorithm, namely RwOMP, as displayed in Algorithm 1. In the following experiments, the weighting times is determined by repeated experiments, the maximum weighting times $n = 9$, and the weighted parameters a_n are set to 1.9 (in Figs. 1, 3 and 4) and 3 (in Fig. 2), respectively. The initial sensing matrix A in (1) is a $\{0, 1\}$ binary random matrix. Compared with the reweighted method of [30], the reweighted method in this paper has a lower computational complexity, i.e., the order of the weighting matrix is much lower than that of [30]. Their reconstruction results are almost the same and not be compared here. For the sake of illustration, we use SVD-WOMP to denote the SVD-based weighted OMP recovery algorithms of [29]. To verify the superiority of the proposed algorithm, we compare the RwOMP algorithm with the OMP [36] and SVD-WOMP algorithms by recovering sparse signals. All the experiments were run on the MATLAB code (if needed, please contact us via email) on a standard PC with a 2.11 GHz Intel Core i5 processor and 16 GB of memory, running on the Windows 11 system.

Algorithm 1 RwOMP Algorithm

Input:
observed data y, Sensing matrix A, sparsity s, the maximum weighting times $\bar{n}$, $T^0 = \emptyset$, $x^0 = 0$.
Reweighting the CS model (1):
for $n = 1, \ldots, \bar{n}$ do
1: the calculation of weighting matrix V_n^{-1} with (9b)
2: the calculations of the matrix A_n and the vector y_n with (9c)
Iteration:
Repeat until a stopping criterion is met at $k = \bar{k}$
1: $T^{k+1} = T^k \bigcup \{j_{k+1}\}$,
2: $j_{k+1} = \arg\max_{j \in [1,\ldots,s]} \left\{ \left| \left(A_n^T \left(y_n - A_n x^k \right) \right)_j \right| \right\}$,
3: $x^{k+1} = \arg\min_{z \in \mathbb{R}^n} \left\{ \|y_n - A_n z\|_2 \right\}$, $\sup(z) \subseteq T^{k+1}$.
Output: $x^{\bar{k}}$

A. ONE-DIMENSIONAL SPARSE SIGNAL RECOVERY

In the following CS experiments, the recovered one-dimensional sparse signal x is with the length $N = 512$ and the sparsity $s = 1, 3, 5 \cdots, 65$. The OMP and RwOMP algorithms are used to recover sparse signals, respectively. We choose the Gaussian random binary matrix, the Toeplitz random binary matrix and the uniform random binary matrix as the initial sensing matrix A in (1), respectively. Fig. 1 shows the average mutual coherence $\bar{\mu}$ and the mutual coherence μ of sensing matrices as a function of the number of measurements $m = 64, 96 \cdots, 384$. We can find that the average mutual coherence $\bar{\mu}$ and the mutual coherence μ of the optimized sensing matrices are better than those of the

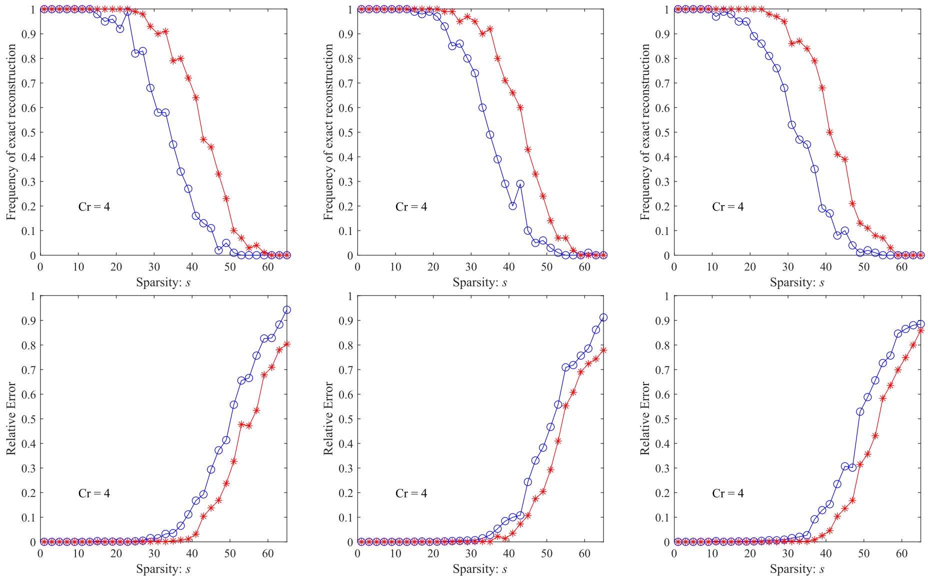

FIGURE 2. **The reconstruction results of OMP (blue) and RwOMP (red: the weighting times $n = 8$ and the weighted parameters $a_k = 3$ for $k = 1, 2, \cdots, 8$). Left: Gaussian random matrix. Middle: Toeplitz random matrix. Right: uniform random matrix.**

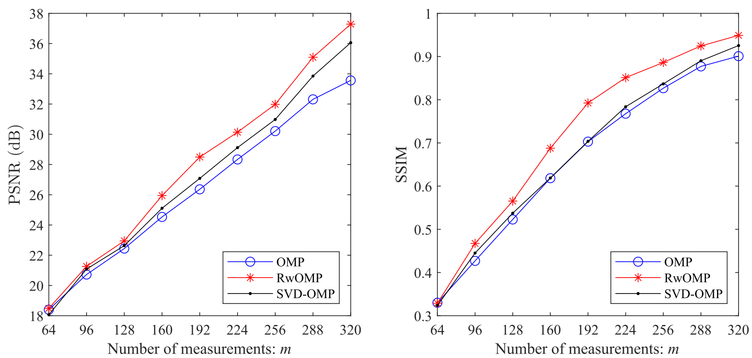

FIGURE 3. **PSNR and SSIM of reconstruction results of the cameraman image as a function of the measurement number m. For each m, the weighting times n is 2, 6, 5, 9, 8, 9, 7, 6 and 9, respectively. The weighted parameters $a_k = 1.9$ for $k = 1, 2, \cdots, 9$.**

initial sensing matrices, and the mutual coherence μ of the optimized sensing matrices is closer to the Welch bound.

In the one-dimensional sparse signal recovery, we use the frequency of exact reconstruction and the relative error as two indexes to evaluate the reconstruction results. For every sparsity s, we reconstruct repeatedly 100 different random s sparse signals obeying Gaussian distribution to compute the frequency of exact reconstruction and the relative error between the original signals x and the reconstructions $\bar{x}$, which is defined as follows

$$\text{Relative Error} := \frac{\| x - \bar{x} \|_2}{\| x \|_2}. \tag{14}$$

Fig. 2 shows the frequency of exact reconstruction and the relative error with different sparsity levels, where the measurement number is fixed to 128, i.e., the compression rate (Cr) is 4. As can be seen from Fig. 2, for a fixed measurement number m, the sparser the signal, the higher the frequency of exact reconstruction and the smaller the relative error. Moreover, the RwOMP algorithm with the weighting times $n = 8$ can improve the frequency of the exact reconstruction and the relative error compared with the direct OMP reconstruction.

B. TWO-DIMENSIONAL IMAGE RECONSTRUCTION

Next, we conduct the compressive coded aperture imaging experiments. In the following experiments, the original image is the cameraman with a size of 512×512, the number of measurements $m = 64, 96 \cdots, 320$. The reconstruction is done column by column. More specifically, each column of the target image is first encoded with m different mask patterns to obtain measurements, and then decoded through CS algorithms to achieve reconstruction. Thus, the sensing matrix A has a fixed number of columns $N = 512$ and the number of rows m varying from 64 to 320. Considering that the original image signal itself is non-sparse, we take the sparse transform, Ψ, as a discrete symlet wavelet transform matrix to give the sparse representation of x and as follows

$$\tilde{x} = \Psi^{-1} x. \tag{15}$$

FIGURE 4. Reconstructed results of different algorithms under the number of measurements $m = 192$. The first is the original cameraman image, the second is the image reconstructed by the OMP (PSNR=26.35db, SSIM = 0.7030), the third is the image reconstructed by the SVD-WOMP (PSNR=27.08db, SSIM = 0.7036), and the fourth is the image reconstructed by the RwOMP with the weighting times $n = 8$ (PSNR=28.49db, SSIM = 0.7936). The weighted parameters $a_k = 1.9$ for $k = 1, 2, \cdots, 8$.

Thus, the sensing process (1) can be rewritten as

$$A\Psi\tilde{x} = y. \tag{16}$$

Now, $A\Psi$ is the sensing matrix, we reconstruct $\tilde{x}$ by (16), and then get x through (15). We use the OMP, SVD-WOMP, and RwOMP algorithms to reconstruct x, respectively.

The peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) are two important indexes for evaluating the quality of reconstruction images. Fig. 3 gives the PSNR and SSIM curves of the reconstruction results of different algorithms, where the number of measurements $m = 64, 96 \cdots, 320$. Fig. 4. shows the reconstruction results of different algorithms, where the number of measurements is 192. We can see that the direct OMP reconstruction performs worst, and the reconstruction results of the RwOMP algorithm in this paper are better than those of the other two algorithms.

V. CONCLUSION

In this paper, we have proposed a method of reweighting the CS models, established its complete theoretical analysis of improving the coherence of sensing matrices, and obtained a coherence improvement CS model (12) equivalent to the original model (1). We have also provided the RwOMP algorithm for recovering sparse signals by incorporating the proposed reweighted method into the idea of the OMP algorithm, and implemented the CS recovery experiments of the one-dimensional sparse signals and the two-dimensional non-sparse images to verify its superiority and effectiveness.

It is noteworthy that our method of improving the coherence of sensing matrices is applicable to general CS models with a linear sampling process. Two of the factors affecting the practical application of CS technology are: 1) the fixed sensing matrix determined by the measurement process is not with a sufficiently small coherence, and 2) it is tough to design a satisfactory binary sensing matrix with a small coherence due to certain limitations. For the two factors, our method can reduce effectively the coherence of sensing matrices and then improve the recovery results. Thus, it is helpful to expand the application field of CS technology.

In a practical application of the reweighted method proposed in this paper, the weighting times n may be increased appropriately according to the actual recovery requirement, and it is noteworthy that the weighting times n cannot be very large to trade off between the improvement of coherence and the errors generated by discretization and accumulated with the increase of n. According to the condition of Theorem 1, $a_n \geq 1 + 1/\mathcal{K}(A_n) + 1/\mathcal{K}^2(A_n)$, the weighted parameters are estimated reasonably, and our method always improves the reconstruction results. A comprehensive analysis of maximum weighting times and optimal sequence parameters is our future work.

APPENDIX A
PROOF OF LEMMA 1

Proof: For the Gram matrix $G(A)$, it is easy to know that its trace $tr\big(G(A)\big) = N$. Then the sum of its eigenvalues $\sum_{k=1}^{\bar{m}} p_k \sigma_k{}^2 = tr\big(G(A)\big) = N$, and $\frac{N}{m}$ is the average of eigenvalues. Expanding the expression $\beta(\sigma)$ yields that $\beta(\sigma) = \|G(A)\|_F - \frac{N^2}{m}$. Then we have that

$$\|G(A)\|_F = \beta(\sigma) + \frac{N^2}{m} \geq \frac{N^2}{m}. \tag{17}$$

Thus, the smaller the value of $\beta(\sigma)$, the smaller the value of $\|G(A)\|_F$, and $\|G(A)\|_F$ reaches the minimum $\frac{N^2}{m}$ if and only if $\beta(\sigma) = 0$, i.e., $\sigma_1{}^2 = \cdots = \sigma_{\bar{m}}{}^2 = \frac{N}{m}$. ∎

APPENDIX B
PROOF OF THEOREM 1

Proof: For $n = 0, 1 \cdots$, denote $\{\sigma_{n,k}\}_{k=1}^{\bar{m}}$ with the algebraic multiplicity p_k as the nonzero singular values of A_n such that $\sigma_{n,1} > \sigma_{n,2} > \ldots > \sigma_{n,\bar{m}}$. By the definition of A_n in (9c), with the same U and V in (7), the matrices A_n and A_{n+1} satisfy

$$U^T A_n V = \left(diag\left(\sigma_{n,1} I_{p_1}, \cdots, \sigma_{n,\bar{m}} I_{p_{\bar{m}}}, O_{m-r}\right), O_{m\times(N-m)}\right), \tag{18}$$

and

$$\begin{aligned} &U^T A_{n+1} V \\ &= \Big(diag\big((a_n\sigma_{n,1}^2 - \sigma_{n,1}^2)\sigma_{n,1} I_{p_1}, \\ &\quad \cdots, (a_n\sigma_{n,1}^2 - \sigma_{n,\bar{m}}^2)\sigma_{n,\bar{m}} I_{p_{\bar{m}}}, O_{m-r}\big), O_{m\times(N-m)}\Big), \end{aligned} \tag{19}$$

respectively. Define the function with respect to $t \in (0, +\infty)$

$$\varphi_n(t) := (a_n\sigma_{n,1}^2 - t^2)t. \tag{20}$$

By (19) and (20), the nonzero singular values of A_{n+1} can be represented as $\{\varphi_n(\sigma_{n,k})\}_{k=1}^{\bar{m}}$, and $\varphi_n(\sigma_{n,k}) > 0$ for $k = 1, 2, \cdots, \bar{m}$, $n = 0, 1, \cdots$. $\{\varphi_n(\sigma_{n,k})\}_{k=1}^{\bar{m}}$ are permuted as $\{\sigma_{n+1,k}\}_{k=1}^{\bar{m}}$ satisfying $\sigma_{n+1,1} > \sigma_{n+1,2} > \cdots \sigma_{n+1,\bar{m}} > 0$. Taking the derivative of a function $\varphi_n(t)$ yields that $\varphi_n(t)' = a_n\sigma_{n,1}^2 - 3t^2$, and it follows that $\arg\max_{t>0}\varphi_n(t) = \sqrt{a_n/3}\sigma_{n,1}$. We therefore have

$$\sigma_{n+1,1} \leq \varphi_n\left(\sqrt{\frac{a_n}{3}}\sigma_{n,1}\right) \tag{21}$$

and

$$\sigma_{n+1,\bar{m}} = \min\left\{\varphi_n(\sigma_{n,1}), \varphi_n(\sigma_{n,\bar{m}})\right\}. \tag{22}$$

If $a_n > 3$, $\sigma_{n,\bar{m}} < \sigma_{n,1} < \sqrt{\frac{a_n}{3}}\sigma_{n,1}$. Then $\sigma_{n+1,1} = \varphi_n(\sigma_{n,1})$ and $\sigma_{n+1,\bar{m}} = \varphi_n(\sigma_{n,\bar{m}})$. So

$$\mathcal{K}(A_{n+1}) = \frac{\sigma_{n+1,1}}{\sigma_{n+1,\bar{m}}} = \frac{(a_n\sigma_{n,1}^2 - \sigma_{n,1}^2)\sigma_{n,1}}{(a_n\sigma_{n,1}^2 - \sigma_{n,\bar{m}}^2)\sigma_{n,\bar{m}}} < \frac{\sigma_{n,1}}{\sigma_{n,\bar{m}}} = \mathcal{K}(A_n). \tag{23}$$

If $1 + \frac{1}{\mathcal{K}(A_n)} + \frac{1}{\mathcal{K}^2(A_n)} \leq a_n \leq 3$, then

$$(a_n - 1)\sigma_{n,1}^2 \geq \sigma_{n,\bar{m}}\sigma_{n,1} + \sigma_{n,\bar{m}}^2. \tag{24}$$

By the above inequality (24) and $\sigma_{n,1} - \sigma_{n,\bar{m}} > 0$, we have

$$\begin{aligned}\varphi_n(\sigma_{n,1}) - \varphi_n(\sigma_{n,\bar{m}}) &= (a_n - 1)\sigma_{n,1}^3 \\ &- (a_n\sigma_{n,1}^2 - \sigma_{n,\bar{m}}^2)\sigma_{n,\bar{m}} = (\sigma_{n,1} - \sigma_{n,\bar{m}}) \\ &\times [(a_n - 1)\sigma_{n,1}^2 - \sigma_{n,\bar{m}}\sigma_{n,1} - \sigma_{n,\bar{m}}^2] \geq 0.\end{aligned} \tag{25}$$

By (22) and (25), then

$$\sigma_{n+1,\bar{m}} = \varphi_n(\sigma_{n,\bar{m}}). \tag{26}$$

From $1 + \frac{1}{\mathcal{K}(A_n)} + \frac{1}{\mathcal{K}^2(A_n)} \leq a_n \leq 3$ and it follows that $a_n\mathcal{K}^2(A_n) - 1 \geq \mathcal{K}^2(A_n) + \mathcal{K}(A_n) \geq 2$. Combining (21) and (26) yields that

$$\begin{aligned}\mathcal{K}(A_{n+1}) &= \frac{\sigma_{n+1,1}}{\sigma_{n+1,\bar{m}}} \leq \frac{\varphi_n\left(\sqrt{\frac{a_n}{3}}\sigma_{n,1}\right)}{\varphi_n(\sigma_{n,\bar{m}})} \\ &= \frac{2a_n\sqrt{a_n}\sigma_{n,1}^3}{3\sqrt{3}(a_n\sigma_{n,1}^2 - \sigma_{n,\bar{m}}^2)\sigma_{n,\bar{m}}} \\ &= \frac{2a_n\sqrt{a_n}\mathcal{K}^2(A_n)}{3\sqrt{3}\left(a_n\mathcal{K}^2(A_n) - 1\right)}\mathcal{K}(A_n) \\ &= \frac{2\sqrt{a_n}}{3\sqrt{3}}\left(1 + \frac{1}{a_n\mathcal{K}^2(A_n) - 1}\right)\mathcal{K}(A_n) \\ &< \frac{2}{3}\left(1 + \frac{1}{2}\right)\mathcal{K}(A_n) = \mathcal{K}(A_n).\end{aligned} \tag{27}$$

Thus the inequality (10) holds for $a_n \geq 1 + \frac{1}{\mathcal{K}(A_n)} + \frac{1}{\mathcal{K}^2(A_n)}$.

The condition number $\mathcal{K}(A_n) \geq 1$ is monotonically decreasing, so the limit of $\mathcal{K}(A_n)$ exists. Assume that

$$\lim_{n\to\infty}\mathcal{K}(A_n) = z \geq 1. \tag{28}$$

For $\bar{a} > 3$, taking the limit on both sides of (23) yields

$$1 \leq z = \frac{(\bar{a} - 1)z^3}{\bar{a}z^2 - 1}. \tag{29}$$

Solving (29) and we have $z = 1$. Then the equality of (11) holds for $\bar{a} > 3$.

For $\bar{a} = 3$, let $\{a_{n_{j1}}\}$ and $\{a_{n_{j2}}\}$ be the two subsequences of $\{a_n\}$ such that $1 + \frac{1}{\mathcal{K}(A_{n_{j1}})} + \frac{1}{\mathcal{K}(A_{n_{j1}})^2} \leq a_{n_{j1}} \leq 3$, $a_{n_{j2}} > 3$, and $\{a_n\} = \{a_{n_{j1}}\} \bigcup \{a_{n_{j2}}\}$. At least one of the two subsequences is an infinite sequence. If $\{a_{n_{j1}}\}$ is an infinite subsequence, replacing n by n_{j1} in (27) and this yields

$$\mathcal{K}(A_{n_{j1}+1}) \leq \frac{2\sqrt{a_{n_{j1}}}}{3\sqrt{3}}\left(1 + \frac{1}{a_{n_{j1}}\mathcal{K}^2(A_{n_{j1}}) - 1}\right)\mathcal{K}(A_{n_{j1}}). \tag{30}$$

Due to that $\mathcal{K}(A_{n_{j1}}) \geq 1$, we can rewrite (30) as

$$\frac{\mathcal{K}(A_{n_{j1}+1})}{\mathcal{K}(A_{n_{j1}})} \leq \frac{2\sqrt{a_{n_{j1}}}}{3\sqrt{3}}\left(1 + \frac{1}{a_{n_{j1}}\mathcal{K}^2(A_{n_{j1}}) - 1}\right). \tag{31}$$

From (28) and it follows that $\lim_{j1\to\infty}\frac{\mathcal{K}(A_{n_{j1}+1})}{\mathcal{K}(A_{n_{j1}})} = 1$, $\lim_{j1\to\infty}2\sqrt{a_{n_{j1}}} = 2\sqrt{3}$, and $\lim_{j1\to\infty}a_{n_{j1}}\mathcal{K}^2(A_{n_{j1}}) = 3z^2$. By the definition of the limit of sequence, for every $\varepsilon > 0$, there exists an integer $\bar{n} > 0$, when $j1 \geq \bar{n}$, it follows that $1 - \varepsilon \leq \frac{\mathcal{K}(A_{n_{j1}+1})}{\mathcal{K}(A_{n_{j1}})} \leq 1 + \varepsilon$, $2\sqrt{3} - \varepsilon \leq \sqrt{2a_{n_{j1}}} \leq 2\sqrt{3} + \varepsilon$ and $3z^2 - \varepsilon \leq a_{n_{j1}}\mathcal{K}^2(A_{n_{j1}}) \leq 3z^2 + \varepsilon$. By (31), then

$$1 - \varepsilon \leq \frac{2\sqrt{3} + \varepsilon}{3\sqrt{3}}\left(1 + \frac{1}{3z^2 - \varepsilon - 1}\right). \tag{32}$$

Solving (32) gives that $1 \leq z^2 \leq 1 + \frac{1}{3}\left(1 + \frac{5+6\sqrt{3}}{\sqrt{3}-(2+3\sqrt{3})\varepsilon}\right)\varepsilon$. So $z^2 = 1$. By (28) and thus $\lim_{j1\to\infty}\mathcal{K}(A_{n_{j1}}) = 1$. If $\{a_{n_{j2}}\}$ is an infinite sequence, we also have $\lim_{j2\to\infty}\mathcal{K}(A_{n_{j2}}) = 1$. The proof is similar and not given here. Then $\lim_{n\to\infty}\mathcal{K}(A_n) = 1$ for $\bar{a} = 3$. Hence, (10) and (11) hold for $a_n \geq 1 + \frac{1}{\mathcal{K}(A_n)} + \frac{1}{\mathcal{K}(A_n)^2}$ and $\lim_{n\to\infty}a_n = \bar{a} \geq 3$. The theorem is established. ■

APPENDIX C
PROOF OF THE THEOREM 2

Proof: From the assumptions, it follows that

$$\begin{aligned}\mathcal{K}(A) - d_0 \leq \mathcal{K}(\bar{A}) \leq \mathcal{K}(A) + d_0, \\ \mathcal{K}(A_{n_1}) - d_{n_1} \leq \mathcal{K}(\bar{A}_{n_1}) \leq \mathcal{K}(A_{n_1}) + d_{n_1}.\end{aligned} \tag{33}$$

If $\mathcal{K}(A) - \mathcal{K}(A_{n_1}) > d_0 + d_{n_1}$, i.e., $\mathcal{K}(A_{n_1}) + d_{n_1} < \mathcal{K}(A) - d_0$, by the inequality (33), then

$$\mathcal{K}(\bar{A}_{n_1}) \leq \mathcal{K}(A_{n_1}) + d_{n_1} < \mathcal{K}(A) - d_0 \leq \mathcal{K}(\bar{A}). \tag{34}$$

Let $\{{\sigma_k}^2\}_{k=1}^{\bar{m}}$ with the algebraic multiplicity p_k and $\{{\sigma_{n_1,k}}^2\}_{k=1}^{\bar{m}}$ with the algebraic multiplicity $p_{n_1,k}$ be the positive eigenvalues of the positive semidefinite Gram matrices $G(\bar{A})$ and $G(\bar{A}_{n_1})$, respectively. Since the matrices $\bar{A}_{n_1}$ and $\bar{A}$ are column-normalized, $\sum_{k=1}^{\bar{m}}p_k{\sigma_k}^2 = \sum_{k=1}^{\bar{m}}p_{n_1,k}{\sigma_{n_1,k}}^2 = N$ and then the mean values of the eigenvalues $\{{\sigma_k}^2\}_{k=1}^{\bar{m}}$

and $\{\sigma_{n_1,k}{}^2\}_{k=1}^{\bar{m}}$ are both $\frac{N}{m}$. Because the sum of eigenvalues of matrix $\bar{A}_{n_1}$ is equal to that of $\bar{A}$, and $\mathcal{K}(\bar{A}_{n_1}) < \mathcal{K}(\bar{A})$, then the sum of the distances between the eigenvalues $\sigma_{n_1,k}{}^2$ and the mean value $\frac{N}{m}$ is smaller than that of $\sigma_k{}^2$ for $k = 1, 2, \cdots, \bar{m}$. More specifically, $\sum_{k=1}^{\bar{m}} p_{n_1,k}(\sigma_{n_1,k}{}^2 - \frac{N}{m})^2 < \sum_{k=1}^{\bar{m}} p_k(\sigma_k{}^2 - \frac{N}{m})^2$, i.e., $\beta(\sigma_{n_1}) < \beta(\sigma)$. From Lemma 2.4 and (6), the inequality (13) is held, and the Theorem 2 is established. ■

ACKNOWLEDGMENT

The authors would like to thank the editor and the reviewers for their detailed comments and valuable suggestions.

REFERENCES

[1] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Boston, MA, USA: Birkh, 2013, pp. 1–8.

[2] M. Rudelson and R. Vershynin, "Sparse reconstruction by convex relaxation: Fourier and Gaussian measurements," in *Proc. 40th Annu. Conf. Inf. Sci. Syst.*, Mar. 2006, pp. 207–212.

[3] R. Wu and D.-R. Chen, "The improved bounds of restricted isometry constant for recovery via ℓ_p-Minimization," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 6142–6147, Sep. 2013.

[4] C.-B. Song, S.-T. Xia, and X.-J. Liu, "Improved analysis for subspace pursuit algorithm in terms of restricted isometry constant," *IEEE Signal Process. Lett.*, vol. 21, no. 11, pp. 1365–1369, Nov. 2014.

[5] J. Wang and P. Li, "Recovery of sparse signals using multiple orthogonal least squares," *IEEE Trans. Signal Process.*, vol. 65, no. 8, pp. 2049–2062, Apr. 2017.

[6] H. Li, J. Zhang, and J. Zou, "Improving the bound on the restricted isometry property constant in multiple orthogonal least squares," *IET Signal Process.*, vol. 12, no. 5, pp. 666–671, Jul. 2018.

[7] H. Ge, L. Wang, J. Wen, and J. Xian, "An RIP condition for exact support recovery with covariance-assisted matching pursuit," *IEEE Signal Process. Lett.*, vol. 26, no. 3, pp. 520–524, Mar. 2019.

[8] H. Liu, J. Peng, and Z. Lin, "A theoretical result of sparse signal recovery via alternating projection method," *Inf. Sci.*, vol. 506, pp. 51–57, Jan. 2020.

[9] A. Ravelomanantsoa, H. Rabah, and A. Rouane, "Compressed sensing: A simple deterministic measurement matrix and a fast recovery algorithm," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 12, pp. 3405–3413, Dec. 2015.

[10] M. Sadeghi and M. Babaie-Zadeh, "Dictionary learning with low mutual coherence constraint," *Neurocomputing*, vol. 407, pp. 163–174, Sep. 2020.

[11] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.

[12] R. A. DeVore, "Deterministic constructions of compressed sensing matrices," *J. Complex.*, vol. 23, nos. 4–6, pp. 918–925, Aug. 2007.

[13] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[14] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[15] T. L. N. Nguyen and Y. Shin, "Deterministic sensing matrices in compressive sensing: A survey," *Sci. World J.*, vol. 2013, pp. 1–6, Oct. 2013.

[16] V. Saligrama, "Deterministic designs with deterministic guarantees: Toeplitz compressed sensing matrices, sequence design and system ID," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 1–15, Jan. 2012.

[17] V. Abolghasemi, S. Ferdowsi, B. Makkiabadi, and S. Sanei, "On optimization of the measurement matrix for compressive sensing," in *Proc. 18th Eur. Signal Process. Conf.*, Aug. 2010, pp. 427–431.

[18] V. Abolghasemi, S. Ferdowsi, and S. Sanei, "A gradient-based alternating minimization approach for optimization of the measurement matrix in compressive sensing," *Signal Process.*, vol. 92, no. 4, pp. 999–1009, Apr. 2012.

[19] J. Pan and Y. Qiu, "An orthogonal method for measurement matrix optimization," *Circuits, Syst., Signal Process.*, vol. 35, no. 3, pp. 837–849, Mar. 2016.

[20] B. Li, L. Zhang, T. Kirubarajan, and S. Rajan, "Projection matrix design using prior information in compressive sensing," *Signal Process.*, vol. 135, pp. 36–47, Jun. 2017.

[21] S. Jin, W. Sun, and L. Huang, "Joint optimization methods for Gaussian random measurement matrix based on column coherence in compressed sensing," *Signal Process.*, vol. 207, Jun. 2023, Art. no. 108941.

[22] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 83–91, Mar. 2008.

[23] X. Yuan, D. J. Brady, and A. K. Katsaggelos, "Snapshot compressive imaging: Theory, algorithms, and applications," *IEEE Signal Process. Mag.*, vol. 38, no. 2, pp. 65–88, Mar. 2021.

[24] J. Ke, P. Wei, and E. Y. Lam, "Binary sensing matrix design for compressive imaging measurements," *Imaging Applied Optics*, vol. 46, no. 10, pp. 2469–2472, Jul. 2014.

[25] J. Ke and E. Y. Lam, "Fast compressive measurements acquisition using optimized binary sensing matrices for low-light-level imaging," *Opt. Exp.*, vol. 24, no. 9, p. 9869, Apr. 2016.

[26] P. Sasmal, R. R. Naidu, C. S. Sastry, and P. Jampana, "Composition of binary compressed sensing matrices," *IEEE Signal Process. Lett.*, vol. 23, no. 8, pp. 1096–1100, Aug. 2016.

[27] R. R. Naidu and C. R. Murthy, "Construction of binary sensing matrices using extremal set theory," *IEEE Signal Process. Lett.*, vol. 24, no. 2, pp. 211–215, Feb. 2017.

[28] W. Lu, T. Dai, and S.-T. Xia, "Binary matrices for compressed sensing," *IEEE Trans. Signal Process.*, vol. 66, no. 1, pp. 77–85, Jan. 2018.

[29] Q. Wang and G. Qu, "Restricted isometry constant improvement based on a singular value decomposition-weighted measurement matrix for compressed sensing," *IET Commun.*, vol. 11, no. 11, pp. 1706–1718, Aug. 2017.

[30] L. Shi, G. Qu, and Q. Wang, "A method of reweighting the sensing matrix for compressed sensing," *IEEE Access*, vol. 9, pp. 21425–21432, 2021.

[31] L. Welch, "Lower bounds on the maximum cross correlation of signals (Corresp.)," *IEEE Trans. Inf. Theory*, vols. IT-20, no. 3, pp. 397–399, May 1974.

[32] M. Fickus and B. R. Mayo, "Mutually unbiased equiangular tight frames," *IEEE Trans. Inf. Theory*, vol. 67, no. 3, pp. 1656–1667, Mar. 2021.

[33] R. Calderbank, S. Howard, and S. Jafarpour, "Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 358–374, Apr. 2010.

[34] A. Barg, A. Mazumdar, and R. Wang, "Restricted isometry property of random subdictionaries," *IEEE Trans. Inf. Theory*, vol. 61, no. 8, pp. 4440–4450, Aug. 2015.

[35] J. M. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1395–1408, Jul. 2009.

[36] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Asilomar Conf. Signals, Syst. Comput.*, 1993, pp. 40–44.

LEI SHI received the Ph.D. degree in computational mathematics from Beijing Jiaotong University, Beijing, China, in 2022. He is currently a Postdoctoral Researcher with the Research Institute of Knowledge and Reason, Luster LightTech Company Ltd. His research interests include image reconstruction and compressed sensing.

GANGRONG QU received the B.S. degree in mathematics from Jilin University, Jilin, China, in 1983. He is currently a Professor with the School of Mathematics and Statistics, Beijing Jiaotong University. His research interests include image reconstruction and compressed sensing.