

CHAPTER III

REWEIGHTED ℓ_1 MINIMIZATION

1. Weighted ℓ_1 -minimization analysis

An idea to enhance the sparse signal recovery of ℓ_1 -minimization is possible, if a relaxation of ℓ_1 -minimization still results *convex optimization* with the same constraints. This was initially proposed in 2007 by Candès and Wakin [5] who designed the counterpart “weighted” ℓ_1 -minimization problem,

$$(WP_1) \quad \min \sum_{i=1}^N w_i |s_i| \quad \text{sub. to } \Theta s = y, \quad (8)$$

where w_i are the positive weights. This optimization is the same as (P_1) , if $w_i = 1$, called “un-weighted” counterpart. For the convenient computing structure, the objective function in (8) is replaced by $\|Ws\|_1$,

$$(WP_1) \quad \min \|Ws\|_1 \quad \text{sub. to } \Theta s = y, \quad (9)$$

where weighting matrix W is the diagonal matrix with w_1, w_2, \dots, w_N on the diagonal and zeros elsewhere, thus $W = \text{diag}([w_1, w_2, \dots, w_N]^T)$.

The point of view is first to consider the relaxation of a weighted ℓ_0 -minimization that can be expressed as,

$$(WP_0) \quad \min \|Ws\|_0 \quad \text{sub. to } \Theta s = y. \quad (10)$$

Assume the solution to ℓ_0 -minimization be unique, it is also the unique solution to weighted ℓ_0 -minimization (P_0) provided that the weights do not vanish. However, the optimizations corresponding to ℓ_1 relaxations (P_1) and (WP_1) might generally present the different solutions.

Consider the geometry of weighted ℓ_1 ball (the possible capacity cost of objective function $\|s\|_1$) in Figure 7 which visualizes the improved shape succeeds where un-weighted ℓ_1 ball fails. In Figure 7(a), suppose the K – sparse entries are on the vertex of ℓ_1 ball with radius $\|s_0\|_1$ and touch the hyperplane \mathcal{H} at the true vector s_0 . In some conditions, the un-weighted ℓ_1 ball spans to touch the hyperplane \mathcal{H} in the fault point for which $\|s\|_1 < \|s_0\|_0$, as Figure 7(b). Thus the weighting function is used to improve the shape of interior ℓ_1 ball for which $\|Ws\|_1 \leq \|Ws_0\|_1$. It is possible that the ℓ_1 ball will touch the hyperplane \mathcal{H} on the correct point, as Figure 7(c).

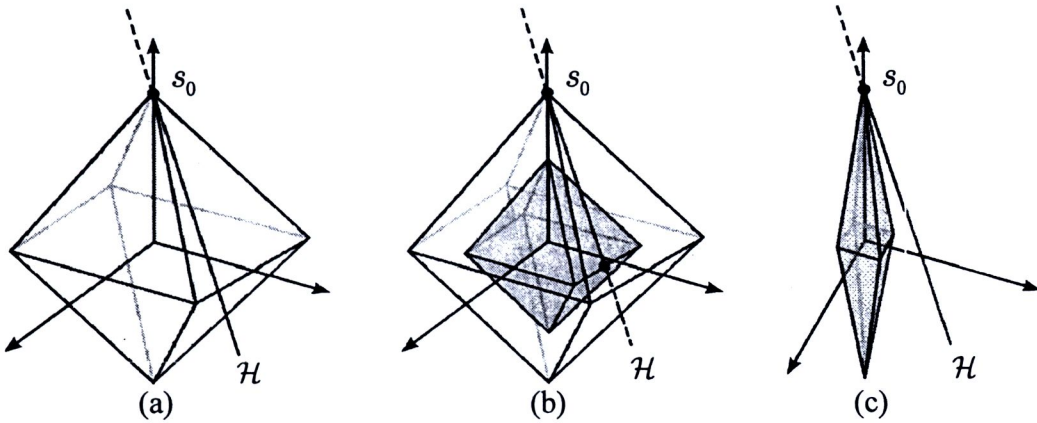
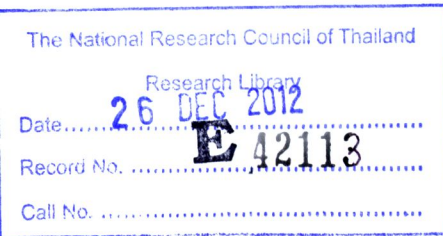


Figure 7 Reconstructed sparse signal on the vertex of ℓ_1 balls: (a) it touches the hyperplane \mathcal{H} . (b) Interior ℓ_1 ball touching the hyperplane \mathcal{H} at the fault solution vector. (c) Enhanced weighted ℓ_1 ball touching the hyperplane \mathcal{H} at the true K – sparse vector.

In sequel, the intermediate stage is to provide the weights w_i which are available to enhance the K – sparse signal reconstruction. One possible use for the weights can be to counteract the influence of the signal magnitude on the ℓ_1 penalty function. Note that as example, the weights are inversely proportional to the true signal magnitude —i.e.,



$$w_i = \begin{cases} 1 & \text{if } s_{0,i} \neq 0 \\ |s_{0,i}| & \\ \infty & \text{if } s_{0,i} = 0 \end{cases}, \quad (11)$$



which $s_{0,i}$ is the true solution to each entry i . If the original signal is K -sparse and $\|s_0\|_0 \leq K$, then reweighted ℓ_1 -minimization algorithm will guarantee to find the correct solution [5]. However, the weights $w_i = 0$ cannot be defined in numerical computation so that the weighting function is basically revised to the equivalent form,

$$w_i = \frac{1}{|s_i| + \varepsilon}, \quad (12)$$

where $\varepsilon > 0$ for ensuring the division to zero-value component in the reconstructed vector \bar{s}_i which might estimate the weighting value w_i to infinite. Figure 8 shows the distribution of weighting function (12) when varying the parameter $\varepsilon = 0.01, 0.05$ and 0.1 (denote “eps” for ε in the Figure). Furthermore, this weighting function does not require the locations to the K -sparse signal in order that the efficiency to search the true solutions might be generally decreased.

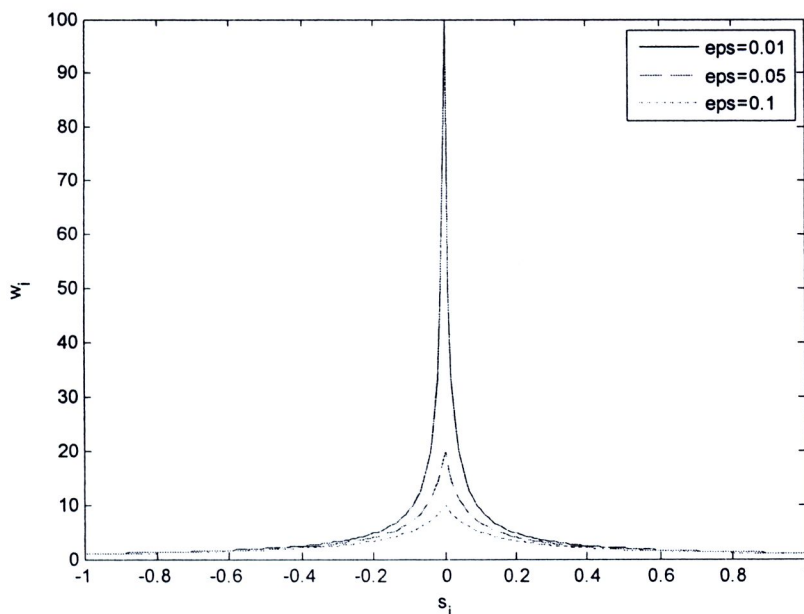


Figure 8 Distribution of original weighting function (when varying the parameter ε on the interval of possible solutions $[-1,1]$).

2. Iterative algorithm for reweighted ℓ_1 -minimization

As Figure 7, the boundary of un-weighted ℓ_1 ball might estimate the wrong solutions. Although, the favorable weighting matrix W helps to fix the range of ℓ_1 ball, the vector magnitudes might not reshape ℓ_1 ball touching the hyperplane at the true vector s_0 on a sole iteration. Thus, the proposed process which enhances the boundary of ℓ_1 ball closer the convergent solution is to refine the weights and repeat the weighted ℓ_1 -minimization, iteratively. A simple reweighted algorithm is illustrated in Figure 9.

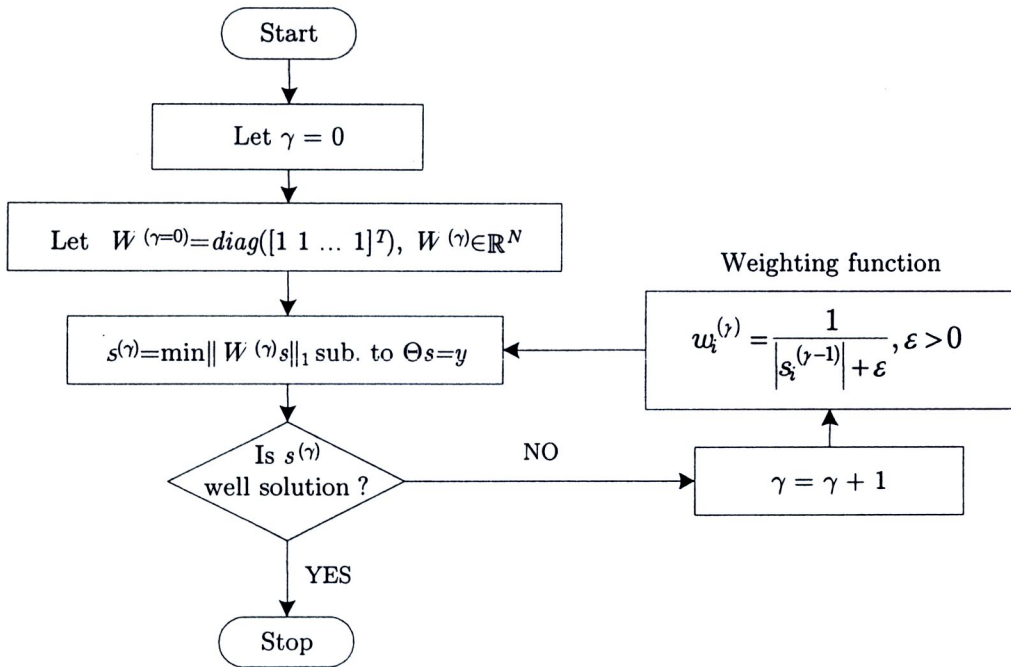


Figure 9 Iterative reweighted ℓ_1 -minimization

The first knowing solution vector $s^{(0)}$ as of un-weighted ℓ_1 -minimization ($w_i = 1$) is used to construct the range of weighting matrix $W^{(1)}$. Then, the next weighting matrix will be estimated by the previous solution vector $s^{(\gamma-1)}$. This algorithm repeats until satisfies the well solution which means terminating on convergence or when γ attains a specific maximum number of iterations γ_{\max} . Otherwise, increment $\gamma = \gamma + 1$ and repeat the algorithm, successively.

In general, the efficiency of reweighted sparse recovery depends on the choice of parameter ε which distributes the different weighting magnitudes when varying its values as shown in Figure 8. However, the smart and robust rules that would automatically select the parameter ε adapting for each K -sparse signals are currently on the experience trials. Thus, the algorithm that ensures the appropriate weighting magnitude is still an open question.

3. Hard selective reweighted ℓ_1 -minimization

Since the parameter ε in weighting function (12) is unbounded for some $\varepsilon > 0$ so that it is difficult to vary its value when undergoing the experiments. Our first intuitive idea is to design an alternative weighting function which is more general to select the controlling parameter bounded in the possible period. Notice Figure 8, the distribution of weighting function (12) is mostly to scale the solutions which their values are close to zero by the large factor. Thus, the achievable weighting function which has the similar characteristic is proposed by *hard selective weighting function*,

$$w_i = \begin{cases} \beta_{\max} & , |s_i| \leq \tau \\ \beta_{\min} & , |s_i| > \tau \end{cases}, \quad (13)$$

where τ is a threshold which its value locates in the period of possible solution divided the solutions into 2 groups; the solutions of first group which are close to zero assumed that they are zero candidates scaled by β_{\max} and the other solutions are scaled by β_{\min} . Furthermore, the scaling factors $\beta_{\max}, \beta_{\min} > 0$ are well-defined because if $\beta_{\max}, \beta_{\min} = 0$, some important solutions which are in the summation $\|W_s\|_1$ will be cancelled. Figure 10 shows the distribution of this weighting function when defining $\beta_{\max} = 100$, $\beta_{\min} = 1$ and varying the value of threshold $\tau = 0.01$, 0.1 and 0.5 (denote “thres” for τ in the Figure).

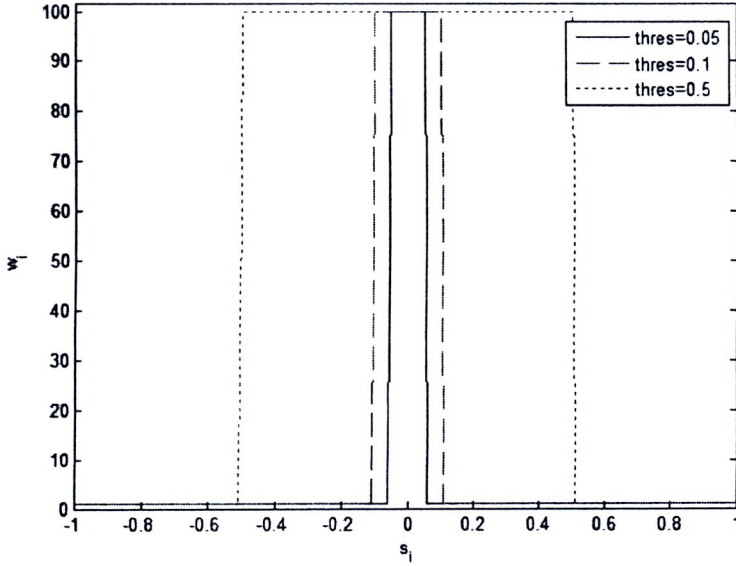


Figure 10 Distribution of hard selective weighting function (when varying the threshold τ in on the interval of possible solutions $[-1,1]$).

Consider a 3-D example in Figure 11 shows the hard selective weighted ℓ_1 ball of radius $\|Ws\|_1 = 1$ centered at the origin. This ball is reshaped inversely scaling factors because the changing of weights w_i effects the solutions s_i by $1/\beta_{\max}$ and $1/\beta_{\min}$ for the scaling factors β_{\max} and β_{\min} , respectively.

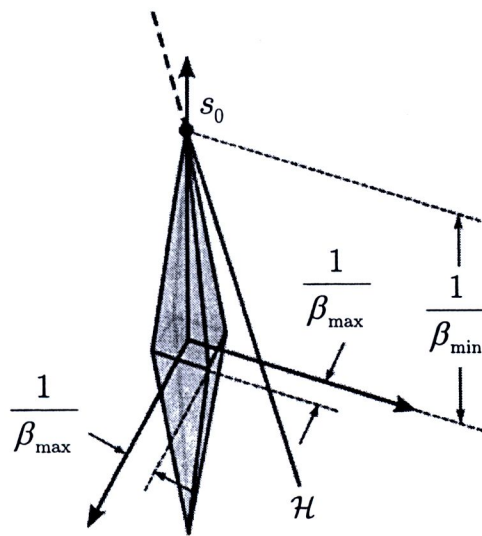


Figure 11 Hard selective weighted ℓ_1 ball

One possible to control the weights, the right ℓ_1 ball has to span the vertex which contains many zero entries possibly on non-zero coordinate axes intersecting hyperplane \mathcal{H} at the true vector before the other vertices catch up. Suppose if all zero candidates (the reconstructed solutions $|s_i| \leq \tau$) are the exact zero entries, the weighting value β_{\max} should be infinitely larger than β_{\min} that is guaranteed to find the correct solutions as *hard selective ratio*,

$$\frac{\beta_{\max}}{\beta_{\min}} = \infty. \quad (14)$$

In numerical experiment, this ratio cannot be defined as infinity in order that the infinite rate is necessarily changed to the recommended ratio,

$$\frac{\beta_{\max}}{\beta_{\min}} = \frac{1}{\mu}, \quad (15)$$

where $\mu \leq 1$ is represented as an expanding rate. The ℓ_1 ball is un-weighted for $\mu = 1$ moreover it can be spanned more correctly touching the hyperplane \mathcal{H} when the limit of μ close to zero. For instruction use, if the considered signal is of the accuracy of floating point fixed point value at 0.01, the parameter $\mu \leq 0.01$ will be sufficient value.

However, the various expanding rates do not sometimes impact the changing of solution vectors because these must exist in the same constraints which the minimizers are possible to be the unique vector especially when decreasing expanding rate μ until reaching a constant number. Consider the weighted ℓ_1 balls in Figure 12 that are reshaped by the different expanding rates μ but these shapes which are pointy with the points along the coordinate axes (it becomes pointier as the ambient dimension N grows) touch the hyperplane \mathcal{H} correctly at the same solution vector. For example, Figure 13(c) shows the results of iterative reweighted reconstructions when using hard selective weighting function and varying expanding rates $\mu = 0.1, 0.01, 0.001$ and 0.0001 .

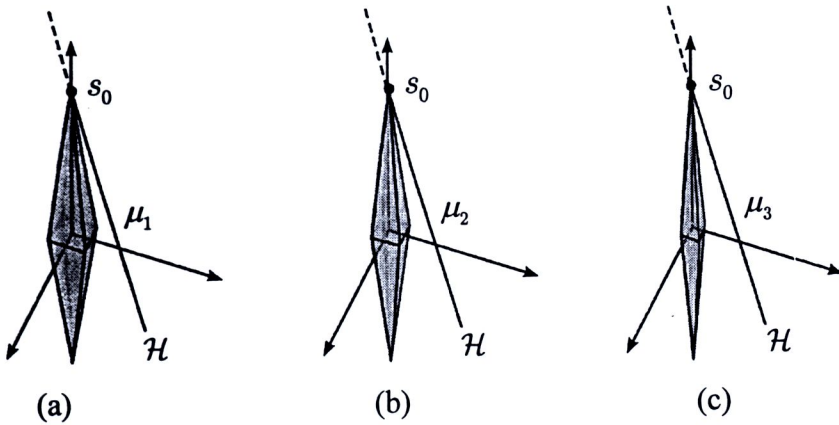


Figure 12 Hard selective weighted ℓ_1 balls when varying the parameter μ (where (a) $\mu_1 >$ (b) $\mu_2 >$ (c) μ_3).

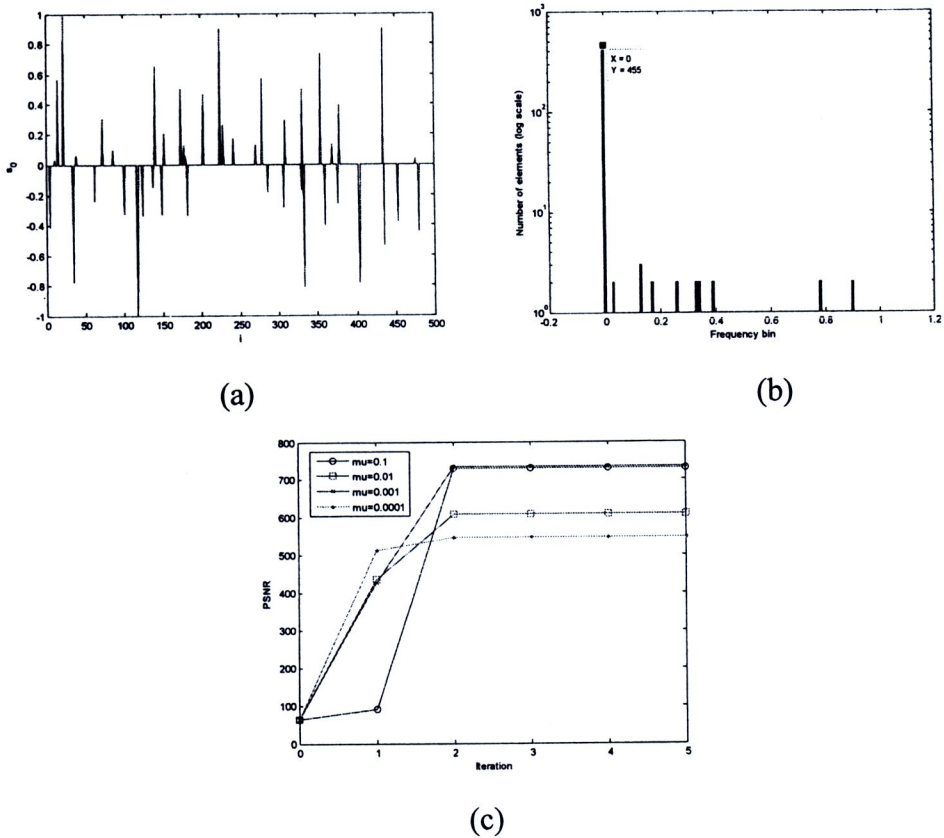


Figure 13 Sparse signal recovery using iterative reweighted ℓ_1 -minimization: (a) original coefficient vector s_0 on the interval $[-1,1]$, length $N = 500$, with 45 spikes, (b) its histogram of absolute coefficient vector $|s_0|$ and (c) its PSNRs (in addition at Experiment I, CHAPTER IV) when using hard selective weighting function with threshold $\tau = 0.05$ and varying expanding rates μ .

In general, a number of zero entries in reconstructed signal must be equal to the original one. Due to the fault of ℓ_1 -minimization, this means that the zero entries in the original signal are shifted to the other indices in reconstructed signal. Thus, if a number of original zero entries are already known, we can classify the groups of zero and non-zero entries by selecting a number of the closest-to-zero entries which are equal to a number of original zero entries to be zero candidates and the other entries are non-zero candidates. We apply the above concept to define the threshold τ for hard selective weighting function as shown in Figure 14.

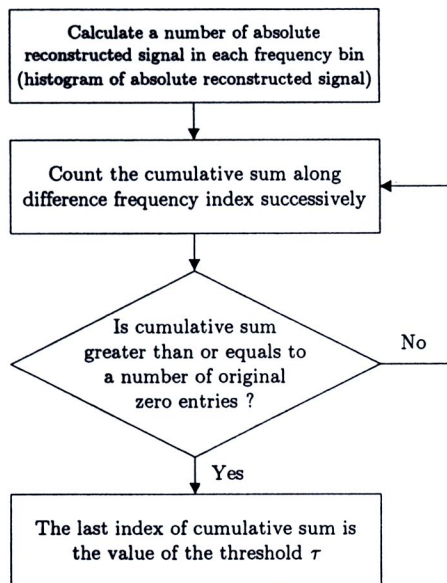


Figure 14 Selecting threshold τ for Hard selective weighting function

For example, the reconstruction of original coefficient vector s_0 (Figure 13(a)) via un-weighted ℓ_1 -minimization recovers the fault signal (Figure 15(a)). However, we know that the original signal contains 455 zero entries (Figure 13(b)) while the zero entries of absolute reconstructed coefficient vector $|s^{(0)}|$ are merely about 371 (Figure 15(b)) so that the threshold τ is defined as 0.06 for containing equally 455 zero candidate solutions (Figure 15(c)).

The reweighted algorithm which includes the selecting threshold τ process in the initial iteration $\gamma = 0$ after solving via ℓ_1 -minimization is called *hard selective reweighted* (HSR) [6, 7] algorithm as shown in Figure 16.

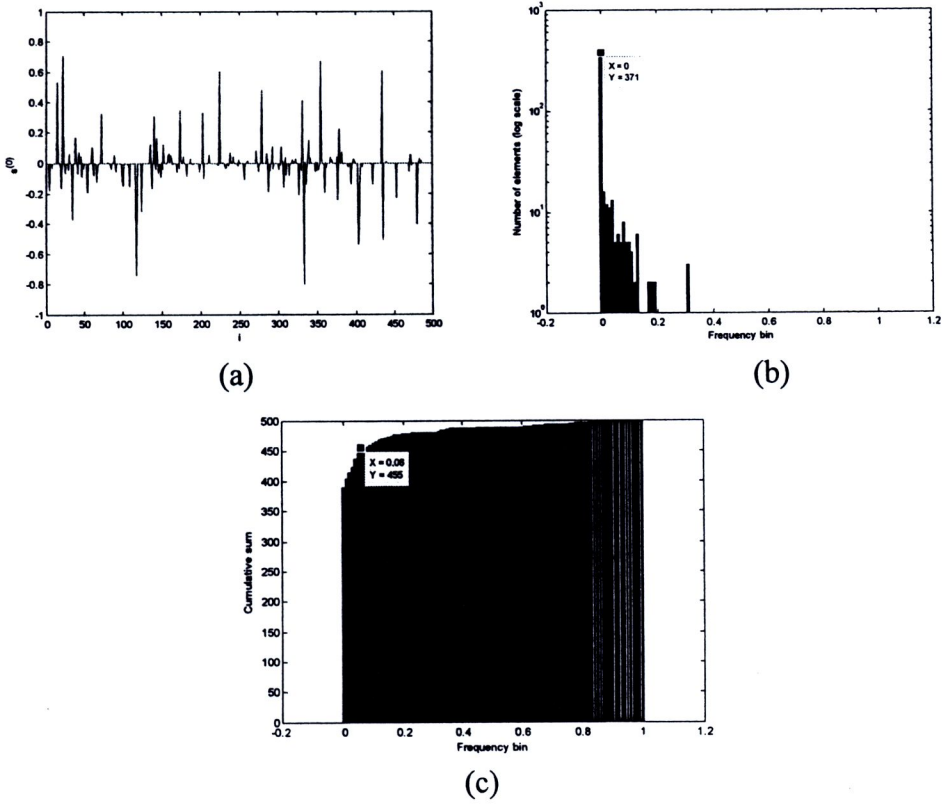


Figure 15. Fault reconstructed signal (original coefficient vector in Figure 13(a)) via un-weighted ℓ_1 -minimization: (a) its signal form, (b) its histogram of absolute coefficient vector $|s^{(0)}|$ and (c) its cumulative sum of histogram.

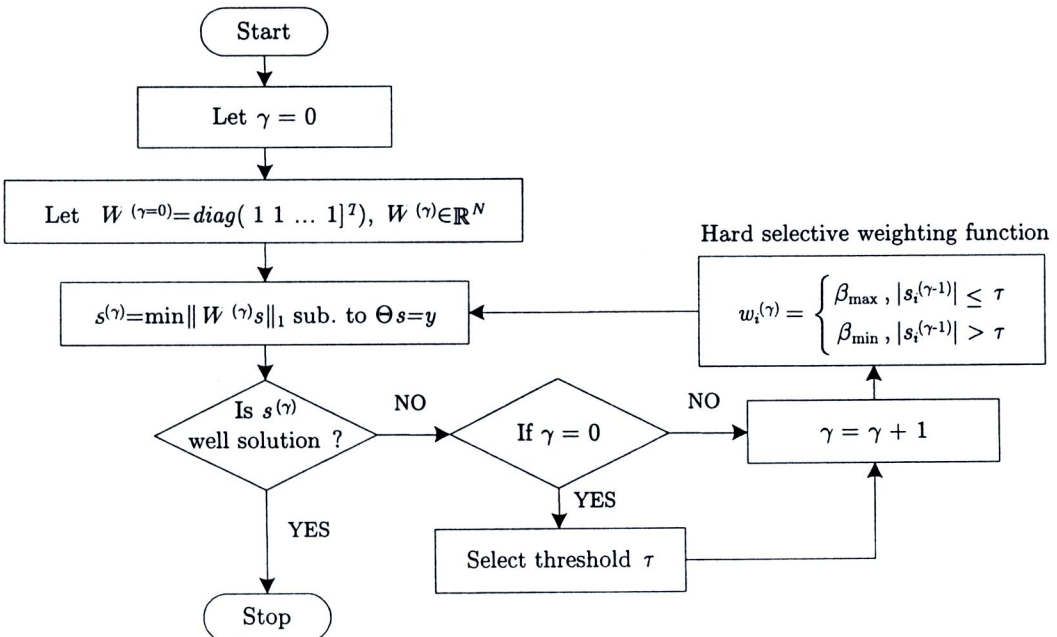


Figure 16 Iterative hard selective reweighted algorithm

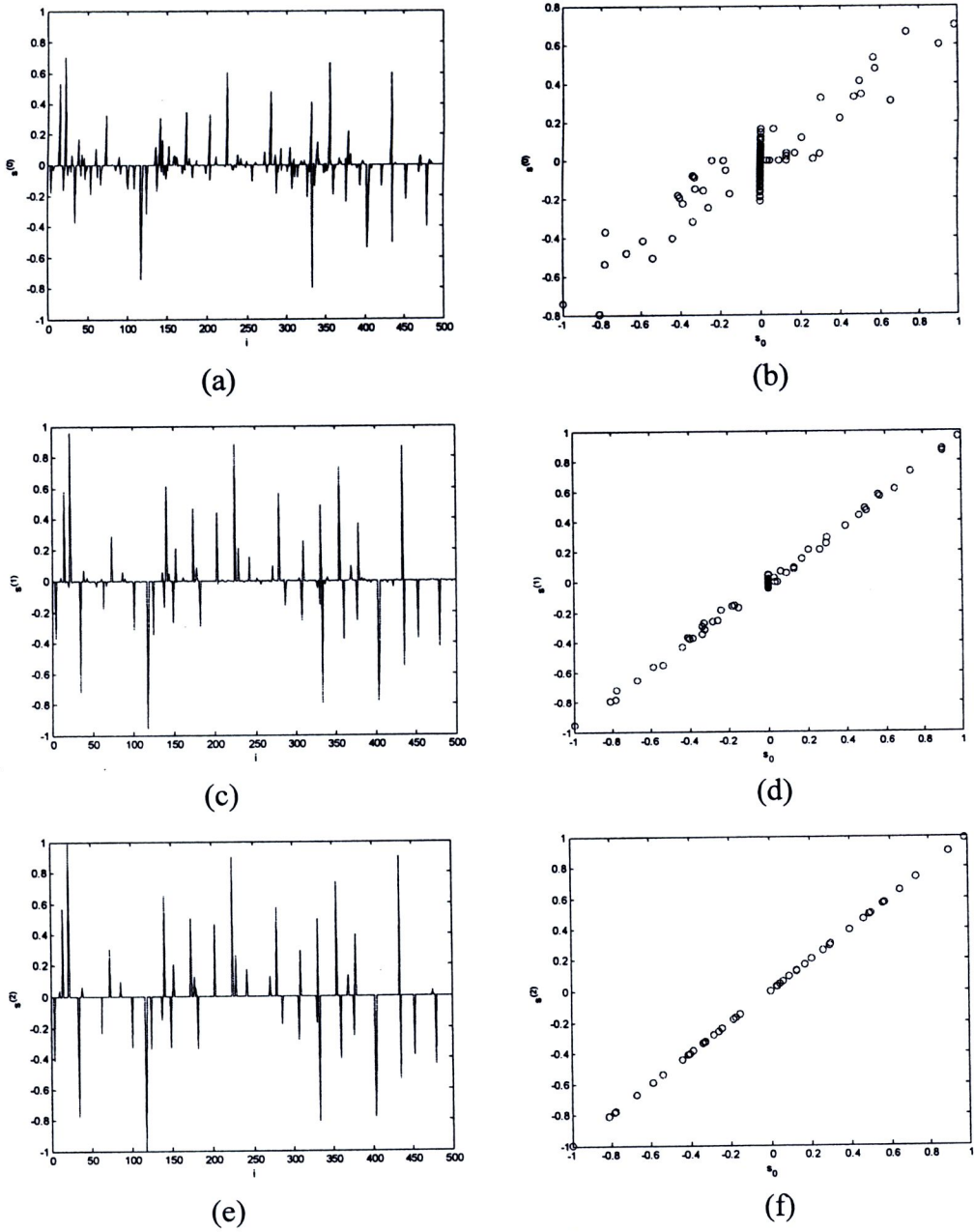


Figure 17 Sparse signal recovery using HSR algorithm (original coefficient vector in Figure 13(a)): (a) Reconstructed coefficient vector $s^{(0)}$ and (b) scatter plot using unweighted ℓ_1 -minimization. (c) Reconstructed coefficient vector $s^{(1)}$ at the first reweighted iteration and (d) its scatter plot. (e) Reconstructed coefficient vector $s^{(2)}$ at the second reweighted iteration and (f) its scatter plot.

4. Automatic adaptive reweighted ℓ_1 -minimization

In general recovery process, a number of original zero entries are not known so that the selecting threshold τ process in iterative reweighted algorithm without knowing a number of original zero entries is proposed by finding only the smallest frequency bin of histogram of absolute reconstructed coefficient vector $|s^{(0)}|$ which is empty, called automatic adaptive reweighted (AAR) algorithm [8, 7] as Figure 18.

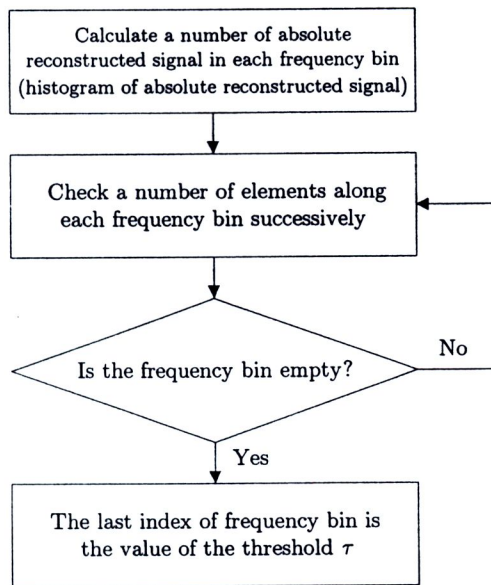


Figure 18 Selecting threshold τ for AAR algorithm

A brief algorithm concept is to assume the near-zero solutions being the zero candidates and their entries are punctuated by the nearest empty frequency bin of histogram of absolute reconstructed coefficient vector $|s^{(0)}|$.

For example, Figure 19(a) shows the histogram of fault reconstructed signal in absolute reconstructed coefficient vector $|s^{(0)}|$. There are not only zero entries in zero index, but also the other zero entries might be shifted to the other values around zero index. The nearest empty frequency bin of this example is about 0.16 so that the AAR algorithm has decided to select the threshold $\tau = 0.16$ for containing the near-zero solutions to be zero candidate entries.

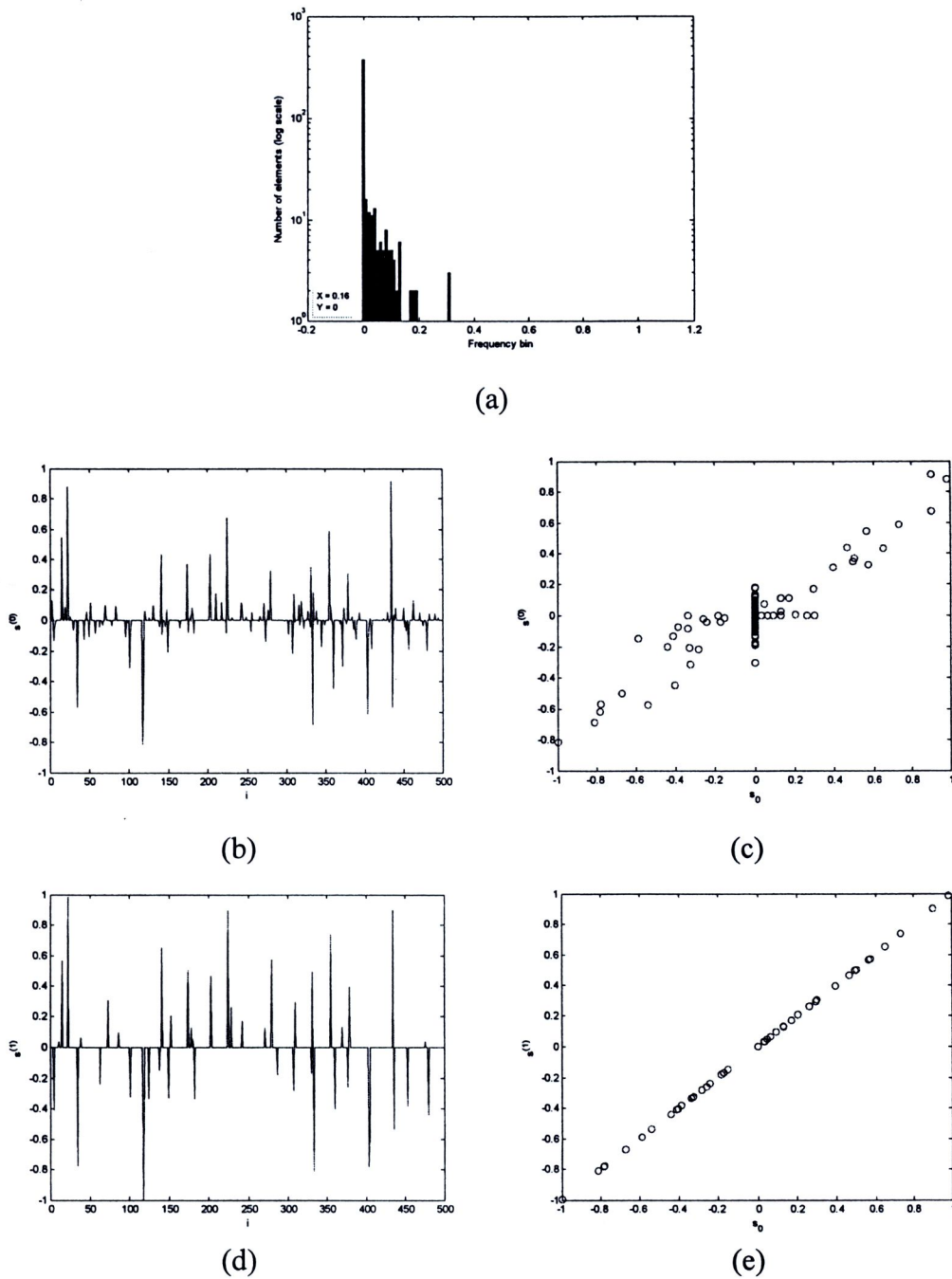


Figure 19 Sparse signal recovery using AAR algorithm (original coefficient vector in Figure 13(a)): (a) Histogram of absolute reconstructed coefficient vector $|s^{(0)}|$, (b) its waveform and (c) scatter plot using unweighted ℓ_1 -minimization. (d) Reconstructed coefficient vector $s^{(1)}$ at the first reweighted iteration and (e) its scatter plot.