*Progress In Electromagnetics Research Letters, Vol. 8, 73–82, 2009*

# A THRESHOLDED LANDWEBER ITERATION BASED ON SENSING DICTIONARY

**A. Huang, Q. Wan, G. Gui, and W. Yang**

School of Electronic Engineering
University of Electronic Science and Technology of China (UESTC)
No. 4, The 2nd Section of North Jianshe Road, Chengdu, China

**Abstract**—Thresholded Landweber Iteration (TLI) is an attractive algorithm since it has the advantage of simplicity for the problem of sparse reconstruction. However, this algorithm depends heavily on the coherence property of the redundant dictionary, and its convergence rate is slow. In this paper, we develop a modified version of TLI by using a sensing dictionary. The proposed algorithm significantly improves the reconstruction performance and the convergence properties when compared to the classical TLI. We provide a sufficient condition for which the modified TLI algorithm can be guaranteed to exactly identify the correct atoms and also discuss the convergence properties for this algorithm. Finally, simulation results are presented to demonstrate the superior performance of the proposed algorithm.

## 1. INTRODUCTION

In the last decades, sparse signal expansion over overcomplete or redundant dictionaries has attracted a lot of attention, such as image reconstruction [1], sources localization [2] and many more. The goal of this problem is to represent a signal as a linear combination of a few elementary waveforms, termed as atoms, selected from a large collection of basis functions. Unfortunately, this problem is known to be NP-hard. In order to get the sparsest solution to this problem, some suboptimal algorithms have been developed. One such algorithm is orthogonal matching pursuit (OMP) [3]. This algorithm is easy to implement since it provides sparse solution through iteratively selecting

Corresponding author: A. Huang (huanganmin@uestc.edu.cn).

an atom at each step. However, OMP is too computationally complex when the sparse reconstruction with large scale is considered.

Recently, thresholded Landweber iteration (TLI), also named as iterative shrinkage thresholding (IST), has been proposed as a simple method to solve this linear inverse problem [4]. This algorithm combines the Landweber iteration with thresholding technique. It has been developed independently through various techniques, such as surrogate-functions [4], marjorization-minimization (MM) algorithm [5], expectation-maximization (EM) algorithm [6] and fixed point strategy [7]. The best advantage of TLI algorithm is its simplicity since this algorithm only requires matrix-vector multiplication at each iteration. However, the performance of this algorithm depends heavily on the coherence property of the redundant dictionary, and the identification of correct atoms can be guaranteed only under very strict conditions [8]. In most situations, redundant dictionaries for the decomposition of practical signals may be highly coherent, and the reconstruction performance decays greatly in these cases. Additionally, the classical TLI is known to converge slowly.

The primary purpose of this paper is to develop a variation of TLI to improve the performance of the classical TLI algorithm. Motivated by the work of Schnass and Vandergheynst [9], we modify the classical TLI algorithm by introducing a sensing dictionary. The proposed algorithm keeps computational complexity similar to the classical TLI while significantly improving the reconstruction performance and the convergence properties, especially in the case of highly coherent dictionaries. Theoretical analysis and simulation results are provided in this paper.

The remaining sections are organized as follows. In Section 2, we formulate the classical TLI algorithm. Section 3 develops a modified TLI algorithm by introducing a sensing dictionary and provides a sufficient condition for this algorithm. Convergence analysis for the modified TLI algorithm is discussed in Section 4. And simulation results are presented in Section 5 to illustrate the improved performance of the proposed algorithm. Finally, conclusions are given in Section 6.

## 2. CLASSICAL THRESHOLDED LANDWEBER ITERATION

Consider a redundant dictionary $\Phi$ composed of $N$ vectors $\phi_i \in \mathbb{R}^{M\times 1}$ $(i \in \Omega = \{1, \ldots, N\}, M < N)$ from a Hilbert space. In general, these vectors are normalized and called atoms. Assume a signal $y \in \mathbb{R}^{M\times 1}$ can be exactly represented as a linear combination of a small number of atoms in this dictionary, i.e., $y = \Phi x = \Phi_{opt} x_{opt}$, where $\Phi_{opt}$ is the matrix containing the optimal collection of atoms with index set $\Lambda_{opt}$. The signal is called $K$-exact sparsity if $|\Lambda_{opt}| = K$, where $K \ll N$ and $|\cdot|$ returns the cardinality of a set.

The classical Landweber iteration is one of the simplest methods to solve the linear inverse problem $y = \Phi x$. It generates a sequence to approximate the true solution. The iterative procedure can be described as

$$x^{(l+1)} = x^{(l)} + \Phi^T \left( y - \Phi x^{(l)} \right), \tag{1}$$

where $T$ denotes transpose. Using the thresholding operation at each iteration step, we can get the TLI as

$$x^{(l+1)} = H \left( x^{(l)} + \Phi^T \left( y - \Phi x^{(l)} \right) \right), \tag{2}$$

where $H(\cdot)$ is thresholding operators.

For a $K$-exact sparsity signal $y$, the $K$-term sparse problem can be formulated as the following constrained optimization

$$y = \Phi x, \quad s.t. \quad \|x\|_0 = K, \tag{3}$$

where $\|\cdot\|_0$ denotes the $l_0$ quasi-norm which counts the number of non-zero coefficients. Using a hard thresholding operator, Thrumber and Davis derived the following TLI algorithm [1]

$$x^{(l+1)} = H_K \left( x^{(l)} + \Phi^T \left( y - \Phi x^{(l)} \right) \right), \tag{4}$$

where $H_K(x)$ is element-wise hard thresholding operator defined as

$$H_K(x_i) = \begin{cases} x_i, & |x_i| \geq \tilde{x}_K \\ 0, & |x_i| < \tilde{x}_K \end{cases}, \tag{5}$$

where $\tilde{x}_K$ denotes the $K$-th largest absolute value of the coefficient vector $x$. In other words, $H_K(x)$ is nonlinear thresholding operator which only retains $K$ coefficients with the largest magnitude. Starting from an initialization $x^{(0)}$, this iterative algorithm can generate a sequential estimation to approximate the solution.

## 3. A THRESHOLDED LANDWEBER ITERATION BASED ON SENSING DICTIONARY

As described in the last section, the classical TLI algorithm updates the coefficients through calculating the correlation between atoms and the residuals. When the dictionary is highly coherent, the reconstruction performance will be deteriorated due to the interference between atoms. Here, we modify the classical TLI algorithm by using a sensing dictionary.

Assume that we can construct a lower coherent dictionary $D = B\Phi$ by using an invertible matrix $B \in \mathbb{R}^{M\times M}$. With the matrix $B$ and the observation data $y$, we can get a new vector $z = By = Dx$. Therefore, the sparse solution can be obtained through equivalently solving the new optimal problem

$$z = Dx, \quad s.t. \quad \|x\|_0 = K. \tag{6}$$

Similarly, we can derive a hard TLI algorithm for $K$-term sparse problem as

$$x^{(l+1)} = H_K\left(x^{(l)} + D^T\left(z - Dx^{(l)}\right)\right). \tag{7}$$

Define the sensing dictionary as $\Psi = B^T B\Phi$, it is easy to obtain $D^T z = \Psi^T y$ and $D^T D = \Psi^T \Phi$. Substituting these relations to (7), we can get a variation of the classical TLI as

$$x^{(l+1)} = H_K\left(x^{(l)} + \Psi^T\left(y - \Phi x^{(l)}\right)\right). \tag{8}$$

This algorithm extends the classical TLI to a more general case and relaxes the strong requirement of incoherence for redundant dictionary. The iteration procedure (8) reduces to the classical TLI if the sensing dictionary is selected as $\Psi = \Phi$ for incoherent dictionary.

Now, we provide a sufficient condition for which the modified TLI can be guaranteed to recover the $K$-term sparse signal. In order to simplify the presentation, the cross cumulative coherence of redundant dictionary is defined as [9]

$$\tilde{\mu}_1(k) = \max_{|J|=k} \max_{i\notin J} \sum_{j\in J} |\langle \psi_i, \phi_j\rangle|, \tag{9}$$

where $\psi_i (i = 1, \ldots, N)$ denote the atoms of the sensing dictionary $\Psi$. The above definition reduces to the cumulative coherence $\mu_1(k)$ if $\Psi = \Phi$ [8]. In this paper, we are concerned with 0–1 sparse signal of which the coefficient vector contains only ones and zeros.

Theorem 1: Assume that a signal can be represented as $y = \Phi_{opt} x_{opt}$ with $|\Lambda_{opt}| = K$ and the nonzero entries of the coefficient vector are ones. With the initialization $x^{(0)} = 0$, the modified TLI based on a sensing dictionary $\Psi$ can identify correct atoms if

$$\tilde{\mu}_1(k) + \tilde{\mu}_1(k-1) < 1. \tag{10}$$

Proof: With the initialization $x^{(0)} = 0$, the modified TLI algorithm at the first step is $x^{(1)} = H_K(\Psi^T y)$. It is easy to bound the inner product of atoms $\psi_i$ and the signal $y$ as

$$i \in \Lambda_{opt}, \quad |\langle \psi_i, y \rangle| = \left| \sum_{j \in \Lambda_{opt}} \langle \psi_i, \phi_j \rangle \right| \geq |\langle \psi_i, \phi_i \rangle| - \sum_{\substack{j \in \Lambda_{opt} \\ j \neq i}} |\langle \psi_i, \phi_j \rangle| \geq 1 - \tilde{\mu}_1(k-1), \tag{11}$$

and

$$i \in \bar{\Lambda}_{opt}, \quad |\langle \psi_i, y \rangle| = \left| \sum_{j \in \Lambda_{opt}} \langle \psi_i, \phi_j \rangle \right| \leq \sum_{j \in \Lambda_{opt}} |\langle \psi_i, \phi_j \rangle| \leq \tilde{\mu}_1(k). \tag{12}$$

If $\tilde{\mu}_1(k) + \tilde{\mu}_1(k-1) < 1$, we have $\tilde{\mu}_1(k) < 1 - \tilde{\mu}_1(k-1)$, and the inner product corresponding to $\Lambda_{opt}$ is larger than others. Therefore, this algorithm can identify the correct atoms by only using one iteration.

The inequality (10) shows that this proposed algorithm can be guaranteed to identify correct atoms when the cross cumulative coherence grows slowly. This result extends previous results by Tropp [8] to the less restrictive condition. If we can construct a sensing dictionary satisfying $\tilde{\mu}_1(k) < \mu_1(k)$, the modified TLI can be more easily guaranteed to identify the correct atoms.

The remaining problem is how to design an effective sensing dictionary. Schnass and Vandergheynst developed a method for designing sensing dictionary based on alternating projection (AP) algorithm [9]. Indeed, an appropriate sensing dictionary for sparse reconstruction is the one of which the gram type matrix $G = \Psi^* \Phi$ is closest to unity matrix as possible, where $*$ represents the complex conjugate transpose. Therefore, we construct a sensing dictionary through solving the following optimal problem

$$\min_{\Psi \in \mathbb{C}^{M \times N}} \|\Psi^* \Phi\|_F^2, \quad s.t. \quad \psi_i^* \phi_i = 1, \quad \text{for} \quad i \in \Omega, \tag{13}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Based on the definition of the matrix norm, we can get

$$\min_{\Psi \in \mathbb{C}^{M \times N}} \|\Psi^* \Phi\|_F^2 = \min_{\psi_i \in \mathbb{C}^{M \times 1}} \sum_{i=1}^{N} \|\Phi^* \psi_i\|_2^2. \tag{14}$$

With (14), the optimization problem (13) can be transformed to minimize the following objective function

$$J(\Psi, \lambda) = \sum_{i=1}^{N} \left[ \frac{1}{2} \left\| \Phi^* \psi_i \right\|_2^2 + \lambda_i (1 - \psi_i^* \phi_i) \right], \tag{15}$$

where $\lambda = [\lambda_1, \ldots, \lambda_N]^T$ is the vector composed of the Lagrange multipliers. The necessary condition for $(\psi_i, \lambda_i)(i = 1, \ldots, N)$ to be the minimizing solutions is that the partial differential of $J(\Psi, \lambda_i)$ at $(\psi_i, \lambda_i)$ satisfies

$$\frac{\partial J(\Psi, \lambda)}{\partial \psi_i} = \Phi\Phi^* \psi_i - \lambda_i \phi_i = 0, \quad i = 1, \ldots, N, \tag{16}$$

and

$$\frac{\partial J(\Psi, \lambda)}{\partial \lambda_i} = 1 - \psi_i^* \phi_i = 0, \quad i = 1, \ldots, N. \tag{17}$$

Combine (16) with (17), we can obtain the column vectors of the sensing dictionary as

$$\psi_i = \frac{R^{-1} \phi_i}{\phi_i^* R^{-1} \phi_i}, \quad i = 1, \ldots, N, \tag{18}$$

where $R = \Phi\Phi^*$.

## 4. CONVERGENCE ANALYSIS

We have shown that the modified TLI can identify the correct atoms if $\tilde{\mu}_1(k) + \tilde{\mu}_1(k-1) < 1$ holds. In this section, we will provide theoretical analysis for the convergence properties of the modified TLI algorithm. Theorem 2: Assume that a signal can be represented as $y = \Phi_{opt} x_{opt}$ with $|\Lambda_{opt}| = K$ and the nonzero entries of the coefficient vector are ones. With the initialization $x^{(0)} = 0$, the TLI algorithm based on a sensing dictionary $\Psi$ is used to recover this sparse signal. If $\tilde{\mu}_1(k) + \tilde{\mu}_1(k-1) < 1$, the coefficients obtained by this algorithm at $n$-th iteration step satisfy

$$1 - (\tilde{\mu}_1(k-1))^n \leq x_i^{(n)} \leq 1 + (\tilde{\mu}_1(k-1))^n, \quad \text{for} \quad i \in \Lambda_{opt}. \tag{19}$$

Proof: With the initialization $x^{(0)} = 0$, after one iteration we can get

$$\begin{aligned} x_i^{(1)} = \langle \psi_i, y \rangle &= \sum_{j \in \Lambda_{opt}} \langle \psi_i, \phi_j \rangle \\ &= \langle \psi_i, \phi_i \rangle + \sum_{\substack{j \in \Lambda_{opt} \\ j \neq i}} \langle \psi_i, \phi_j \rangle, \quad \text{for} \quad i \in \Lambda_{opt}. \end{aligned} \tag{20}$$

With (9) and (13), we have

$$1 - \tilde{\mu}_1(k-1) \leq x_i^{(1)} \leq 1 + \tilde{\mu}_1(k-1), \quad \text{for} \quad i \in \Lambda_{opt}. \tag{21}$$

Assume that the nonzero coefficients at $l$-th iteration satisfies

$$1 - (\tilde{\mu}_1(k-1))^l \leq x_i^{(l)} \leq 1 + (\tilde{\mu}_1(k-1))^l, \quad \text{for} \quad i \in \Lambda_{opt}. \tag{22}$$

Theorem 1 has proved that the correct atoms are identified at only one iteration if $\tilde{\mu}_1(k) + \tilde{\mu}_1(k-1) < 1$. Then, the residual after $l$ steps can be represented as a linear combination of atoms $\Phi_{opt}$, that is $y - \Phi x^{(l)} = \Phi_{opt} h^{(l)}$, where $h$ can be calculated as $h_i^{(l)} = 1 - x_i^{(l)}$. With (22), $h_i^{(l)}$ can be bounded as

$$-(\tilde{\mu}_1(k-1))^l \leq h_i^{(l)} \leq (\tilde{\mu}_1(k-1))^l, \quad \text{for} \quad i \in \Lambda_{opt}. \tag{23}$$

Then, at the $(l{+}1)$-th iteration, we can obtain

$$\begin{aligned} x_i^{(l+1)} &= x_i^{(l)} + \psi_i^T (y - \Phi_{opt}) \\ &= \phi_i^T \psi_i + \sum_{\substack{j \in \Lambda_{opt} \\ j \neq i}} \langle \psi_i, \phi_j \rangle h_j, \quad \text{for} \quad i \in \Lambda_{opt}. \end{aligned} \tag{24}$$

With (23), we can get

$$1 - (\tilde{\mu}_1(k-1))^{l+1} \leq x_i^{(l+1)} \leq 1 + (\tilde{\mu}_1(k-1))^{l+1}, \quad \text{for} \quad i \in \Lambda_{opt}. \tag{25}$$

The above results show that the sequence of coefficients generated by the modified TLI converges to the true value while the iteration procedure put forward. Through constructing an appropriate sensing dictionary with cross cumulative coherence small enough, the modified TLI can obtain a sparse approximation with any expected tolerance in a few iterations.

Theorem 3: Assume that a signal can be represented as $y = \Phi_{opt} x_{opt}$ with $|\Lambda_{opt}| = K$ and the nonzero entries of the coefficient vector are ones. With the initialization $x^{(0)} = 0$, the modified TLI algorithm based on a sensing dictionary $\Psi$ is used to recover this sparse signal. If $\tilde{\mu}_1(k) + \tilde{\mu}_1(k-1) < 1$, the residual at $n$-th iteration step satisfies

$$\left\| y - \Phi x^{(n)} \right\|_2 \leq C (\tilde{\mu}_1(k-1))^n, \tag{26}$$

where $C = \sqrt{K} \sigma_{\max}$.

Proof: With Theorem 2, we can get

$$\left\| y - \Phi x^{(n)} \right\|_2 = \left\| \Phi_{opt} h^{(n)} \right\|_2 \leq \| \Phi_{opt} \|_{2,2} \left\| h^{(n)} \right\|_2 = \sigma_{\max} \left\| h^{(n)} \right\|_2, \tag{27}$$

where $\sigma_{\max}$ denotes the maximal singular value of the sub-dictionary $\Phi_{opt}$. With (23), we can obtain $\left\|h^{(n)}\right\|_2 \leq \sqrt{K}(\tilde{\mu}_1(k-1))^n$. Substitute this relation to (27), we complete the proof.

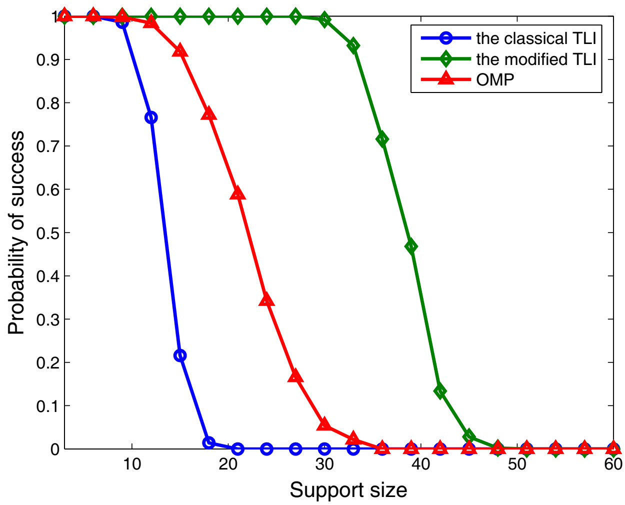


**Figure 1.** Probability of exact recovery via different support size.

## 5. SIMULATION RESULTS

To illustrate the performance of the proposed algorithm, numerical simulations are presented in this section. We generate a redundant dictionary with the dimensions $M = 128$ and $N = 256$. The entries of this dictionary were drawn independently from i.i.d. normal distribution. And the coefficient vector contains only ones and zeros. In our experiments, we fix the number of iteration for both TLI algorithms as 10.

In the first experiment, we compare the performance of exact recovery obtained by these methods. Simulation results are obtained over independent 500 Monte-Carlo trails. Fig. 1 shows the probability of exact recovery via different support size for different algorithms. As shown in this plot, the exact reconstruction performance of the modified TLI algorithm is better than both the classical TLI and OMP. In the second experiment, the convergence rate for both TLI algorithms is considered. We set the sparsity $K = 6$ since both algorithms can identify the correct atoms in this case. Fig. 2 shows the residuals via the number of iteration for both TLI algorithms. As shown in this

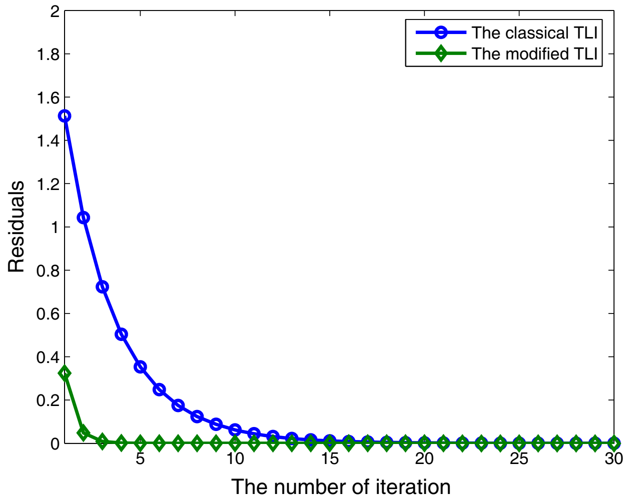


**Figure 2.** Residuals via the number of iteration.

result, the convergence rate of the proposed algorithm is faster than the classical TLI.

## 6. CONCLUSIONS

In this paper, we proposed a modified TLI algorithm for the problem of sparse representation. Through constructing a sensing dictionary, the proposed algorithm can significantly improve the performance of the classical TLI. Theoretical analysis and numerical simulations are presented to illustrate the superior performance of the proposed algorithm.

## ACKNOWLEDGMENT

This work was supported by a grant from the National Natural Science Foundation of China (No. 60772146) and the National High Technology Research and Development Program of China (863 Program, No. 2008AA12Z306).

## REFERENCES

1. Blumensath, T. and M. E. Davies, "Iterative thresholding for sparse approximations," *The Journal of Fourier Analysis and*

*Applications*, Vol. 14, No. 5, 629–654, Dec. 2008.

2. Zhang, Y., Q. Wan, and A. Huang, "Localization of narrow band sources in the presence of mutual coupling sparse solution finding," *Progress In Electromagnetics Resaerch*, PIER 86, 243–257, 2008.
3. Pati, Y. C., R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," *Proc. 27th Annu. Asilomar Conf. Signals, Systems, and Computers*, Vol. 1, 40–44, Pacific Grove, CA, Nov. 1993.
4. Daubechies, I., M. Defrise, and C. De Mol, "An iterative thresholding algorothm for linear inverse problems," *Comm. Pure Appl. Math.*, Vol. 57, No. 11, 1413–1457, Aug. 2004.
5. Hunter, D. and K. Lange, "A tutorial on MM algorithms," *Amer. Statist.*, Vol. 58, 30–37, 2004.
6. Bioucas-Dias, J., "Bayesian wavelet-based image deconvolution: A GEM algorithm exploiting a class of heavy-tailed priors," *IEEE Trans. Image Process.*, Vol. 15, No. 4, 937–951, Apr. 2006.
7. Adeyemi, T. and M. E. Davies, "Sparse representations of images using overcomplete complex wavelets," *IEEE SP 13th Workshop on Statistical Signal Processing*, 805–809, Jul. 2006.
8. Herrity, K. K., A. C. Gilbert, and J. A. Tropp, "Sparse approximation via iterative thresholding," *IEEE International Conference on ICASSP*, Vol. 3, 14–19, May 2006.
9. Schnass, K. and P. Vandergheynst, "Dictionary preconditioning for greedy algorithms," *IEEE Trans. Signal Process.*, Vol. 56, No. 5, 1994–2002, May 2008.