

Convergence Analysis of Iterative Thresholding Algorithms

Arian Maleki

Department of Electrical Engineering and Statistics,
Stanford University
arianm@stanford.edu

Abstract—There is a recent surge of interest in developing algorithms for finding sparse solutions of underdetermined systems of linear equations $y = \Phi x$. In many applications, extremely large problem sizes are envisioned, with at least tens of thousands of equations and hundreds of thousands of unknowns. For such problem sizes, low computational complexity is paramount. The best studied ℓ_1 minimization algorithm is not fast enough to fulfill this need. Iterative thresholding algorithms have been proposed to address this problem. In this paper we want to analyze two of these algorithms theoretically, and give sufficient conditions under which they recover the sparsest solution.

I. INTRODUCTION

Finding the sparsest solution of an underdetermined system of linear equations $y = \Phi x$, is a problem of interest in signal processing, data transmission, biology and statistics—just to name a few. Unfortunately, this problem is NP-hard and in general can not be solved by a polynomial time algorithm. Chen et al. [1] proposed the following optimization for recovering the sparsest solution;

$$(Q_1) \quad \min \|x\|_1 \quad \text{s.t. } \Phi x = y,$$

where ℓ_p -norm is defined as $\|x\|_p = \sqrt[p]{\sum_i |x_i|^p}$. Greedy methods have also been proposed as another alternative for solving such a problem. One of the best known algorithms of this class is orthogonal matching pursuit (OMP) [2]. Intuitively speaking at each iteration, OMP finds a column of Φ which has the maximum correlation with the error of the approximation up to this step, and adds it to the active set and projects y onto the range of the active set to get a new estimate. The third class of algorithms that has drawn a lot of attention recently is the class of iterative thresholding algorithms. This class has the least computational complexity and is the most suitable class for very large scale problems [3]. There are many theoretical results that prove the optimality of the first two classes of algorithms under certain conditions, but there are much less rigorous results for thresholding algorithms. Before mentioning some of the results, we first set up the notation we are going to use in the paper. Suppose that $x_o \in \mathbb{R}^N$ is a k sparse vector (i.e. it has at most k non-zero elements). We observe the measurement vector $y = \Phi x_o$ which is in \mathbb{R}^n ($n < N$) and the goal is to reconstruct the original vector x_o . Without loss of generality, we assume that the columns of Φ have unit ℓ_2 norm. Another notation that is used in the paper is the notion of restricted submatrices. For a subset of columns of Φ called J , Φ_J includes all the columns of Φ whose indices

are in J , and x_J all the elements of x whose indices are in J . The coherence of Φ is defined as,

$$\mu = \max_{\{i,j:1 \leq i,j \leq N, i \neq j\}} |\langle \phi_i, \phi_j \rangle|. \quad (1)$$

where ϕ_i is the i^{th} column of the matrix Φ . In the following, a summary of the results proved for ℓ_1 minimization and OMP algorithms in [4] and [5] respectively, are presented.

Theorem I.1. *If $k \leq \frac{1}{2}(1+\mu^{-1})$, then both the ℓ_1 minimization and the OMP recover the sparsest solution.*

When the matrix Φ is drawn from a random ensemble [6], [7], we can bound the coherence [8], and find conditions for the exact sparse signal recovery. In this random setting, however, the results can be improved [9]. Although the theoretical results are basically focused on ℓ_1 relaxation and greedy methods, many large scale applications have already moved toward the thresholding algorithms [10], [11]. In a recent paper, we considered a few thresholding policies and showed that the results of these algorithms are very impressive in practical situations such as compressed sensing [3]. In this paper we focus on the theoretical aspects of these algorithms. The organization of the paper is as follows. In Section II, we discuss the thresholding algorithms and the thresholding policy considered in the paper; The main results of the paper will also be reviewed. Section III presents the convergence proof of the thresholding algorithms. In Section V, we will briefly review the existing literature on iterative thresholding algorithms and compare those results to ours. Finally Section VI concludes the paper.

II. ITERATIVE THRESHOLDING ALGORITHMS

A. Abstracted thresholding Algorithm

Consider two threshold functions $\eta_t(x)$ to be applied elementwise to vectors: hard thresholding $\eta_\mu^H(x) = x \mathbf{1}_{\{|x|>\mu\}}$ and soft thresholding $\eta_\mu^S(x) = \text{sgn}(x)(|x| - \mu)_+$, where $\mathbf{1}$ is the indicator function and $(a)_+$ is equal to a if $a > 0$, and zero otherwise. Iterative hard thresholding (IHT) and iterative soft thresholding (IST) algorithms are defined with the following iteration,

$$x^{t+1} = \eta_{\lambda_t}^*(x^t + \Phi^T(y - \Phi x^t)), \quad (2)$$

where λ_t is the threshold value at time t , $*$ $\in \{H, S\}$ represents hard or soft thresholding, Φ^T is the transpose

of the matrix Φ and x^t is our estimate at time t . Note that the threshold value may depend on the iteration. The basic intuition is that since the solution satisfies the equation $y = \Phi x$, algorithm makes progress by moving in the direction of the gradient of $\|y - \Phi x\|_2^2$ and then by thresholding the result, it tries to get a sparse vector closer to the hyperplane $y = \Phi x$. Another intuition for this algorithm comes from [12] and is as follows. Suppose that we want to solve the following optimization problem,

$$(\mathcal{P}_q) \quad \min_x \|y - \Phi x\|_2^2 + 2\lambda \|x\|_q.$$

It has been proved that the following IST algorithm converges to the solution of \mathcal{P}_1 in case $\|\Phi^T \Phi - I\|_{2,2} < 1$,

$$x^{t+1} = \eta_\lambda^S(x^t + \Phi^T(y - \Phi x^t)), \quad (3)$$

where $\|A\|_{2,2}$ is the spectral norm of the matrix A . It may be noted that λ is fixed here and does not depend on the iteration. It is also well-known that as $\lambda \rightarrow 0$ the solution of \mathcal{P}_1 converges to the solution of \mathcal{Q}_1 . But it is easy to see if Φ is a fat matrix, setting λ to a very small value in (3) will not work. A proper thresholding policy is to set the threshold to a large value and gradually decrease it as the algorithm proceeds. The following theorem justifies this intuition.

Consider the iterative soft or hard thresholding algorithms introduced in equation (2). Suppose $\lambda_t \rightarrow 0$ as $t \rightarrow \infty$, and λ_t is a decreasing sequence (this condition may not hold, but for the simplicity of the proof we assume it is true). Let J_t denote the union of the support of x^t and x_o and define $L_t := J_{t+1} \cup J_t$. Assume that L_t satisfies, $\sup_t \|I - \Phi_{L_t}^T \Phi_{L_t}\|_{2,2} = \gamma < 1$. Under these conditions:

Theorem II.1. *The iterative thresholding algorithm will converge to the sparsest solution.*

Proof:

$$\begin{aligned} \|x^{t+1} - x_o\|_2 &= \|x_{L_{t+1}}^{t+1} - x_{o_{L_{t+1}}}\|_2 \\ &\leq \|\eta_{\lambda_{t+1}}^*(x_{L_t}^t + \Phi_{L_t}^T(\Phi_{L_t} x_{o_{L_t}} - \Phi_{L_t} x_{L_t}^t)) - x_{o_{L_t}}\|_2, \\ &\leq \|(x_{L_t}^t + \Phi_{L_t}^T(\Phi_{L_t} x_{o_{L_t}} - \Phi_{L_t} x_{L_t}^t)) + \epsilon_{t+1} - x_{o_{L_t}}\|_2, \\ &\stackrel{1}{\leq} \|(I - \Phi_{L_t}^T \Phi_{L_t})(x_{L_t}^t - x_{o_{L_t}})\|_2 + \sqrt{n} \lambda_{t+1}, \\ &\leq \|(I - \Phi_{L_t}^T \Phi_{L_t})\|_{2,2} \|x_{L_t}^t - x_{o_{L_t}}\|_2 + \sqrt{n} \lambda_{t+1}, \end{aligned}$$

where ϵ_{t+1} is an extra error introduced by the thresholding process and therefore each element of this vector is less than λ_{t+1} . Also all the elements that are not in L_t are zero. Inequality (1) is just the triangle inequality for ℓ_2 norm. For any $\epsilon > 0$, choose T_0 such that $\sqrt{n} \lambda_{T_0+1} < \frac{\epsilon(1-\gamma)}{2}$, and let $\|x^{T_0+1} - x_o\|_2 = e$. Then, find T_1 such that $\gamma^{T_1} e < \epsilon/2$. Now it is easy to prove that at $t = T_0 + T_1$, the error is less than ϵ and therefore the total error goes to zero. ■

This theorem is not useful for practical purposes since we should have information on the size of L_t . In the next section we mention a practical thresholding policy that may be used in practice.

B. Thresholding Policy

Suppose that an oracle tells us the true underlying k . Then since the final solution is k sparse, the threshold can be set to

the magnitude of the $(k+1)^{\text{th}}$ largest coefficient. This type of thresholding policy has also been used in [13],[14], [15]. The only problem is how to get the oracle information. In a recent paper, we showed how one can de-oracle such algorithms for compressed sensing problems [3]. For other types of problems, k may be estimated using cross validation. If neither of these two methods is applicable, the bounds derived in this paper for the sparsity may be used for setting k . From now on, whenever we refer to IHT or IST, the thresholding policy is the k largest element thresholding policy unless otherwise stated.

C. Main Results

We will prove two main theorems for the two thresholding algorithms that have been mentioned in the last section.

Theorem II.2. *Suppose that $k < \frac{1}{3.1} \mu^{-1}$ and $\frac{|x_o(i)|}{|x_o(i+1)|} < 3^{\ell_i-4}, \forall i, 1 \leq i < k$. Then IHT finds the correct active set in at most $\sum_{i=1}^k \ell_i + k$ steps. After this step all of these elements will remain in the active set and the error will go to zero exponentially fast.*

Theorem II.3. *Suppose that $k < \frac{1}{4.1} \mu^{-1}$ and $\forall i, 1 \leq i < k$, we have $\frac{|x_o(i)|}{|x_o(i+1)|} < 2^{\ell_i-5}$. Then IST recovers the correct active set in at most $\sum_{i=1}^k \ell_i + k$ steps. After that all these coefficients will remain in the active set and the error will go to zero exponentially fast.*

The sufficient conditions provided here are slightly weaker than the conditions mentioned for ℓ_1 or OMP. Simulation results also confirm that IHT and IST are weaker than ℓ_1 in practice [3]. Another interesting fact is that the number of iterations needed, depends on the ratio of the coefficients but this dependency is roughly logarithmic and therefore it will work well in practice. Also, the algorithms find the correct active set in a finite number of iterations and once the algorithms find the correct active set, they converge to the exact solution exponentially fast.

III. PROOF OF CONVERGENCE FOR THE IHT ALGORITHM

The goal of this section is to give an outline of the proof of Theorem II.2. We define the following two variables,

$$z^i = x^i + \Phi^T(\Phi x_o - \Phi x^i), \quad (4)$$

$$w^i = x_o - x^i, \quad (5)$$

where x_o is the optimal value and x^i is our estimate at the i^{th} step. The j^{th} element of these two vectors will be denoted by $z^i(j)$ and $w^i(j)$. The active set of x^i is called I^i . Finally, $x_o(i)$ denotes the i^{th} element of x_o . Without loss of generality we assume that $x_o(i)$'s are sorted in descending order of their absolute values and therefore the only non-zero elements of x_o are the first k elements. The next lemma will be useful later when we try to bound the error at each iteration.

Lemma III.1. *Consider the following sequence for $s \geq 0$,*

$$f_s = \alpha^1 + \dots + \alpha^s + \beta \alpha^{s+1},$$

where $0 < \alpha < 1$. The following statements are true;

- 1) If $\beta(1 - \alpha) < 1$, then for every s , $f_s < \frac{\alpha}{1-\alpha}$.

- 2) If $\beta(1 - \alpha) > 1$, then for every s , $f_s < \beta\alpha$.
 3) If $\beta(1 - \alpha) = 1$, then f_s is a constant sequence and is always equal to $\frac{\alpha}{1-\alpha}$.

It is easy to see that the sequence is either increasing or decreasing or constant depending on the values of α and β . The proof is simple and is omitted for the sake of brevity.

Lemma III.2. Suppose that $x_o(1), x_o(2), \dots, x_o(r-1)$, $r-1 < k$, are in the active set at the m^{th} step. Also assume that,

$$|z^m(j) - x_o(j)| \leq 1.5k\mu|x_o(r-1)| \quad \forall j.$$

If $k\mu < \frac{1}{3.1}$, then at stage $m+s$ and for every j we will have the following upper bound for $|z^{m+s}(j) - x_o(j)|$,

$$|x_o(r)|(k\mu + \dots + (k\mu)^s) + 1.5(k\mu)^{s+1}|x_o(r-1)|. \quad (6)$$

Moreover, $x_o(1), x_o(2), \dots, x_o(r-1)$ will remain in the active.

Proof: We prove this by induction; Assuming that the bound holds at stage $m+s$ and $x_o(1), x_o(2), \dots, x_o(r-1)$ are in the active set, we show that the upper bound holds at stage $m+s+1$ and the first $r-1$ elements will remain in the active set.

$$\begin{aligned} & |z^{m+s+1}(i) - x_o(i)| \\ & \leq \sum_{j \in I^{m+s} \setminus \{i\}} |\langle \phi_i, \phi_j \rangle w^{m+s}(j)| + \sum_{j \in \{1, 2, \dots, k\} \setminus I^{m+s} \cup \{i\}} |\langle \phi_i, \phi_j \rangle w^{m+s}(j)|, \\ & \stackrel{1}{=} \sum_{j \in I^{m+s} \setminus \{i\}} |\langle \phi_i, \phi_j \rangle w^{m+s}(j)| + \sum_{j \in \{r, \dots, k\} \setminus I^{m+s} \cup \{i\}} |\langle \phi_i, \phi_j \rangle w^{m+s}(j)|, \\ & \stackrel{2}{\leq} \sum_{j \in I^{m+s} \setminus \{i\}} |\langle \phi_i, \phi_j \rangle (z^{m+s}(j) - x_o(j))| + k\mu x_o(r), \\ & \leq k\mu|x_o(r)|(k\mu + \dots + (k\mu)^s) + 1.5(k\mu)^{s+2}|x_o(r-1)| \\ & \quad + k\mu|x_o(r)|, \\ & \leq |x_o(r)|(k\mu + \dots + (k\mu)^{s+1}) + 1.5(k\mu)^{s+2}|x_o(r-1)|. \end{aligned}$$

In these calculations equality (1) is due to the assumptions of the induction, i.e. the first $r-1$ elements are in the active set at stage $m+s$. To get inequality (2) we have used two different facts. The first one is that when $j \in I^{m+s}$, $w^{m+s}(j) = x_o(j) - z^{m+s}(j)$ and the second one is that when $j \in \{r, \dots, k\} \setminus I^{m+s}$ then $w^{m+s}(j) = x_o(j)$ and therefore $|x_o(j)| \leq |x_o(r)|$. The last step is to prove that all the first $r-1$ elements remain in the active set. For $i \in \{1, 2, \dots, r-1\}$,

$$\begin{aligned} & |z^{m+s+1}(i)| \geq |x_o(i)| - |z^{m+s+1}(i) - x_o(i)|, \\ & \stackrel{1}{\geq} |x_o(i)| - (k\mu|x_o(r-1)| + \dots + (k\mu)^{s+1}|x_o(r-1)|) \\ & \quad - 1.5(k\mu)^{s+2}|x_o(r-1)| \stackrel{2}{\geq} |x_o(i)| - \frac{|x_o(r-1)|}{2.05} \\ & \geq |x_o(r-1)| - \frac{|x_o(r-1)|}{2.05}. \end{aligned}$$

In inequality (1) we have used the bound in (6) by replacing $x_o(r)$ with $x_o(r-1)$. Inequality (2) is the result of Lemma III.1. For $i \notin \{1, 2, \dots, k\}$, we have

$$|z^{m+s+1}(i)| \leq \frac{|x_o(r-1)|}{2.05},$$

and since $\min_{\{i:i \leq r-1\}} |z^{m+s+1}(i)| > \max_{\{i:i > k\}} |z^{m+s+1}(i)|$, the first $r-1$ elements will remain in the active set. The base of the induction is the same as the assumptions of this lemma and the proof is complete. ■

Lemma III.3. Suppose that $k < \frac{1}{3.1}\mu^{-1}$, and $x_o(1), x_o(2), \dots, x_o(r)$, $r < k$, are in the active set at the m^{th} step. Also assume that $\frac{|x_o(r)|}{|x_o(r+1)|} \leq 3^{\ell_r - 4}$. If

$$|z^m(j) - x_o(j)| \leq 1.5k\mu|x_o(r)| \quad \forall j,$$

after ℓ_r more steps $x_o(r+1)$ will get into the active set, and

$$|z^{m+\ell_r+1}(j) - x_o(j)| \leq 1.5k\mu|x_o(r+1)| \quad \forall j.$$

Proof: By setting $q = \ell_r$ in the upper bound we get,

$$|z^{m+\ell_r}(j) - x_o(j)| \leq \frac{1.5|x_o(r+1)|}{273} + \frac{|x_o(r+1)|}{2.1}.$$

Similar to the last lemma it is also not difficult to see that

$$\begin{aligned} |z^{m+\ell_r}(r+1)| &= |z^{m+\ell_r}(r+1) - x_o(r+1) + x_o(r+1)| \\ &\geq |x_o(r+1)| - |z^{m+\ell_r}(r+1) - x_o(r+1)| \\ &\geq |x_o(r+1)| - \frac{|1.5x_o(r+1)|}{273} - \frac{|x_o(r+1)|}{2.1}. \end{aligned}$$

But,

$$|z^{m+\ell_r}(r+1)| > \max_{\{i:i > k\}} |z^{m+\ell_r}(i)|,$$

and therefore $x_o(r+1)$ will be detected at this step. It may also be noted that at this stage the error is less than $|x_o(r+1)|/2$. For the next stage we will have at most k active elements the error of each is less than $|x_o(r+1)|/2$ and at most $k-r$ non-zero elements of x_o that have not passed the threshold and whose magnitudes are smaller than $|x_o(r+1)|$. Therefore, the error of the next step is less than $1.5k\mu|x_o(r+1)|$. ■

Our goal is to prove the correctness of IHT by induction and we have to know the correctness of IHT at the first stage. The following lemma provides this missing step.

Lemma III.4. Suppose that $k < \frac{1}{3.1}\mu^{-1}$, then at the first stage of the IHT, $x_o(1)$ will be in the active set¹ and $|z^1(j) - x_o(j)| \leq k\mu|x_o(1)|$.

Proof:

$$|z^1(1)| \geq |x_o(1)| - k\mu|x_o(1)|.$$

On the other,

$$\max_{\{i:k < i\}} |z^1(i)| = \max_{\{i:k < i\}} \left| \sum_{j=1}^k \langle \phi_i, \phi_j \rangle x_o(j) \right| \leq k\mu|x_o(1)|.$$

Therefore, since $k\mu < 1 - k\mu$, the index of the first element will be in the active set after the first step. The last claim of the Lemma is also clear. ■

Finally the following lemma describes the performance of the algorithm after detecting all the non-zero elements.

¹This result holds even if $k\mu < \frac{1}{2}$. For the sake of consistency with the other parts of the proof we state it in this way

Lemma III.5. *Suppose that $x_o(1), x_o(2), \dots, x_o(k)$, are in the active set at the m^{th} step. Also assume that,*

$$|z^m(j) - x_o(j)| \leq 1.5k\mu|x_o(k)| \quad \forall j.$$

If $k\mu < \frac{1}{3.1}$, then at stage $m+s$ and for every j we will have,

$$|z^{m+s}(j) - x_o(j)| \leq 1.5(k\mu)^{s+1}|x_o(k)|.$$

Since the proof of this lemma is very similar to the proof of Lemma III.2, it is omitted.

Outline of the proof of Theorem II.2: The proof is an induction that combines the above lemmas. Suppose that $x_o(1), x_o(2), \dots, x_o(r)$ are already in the active set. According to Lemma III.2 all these terms will remain in the active set, and according to Lemma III.3 after ℓ_r steps $x_o(r+1)$ will also get into the active set. In one more step, the error on each element gets smaller than $1.5k\mu|x_o(r+1)|$, and everything can be repeated. Lemma III.4 provides the first step of the induction. Finally when all the elements are in the active set lemma III.5 tells us that the error goes to zero exponentially fast. ■

Since the proof of the convergence of IST is very similar to IHT we do not repeat it here. You may refer to [16] for more details.

IV. PROOF OF CONVERGENCE FOR THE IST ALGORITHM

As mentioned before the main ideas of the proof of the IST algorithm are very similar to those of the IHT. We will mention the proof in detail but will try to emphasize more on the differences. The following lemma helps us find some bounds on the error of the algorithm at each step.

Lemma IV.1. *Suppose that $x_o(1), x_o(2), \dots, x_o(r)$, $r \leq k$, are in the active set at the m^{th} step. Also assume that*

$$|x^m(j) - x_o(j)| \leq 4k\mu|x_o(r)|, \quad \forall j \in I^m,$$

and $k\mu < \frac{1}{4.1}$. Then at stage $m+s$, $\forall i \in I^{m+s}$ we have the following upper bound for $|x^{m+s}(i) - x_o(i)|$,

$$|x_o(r+1)|(2k\mu + \dots + (2k\mu)^s) + 2(2k\mu)^{s+1}|x_o(r)|.$$

Moreover, $x_o(1), x_o(2), \dots, x_o(r)$ remain in the active set.

Proof: As before, this can be proved by induction. We assume that at step $m+s$ the upper bound holds and $x_o(1), x_o(2), \dots, x_o(r)$ are in the active set and we prove the same things for $m+s+1$. Similar to what we saw before,

$$\begin{aligned} & |z^{m+s+1}(i) - x_o(i)| \\ & \leq \sum_{j \in I^{m+s} \setminus \{i\}} |\langle \phi_i, \phi_j \rangle w^{m+s}(j)| + \sum_{j \in \{1, 2, \dots, k\} \setminus I^{m+s} \cup \{i\}} |\langle \phi_i, \phi_j \rangle w^{m+s}(j)|, \\ & \frac{1}{2} \sum_{j \in I^{m+s} \setminus \{i\}} |\langle \phi_i, \phi_j \rangle w^{m+s}(j)| + \sum_{j \in \{r+1, \dots, k\} \setminus I^{m+s} \cup \{i\}} |\langle \phi_i, \phi_j \rangle w^{m+s}(j)|, \\ & \stackrel{2}{\leq} (k-1)\mu(2k\mu|x_o(r+1)| + \dots + (2k\mu)^s|x_o(r+1)| \\ & + 2(2k\mu)^{s+1}|x_o(r)|) + k\mu|x_o(r+1)| := \alpha_s. \end{aligned}$$

Equality (1) is using the assumption that the first r elements are in the active set at stage $m+s$. Inequality (2) is also

due to the assumptions of the induction and the fact that $w^{m+s}(j) = x_o(j) - x^{m+s}(j)$.

At least one of the largest $k+1$ coefficients of z , corresponds to an element whose index is not in $\{1, 2, \dots, k\}$, and the magnitude of this coefficient is less than α_s . Therefore the threshold value is less than or equal to α_s . Applying the soft thresholding to z will at most add α_s to the distance of $z^{s+1}(i)$ and $x_o(i)$, and this completes the proof of the upper bound. The main thing that should be checked is whether the first r elements will remain in the active set or not. For $i \in \{1, 2, \dots, r\}$ we have,

$$\begin{aligned} & |z^{m+s+1}(i)| \geq |x_o(i)| - |z^{m+s+1}(i) - x_o(i)|, \\ & \geq |x_o(i)| - k\mu|x_o(r)|(1 + 2k\mu + \dots + (2k\mu)^{s+1}) \\ & \quad - 2k\mu(2k\mu)^{s+1}|x_o(r)| \geq |x_o(i)| - \frac{|x_o(r)|}{2.05} \\ & \geq |x_o(r)| - \frac{|x_o(r)|}{2.05}. \end{aligned} \quad (7)$$

If the sequence in the above expression is multiplied by 2, the result will be a sequence in the form of the sequences mentioned in lemma III.1 for $\alpha = 2k\mu$, $\beta = 2$ and the last equality is based on that lemma.

If $i \notin \{1, 2, \dots, k\}$,

$$\begin{aligned} & |z^{m+s+1}(i)| \leq k\mu|x_o(r)|(1 + 2k\mu + \dots + (2k\mu)^{s+1}) \\ & \quad + 2k\mu(2k\mu)^{s+1}|x_o(r)| \leq \frac{|x_o(r)|}{2.05}. \end{aligned}$$

Since $\min_{\{i:i \leq r\}} |z^{m+s+1}(i)| > \max_{\{i:i > k\}} |z^{m+s+1}(i)|$, the first r elements remain in the active set. The base of the induction is also clear since it is the same as the assumptions of the lemma. ■

Lemma IV.2. *Suppose that $k \leq \frac{\mu^{-1}}{4.1}$, and $x_o(1), x_o(2), \dots, x_o(r)$, $r \leq k$, are in the active set at the m^{th} step. Also, assume that $\frac{|x_o(r)|}{|x_o(r+1)|} \leq 2^{\ell_r - 5}$. If*

$$|x^m(j) - x_o(j)| \leq 4k\mu|x_o(r)|, \quad \forall j \in I^m,$$

then after ℓ_r steps $x_o(r+1)$ will get into the active set, and

$$|x^{m+\ell_r+1}(j) - x_o(j)| \leq 4k\mu|x_o(r+1)|, \quad \forall j \in I^{m+\ell_r+1}.$$

Proof: As before we try to find a bound for the error at time $m+\ell_r$. For $i \in \{1, 2, \dots, k\}$,

$$\begin{aligned} & |z^{m+\ell_r}(i) - x_o(i)| \leq \frac{1}{2}|x_o(r+1)|(2k\mu + \dots + (2k\mu)^{\ell_r}) \\ & \quad + (2k\mu)^{\ell_r+1}|x_o(r)| \leq \frac{|x_o(r+1)|}{2.1} + \frac{|x_o(r+1)|}{64} \end{aligned}$$

and therefore for $i = r+1$,

$$\begin{aligned} & |z^{m+\ell_r}(r+1)| \geq |x_o(r+1)| - |z^{m+\ell_r}(i) - x_o(i)| \geq \\ & |x_o(r+1)| - \frac{|x_o(r+1)|}{2.1} - \frac{|x_o(r+1)|}{64} \end{aligned} \quad (8)$$

Since $|z^{m+\ell_r}(r+1)| > \max_{\{i:k < i\}} |z^{m+\ell_r}(i)|$, the $r+1^{\text{th}}$ element will get into the active set at this stage. On the other hand for any $i \in I^{m+\ell_r}$ we have $|x^{m+\ell_r}(i) - x_o(i)| \leq x_o(r+1)$

1). For the next stage of the algorithm we will have at most $2k$ non-zero $x^{m+\ell_r}(i) - x_o(i)$ and absolute value of each of them is less than $|x_o(r+1)|$. Therefore $|z^{m+\ell_r+1}(i) - x_o(i)| \leq 2k\mu|x_o(r+1)|$ and after thresholding we have, $|x^{m+\ell_r+1}(i) - x_o(i)| \leq 4k\mu|x_o(r+1)|$ for $i \in I^{m+\ell_r+1}$.

The base of the induction is also clear from the assumptions of this lemma and the proof is complete. ■

For the IHT algorithm we proved that at the first step the first element will pass the threshold. Since the selection step of IST and IHT is exactly the same, we can claim that the same thing is true for IST, i.e. the largest magnitude coefficient will pass the threshold. Also, as we saw for IHT, the error was less than $k\mu|x_o(1)|$. Therefore, for the IST we have, $|x^1(j) - x_o(j)| < 2k\mu|x_o(1)|$. These bounds are even better than the bounds we need for IV.1 and IV.2 and IV.3.

The following lemma will explain what happens when the algorithm detects all the non-zero elements.

Lemma IV.3. *Suppose that $x_o(1), \dots, x_o(k)$, are in the active set at the m^{th} step. Also assume that,*

$$|x^m(j) - x_o(j)| \leq 4k\mu|x_o(k)|.$$

If $k\mu < \frac{1}{4.1}$, at stage $m+s$ all the elements remain in the active set and for every j we will have,

$$|z^{m+s}(j) - x_o(j)| \leq 2(2k\mu)^{s+1}|x_o(k)|$$

The proof of this lemma is very similar to the other lemmas and is omitted.

Outline of the proof of Theorem II.3: The proof is a simple induction by combining the above lemmas. Suppose that $x_o(1), x_o(2), \dots, x_o(r)$ are already in the active set. According to Lemma IV.1 all these terms will remain in the active set, and according to Lemma IV.2 after ℓ_r steps $x_o(r+1)$ will also get into the active set. In one more step, the error on each element gets smaller than $4k\mu|x_o(r+1)|$, and everything can be repeated. Although we have not mentioned the first step of the induction it is not difficult to see that step is also true and it is very similar to the first step of IHT. Finally when all the elements are in the active set lemma IV.3 tells us that the error goes to zero exponentially fast. ■

V. DISCUSSION AND COMPARISON WITH OTHER WORK

There is a huge amount of work on iterative thresholding algorithms, and we cannot mention all of them here; The interested reader is referred to [3]. Most of these papers are dealing with a fixed threshold that does not depend on iteration. In that case, there are rigorous results that give sufficient conditions for the IST algorithm to converge to the solution of \mathcal{P}_1 [12], and for the IHT algorithm to a local minimum of \mathcal{P}_0 [17]. The idea of choosing iteration dependent thresholds is also not new, and some simple variations were introduced in [11]. The k largest element thresholding policy was first introduced in [13] and was first used for IHT in [15]. It was also shown that if the Φ matrix satisfies restricted isometry property (RIP) of order $3k$, the IHT converges to the sparsest solution. There are some basic differences in our approach. First, we are dealing with deterministic settings, and in these settings RIP conditions they have provided are

much weaker than ours ($k\mu < \frac{1}{3\sqrt{32}}$ compared to $k\mu < \frac{1}{3.1}$). Under these more general conditions, as we observed, the performance of IHT is not as simple as what is mentioned in [15], and it may not recover x_o in just k steps. But it will finally recover the sparsest signal and we give bounds on the number of iterations it needs to converge. Secondly, as discussed in the last section, our approach was easily adapted to IST, and can be adapted to the other types of thresholds. Moreover, our method gives us an ordering among ℓ_1 , OMP, IHT and IST which may be useful for deciding on the choice of the algorithm. Finally there is another effort on analyzing the performance of IST by coherence that shows the possibility of success of such an algorithm at the first iteration [18]. But this result does not have any conclusion about the next iterations of IST in case it does not recover all the non-zero elements at the first step.

VI. CONCLUSION

In this paper, we analyzed iterative hard and soft thresholding, and proved that under certain conditions they work properly. These conditions are slightly weaker than their counterparts for ℓ_1 and OMP. But these algorithms are very simple to implement and much faster than both convex relaxation and greedy methods, and they are much more desirable for large scale problems.

VII. ACKNOWLEDGEMENT

The author would like to thank David L. Donoho for helpful discussion and valuable suggestions on the early version of this manuscript. This work was partially supported by NSF DMS 05-05303.

REFERENCES

- [1] S.S. Chen, D.L. Donoho and M.A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, Vol. 20, pp. 33-61, 1998.
- [2] Y. C. Pati, R. Rezaifar, P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Asilomar Conference on Signals, Systems and Computers*, A. Singh, ed., *IEEE Comput. Soc. Press*, Los Alamitos, CA, 1993.
- [3] A. Maleki, D. L. Donoho, "Optimally Tuned Iterative Thresholding Algorithms," submitted to *IEEE journal on selected areas in signal processing*, 2009.
- [4] D. L. Donoho, M. Elad, "Maximal sparsity representation via minimization," *Proc. Natl. Acad. Sci.*, vol. 100, pp. 2197-2202, Mar. 2003.
- [5] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Info. Theory*, vol. 50, num. 10, pp. 2231-2242, Oct. 2004.
- [6] D. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, Vol. 52, pp. 489-509, April 2006.
- [7] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Information Theory*, Vol. 52(2), pp. 489-509, February 2006.
- [8] Joel A. Tropp, Anna C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Info. Theory* 53(12), pp. 4655-4666, 2007
- [9] D. L. Donoho, J. Tanner, "Phase transitions as 'sparse sampling theorems' " submitted to *IEEE Trans. on Information Theory*
- [10] M. Figueiredo and R. Nowak, "An EM Algorithm for Wavelet-Based Image Restoration," *IEEE Transactions on Image Processing*, Vol.12, no.8, pp. 906-916, August 2003.

- [11] J.L. Starck, M. Elad, and D.L. Donoho, "Image decomposition via the combination of sparse representations and a variational approach", *IEEE Trans. On Image Processing*, Vol. 14, No. 10, pp. 1570-1582, October 2005.
- [12] I. Daubechies, M. Defrise and C. De Mol "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, Vol. 75, pp. 1412-1457, 2004.
- [13] D. Needel, J. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," Accepted to *Appl. Comp. Harmonic Anal*, 2008.
- [14] W. Dai, O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction", submitted to *IEEE Transactions on Information Theory*, 2009.
- [15] T. Blumensath, M. E. Davies, "Iterative hard thresholding for compressed sensing," arXiv:0805.0510v1.
- [16] A. Maleki, "Coherence Analysis of Iterative Thresholding Algorithms," Technical Report, Department of Statistics, Stanford University, 2009.
- [17] T. Blumensath and M. Davies, "Iterative thresholding for sparse approximations," to appear in *Journal of Fourier Analysis and Applications*, special issue on sparsity, 2008.
- [18] K. K. Herrity, A. C. Gilbert, and J. A. Tropp, "Sparse approximation via iterative thresholding," *Proc. ICASSP*, Vol. 3, pp. 624-627, Toulouse, May 2006.