A New Nonconvex Sparse Recovery Method for Compressive Sensing

Zhiyong Zhou* and Jun Yu

Department of Mathematics and Mathematical Statistics, Umeå University, Umeå, Sweden

As an extension of the widely used ℓ_r-minimization with $0 < r \leq 1$, a new non-convex weighted $\ell_r - \ell_1$ minimization method is proposed for compressive sensing. The theoretical recovery results based on restricted isometry property and q-ratio constrained minimal singular values are established. An algorithm that integrates the iteratively reweighted least squares algorithm and the difference of convex functions algorithm is given to approximately solve this non-convex problem. Numerical experiments are presented to illustrate our results.

Keywords: compressive sensing, nonconvex sparse recovery, iteratively reweighted least squares, difference of convex functions, *q*-ratio constrained minimal singular values

1. INTRODUCTION

Compressive sensing (CS) has attracted a great deal of interests since its advent [1, 2], see the monographs [3, 4] and the references therein for a comprehensive view. Basically, the goal of CS is to recover an unknown (approximately) sparse signal $x \in \mathbb{R}^N$ from the noisy underdetermined linear measurements

$$y = Ax + e \in \mathbb{R}^m, \tag{1}$$

with $m \ll N$, $A \in \mathbb{R}^{m \times N}$ being the pre-given measurement matrix and $e \in \mathbb{R}^m$ being the noise vector. If the measurement matrix A satisfies some kinds of incoherence conditions (e.g., mutual coherence condition [5, 6], restricted isometry property (RIP) [7, 8], null space property (NSP) [9, 10], or constrained minimal singular values (CMSV) [11, 12]), then stable (w.r.t. sparsity defect) and robust (w.r.t. measurement error) recovery results can be guaranteed by using the constrained ℓ_1-minimization [13]:

$$\min_{z \in \mathbb{R}^N} \|z\|_1 \quad \text{subject to} \quad \|Az - y\|_2 \leq \eta. \tag{2}$$

Here the ℓ_1-minimization problem works as a convex relaxation of ℓ_0-minimization problem, which is NP-hard to solve [14].

Meanwhile, non-convex recovery algorithms such as the ℓ_r-minimization ($0 < r < 1$) have been proposed to enhance sparsity [15–20]. ℓ_r-minimization enables one to reconstruct the sparse signal from fewer number of measurements compared to the convex ℓ_1-minimization, although it is more challenging to solve because of its non-convexity. Fortunately, an iteratively reweighted least squares (IRLS) algorithm can be applied to approximately solve this non-convex problem in practice [21, 22].

As an extension of the ℓ_r-minimization, we study in this paper the following weighted $\ell_r - \ell_1$ minimization problem for sparse signal recovery:

$$\min_{z \in \mathbb{R}^N} \|z\|_r^r - \alpha \|z\|_1^r \quad \text{subject to} \quad \|Az - y\|_2 \leq \eta, \quad (3)$$

where $y = Ax + e$ with $\|e\|_2 \leq \eta$, $0 \leq \alpha \leq 1$, and $0 < r \leq 1$. Throughout the paper, we assume that $\alpha \neq 1$ when $r = 1$. Obviously, it reduces to the traditional ℓ_r-minimization problem when $\alpha = 0$. This hybrid norm model is inspired by the non-convex Lipschitz continuous $\ell_1 - \ell_2$ model (minimizing the difference of ℓ_1 norm and ℓ_2 norm) proposed in Lou et al. [23] and Yin et al. [24], which improves the ℓ_1-minimization in a robust manner, especially for the highly coherent measurement matrices. Roughly speaking, the underlying logic of adopting these kinds of norm differences or the ratios of norms [25] comes from the fact that they can be viewed as sparsity measures, see the effective sparsity measure called q-ratio sparsity (involving the ratio of ℓ_1 norm and ℓ_q norm) defined later in Definition 2 of section 2.2. Other recent related literatures include [26–29], to name a few.

To illustrate these weighted $\ell_r - \ell_1$ norms ($\|\cdot\|_r^r - \alpha\|\cdot\|_1^r$), we present their corresponding contour plots in **Figure 1**[1]. As is shown, different non-convex patterns arise while varying the difference weight α or the norm order r. And the level curves of weighted $\ell_r - \ell_1$ norms approach the x and y axes as the norm values get small, which reflects their ability to promote sparsity. In the present paper, we shall focus on both the theoretical aspects and the computational study for this non-convex sparse recovery method.

This paper is organized as follows. In section 2, we derive the theoretical performance bounds for the weighted $\ell_r - \ell_1$ minimization based on both r-RIP and q-ratio CMSV. In section 3, we give an algorithm to approximately solve the unconstrained version of the weighted $\ell_r - \ell_1$ minimization problem. Numerical experiments are provided in section 4. Section 5 concludes with a brief summary and an outlook on future extensions.

2. RECOVERY ANALYSIS

In this section, we establish the theoretical performance bounds for the reconstruction error of the weighted $\ell_r - \ell_1$ minimization problem, based on both r-RIP and q-ratio CMSV. Hereafter, we say a signal $x \in \mathbb{R}^N$ is s-sparse if $\|x\|_0 = \sum_{i=1}^N 1\{x_i \neq 0\} \leq s$, and denote by x_S the vector that coincides with x on the indices in $S \subseteq [N] := \{1, 2, \cdots, N\}$ and takes zero outside S.

2.1. r-RIP
We start with the definition of the s-th r-restricted isometry constant, which was introduced in Chartrand and Staneva [30].

Definition 1. ([30]) *For integer $s > 0$ and $0 < r \leq 1$, the s-th r-restricted isometry constant (RIC) $\delta_s = \delta_s(A)$ of a matrix*

$A \in \mathbb{R}^{m \times N}$ *is defined as the smallest $\delta \geq 0$ such that*

$$(1 - \delta)\|x\|_2^r \leq \|Ax\|_r^r \leq (1 + \delta)\|x\|_2^r \quad (4)$$

for all s-sparse vectors $x \in \mathbb{R}^N$.

Then, the r-RIP means that the s-th r-RIC δ_s is small for reasonably large s. In Chartrand and Staneva [30], the authors established the recovery analysis result for ℓ_r-minimization problem based on this r-RIP. To extend this to the weighted $\ell_r - \ell_1$ minimization problem, the following lemma plays a crucial role.

Lemma 1. *Suppose $x \in \mathbb{R}^N$, $0 \leq \alpha \leq 1$ and $0 < r \leq 1$, then we have*

$$(N - \alpha N^r)\left(\min_{i \in [N]} |x_i|\right)^r \leq \|x\|_r^r - \alpha\|x\|_1^r \leq (N^{1-r} - \alpha)\|x\|_1^r. \quad (5)$$

In particular, when $S = \text{supp}(x) \subseteq [N]$ and $|S| = s$, then

$$(s - \alpha s^r)\left(\min_{i \in S} |x_i|\right)^r \leq \|x\|_r^r - \alpha\|x\|_1^r \leq (s^{1-r} - \alpha)\|x\|_1^r. \quad (6)$$

Proof. The right hand side of (5) follows immediately from the norm inequality $\|x\|_r \leq N^{1/r-1}\|x\|_1$ for any $x \in \mathbb{R}^N$ and $0 < r \leq 1$. As for the left hand side, it holds trivially if $\min_{i \in [N]} |x_i| = 0$. When $\min_{i \in [N]} |x_i| \neq 0$, by dividing $\left(\min_{i \in [N]} |x_i|\right)^r$ on both sides, it is equivalent to show that

$$\sum_{j=1}^N \left(\frac{|x_j|}{\min_{i \in [N]} |x_i|}\right)^r - \alpha \left(\sum_{j=1}^N \frac{|x_j|}{\min_{i \in [N]} |x_i|}\right)^r \geq N - \alpha N^r. \quad (7)$$

By denoting $a_j = \frac{|x_j|}{\min_{i \in [N]} |x_i|}$, we have $a_j \geq 1$ for any $1 \leq j \leq N$, and to show (7) it suffices to show

$$\sum_{j=1}^N a_j^r - \alpha \left(\sum_{j=1}^N a_j\right)^r \geq N - \alpha N^r.$$

Assume the function $f(a_1, a_2, \cdots, a_N) = \sum_{j=1}^N a_j^r - \alpha \left(\sum_{j=1}^N a_j\right)^r$. Then, as a result of

$$\frac{\partial f}{\partial a_k} = r a_k^{r-1} - \alpha r \left(\sum_{j=1}^N a_j\right)^{r-1} > 0 \quad \text{for any } 1 \leq k \leq N,$$

we have $f(a_1, a_2, \cdots, a_N) \geq f(1, 1, \cdots, 1) = N - \alpha N^r$. Thus, the left hand side of (5) holds and the proof is completed. (6) follows as we apply (5) to x_S.

Now, we are ready to present the r-RIP based bound for the ℓ_2 norm of the reconstruction error.

---

[1] All figures can be reproduced from the code available at https://github.com/zzy583661/Weighted-l_r-l_1-minimization
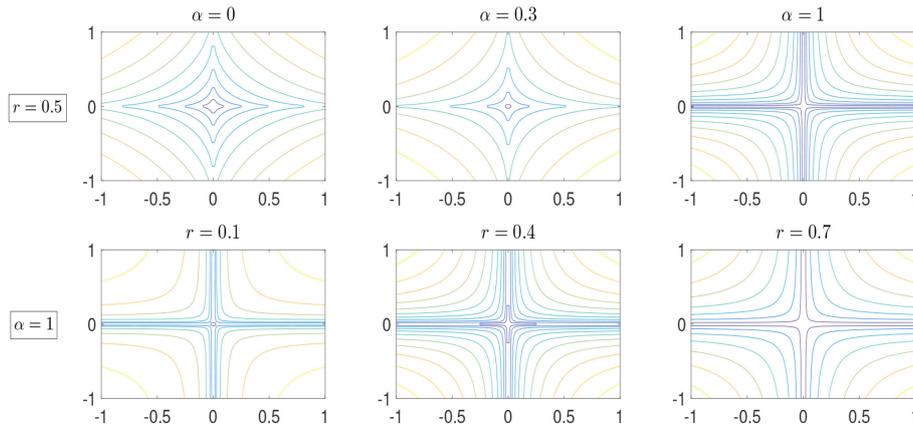
FIGURE 1 | The contour plots for weighted $\ell_r - \ell_1$ norms with different α and r. The first row corresponds to the cases that $r = 0.5$ with different $\alpha = 0, 0.3, 1$, while the second row shows the cases that $r = 0.1, 0.4, 0.7$ with fixed $\alpha = 1$.

Theorem 1. *Let the ℓ_r-error of best s-term approximation of x be $\sigma_s(x)_r = \inf\{\|x - z\|_r, z \in \mathbb{R}^N is\ s\text{-sparse}\}$. We assume that $a > 0$ is properly chosen so that as is an integer. If*

$$b = \frac{(as)^{1-\frac{r}{2}} - \alpha(as)^{\frac{r}{2}}}{s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}}} > 1 \qquad (8)$$

and suppose the measurement matrix A satisfies the condition

$$\delta_{as} + b\delta_{(a+1)s} < b - 1, \qquad (9)$$

then any solution $\hat{x}$ to the minimization problem (3) obeys

$$\|\hat{x} - x\|_2 \leq C_1 m^{1/r-1/2}\eta + C_2(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})^{-1/r}\sigma_s(x)_r \quad (10)$$

with $C_1 = \frac{2^{1/r}(1+b^{1/r})}{(b-b\delta_{(a+1)s}-1-\delta_{as})^{1/r}}$ *and* $C_2 = \frac{2^{2/r-1}[(1+\delta_{as})^{1/r}+(1-\delta_{(a+1)s})^{1/r}]}{(b-b\delta_{(a+1)s}-1-\delta_{as})^{1/r}}$.

Proof. We assume that S is the index set that contains the largest s absolute entries of x so that $\sigma_s(x)_r = \|x_{S^c}\|_r$ and let $h = \hat{x} - x$. Then we have

$$\begin{aligned}
&\|x_S\|_r^r + \|x_{S^c}\|_r^r - \alpha\|x\|_1^r \\
&= \|x\|_r^r - \alpha\|x\|_1^r \\
&\geq \|\hat{x}\|_r^r - \alpha\|\hat{x}\|_1^r \\
&= \|x_S + x_{S^c} + h_S + h_{S^c}\|_r^r - \alpha\|x + h_S + h_{S^c}\|_1^r \\
&\geq \|x_S + h_S\|_r^r + \|x_{S^c} + h_{S^c}\|_r^r - \alpha(\|x + h_S\|_1 + \|h_{S^c}\|_1)^r \\
&\geq \|x_S\|_r^r - \|h_S\|_r^r + \|h_{S^c}\|_r^r - \|x_{S^c}\|_r^r - \alpha\|x + h_S\|_1^r - \alpha\|h_{S^c}\|_1^r \\
&\geq \|x_S\|_r^r - \|h_S\|_r^r + \|h_{S^c}\|_r^r - \|x_{S^c}\|_r^r \\
&\quad - \alpha\|x\|_1^r - \alpha\|h_S\|_1^r - \alpha\|h_{S^c}\|_1^r,
\end{aligned}$$

which implies

$$\|h_{S^c}\|_r^r - \alpha\|h_{S^c}\|_1^r \leq \|h_S\|_r^r + \alpha\|h_S\|_1^r + 2\|x_{S^c}\|_r^r. \quad (11)$$

Using the Holder's inequality, we obtain

$$\|Ah\|_r^r \leq \left(\sum_{i=1}^m (|(Ah)_i|^r)^{2/r}\right)^{r/2} \cdot \left(\sum_{i=1}^m 1\right)^{1-r/2} = m^{1-r/2}\|Ah\|_2^r.$$

By $\|Ax - y\|_2 = \|e\|_2 \leq \eta$ and the triangular inequality,

$$\begin{aligned}
\|Ah\|_2 &= \|(A\hat{x} - y) - (Ax - y)\|_2 \\
&\leq \|A\hat{x} - y\|_2 + \|Ax - y\|_2 \leq 2\eta. \quad (12)
\end{aligned}$$

Thus,

$$\|Ah\|_r^r \leq m^{1-r/2}\|Ah\|_2^r \leq m^{1-r/2}(2\eta)^r. \quad (13)$$

Arrange $S^c = S_1 \cup S_2 \cup \cdots$, where S_1 is the index set of $M = as$ largest absolute entries of h in S^c, S_2 is the index set of M largest absolute entries of h in $(S \cup S_1)^c$, etc. And we denote $S_0 = S \cup S_1$. Then, by adopting Lemma 1, for each $i \in S_k$, $k \geq 2$,

$$|h_i| \leq \min_{j \in S_{k-1}} |h_j| \Rightarrow |h_i|^r \leq \left(\min_{j \in S_{k-1}} |h_j|\right)^r \leq \frac{\|h_{S_{k-1}}\|_r^r - \alpha\|h_{S_{k-1}}\|_1^r}{M - \alpha M^r}. \quad (14)$$

Thus we have $\|h_{S_k}\|_2^r = \left(\sum_{i \in S_k} |h_i|^2\right)^{r/2} \leq M^{r/2}\frac{\|h_{S_{k-1}}\|_r^r - \alpha\|h_{S_{k-1}}\|_1^r}{M - \alpha M^r} = \frac{\|h_{S_{k-1}}\|_r^r - \alpha\|h_{S_{k-1}}\|_1^r}{M^{1-\frac{r}{2}} - \alpha M^{\frac{r}{2}}}$. Hence it follows that

$$\sum_{k\geq 2}\|h_{S_k}\|_2^r \leq \frac{\sum_{k\geq 1}(\|h_{S_k}\|_r^r - \alpha\|h_{S_k}\|_1^r)}{M^{1-\frac{r}{2}} - \alpha M^{\frac{r}{2}}} = \frac{\sum_{k\geq 1}\|h_{S_k}\|_r^r - \alpha\sum_{k\geq 1}\|h_{S_k}\|_1^r}{M^{1-\frac{r}{2}} - \alpha M^{\frac{r}{2}}}. \quad (15)$$

Note that

$$\sum_{k\geq 1}\|h_{S_k}\|_r^r = \|h_{S^c}\|_r^r \text{ and } \sum_{k\geq 1}\|h_{S_k}\|_1^r \geq \left\|\sum_{k\geq 1}h_{S_k}\right\|_1^r = \|h_{S^c}\|_1^r,$$

therefore, with (11), it holds that

$$
\begin{aligned}
\sum_{k \geq 2} \|h_{S_k}\|_2^r &\leq \frac{\|h_{S^c}\|_r^r - \alpha\|h_{S^c}\|_1^r}{M^{1-\frac{r}{2}} - \alpha M^{\frac{r}{2}}} \\
&\leq \frac{\|h_S\|_r^r + \alpha\|h_S\|_1^r + 2\|x_{S^c}\|_r^r}{M^{1-\frac{r}{2}} - \alpha M^{\frac{r}{2}}} \\
&\leq \frac{s^{1-\frac{r}{2}}\|h_S\|_2^r + \alpha s^{\frac{r}{2}}\|h_S\|_2^r + 2\|x_{S^c}\|_r^r}{M^{1-\frac{r}{2}} - \alpha M^{\frac{r}{2}}} \\
&\leq \frac{(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})\|h_{S_0}\|_2^r + 2\|x_{S^c}\|_r^r}{M^{1-\frac{r}{2}} - \alpha M^{\frac{r}{2}}}.
\end{aligned} \tag{16}
$$

Meanwhile, according to the definition of r-RIC, we have

$$
\begin{aligned}
\|Ah\|_r^r &= \|Ah_{S_0} + \sum_{k \geq 2} Ah_{S_k}\|_r^r \\
&\geq \|Ah_{S_0}\|_r^r - \|\sum_{k \geq 2} Ah_{S_k}\|_r^r \\
&\geq \|Ah_{S_0}\|_r^r - \sum_{k \geq 2}\|Ah_{S_k}\|_r^r \\
&\geq (1 - \delta_{M+s})\|h_{S_0}\|_2^r - (1 + \delta_M)\sum_{k \geq 2}\|h_{S_k}\|_2^r.
\end{aligned}
$$

Thus by using (16), it follows that

$$
\begin{aligned}
\|Ah\|_r^r &\geq (1 - \delta_{M+s})\|h_{S_0}\|_2^r - (1 + \delta_M) \cdot \frac{(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})\|h_{S_0}\|_2^r + 2\|x_{S^c}\|_r^r}{M^{1-\frac{r}{2}} - \alpha M^{\frac{r}{2}}} \\
&= \left(1 - \delta_{M+s} - \frac{1 + \delta_M}{b}\right)\|h_{S_0}\|_2^r - \frac{2(1 + \delta_M)}{M^{1-\frac{r}{2}} - \alpha M^{\frac{r}{2}}}\|x_{S^c}\|_r^r,
\end{aligned}
$$

where $b = \frac{M^{1-\frac{r}{2}} - \alpha M^{\frac{r}{2}}}{s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}}} = \frac{(as)^{1-\frac{r}{2}} - \alpha(as)^{\frac{r}{2}}}{s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}}}$. Therefore, if $\delta_M + b\delta_{M+s} < b - 1$, then it yields that

$$
\begin{aligned}
\|h_{S_0}\|_2^r &\leq \frac{b}{b - b\delta_{M+s} - 1 - \delta_M}\|Ah\|_r^r + \frac{2(1 + \delta_M)(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})^{-1}}{b - b\delta_{M+s} - 1 - \delta_M}\|x_{S^c}\|_r^r \\
&\leq \frac{bm^{1-r/2}(2\eta)^r}{b - b\delta_{M+s} - 1 - \delta_M} + \frac{2(1 + \delta_M)(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})^{-1}}{b - b\delta_{M+s} - 1 - \delta_M}\|x_{S^c}\|_r^r.
\end{aligned} \tag{17}
$$

On the other hand,

$$
\begin{aligned}
\left(\sum_{k \geq 2}\|h_{S_k}\|_2\right)^r &\leq \sum_{k \geq 2}\|h_{S_k}\|_2^r \\
&\leq \frac{(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})\|h_{S_0}\|_2^r + 2\|x_{S^c}\|_r^r}{M^{1-\frac{r}{2}} - \alpha M^{\frac{r}{2}}} \\
&= \frac{1}{b}\|h_{S_0}\|_2^r + \frac{2(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})^{-1}}{b}\|x_{S^c}\|_r^r \\
&\leq \frac{1}{b}\left(\frac{b}{b - b\delta_{M+s} - 1 - \delta_M}\|Ah\|_r^r \right. \\
&\quad \left. + \frac{2(1 + \delta_M)(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})^{-1}}{b - b\delta_{M+s} - 1 - \delta_M}\|x_{S^c}\|_r^r\right)
\end{aligned}
$$

$$
\begin{aligned}
&\quad + \frac{2(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})^{-1}}{b}\|x_{S^c}\|_r^r \\
&\leq \frac{1}{b - b\delta_{M+s} - 1 - \delta_M}\|Ah\|_r^r \\
&\quad + \frac{2(1 - \delta_{M+s})(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})^{-1}}{b - b\delta_{M+s} - 1 - \delta_M}\|x_{S^c}\|_r^r \\
&\leq \frac{m^{1-r/2}(2\eta)^r}{b - b\delta_{M+s} - 1 - \delta_M} \\
&\quad + \frac{2(1 - \delta_{M+s})(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})^{-1}}{b - b\delta_{M+s} - 1 - \delta_M}\|x_{S^c}\|_r^r.
\end{aligned} \tag{18}
$$

Since $(v_1^r + v_2^r)^{1/r} \leq 2^{1/r-1}(v_1 + v_2)$ for any $v_1, v_2 \geq 0$, combining (17) and (18) gives

$$
\begin{aligned}
\|h\|_2 &\leq \|h_{S_0}\|_2 + \sum_{k \geq 2}\|h_{S_k}\|_2 \\
&\leq 2^{1/r-1}\left(\frac{2b^{1/r}m^{1/r-1/2}\eta}{(b - b\delta_{M+s} - 1 - \delta_M)^{1/r}} \right. \\
&\quad \left. + \frac{2^{1/r}(1 + \delta_M)^{1/r}(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})^{-1/r}}{(b - b\delta_{M+s} - 1 - \delta_M)^{1/r}}\|x_{S^c}\|_r\right) \\
&\quad + 2^{1/r-1}\left(\frac{2m^{1/r-1/2}\eta}{(b - b\delta_{M+s} - 1 - \delta_M)^{1/r}} \right. \\
&\quad \left. + \frac{2^{1/r}(1 - \delta_{M+s})^{1/r}(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})^{-1/r}}{(b - b\delta_{M+s} - 1 - \delta_M)^{1/r}}\|x_{S^c}\|_r\right) \\
&= \frac{2^{1/r}m^{1/r-1/2}(1 + b^{1/r})}{(b - b\delta_{(a+1)s} - 1 - \delta_{as})^{1/r}}\eta \\
&\quad + \frac{2^{2/r-1}[(1 + \delta_{as})^{1/r} + (1 - \delta_{(a+1)s})^{1/r}]}{(b - b\delta_{(a+1)s} - 1 - \delta_{as})^{1/r}} \\
&\quad \times (s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})^{-1/r}\|x_{S^c}\|_r \\
&= C_1 m^{1/r-1/2}\eta + C_2(s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}})^{-1/r}\sigma_s(x)_r. \tag{19}
\end{aligned}
$$

The proof is completed.

Based on this theorem, we can obtain the following corollary by assuming that the original signal x is s-sparse ($\sigma_s(x)_r = 0$) and the measurement vector is noise free ($e = 0$ and $\eta = 0$), which acts as a natural generalization of Theorem 2.4 in Chartrand and Staneva [30] from the case $\alpha = 0$ to any $\alpha \in [0, 1]$.

Corollary 1. *For any s-sparse signal x, if the conditions in Theorem 1 hold, then the unique solution of (3) with $\eta = 0$ is exactly x.*

Remarks. Observe that r-RIP based condition for exact sparse recovery given in Chartrand and Staneva [30] reads

$$
\delta_{as} < a^{1-\frac{r}{2}}(1 - \delta_{(a+1)s}) - 1,
$$

while ours goes to

$$
\delta_{as} < b(1 - \delta_{(a+1)s}) - 1
$$

with $b = \frac{(as)^{1-\frac{r}{2}} - \alpha(as)^{\frac{r}{2}}}{s^{1-\frac{r}{2}} + \alpha s^{\frac{r}{2}}} < a^{1-\frac{r}{2}}$ when $\alpha \in (0, 1]$. Thus, the sufficient condition established here is slightly stronger than that for the traditional ℓ_r-minimization in Chartrand and Staneva [30] if $\alpha \in (0, 1]$.

2.2. q-Ratio CMSV

Before the discussion of q-ratio CMSV, we start with presenting the definition of q-ratio sparsity as a kind of effective sparsity measure. We list the detailed statement here for the sake of completeness.

Definition 2. ([12, 31, 32]) *For any non-zero $z \in \mathbb{R}^N$ and non-negative $q \notin \{0, 1, \infty\}$, the q-ratio sparsity level of z is defined as*

$$s_q(z) = \left(\frac{\|z\|_1}{\|z\|_q} \right)^{\frac{q}{q-1}}. \tag{20}$$

The cases of $q \in \{0, 1, \infty\}$ are evaluated as limits: $s_0(z) = \lim_{q \to 0} s_q(z) = \|z\|_0$, $s_\infty(z) = \lim_{q \to \infty} s_q(z) = \frac{\|z\|_1}{\|z\|_\infty}$, and $s_1(z) = \lim_{q \to 1} s_q(z) = \exp(H_1(\pi(z)))$, where $\pi(z) \in \mathbb{R}^N$ with entries $\pi_i(z) = |z_i|/\|z\|_1$ and H_1 is the ordinary Shannon entropy $H_1(\pi(z)) = -\sum_{i=1}^N \pi_i(z) \log \pi_i(z)$.

We are able to establish the performance bounds for both the ℓ_q norm and ℓ_r norm of the reconstruction error via a recently developed computable incoherence measure of the measurement matrix, called q-ratio CMSV. Its definition is given as follows.

Definition 3. ([12, 32]) *For any real number $s \in [1, N]$, $q \in (1, \infty]$, and matrix $A \in \mathbb{R}^{m \times N}$, the q-ratio constrained minimal singular value (CMSV) of A is defined as*

$$\rho_{q,s}(A) = \min_{z \neq 0, s_q(z) \leq s} \frac{\|Az\|_2}{\|z\|_q}. \tag{21}$$

Then, when the signal is exactly sparse, we have the following q-ratio CMSV based sufficient condition for valid upper bounds of the reconstruction error, which are much more concise to obtain than the r-RIP based ones.

Theorem 2. *For any $1 < q \leq \infty$, $0 \leq \alpha \leq 1$, and $0 < r \leq 1$, if the signal x is s-sparse and the measurement matrix A satisfies the condition*

$$\rho_{q, \left(\frac{2}{1-\alpha} \right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A) > 0, \tag{22}$$

then any solution $\hat{x}$ to the minimization problem (3) obeys

$$\|\hat{x} - x\|_q \leq \frac{2\eta}{\rho_{q, \left(\frac{2}{1-\alpha} \right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A)}, \tag{23}$$

$$\|\hat{x} - x\|_r \leq \left(\frac{2^{r+1}}{1-\alpha} \right)^{1/r} \cdot \frac{s^{1/r-1/q}\eta}{\rho_{q, \left(\frac{2}{1-\alpha} \right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A)}. \tag{24}$$

Proof. Suppose the support of x to be S with $|S| \leq s$ and $h = \hat{x} - x$, then, based on (11), we have

$$\|h_{S^c}\|_r^r - \alpha\|h_{S^c}\|_1^r \leq \|h_S\|_r^r + \alpha\|h_S\|_1^r. \tag{25}$$

Hence, for any $1 < q \leq \infty$, it holds that

$$\begin{aligned} \|h\|_r^r - \alpha\|h\|_1^r &\leq \|h_S\|_r^r + \|h_{S^c}\|_r^r - \alpha\|h_S\|_1^r - \alpha\|h_{S^c}\|_1^r \\ &\leq \|h_S\|_r^r - \alpha\|h_S\|_1^r + \|h_S\|_r^r + \alpha\|h_S\|_1^r \\ &\leq 2\|h_S\|_r^r \leq 2s^{1-r/q}\|h_S\|_q^r \leq 2s^{1-r/q}\|h\|_q^r. \end{aligned}$$

Then since $(1-\alpha)\|h\|_r^r \leq \|h\|_r^r - \alpha\|h\|_1^r$, it implies that $(1-\alpha)\|h\|_1^r \leq 2s^{1-r/q}\|h\|_q^r$. As a consequence,

$$\begin{aligned} s_q(h) = \left(\frac{\|h\|_1}{\|h\|_q} \right)^{\frac{q}{q-1}} &\leq \left(\frac{2s^{1-r/q}}{1-\alpha} \right)^{\frac{q}{r(q-1)}} \\ &= \left(\frac{2}{1-\alpha} \right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}. \end{aligned} \tag{26}$$

Therefore, according to the definition of q-ratio CMSV the condition (22), and the fact that $\|Ah\|_2 \leq 2\eta$ [see (12)], we can obtain that

$$\begin{aligned} \rho_{q, \left(\frac{2}{1-\alpha} \right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A) &\leq \frac{\|Ah\|_2}{\|h\|_q} \Rightarrow \|h\|_q \leq \frac{\|Ah\|_2}{\rho_{q, \left(\frac{2}{1-\alpha} \right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A)} \\ &\leq \frac{2\eta}{\rho_{q, \left(\frac{2}{1-\alpha} \right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A)}, \end{aligned} \tag{27}$$

which completes the proof of (23). In addition, $(1-\alpha)\|h\|_r^r \leq \|h\|_r^r - \alpha\|h\|_1^r \leq 2s^{1-r/q}\|h\|_q^r$ yields

$$\begin{aligned} \|h\|_r &\leq \left(\frac{2}{1-\alpha} \right)^{1/r} s^{1/r-1/q}\|h\|_q \\ &\leq \left(\frac{2^{r+1}}{1-\alpha} \right)^{1/r} \cdot \frac{s^{1/r-1/q}\eta}{\rho_{q, \left(\frac{2}{1-\alpha} \right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A)}. \end{aligned} \tag{28}$$

Therefore, (24) holds and the proof is completed.

Remarks. Note that the results (11) and (12) in Theorem 1 of Zhou and Yu [12] correspond to the special case of $\alpha = 0$ and $r = 1$ in this result. As a by-product of this theorem, we have that the perfect recovery can be guaranteed for any s-sparse signal x via (3) with $\eta = 0$, if there exists some $q \in (1, \infty]$ such that the q-ratio CMSV of the measurement matrix A fulfils $\rho_{q, \left(\frac{2}{1-\alpha} \right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A) > 0$. As studied in Zhou and Yu [12, 32], this kind of q-ratio CMSV based sufficient conditions holds with high probability for subgaussian and a class of structured random matrices as long as the number of measurements is reasonably large.

Next, we extend the result to the case that x is compressible (i.e., not exactly sparse but can be well-approximated by an exactly sparse signal).

Theorem 3. *For any $1 < q \le \infty$, $0 \le \alpha \le 1$ and $0 < r \le 1$, if the measurement matrix A satisfies the condition*

$$\rho_{q,\left(\frac{4}{1-\alpha}\right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A) > 0, \tag{29}$$

then any solution $\hat{x}$ to the minimization problem (3) fulfils

$$\|\hat{x} - x\|_q \le \frac{2\eta}{\rho_{q,\left(\frac{4}{1-\alpha}\right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A)} + s^{1/q-1/r}\sigma_s(x)_r, \tag{30}$$

$$\|\hat{x} - x\|_r \le \left(\frac{4}{1-\alpha}\right)^{1/r} \frac{s^{1/r-1/q}\eta}{\rho_{q,\left(\frac{4}{1-\alpha}\right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A)}$$

$$+ \left(\frac{4}{1-\alpha}\right)^{1/r} \sigma_s(x)_r. \tag{31}$$

Proof. We assume that S is the index set that contains the largest s absolute entries of x so that $\sigma_s(x)_r = \|x_{S^c}\|_r$ and let $h = \hat{x} - x$. Then we still have (11), that is,

$$\|h_{S^c}\|_r^r - \alpha\|h_{S^c}\|_1^r \le \|h_S\|_r^r + \alpha\|h_S\|_1^r + 2\|x_{S^c}\|_r^r. \tag{32}$$

As a result,

$$\begin{aligned}
(1-\alpha)\|h\|_r^r &\le \|h\|_r^r - \alpha\|h\|_1^r \\
&\le \|h_S\|_r^r + \|h_{S^c}\|_r^r - \alpha\|h_S\|_1^r - \alpha\|h_{S^c}\|_1^r \\
&\le 2\|h_S\|_r^r + 2\|x_{S^c}\|_r^r \\
&\le 2s^{1-r/q}\|h\|_q^r + 2\|x_{S^c}\|_r^r
\end{aligned} \tag{33}$$

holds for any $1 < q \le \infty$, $0 \le \alpha \le 1$ and $0 < r \le 1$.

To prove (30), we assume $h \ne 0$ and $\|h\|_q > \frac{2\eta}{\rho_{q,\left(\frac{4}{1-\alpha}\right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A)}$, otherwise it holds trivially. Then

$$\|h\|_q > \frac{\|Ah\|_2}{\rho_{q,\left(\frac{4}{1-\alpha}\right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A)} \Rightarrow s_q(h) > \left(\frac{4}{1-\alpha}\right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}$$

$$\Rightarrow \|h\|_1^r > \frac{4}{1-\alpha} s^{1-r/q}\|h\|_q^r,$$

which implies that $(1-\alpha)\|h\|_r^r \ge (1-\alpha)\|h\|_1^r > 4s^{1-r/q}\|h\|_q^r$. Then combining with (33), it yields that

$$\|h\|_q \le (s^{r/q-1})^{1/r}\|x_{S^c}\|_r = s^{1/q-1/r}\|x_{S^c}\|_r = s^{1/q-1/r}\sigma_s(x)_r. \tag{34}$$

Therefore, we have

$$\|h\|_q \le \frac{2\eta}{\rho_{q,\left(\frac{4}{1-\alpha}\right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A)} + s^{1/q-1/r}\sigma_s(x)_r, \tag{35}$$

which completes the proof of (30).

Moreover, by using (33) and the inequality $(v_1^r + v_2^r)^{1/r} \le 2^{1/r-1}(v_1 + v_2)$ for any $v_1, v_2 \ge 0$, we obtain that

$$\|h\|_r \le \left(\frac{1}{1-\alpha}\right)^{1/r} (2s^{1-r/q}\|h\|_q^r + 2\|x_{S^c}\|_r^r)^{1/r}$$

$$\le \left(\frac{1}{1-\alpha}\right)^{1/r} 2^{2/r-1}(s^{1/r-1/q}\|h\|_q + \|x_{S^c}\|_r)$$

$$\le \left(\frac{1}{1-\alpha}\right)^{1/r} 2^{2/r-1}\left(s^{1/r-1/q}\frac{2\eta}{\rho_{q,\left(\frac{4}{1-\alpha}\right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A)} + 2\sigma_s(x)_r\right)$$

$$\le \left(\frac{4}{1-\alpha}\right)^{1/r} \frac{s^{1/r-1/q}\eta}{\rho_{q,\left(\frac{4}{1-\alpha}\right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}}}(A)} + \left(\frac{4}{1-\alpha}\right)^{1/r}\sigma_s(x)_r. \tag{36}$$

Hence, (31) holds and the proof is completed.

Remarks. When we select $\alpha = 0$ and $r = 1$, our results reduce to the corresponding results for the ℓ_1-minimization or Basis Pursuit in Theorem 2 of Zhou and Yu [12]. In general, the sufficient condition provided here and that in Theorem 2 are slightly stronger than those established for the ℓ_1-minimization in Zhou and Yu [12], noticing that $\left(\frac{2}{1-\alpha}\right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}} \ge 2^{\frac{q}{q-1}} s$ and $\left(\frac{4}{1-\alpha}\right)^{\frac{q}{r(q-1)}} s^{\frac{q-r}{r(q-1)}} \ge 4^{\frac{q}{q-1}} s$ for any $1 < q \le \infty$, $0 \le \alpha \le 1$, and $0 < r \le 1$. This is caused by the fact that the technical inequalities used like (25) and (32) are far from tight. And this is also the case in the r-RIP based analysis. In fact, both r-RIP and q-ratio CMSV based conditions are loose. The discussion on much tighter sufficient conditions such as the NSP based conditions investigated in Tran and Webster [33], is left for future work.

3. ALGORITHM

In this section, we discuss the computational approach for the unconstrained version of (3), i.e.,

$$\min_{x\in\mathbb{R}^N} \frac{1}{2}\|Ax - y\|_2^2 + \lambda(\|x\|_r^r - \alpha\|x\|_1^r), \tag{37}$$

with $\lambda > 0$ being the regularizer parameter.

We integrate the iteratively reweighted least squares (IRLS) algorithm [21, 22] and the difference of convex functions algorithm (DCA) [34, 35] to solve this problem. In the outer loop, we use the IRLS to approximate the term $\|x\|_r^r$, and use an iteratively reweighted ℓ_1 norm to approximate $\|x\|_1^r$. Specifically, we begin with $x^0 = \arg\min_{x\in\mathbb{R}^N} \|y - Ax\|_2^2$ and $\varepsilon_0 = 1$, for $n = 0, 1, \cdots$,

$$x^{n+1} = \arg\min_{x\in\mathbb{R}^N} \frac{1}{2}\|Ax - y\|_2^2 + \lambda\|W^n x\|_2^2 - \alpha\lambda v^n\|x\|_1, \tag{38}$$

where $W^n = \text{diag}\{((x_i^n)^2 + \varepsilon_n)^{r/4-1/2}\}$ and $v^n = \|x^n\|_1^{r-1}$. We let $\varepsilon_{n+1} = \varepsilon_n/10$ if the error $\|x^{n+1} - x^n\|_2 < \sqrt{\varepsilon_n}/100$. The algorithm is stopped when $\varepsilon_{n+1} < 10^{-8}$ for some n.
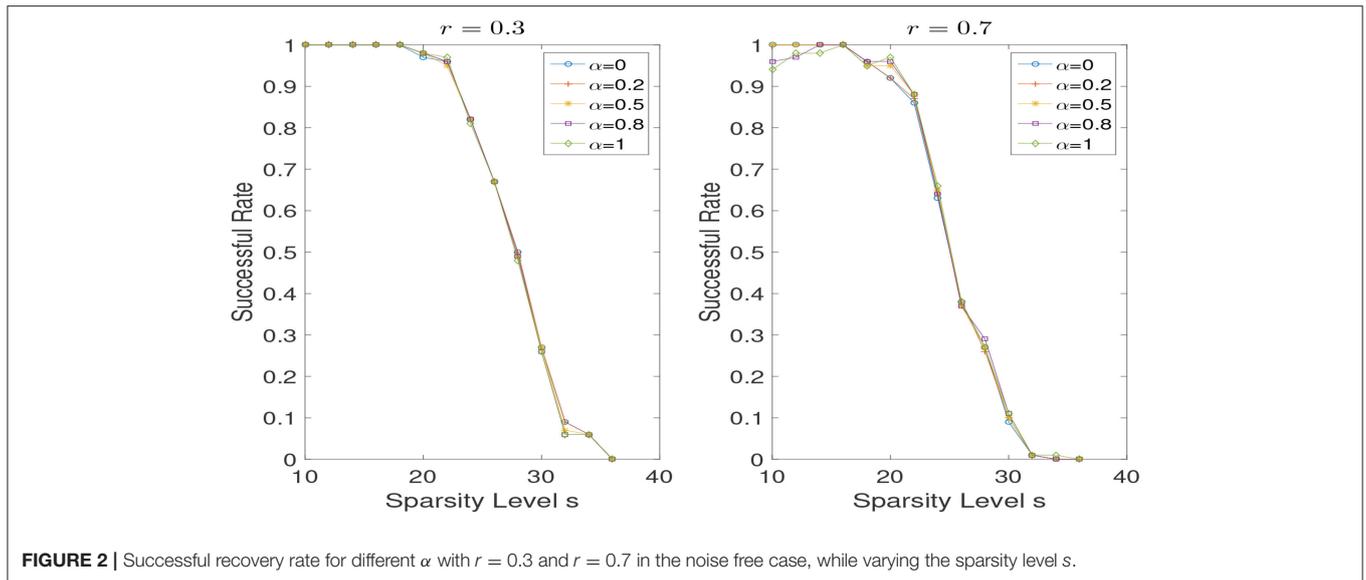
FIGURE 2 | Successful recovery rate for different α with $r = 0.3$ and $r = 0.7$ in the noise free case, while varying the sparsity level s.

As for the inner loop used to solve (38), we view it as a minimization problem of a difference of two convex functions, that is, the objective function $F(x) = \left(\frac{1}{2} \|Ax - y\|_2^2 + \lambda \|W^n x\|_2^2 \right) - \alpha \lambda v^n \|x\|_1 = : G(x) - H(x)$. We start with $x^{n+1,0} = 0$. For $k = 0, 1, 2, \cdots$, in the $k + 1$ step, by linearizing $H(x)$ with the approximation $H(x^{n+1,k}) + \langle y^{n+1,k}, x - x^{n+1,k} \rangle$ where $y^{n+1,k} \in \partial H(x^{n+1,k})$, i.e. $y^{n+1,k}$ is a subgradient of $H(x)$ at $x^{n+1,k}$. Then we have

$$
\begin{aligned}
x^{n+1,k+1} &= \arg\min_{x \in \mathbb{R}^N} \frac{1}{2} \|Ax - y\|_2^2 + \lambda \|W^n x\|_2^2 \\
&\quad - \left(\alpha \lambda v^n \|x^{n+1,k}\|_1 + \langle \alpha \lambda v^n \text{sign}(x^{n+1,k}), x - x^{n+1,k} \rangle \right) \\
&= \arg\min_{x \in \mathbb{R}^N} \frac{1}{2} \|Ax - y\|_2^2 + \lambda \|W^n x\|_2^2 - \langle \alpha \lambda v^n \text{sign}(x^{n+1,k}), x \rangle \\
&= (A^T A + 2\lambda (W^n)^T W^n)^{-1} [A^T y + \alpha \lambda v^n \text{sign}(x^{n+1,k})],
\end{aligned}
$$

where $\text{sign}(\cdot)$ is the sign function. The termination criterion for the inner loop is set to be

$$
\frac{\|x^{n+1,k+1} - x^{n+1,k}\|_2}{\max\{\|x^{n+1,k}\|_2, 1\}} < \delta
$$

for some given parameter tolerance parameter $\delta > 0$. Basically, this algorithm can be regarded as a generalized version of IRLS algorithm. Obviously, when $\alpha = 0$, it exactly reduces to the traditional IRLS algorithm used for solving the ℓ_r-minimization problem.

4. NUMERICAL EXPERIMENTS

In this section, some numerical experiments on the proposed algorithm in section 3 are conducted to illustrate the performance of the weighted $\ell_r - \ell_1$ minimization in simulated sparse signal recovery.

4.1. Successful Recovery

First, we focus on the weighted $\ell_r - \ell_1$ minimization itself. In this set of experiments, the s-sparse signal x is of length $N = 256$, which is generated by choosing s entries uniformly at random, and then choosing the non-zero values from the standard normal distribution for these s entries. The underdetermined linear measurements $y = Ax + e \in \mathbb{R}^m$, where $A \in \mathbb{R}^{m \times N}$ is a standard Gaussian random matrix and the entries of the noise vector $\{e_i, i = 1, 2, \cdots, m\} \overset{i.i.d.}{\sim} N(0, \sigma^2)$. Here we fix the number of measurements $m = 64$ and select a sequence of s as $10, 12, \cdots, 36$. We run the experiments for both noiseless and noisy cases. In all the experiments, we let the tolerance parameter $\delta = 10^{-3}$. And all the results are averaged over 100 repetitions.

In the noiseless case, i.e., $\sigma = 0$, we set $\lambda = 10^{-6}$. In **Figure 2**, we show the results of successful recovery rate for different α (i.e., $\alpha = 0, 0.2, 0.5, 0.8, 1$) while fixing r but varying the sparsity level s. We view it as a successful recovery if $\|\hat{x} - x\|_2/\|x\|_2 < 10^{-3}$. We do the experiments for $r = 0.3$ and $r = 0.7$, respectively. As we can see, when r is fixed, the influence of the weight α is negligible, especially in the case that r is relatively small. But the performance does improve in some scenarios when a proper weight α is used. However, the problem of adaptively selecting the optimal α seems to be challenging and is left for future work. In addition, we present the reconstruction performances for different r (i.e., $r = 0.01, 0.2, 0.5, 0.8, 1$) while the weight α is fixed to be 0.2 and 0.8 in **Figure 3**. Note that small r is favored when the weight α is fixed. And a non-convex recovery with $0 < r < 1$ performs much better than the convex case ($r = 1$).

Next, we consider the noisy case, that is $\sigma = 0.01$. We set $\lambda = 10^{-4}$. And we evaluate the recovery performance by the signal to noise ratio (SNR), which is given by

$$
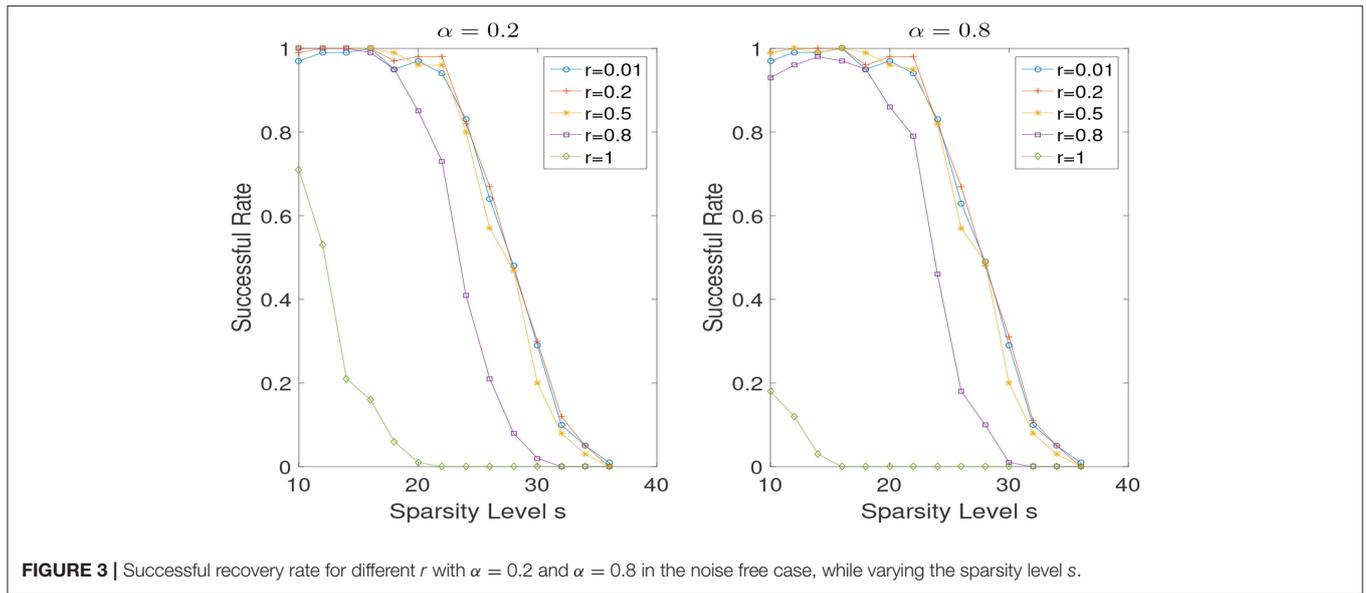\text{SNR} = 20 \log_{10} \left(\frac{\|x\|_2}{\|\hat{x} - x\|_2} \right).
$$

FIGURE 3 | Successful recovery rate for different r with $\alpha = 0.2$ and $\alpha = 0.8$ in the noise free case, while varying the sparsity level s.
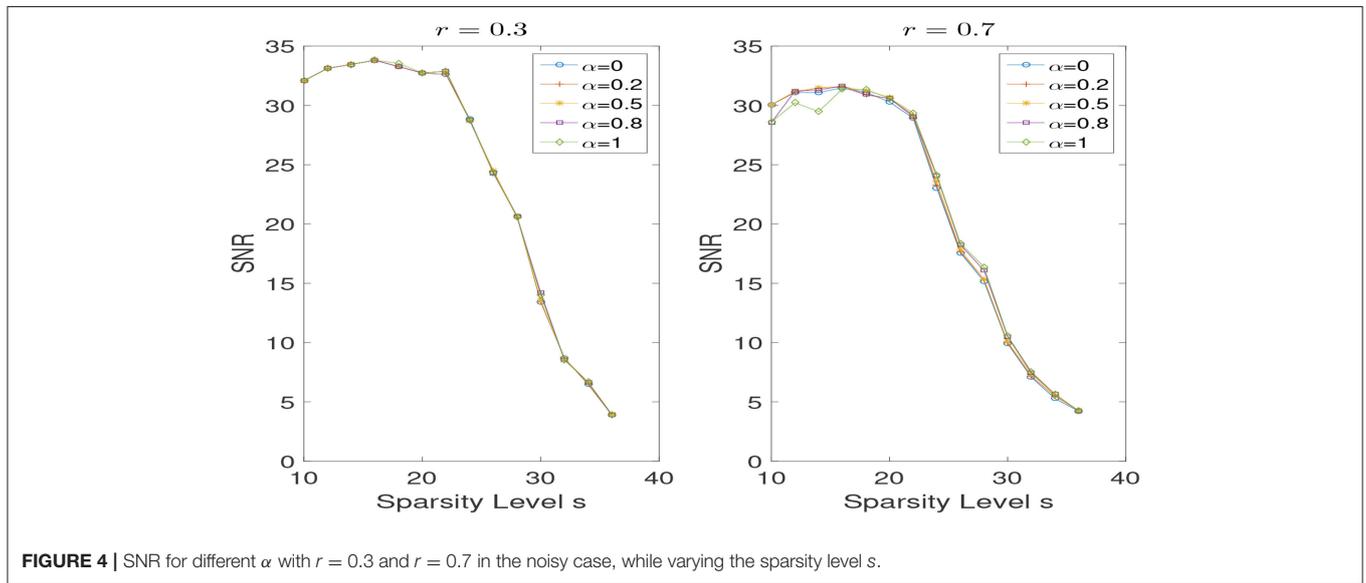

FIGURE 4 | SNR for different α with $r = 0.3$ and $r = 0.7$ in the noisy case, while varying the sparsity level s.

As shown in **Figures 4**, **5**, the findings aforementioned can still be seen here.

4.2. Algorithm Comparisons

Second, we compare the weighted $\ell_r - \ell_1$ minimization with some well-known algorithms. The following state-of-the-art recovery algorithms are operated:

- ADMM-Lasso, see Boyd et al. [36].
- CoSaMP, see Needell and Tropp [37].
- Iterative Hard Thresholding (IHT), see Blumensath and Davies [38].
- $\ell_1 - \ell_2$ minimization, see Yin et al. [24].

The tuning parameters used for these algorithms are the same as those adopted in section 5.2 of Yin et al. [24]. Specifically, for ADMM-Lasso, we choose $\lambda = 10^{-6}$, $\beta = 1$, $\rho = 10^{-5}$, $\varepsilon^{abs} = 10^{-7}$, $\varepsilon^{rel} = 10^{-5}$, and the maximum number of iterations maxiter = 5,000. For CoSaMP, maxiter=50 and the tolerance is set to be 10^{-8}. The tolerance for IHT is 10^{-12}. For $\ell_1 - \ell_2$ minimization, we choose the parameters as $\varepsilon^{abs} = 10^{-7}$, $\varepsilon^{rel} = 10^{-5}$, $\varepsilon = 10^{-2}$, MAXoit = 10, and MAXit = 500. For our weighted $\ell_r - \ell_1$ minimization, we choose $\lambda = 10^{-6}$, $r = 0.5$ but with two different weights $\alpha = 0$ (denoted as $\ell_{0.5}$) and $\alpha = 1$ (denoted as $\ell_{0.5} - \ell_1$).

We only consider the exactly sparse signal recovery in the noiseless case, and conduct the experiments under the same settings as in section 4.1. We present the successful
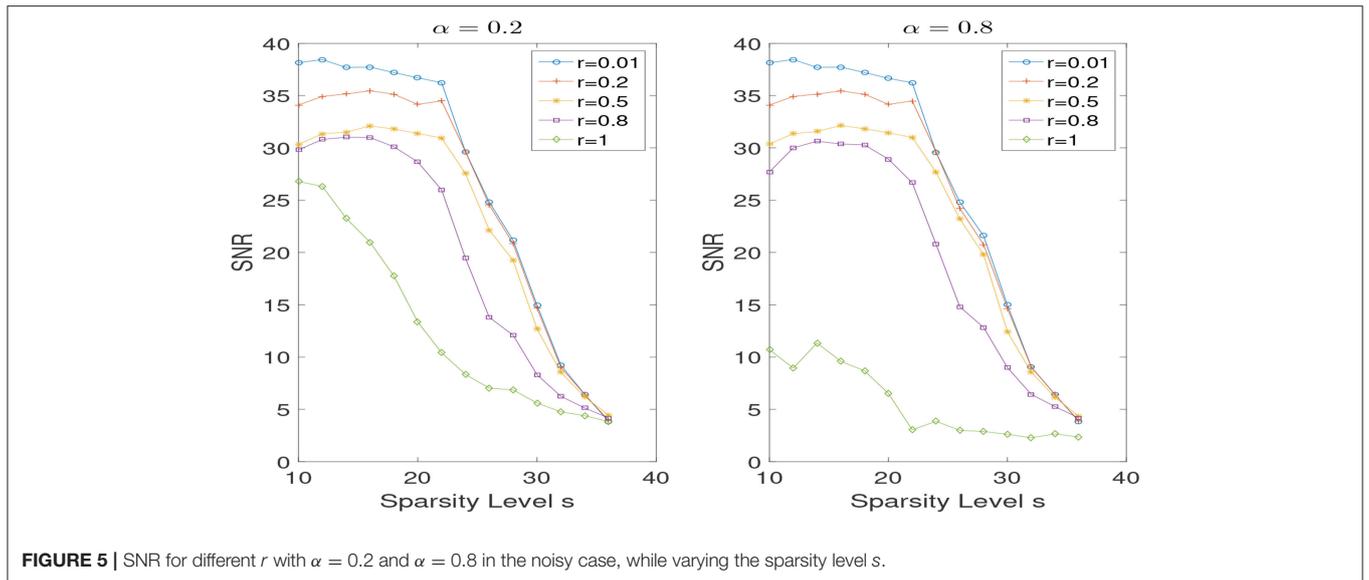
FIGURE 5 | SNR for different r with $\alpha = 0.2$ and $\alpha = 0.8$ in the noisy case, while varying the sparsity level s.
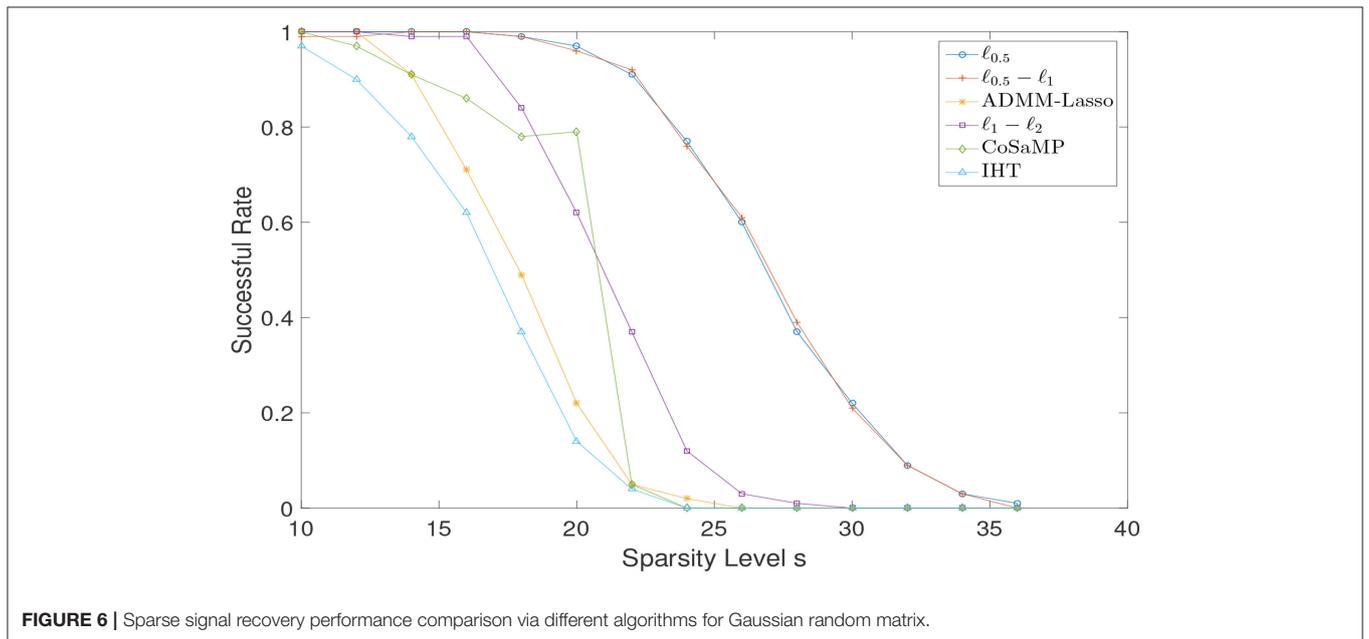

FIGURE 6 | Sparse signal recovery performance comparison via different algorithms for Gaussian random matrix.

recovery rates for different reconstruction algorithms while varying the sparsity level s in **Figure 6**. It can be observed that both $\ell_{0.5}$ and $\ell_{0.5} - \ell_1$ outperform over other algorithms, although their own performances are almost the same.

5. CONCLUSION

In this paper, we studied a new non-convex recovery method, developed as minimizing a weighted difference of ℓ_r ($0 < r \leq 1$) norm and ℓ_1 norm. We established the performance bounds

for this problem based on both r-RIP and q-ratio CMSV. An algorithm was proposed to approximately solve the non-convex problem. Numerical experiments show that the proposed algorithm provides superior performance compared to the existing algorithms such as ADMM-Lasso, CoSaMP, IHT and $\ell_1 - \ell_2$ minimization.

Besides, there are some open problems left for future work. One is the convergence study of the proposed algorithm in section 3. Another one is the generalization of this 1-D non-convex version to 2-D non-convex total variation minimization as done in Lou et al. [39] and the exploration of its application to medical imaging. Moreover, analogous to the non-convex

block-sparse compressive sensing studied in Wang et al. [40], the study of the following non-convex block-sparse recovery minimization problem:

$$\min_{z \in \mathbb{R}^N} \|z\|_{2,r}^r - \alpha \|z\|_{2,1}^r \ \text{ subject to } \ \|Az - y\|_2 \leq \eta, \quad (39)$$

where $\|z\|_{2,r} = (\sum_{i=1}^p \|z[i]\|_2^r)^{1/r}$ with $z[i]$ denoting the i-th block of z, $0 \leq \alpha \leq 1$, and $0 < r \leq 1$, is also an interesting topic for further investigation.

AUTHOR CONTRIBUTIONS

ZZ contributed to the initial idea and wrote the first draft. JY provided critical feedback and helped to revise the manuscript.

FUNDING

REFERENCES

1. Candes EJ, Tao T. Decoding by linear programming. *IEEE Trans Inf Theory.* (2005) **51**:4203–15. doi: 10.1109/TIT.2005.858979

2. Donoho DL. Compressed sensing. *IEEE Trans Inf Theory.* (2006) **52**:1289–306. doi: 10.1109/TIT.2006.871582

3. Eldar YC, Kutyniok G. *Compressed Sensing: Theory and Applications.* Cambridge: Cambridge University Press (2012).

4. Foucart S, Rauhut H. *A Mathematical Introduction to Compressive Sensing.* Vol. 1. New York, NY: Birkhäuser (2013).

5. Gribonval R, Nielsen M. Sparse representations in unions of bases. *IEEE Trans Inf Theory.* (2003) **49**:3320–5. doi: 10.1109/TIT.2003.820031

6. Tropp JA. Greed is good: algorithmic results for sparse approximation. *IEEE Trans Inf Theory.* (2004) **50**:2231–42. doi: 10.1109/TIT.2004.834793

7. Candes EJ. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Math.* (2008) **346**:589–92. doi: 10.1016/j.crma.2008.03.014

8. Candes EJ, Romberg JK, Tao T. Stable signal recovery from incomplete and inaccurate measurements. *Commun Pure Appl Math.* (2006) **59**:1207–23. doi: 10.1002/cpa.20124

9. Cohen A, Dahmen W, DeVore R. Compressed sensing and best k-term approximation. *J Am Math Soc.* (2009) **22**:211–31. doi: 10.1090/S0894-0347-08-00610-3

10. Dirksen S, Lecué G, Rauhut H. On the gap between restricted isometry properties and sparse recovery conditions. *IEEE Trans Inf Theory.* (2018) **64**:5478–87. doi: 10.1109/TIT.2016.2570244

11. Tang G, Nehorai A. Performance analysis of sparse recovery based on constrained minimal singular values. *IEEE Trans Signal Process.* (2011) **59**:5734–45. doi: 10.1109/TSP.2011.2164913

12. Zhou Z, Yu J. Sparse recovery based on q-ratio constrained minimal singular values. *Signal Process.* (2019) **155**:247–58. doi: 10.1016/j.sigpro.2018.10.002

13. Chen SS, Donoho DL, Saunders MA. Atomic decomposition by basis pursuit. *SIAM J Sci Comput.* (1998) **20**:33–61.

14. Natarajan BK. Sparse approximate solutions to linear systems. *SIAM J Sci Comput.* (1995) **24**:227–34. doi: 10.1137/S0097539792240406

15. Chartrand R. Exact reconstruction of sparse signals via nonconvex minimization. *IEEE Signal Process Lett.* (2007) **14**:707–10. doi: 10.1109/LSP.2007.898300

16. Foucart S, Lai MJ. Sparsest solutions of underdetermined linear systems via ℓ_q-minimization for $0 < q \leq 1$. *Appl Comput Harmon Anal.* (2009) **26**:395–407. doi: 10.1016/j.acha.2008.09.001

17. Li S, Lin J. Compressed Sensing with coherent tight frames via l_q-minimization for $0 < q \leq 1$. *Inverse Probl Imaging.* (2014) **8**:761–77. doi: 10.3934/ipi.2014.8.761

18. Lin J, Li S. Restricted q-isometry properties adapted to frames for nonconvex l_q-analysis. *IEEE Trans Inf Theory.* (2016) **62**:4733–47. doi: 10.1109/TIT.2016.2573312

19. Shen Y, Li S. Restricted p–isometry property and its application for nonconvex compressive sensing. *Adv Comput Math.* (2012) **37**:441–52. doi: 10.1007/s10444-011-9219-y

20. Xu Z, Chang X, Xu F, Zhang H. $L_{1/2}$ regularization: a thresholding representation theory and a fast solver. *IEEE Trans Neural Netw Learn Syst.* (2012) **23**:1013–27. doi: 10.1109/TNNLS.2012.2197412

21. Chartrand R, Yin W. Iteratively reweighted algorithms for compressive sensing. In: *IEEE International Conference on Acoustics, Speech and Signal Processing, 2008.* (2008). p. 3869–72.

22. Lai MJ, Xu Y, Yin W. Improved iteratively reweighted least squares for unconstrained smoothed ℓ_q minimization. *SIAM J Numer Anal.* (2013) **51**:927–57. doi: 10.1137/110840364

23. Lou Y, Yin P, He Q, Xin J. Computing sparse representation in a highly coherent dictionary based on difference of L1 and L2. *J Sci Comput.* (2015) **64**:178–96. doi: 10.1007/s10915-014-9930-1

24. Yin P, Lou Y, He Q, Xin J. Minimization of ℓ_{1-2} for compressed sensing. *SIAM J Sci Comput.* (2015) **37**:A536–63. doi: 10.1137/140952363

25. Yin P, Esser E, Xin J. Ratio and difference of l_1 and l_2 norms and sparse representation with coherent dictionaries. *Commun Inf Syst.* (2014) **14**:87–109. doi: 10.4310/CIS.2014.v14.n2.a2

26. Lou Y, Yan M. Fast L1–L2 minimization via a proximal operator. *J Sci Comput.* (2018) **64**:767–85. doi: 10.1007/s10915-017-0463-2

27. Wang Y. *New Improved Penalty Methods for Sparse Reconstruction Based on Difference of Two Norms.* Optimization Online, the Mathematical Optimization Society (2015). Available online at: http://www.optimization-online.org/DB_HTML/2015/03/4849.html

28. Wang D, Zhang Z. Generalized sparse recovery model and its neural dynamical optimization method for compressed sensing. *Circ Syst Signal Process.* (2017) **36**:4326–53. doi: 10.1007/s00034-017-0532-7

29. Zhao Y, He X, Huang T, Huang J. Smoothing inertial projection neural network for minimization L_{p-q} in sparse signal reconstruction. *Neural Netw.* (2018) **99**:31–41. doi: 10.1016/j.neunet.2017.12.008

30. Chartrand R, Staneva V. Restricted isometry properties and nonconvex compressive sensing. *Inverse Probl.* (2008) **24**:035020. doi: 10.1088/0266-5611/24/3/035020

31. Lopes ME. Unknown sparsity in compressed sensing: denoising and inference. *IEEE Trans Inf Theory.* (2016) **62**:5145–66. doi: 10.1109/TIT.2016.2587772

32. Zhou Z, Yu J. On q-ratio CMSV for sparse recovery. *arXiv [Preprint]. arXiv:180512022.* (2018). Available online at: https://arxiv.org/abs/1805.12022

33. Tran H, Webster C. Unified sufficient conditions for uniform recovery of sparse signals via nonconvex minimizations. *arXiv preprint arXiv:171007348.* (2017).

34. Tao PD, An LTH. Convex analysis approach to dc programming: theory, algorithms and applications. *Acta Math Vietnam.* (1997) **22**:289–355.

35. Tao PD, An LTH. A DC optimization algorithm for solving the trust-region subproblem. *SIAM J Optim.* (1998) **8**:476–505. doi: 10.1137/S1052623494274313

36. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found Trends Mach Learn.* (2011) **3**:1–122. doi: 10.1561/2200000016

37. Needell D, Tropp JA. CoSaMP: iterative signal recovery from incomplete and inaccurate samples. *Appl Comput Harmon Anal.* (2009) **26**:301–21. doi: 10.1016/j.acha.2008.07.002

38. Blumensath T, Davies ME. Iterative hard thresholding for compressed sensing. *Appl Comput Harmon Anal.* (2009) **27**:265–74. doi: 10.1016/j.acha.2009.04.002

39. Lou Y, Zeng T, Osher S, Xin J. A weighted difference of anisotropic and isotropic total variation model for image processing. *SIAM J Imaging Sci.* (2015) 8:1798–23. doi: 10.1137/140 98435X

40. Wang Y, Wang J, Xu Z. Restricted *p*-isometry properties of nonconvex block-sparse compressed sensing. *Signal Process.* (2014) **104**:188–96. doi: 10.1016/j.sigpro.2014.03.040

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.