Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 9407503, 10 pages
http://dx.doi.org/10.1155/2016/9407503

Research Article

Incoherent Dictionary Learning Method Based on Unit Norm Tight Frame and Manifold Optimization for Sparse Representation

HongZhong Tang,[1,2,3] Xiaogang Zhang,[1] Hua Chen,[4] Ling Zhu,[2,3] Xiang Wang,[2,3] and Xiao Li[2,3]

[1]*College of Electrical and Information Engineering, Hunan University, Changsha 410082, China*
[2]*College of Information Engineering, Xiangtan University, Xiangtan 411105, China*
[3]*Institute of Control Engineering, Xiangtan University, Xiangtan 411105, China*
[4]*School of Information Science and Engineering, Hunan University, Changsha 410082, China*

Correspondence should be addressed to Hua Chen; anneychen@126.com

Received 28 March 2016; Revised 24 June 2016; Accepted 31 July 2016

Academic Editor: Vladan Papic

Copyright © 2016 HongZhong Tang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Optimizing the mutual coherence of a learned dictionary plays an important role in sparse representation and compressed sensing. In this paper, a efficient framework is developed to learn an incoherent dictionary for sparse representation. In particular, the coherence of a previous dictionary (or Gram matrix) is reduced sequentially by finding a new dictionary (or Gram matrix), which is closest to the reference unit norm tight frame of the previous dictionary (or Gram matrix). The optimization problem can be solved by restricting the tightness and coherence alternately at each iteration of the algorithm. The significant and different aspect of our proposed framework is that the learned dictionary can approximate an equiangular tight frame. Furthermore, manifold optimization is used to avoid the degeneracy of sparse representation while only reducing the coherence of the learned dictionary. This can be performed after the dictionary update process rather than during the dictionary update process. Experiments on synthetic and real audio data show that our proposed methods give notable improvements in lower coherence, have faster running times, and are extremely robust compared to several existing methods.

1. Introduction

In recent years, the research of dictionary learning has attracted a lot of attention because a learned dictionary captures some of the intrafeatures of training samples in many applications like denoising [1], compressed sensing [2], pattern recognition, and classification tasks [3–5]. A learned dictionary allows an interesting signal to be represented as a linear combination with relatively few atoms, and the representation coefficients are as sparse as possible. Hence, the problem of dictionary learning can be stated as follows [6]:

$$\begin{aligned} \underset{D\in\mathbb{D},X\in\mathbb{X}}{\operatorname{argmin}} \quad & \frac{1}{2}\|Y - DX\|_F^2 \\ \text{s.t.} \quad & \|\mathbf{x}_i\|_0 \le S, \quad i = 1, 2, \ldots, K, \end{aligned} \tag{1}$$

where $Y = \{\mathbf{y}_i\}_{i=1}^N \in \mathbb{R}^{M\times N}$, $\mathbb{D}$ is the admissible set of all column-normalized dictionaries, $D = \{\mathbf{d}_k\}_{k=1}^K \in \mathbb{R}^{M\times K}$ is an overcomplete dictionary ($M < K$), and each column of D is referred to an atom. $\mathbb{X}$ represents the admissible set of all sparse coefficient matrices (i.e., most of the entries are either zero or are sufficiently small in magnitude), and $X = \{\mathbf{x}_i\}_{i=1}^N \in \mathbb{R}^{K\times N}$. S represents nonzero numbers in $\mathbf{x}_i$.

Equation (1) is not a convex problem regarding the pair (D, X), so most dictionary learning methods employ alternating optimization over D and X. The following two stages are repeated until convergence:

(1) Sparse coding

$$X^{(\text{iter}+1)} = \underset{X\in\mathbb{X}}{\operatorname{argmin}} \quad \frac{1}{2}\left\|Y - D^{(\text{iter})}X\right\|_F^2 \quad \text{s.t.} \quad \left\|\mathbf{x}_i\right\|_0 \le S, \quad i = 1, 2, \ldots, K. \tag{2}$$

(2) Dictionary update

$$D^{(\text{iter}+1)} = \underset{D\in\mathbb{D}}{\operatorname{argmin}} \frac{1}{2}\left\|Y - DX^{(\text{iter}+1)}\right\|_F^2. \tag{3}$$

The first stage is a sparse coding with D fixed, and the second stage is a dictionary update that updates partial atoms with X fixed.

(1) Related Work. Different applications tend to use different optimization algorithms for learning sparsifying dictionaries to obtain the desired characteristics. Traditional dictionary learning methods, such as the method of optimal directions (MOD) [7] and K-means singular value decomposition (K-SVD) [8], aim at optimizing a dictionary to represent all training samples sparsely, but the coherence between atoms is ignored. However, many studies of compressed sensing focus on the mutual coherence of an effective dictionary (the multiplication of a sensing matrix and dictionary) [9–12], which is a key factor in controlling the support of solutions of the least-squares with l_1 penalized and greedy problems. Furthermore, highly incoherent dictionaries tend to avoid ambiguity and improve noise stability when sparse coding is enforced. Therefore, an incoherent frame is applied typically to optimize the sensing matrix in compressed sensing. Tsiligianni et al. [13] constructed an incoherent frame to optimize a sensing matrix by the averaged projections onto a Gram matrix and obtained better performance of sparse signal recovery. Rusu and González-Prelcic [14] directly optimized the maximum inner product between pairs of atoms to construct incoherent frames using convex optimization. It is unlikely that we will focus on learning an incoherent dictionary for sparse representation.

Much research has concentrated on reducing the coherence of a learned dictionaries. Yaghoobi et al. [15, 16] introduced the optimization of dictionary coherence by imposing a minimal coherence constraint to design a parametric dictionary to deal in advance with a relatively well-known signal model. The penalty term on the coherence is added in the dictionary learning; therefore, (1) can be reformulated as follows [17–20]:

$$\underset{D\in\mathbb{D},X\in\mathbb{X}}{\operatorname{argmin}} \quad \frac{1}{2}\|Y - DX\|_F^2 + \frac{1}{2}\left\|D^TD - I\right\|_F^2, \quad \text{s.t.} \quad \left\|\mathbf{x}_i\right\|_0 \le S, \ i = 1, 2, \ldots, K. \tag{4}$$

Inspired by MOD [7], Ramirez et al. [17] proposed the method of optimal coherence-constrained directions (MOCOD) to learn a dictionary. The limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm was used to cooptimized the coherence between atoms and the performance of sparse representation [18]. Abolghasemi et al. [19] constructed an incoherent dictionary by using steepest-gradient descent in the iteration of K-SVD. Bao et al. [20] proposed a hybrid alternating proximal algorithm for incoherent dictionary learning accompanied by a convergent analysis and a proof. The abovementioned methods cannot reduce the coherence obviously because the sum of the squared inner product of all atom pairs is minimized in the second term of (4). Meanwhile, a learned dictionaries cannot obtain the target coherence arbitrarily. Mailhé et al. [21] proposed an incoherent K-SVD (INKSVD) algorithm, in which each pair of atoms, having a higher value than the target coherence, is decorrelated in the dictionary update after K-SVD is performed. The key idea is to cluster atoms and symmetrically decrease the correlation of each pair of atoms based on a greedy method. The main drawback is that if the target coherence is set too low, the work will not perform well on sparse representation (Figure 5), and the computation will rise dramatically (Table 2). Barchiesi and Plumbley [22] proposed an incoherent dictionary learning method that enforces the iterative projection (IP) onto the spectral and structural constraint set in order to obtain the optimal Gram matrix. As a result, dictionary optimization was performed based on the orthogonal Procrustes problem (OPP) for better sparse representation performance. More recently, Rusu and González-Prelcic [14] directly constructed an incoherent dictionary followed by orthogonal constraints such as in [22] for sparse representation. Similar work has been done with regard to the dictionary optimization by [14, 22]. We are only reporting results obtained using methods from [21, 22], as they seem to provide better performance of incoherent dictionaries and sparse representation. Note that it is a difficult task to obtain an arbitrarily low coherence, and this is not approximating the flat spectrum of an equiangular tight frame (ETF) using the methods from the literature [21, 22] (see Section 4). Additionally, our proposed methods improved [21, 22]. Manopt is employed to solve orthogonal Procrustes problem in the dictionary optimization because only reducing the coherence of a learned dictionary will degenerate the performance of sparse representation, which is very different from the work done in [14, 21, 22] and is also the major contribution of our work.

(2) Our Contributions. There are three specific characteristics of our proposed incoherent dictionary learning methods that distinguish them from prior methods.

(1) Rather than general dictionary learning methods, an efficient framework based on a unit norm tight frame (UNTF) is developed for solving the incoherent dictionary learning problem, which constrains the dictionary to approximate to ETF.

(2) The mutual coherence of the dictionary is reduced by alternately restricting the tightness and coherence, which gives a significantly higher incoherence than those reports in [21, 22].

(3) We use manifold optimization (Manopt) to solve the problem of optimization with orthogonal constraints, that is, (14), which aims to obtain better performance

from incoherent dictionaries and sparse representation. Experiments are carried out on synthetic data and real audio data to illustrate the better performance of our proposed methods.

(3) Organization of Paper. The rest of this paper is organized as follows. Section 2 gives the definitions of mutual coherence and ETFs, after which our proposed algorithms are presented in Section 3. Section 4 gives the details of dictionary optimization by employing Manopt. Section 5 reports on extensive experiments that were carried on synthetic data and real audio data. Finally, conclusions are drawn in Section 6.

2. Incoherent Dictionary

2.1. The Mutual Coherence. The mutual coherence of a dictionary is defined as the maximum inner product between atoms [23]:

$$\mu(D) = \max \left| \left\langle \mathbf{d}_k \cdot \mathbf{d}_j \right\rangle \right|, \quad k \neq j, \tag{5}$$

where $\mathbf{d}_k$ and $\mathbf{d}_j$ denote two different atoms.

The coherence measures the similarities between atoms. We have that $\mu(D) \in (0, 1)$, and a dictionary is considered incoherent if $\mu(D)$ is small. Mutual coherence is an important sufficient condition to provide a theoretical guarantee for exactly sparse signal recovery.

Theorem 1 (see [9, 10]). *Let $D \in \mathbb{R}^{M\times K}$ be overcomplete dictionary with mutual coherence $\mu(D)$, if condition (6) is satisfied*

$$S < \frac{1}{2}\left(1 + \frac{1}{\mu(D)}\right), \tag{6}$$

where S is nonzero numbers in $\mathbf{x}$. Consider the system $\mathbf{y} = D\mathbf{x}$, in which case $\mathbf{x}$ can be recovered using basis pursuit (BP) and orthogonal matching pursuit (OMP). Theorem 1 shows that an incoherent dictionary is desirable; here, the best expectation is that the mutual coherence can reach the Welch bound.

Theorem 2 (see [25]). *Consider overcomplete dictionary $D \in \mathbb{R}^{M\times K}$ with normalized columns. The coherence satisfies*

$$\mu(D) \geq \sqrt{\frac{K - M}{M(K-1)}}. \tag{7}$$

The bound is achieved if and only if matrix D is an equiangular tight frame (ETF).

Therefore, optimizing a dictionary to approximate ETF is an effective method to reduce the coherence in sparse representation.

2.2. Equiangular Tight Frames. ETF can be defined as follows.

Definition 3 (see [26]). Let F be a $M \times K$ matrix, where $M < K$, whose columns are $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_K$. Matrix F is named an equiangular tight frame if the following conditions are met:

(1) Each column has unit norm: $\|\mathbf{f}_k\|_2 = 1$ for $k = 1, 2, \ldots, K$.

(2) The columns are equiangular. For some nonnegative c, we have $|\langle \mathbf{f}_k, \mathbf{f}_j \rangle| = c$ when $1 \leq k \leq j \leq K$.

(3) The columns form a tight frame. That is to say, $FF^T = (K/M)\mathrm{I}_M$,

where I_M is a unit matrix with $M \times M$. It follows that c is the lowest coherence, and matrix F is full row rank, and M nonzero singular values are equal to $\sqrt{K/M}$.

3. Our Proposed Incoherent Dictionary Learning

Frames are an overcomplete version of a basis set and tight frames are an overcomplete version of an orthogonal basis set. ETFs generalize the geometric properties of an orthogonal basis [26]. However, ETF is difficult to construct. In particular, Tropp et al. [27] have demonstrated that α-tight frame is the closest design in the Frobenius-norm sense to the solution of the relaxed problem.

Theorem 4 (see [27]). *Given the matrix $D \in \mathbb{R}^{M\times K}$ with $M < K$, suppose that it has singular value decomposition $U\Sigma V^T$, then the matrix UV^T is called orthogonal polar factor. With regard to Frobenius norm, αUV^T is proximate α-tight frame to the matrix D, and it is also obtained by computing $\alpha(DD^T)^{-1/2}D$.*

We call the given α-tight frame UNTF if all columns $\|f_k\|_2 = 1$, in which case $\alpha = \sqrt{K/M}$. UNTF is employed in our proposed methods, because it is closest to the computed low-coherent dictionary in terms of its Frobenius norm.

3.1. Our Proposed Incoherent Dictionary Learning Algorithms. To constrain the coherence between atoms, (3) can be reformulated as

$$\begin{aligned} \underset{D\in\mathbb{D}}{\operatorname{argmin}} \quad & \frac{1}{2}\|Y - DX\|_F^2 \\ \text{s.t.} \quad & \mu(D) \leq \mu_t, \end{aligned} \tag{8}$$

where μ_t is the target coherence. Next, we modify the INKSVD [21] and IP [22] algorithms according to Theorem 4 in the expectation that the new dictionary will be proximal to ETFs. Following these modification, the proposed algorithms are named UNTF-INKSVD and UNTF-IP, respectively, in order to emphasize our different framework as well as prior work.

3.1.1. The Improvement of INKSVD Algorithm. In the INKSVD algorithm [21], coherent optimization is added after K-SVD is performed. It is expressed as follows:

$$\begin{aligned} \underset{D\in\mathbb{D}}{\operatorname{argmin}} \quad & \frac{1}{2}\left\|D - \overline{D}\right\|_F^2 \\ \text{s.t.} \quad & \mu(D) \leq u_t, \end{aligned} \tag{9}$$

Input: initial dictionary D, μ_t, ϵ, iterations
Output: D_{opt}
(1) Initialize:
(2) $q = 1$;
(3) $[U\Sigma V] = \text{SVD}(D)$;
(4) $\Phi_0 = U\Sigma_I V^T$;
(5) normalize the columns of Φ_0;
(6) **while** $q \leq$ iterations **do**
(7) Obtain the new dictionary D_q by solving (10) based on the Φ_{q-1};
(8) $\Phi_q = \sqrt{K/M}(D_q D_q^T)^{-1/2} D_q$;
(9) Normalize the columns of Φ_q, D_q;
(10) Compute the new mutual coherence $\mu(\Phi_q)$ and $\mu(D_q)$;
(11) **if** $\mu(\Phi_q) \leq \mu(\Phi_{q-1})$ **then**
(12) $\epsilon = \epsilon/2$;
(13) **end**
(14) Update D, $D_{\text{opt}} = D_q$;
(15) **if** $\mu(D_q) \leq \mu_t$ **then**
(16) break;
(17) **end**
(18) **end**
(19) return D_{opt};

Algorithm 1: Improved INKSVD algorithm: UNTF-INKSVD.

where $\overline{D}$ is the given dictionary and the minimization of matrix nearest problem is employed to resolve (9).

In the first algorithm, the coherence of an initial dictionary is reduced sequentially by finding a new dictionary, which has a lower coherence and is nearest to the previous one. Accordingly, we modify the objective function based on Theorem 4:

$$\begin{aligned} \underset{D\in\mathbb{D}}{\text{argmin}} \quad & \left\| D^T\Phi - I \right\|_\infty \\ \text{s.t.} \quad & \|D - \Phi\|_F^2 < \epsilon, \\ & \text{Tr}\left(D^T\Phi\right) = M, \end{aligned} \tag{10}$$

where Φ is the reference UNTF of the previous dictionary. Equation (10) can be resolved based on local convex problems by using convex-optimization toolbox (CVX) (http://cvxr.com/cvx/doc/CVX.pdf).

The proposed algorithm is called UNTF-INKSVD, as discussed in relation to Algorithm 1. Firstly, we take the normalized Φ_0 as the initial UNTF; then (10) is used to seek for a new dictionary with a lower coherence, which is proximal to the reference UNTF of the previous one. That is to say, Φ_{q-1} can be viewed as the reference UNTF in the qth iteration. Lastly we project the new dictionary onto the UNTF manifold, $\sqrt{K/M}(DD^T)^{-1/2}D$, achieving an incoherent tight frame. Thus, the constraints between optimizing coherence and projection onto the UNTF manifold are performed alternately in the iterative dictionary update, yielding a tightness and lower coherence between atoms.

3.1.2. The Improvement of IP Algorithm. The off-diagonal entries g_{kj} of the Gram matrix $G = D^T D$ represent the coherence between atoms, so another technique for reducing coherence is to operate on the entries of the Gram matrix also in order to meet the following property:

$$\begin{aligned} &\left|g_{kj}\right| \leq \mu_t, \quad 1 \leq k,\ j \leq K,\ k \neq j, \\ &g_{kk} = 1, \quad 1 \leq k \leq K, \end{aligned} \tag{11}$$

where μ_t is the target coherence.

Barchiesi and Plumbley [22] proposed iterative projections (IP) onto Gram matrix to reduce the correlation between atoms. Shrinkage is performed on the off-diagonal entries of the Gram matrix based on the following function:

$$g_{kj} = \begin{cases} 1 & k = j \\ g_{kj} & \left|g_{kj}\right| < \mu_t \\ \text{sign}\left(g_{kj}\right)\mu_t & \left|g_{kj}\right| > \mu_t. \end{cases} \tag{12}$$

Unfortunately, the rank of shrunken Gram matrix may be greater than M. Therefore, SVD is used to keep the best rank-M approximation. The decomposition can be used further to extract the square root of the new Gram matrix, thus obtaining the optimal dictionary D.

In the second algorithm, the coherence of the initial Gram matrix decreases sequentially upon finding a new Gram matrix that has a lower coherence and is nearest to the

Input: initial dictionary D, μ_t, iterations;
Output: D_{opt}
(1) Initialize $q = 1$;
(2) $[U\Sigma V] = \text{SVD}(D)$;
(3) $\Phi_0 = U\Sigma_I V^T$;
(4) normalize the columns of Φ_0;
(5) **while** $q \leq$ iterations **do**
(6) Compute the reference Gram matrix $G_q = \Phi_{q-1}^T \Phi_{q-1}$;
(7) Apply (12) to G_q for decreasing the coherence;
(8) Apply SVD to G_q to obtain the matrix G_e which rank is equal to be M;
(9) Building the squared-root of $G_e = D_q^T D_q$ to obtain a new dictionary;
(10) Compute $\Phi_q = \sqrt{K/M}(D_q D_q^T)^{-1/2} D_q$ to obtain the next closest UNTF;
(11) Normalize the columns of D_q and Φ_q;
(12) Update D, $D_{\text{opt}} = D_q$;
(13) Compute the new mutual coherence $\mu(D_{\text{opt}})$;
(14) **if** $\mu(D_{\text{opt}}) \leq \mu_t$ **then**
(15) break;
(16) **end**
(17) **end**
(18) return D_{opt};

Algorithm 2: Improved IP algorithm: UNTF-IP.

previous one. The optimization of (9) can be modified into the following problem:

$$\begin{aligned} \text{argmin} \quad & \left\| G_e - G_q \right\|_\infty \\ \text{s.t.} \quad & \text{rank}\left(G_e\right) = M, \\ & \text{diag}\left(G_e\right) = 1, \\ & G_e \succeq 0, \end{aligned} \tag{13}$$

where $G_q = \Phi^T \Phi$ is called to be reference Gram matrix.

The core methodology is to operate the reference Gram matrix G_q, instead of $D^T D$. Our modified algorithm is referred to as UNTF-IP, and the optimization process is described as Algorithm 2.

Firstly, the closest α-tight frame Φ_0 is obtained. Normalization is then executed, after which the Gram matrix $G_1 = \Phi_0^T \Phi_0$ is computed. In the qth iteration, Φ_{q-1} can be viewed as the best coherence set over the current dictionary by employing Theorem 4. The above shrinkage operation equation (12) is performed, and SVD is enforced to obtain the rank M. The updated Gram matrix is then decomposed to obtain the new dictionary. Lastly, we project the new dictionary onto the UNTF manifold, $\sqrt{K/M}(DD^T)^{-1/2}D$, achieving the next reference UNTF. Consequently, we obtain an effectively tighter and lower coherence dictionary that those obtained with the IP algorithms [22].

4. Dictionary Optimization with Manopt

Only reducing the coherence in (10) and (13) will result in poor performance of the sparse representation. Hence, after (10) and (13) are resolved, we add dictionary optimization to maintain good performance of the sparse representation based on the OPP. Equation (1) can be formulated equivalently as an orthogonal-constraint minimization as follows:

$$\begin{aligned} \underset{R \in \mathbb{R}^{M \times M}}{\text{argmin}} \quad & \frac{1}{2} \left\| Y - RDX \right\|_F^2 \\ \text{s.t.} \quad & R^T R = I. \end{aligned} \tag{14}$$

It is clear that $(RD)^T(RD) = D^T R^T RD = D^T D$. So the dictionary optimization in (14) has two advantages: (I) good representation performance can be obtained; (II) incoherence remains unchanged.

In [14, 22], dictionary rotation (DR) is employed to solve (14), but this is performed in the iterative dictionary update of (10) and (13). As demonstrated in [28], Manopt provides efficient algorithms to find an optimal solution of the OPP. In the next section, we introduce an optimization framework based on Manopt.

Let $f(R) = (1/2)\|Y - RDX\|_F^2$. We consider the special orthogonal constraint as a Riemannian submanifold of $R \in \mathbb{R}^{M \times M}$. Hence the purpose of manifold optimization is to find an optimal solution of R for the following model:

$$\underset{R \in \mathscr{M}}{\text{argmin}} \quad f(R), \tag{15}$$

where the search space $\mathscr{M}$ is a Riemannian manifold that can be linearized locally at each point R as a tangent space $T_R\mathscr{M}$.

The inner problem at the current iterate $R_k \in \mathcal{M}$ is defined as follows:

$$\begin{aligned} \min_{\eta_k \in T_{R_k}\mathcal{M}} \quad & m_{R_k}(\eta_k) \\ & = f(R_k) + \langle \eta_k, \operatorname{grad} f(R_k) \rangle \\ & \quad + \frac{1}{2} \langle \eta_k, \operatorname{Hess} f(R_k)[\eta_k] \rangle \\ \text{s.t.} \quad & \|\eta_k\|_M \le \Delta_k, \end{aligned} \tag{16}$$

where $\operatorname{grad} f(R_k)$ and $\operatorname{Hess} f(R_k)$ are the Riemannian gradient and the Hessian of the cost function at R_k, respectively.

The Riemannian gradient of f at R is defined as follows:

$$\operatorname{grad} f(R) = \operatorname{skew}\left(R^T \nabla f(R)\right), \tag{17}$$

where $\nabla f(R) = (RDX - Y)(DX)^T$ is the gradient of f as a function in $\mathbb{R}^{M \times M}$ and $\operatorname{skew}(A) = (A - A^T)/2$.

Intuitively, we also define the Riemannian Hessian of f at R along $\operatorname{grad} f(R)$:

$$\begin{aligned} & \operatorname{Hess} f(R) \\ & = \operatorname{skew}\left(R^T \nabla^2 f(R) - \operatorname{grad} f(R) \operatorname{sym}\left(R^T \nabla f(R)\right)\right), \end{aligned} \tag{18}$$

where $\nabla^2 f(R) = R \operatorname{grad} f(R)(DX)(DX)^T$ is the Hessian matrix of f at R along $\operatorname{grad} f(R)$ as a function in $\mathbb{R}^{M \times M}$ and $\operatorname{sym}(A) = (A + A^T)/2$.

Next, η_k is calculated based on inner iterations with Steihaug-Toint truncated conjugate gradient (tCG) [29]; a candidate next iteration is produced by

$$R_k^+ = P_{R_k}(\eta_k). \tag{19}$$

The term P_{R_k} is a retraction function on the manifold $\mathcal{M}$ and describes the mapping between the tangent space $T_x M$ and $\mathcal{M}$ for any point R_k. A simpler mapping is selected: $P_{R_k}(\eta_k) = R_k + R_k \eta_k$. Let R_k^+ be orthogonalized. The selection of whether to receive or discard the candidate and quotient is used to update the trust-region radius:

$$\begin{aligned} & \rho_k(R_k) \\ & = \frac{f(R_k) - f(R_k^+)}{-\langle \operatorname{grad} f(R_k), \eta_k \rangle - (1/2) \langle \operatorname{Hess} f(R_k)[\eta_k], \eta_k \rangle}. \end{aligned} \tag{20}$$

We optimize R using the Manopt toolbox [29] while D is fixed. Algorithm 3 presents the procedure for this optimization. Afterwards, the optimal D is obtained by $D = RD$. A better performance can be achieved for sparse representation, and the dictionary coherence remains unaffected. Furthermore, this optimization leads to a faster algorithm because it can be performed after the dictionary update process, which is in contrast to [22].

Input: Y, D, X and iterations
Output: D
(1) Initialize $k = 0$, $R_0 = \text{eye}(M)$;
(2) $\rho_0 = 0.1$, $\Delta_0 = 45.6889$, $\overline{\Delta} = 365.5114$, $\varepsilon = 1e-9$;
(3) **while** $k <$ iterations and $\|Y - RDX\|_F^2 \ge \varepsilon$ **do**
(4) Apply (16) $\eta_k \leftarrow \text{tCG}(R_k, \Delta_k)$;
(5) Apply (19) to compute the next iterate R_k^+;
(6) Apply (20) to compute the trust region radius ρ_k;
(7) **if** $\rho_k(R_k) < 1/4$ **then**
(8) $\Delta_{k+1} = (1/4)\Delta_k$;
(9) **else if** $\rho_k(R_k) > 3/4$ and $\|\eta_k\| = \Delta_k$ **then**
(10) $\Delta_{k+1} = \min(2\Delta_k, \overline{\Delta})$;
(11) **else**
(12) $\Delta_{k+1} = \Delta_k$
(13) **end**
(14) **end**
(15) **end**
(16) **if** $\rho_k(R_k) > \rho_0$ **then**
(17) $R_{k+1} = R_k^+$;
(18) **else**
(19) $R_{k+1} = R_k$;
(20) **end**
(21) **end**
(22) **end**
(23) $D = RD$;
(24) return D;

Algorithm 3: Dictionary optimization with Manopt.

5. Experiment Results

In this section, we report on the experiments with synthetic data and real audio data that were intended to compare our proposed incoherent dictionary learning with the prior methods. All the experiments were performed on a Dell computer with 4 GB of memory and a 2-core 2.6 GHz Intel Pentium processor. All the codes were written in MATLAB.

5.1. Incoherent Dictionary Construction. In this experiment, incoherent dictionaries are constructed without learning from training samples, and we aim to reduce directly the mutual coherence of a given dictionary. We set the initial dictionary $D \in \mathbb{R}^{M \times K}$ randomly, and $M = 20$ and $K = 50$ are chosen according to the condition that $M < K$. Each atom is normalized as a unit norm, and the Welch bound is 0.1750. In order to observe the benefits of our proposed methods, the dictionary update is taken as follows: (I) INKSVD [21]; (II) IP [22]; (III) UNTF-INKSVD; (IV) UNTF-IP. The INKSVD and IP are taken from the web (http://code.soundsoftware.ac.uk/). Each algorithm is executed for ten times, and average results are taken. Specifically, the same initial dictionary and iterations are used for the measurements, and we evaluate the mutual coherence of the constructed dictionaries using each algorithm.

Figure 1 shows the mutual coherence of the constructed dictionaries. We note that our proposed algorithms exhibit significantly lower coherences, with the performance of the UNTF-IP algorithm slightly exceeding those of IP, UNTF-INKSVD, and INKSVD. A standard line is $\sqrt{K/M} = 1.5811$ in Figure 2, which indicates that ETF has 20 nonzero singular values that are same. As can be seen, the UNTF-IP and

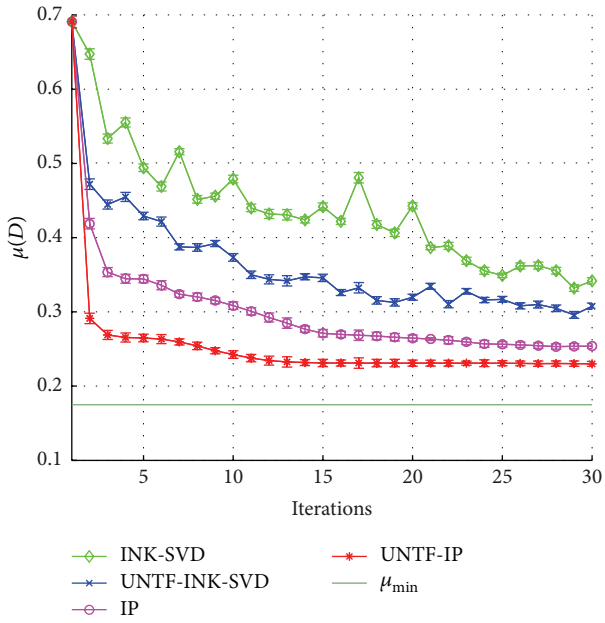


Figure 1: The mutual coherence of constructed dictionaries.

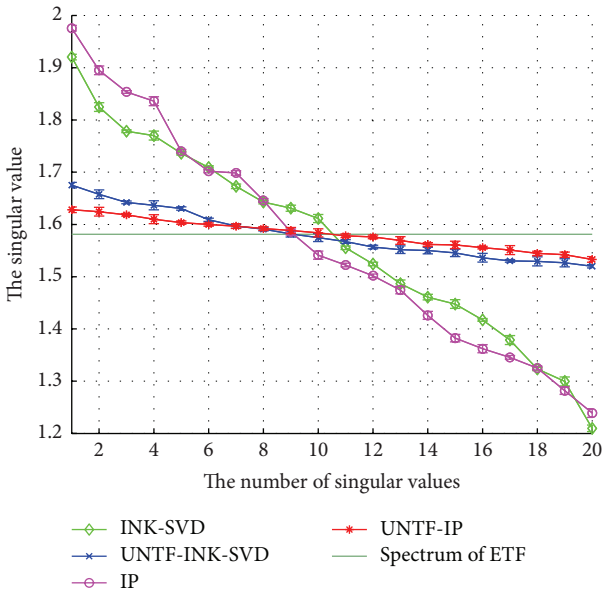


Figure 2: The singular values of constructed dictionaries.

UNTF-INKSVD algorithms give approximately flat spectra and approximate the properties of ETF. This is a better outcome than with the IP or INKSVD algorithms, because alternating the constraints on tightness and coherence has a beneficial effect on incoherent dictionary construction. As a result, the incoherent dictionary constructed with the UNTF-IP algorithm possess an apparent property with approximation of ETFs. The error bars in Figures 1 and 2 show the standard deviation based on running each test 10 times and demonstrate the consistency of the results.

Table 1: Our proposed methods for learning incoherent dictionary.

Method	Sparse coding [24]	Dictionary learning [8]	Dictionary update	Dictionary optimization
Literature [21]	OMP	K-SVD	INKSVD	—
Literature [22]	OMP	K-SVD	IP	DR
The proposed method I	OMP	K-SVD	UNTF-INKSVD	Manopt
The proposed method II	OMP	K-SVD	UNTF-IP	Manopt

5.2. Incoherent Dictionary Learning for Sparse Representation with Synthetic Data. In this section, we have investigated the incoherent dictionary learning performance for sparse representation of synthetic data. The training samples are generated via underdetermined $Y = RDX$, where $R \in \mathbb{R}^{M\times M}$ and $D \in \mathbb{R}^{M\times K}$ are generated randomly. The dictionary is enforced by a column normalization, where $M = 64$ and $K = 256$. The matrix $X \in \mathbb{R}^{K\times N}$ is a sparse matrix with $N = 20{,}000$. The nonzero coefficients are distributed randomly, and their values are determined according to a standard Gaussian distribution. The target coherence μ_t is set as a Welch bound of 0.1085. Table 1 summarizes the tested methods, which are executed for 30 iterations between dictionary update and optimization. Each algorithm is executed for 10 times, and average results are taken.

The error bars in Figures 3 and 4 show the standard deviation based on running each test 10 times, and they demonstrate the consistency of the results. Figure 3(a) shows significant coherence of each learned dictionary. The UNTF-IP and UNTF-INKSVD algorithms have better mutual coherence on average than those of IP [22] and INKSVD [21] algorithms. In particular, the coherence of the learned dictionary with the UNTF-IP algorithm is closest to the Welch bound. Note that we have used Manopt to achieve a better performance of sparse representation. A signal-to-noise (SNR) value of $20\log_{10}(\|Y\|_F^2/\|Y-RDX\|_F^2)$ is computed in order to evaluate the performance of sparse representation. The SNR value is showed in Figure 3(b), where it can be seen that Manopt gives a better performance of sparse representation with compared to [21, 22], while the coherence is reduced.

Figure 4 shows the ratio between the SNR and the coherence, $r = \text{SNR}/\mu(D)$. The experimental results show that our proposed algorithms perform well equilibrium between the coherence and sparse representation and exhibit a generalized performance of learned incoherent dictionaries.

5.3. Application on Audio Data. To verify the efficiency of our proposed methods, experiments are reported in this section on real audio data via $Y = DX + \varepsilon$, where ε is a predetermined noise. For the purposes of comparison and analysis, the audio dataset that we use is the one adopted by [21, 22], in which the data comprise an audio sample

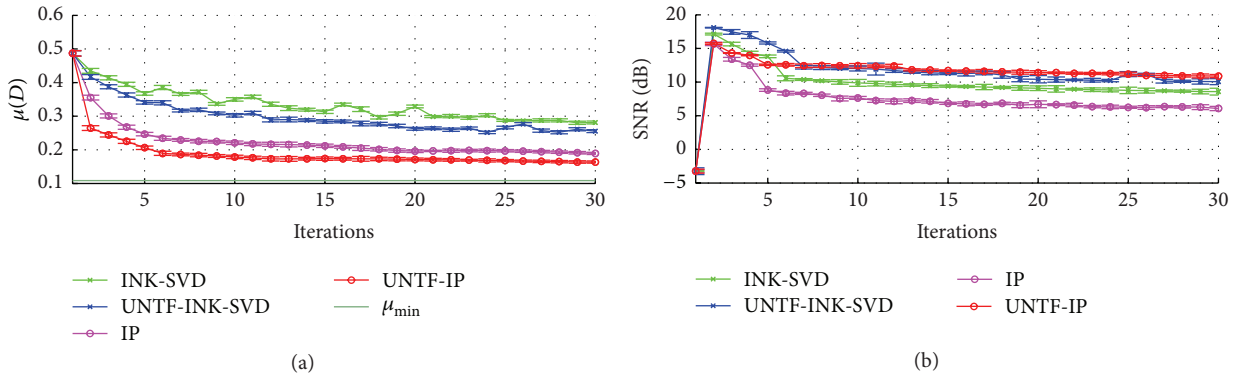


Figure 3: The mutual coherence and SNR with incoherent dictionary learning methods.

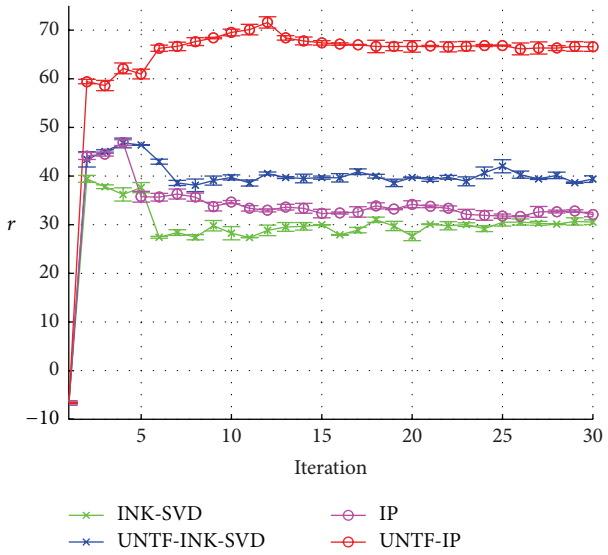


Figure 4: The ratio between the SNR and $\mu(D)$ with incoherent dictionary learning methods.

Table 2: The overall execution times (in seconds) with different incoherent dictionary learning methods.

	μ_t						
	0.05	0.1	0.15	0.2	0.25	0.3	0.35
Literature [21]	1.82×10^4	8252	2309	1127	942	733	623
Literature [22]	3628	3881	3732	3455	3362	3459	3437
Proposed method I	1533	1076	1093	988	936	871	942
Proposed method II	992	948	909	902	744	835	725

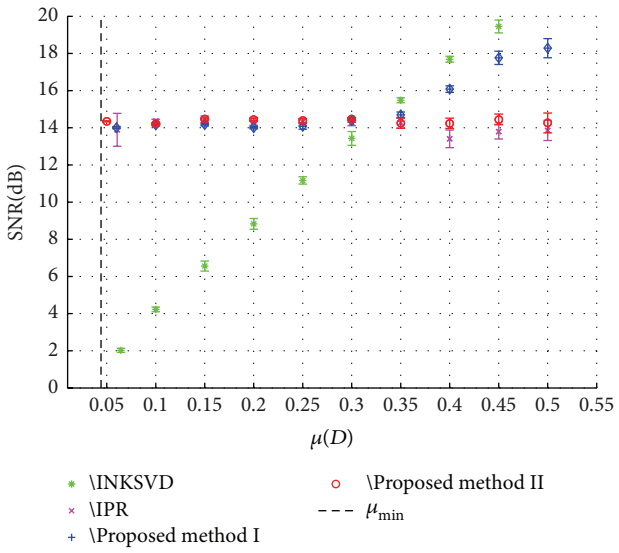


Figure 5: Comparing incoherent dictionary learning methods on audio data.

from a 16 kHz guitar recording. Furthermore, all columns in the initial dictionary are selected randomly from training samples and are normalized.

In this simulation, the target coherence μ_t is set in a range from 0.05 to 0.5, and the step size is 0.05. The tested methods are the same as those in Table 1, which are executed 10 times, and average results are taken. The termination criterion is that the target coherence is satisfied. We then evaluate our proposed incoherent dictionary learning methods by computing the mutual coherence and SNR.

As shown in Figure 5, the standard deviation based on running 10 times is showed, and the consistency in many tests is obtained. And when the target coherence is less than 0.3, the proposed method II in Table 1—employing UNTF-IP followed by Manopt—generates the best effect compared to other methods and approximates the lowest bound. However, if the target coherence is greater than 0.3, the SNR of [21] is the highest, followed by that of our proposed method I. Table 2 shows the computational running times. The key idea behind [21] is to decrease symmetrically the correlation of each pair of atoms having higher coherence based on a greedy method. Therefore, when the target coherence is higher, the number of pairs of atoms to be decorrelated will decrease dramatically, and the computation decreases dramatically as shown in the first row of Table 2. Unlike [22], the most important benefit of our proposed methods is to obtain a better computational efficiency when the target mutual coherence is very low, because Manopt can be performed after the dictionary update

process rather than during the dictionary update process. Compared with the prior methods, the present experimental results indicate that our learned dictionaries have a lower coherence, and with a certain degree of sparse representation.

6. Conclusion

In this paper, we have proposed two methods of learning an incoherent dictionary for sparse representation, adding the dictionary update and dictionary optimization in the traditional dictionary learning.

First, UNTF-INKSVD and UNTF-IP algorithms were developed to solve the problem of the higher incoherent and tighter dictionary effectively. Unlike other dictionary learning algorithms, our proposed algorithms learned an incoherent dictionary based on a unit norm tight frame in the dictionary update. An efficient framework was developed for sequentially reducing the coherence of an initial dictionary (or Gram matrix) by finding a new dictionary (or Gram matrix), which has a lower coherence and is nearest to the previous one. Hence, our learned incoherent dictionaries approximate the properties of ETFs, and the support of sparse coding is maximized.

Second, Manopt was employed to solve the orthogonal Procrustes problem in dictionary optimization, because only reducing the coherence of a learned dictionary will degenerate the performance of sparse representation. Meanwhile, we compared our proposed methods with the other methods, and the experimental results showed that our proposed methods balance the performance between incoherence and sparse representation. In particular, our proposed methods provide state-of-the-art results when μ_t is too low and have higher running speeds and better representation performances when compared to [21, 22]. This is because Manopt is performed after the dictionary update rather than during the dictionary update process.

However, a drawback is that our proposed methods are mainly suitable for learning an incoherent dictionary for sparse representation. Traditional dictionary learning seems to work well if the coherence of a learned dictionary is not restricted. In our work, more general, objective functions are proposed (see (10) and (13)) to construct an incoherent dictionary where tightness and coherence are restricted alternately at each iteration of the algorithm, and this method is similar to each alternating minimization. The theoretical proof of convergence in alternating minimization on more than two sets is still an open issue in [13, 15, 22]. Nevertheless, the experiments in our work show that incoherent dictionary learning methods can converge with a set of accumulation points under certain conditions. Our proposed algorithms can result in approximate converge to the values of objective coherence as in Figures 1, 3(a), and 5. Constructed dictionaries with our proposed algorithms give approximately flat spectrum of ETF in Figure 2. The SNR value is shown in Figures 3(b) and 5, which prove the effectiveness of our proposed algorithm when compared to [21, 22]. The convergence of the objective value does not prove the convergence of our proposed algorithms. Therefore, we will continue to work to prove the convergence of our proposed algorithms. And apply our proposed methods to other domains.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This article is supported by National Natural Science Foundation of China (61573299, 61673162, and 61672216), Scientific Research Project of Hunan Province Education Department (15C1328), and Control Science and Engineering Disciplinary Construction Funds of Xiangtan University.

References

[1] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736–3745, 2006.

[2] J. M. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: simultaneous sensing matrix and sparsifying dictionary optimization," *IEEE Transactions on Image Processing*, vol. 18, no. 7, pp. 1395–1408, 2009.

[3] S. Zubair, F. Yan, and W. Wang, "Dictionary learning based sparse coefficients for audio classification with max and average pooling," *Digital Signal Processing*, vol. 23, no. 3, pp. 960–970, 2013.

[4] Z. L. Jiang, Z. Lin, and L. S. Davis, "Label consistent K-SVD: learning a discriminative dictionary for recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 11, pp. 2651–2664, 2013.

[5] Q. A. Zhang and B. X. Li, "Discriminative K-SVD for dictionary learning in face recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 2691–2698, IEEE, San Francisco, Calif, USA, June 2010.

[6] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1045–1057, 2010.

[7] K. Engan, S. O. Aase, and J. H. Husoy, "Method of optimal directions for frame design," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '99)*, pp. 2443–2446, March 1999.

[8] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[9] D. L. Donoho and M. Elad, "On the stability of the basis pursuit in the presence of noise," *Signal Processing*, vol. 86, no. 3, pp. 511–532, 2006.

[10] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.

[11] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[12] M. P. Pereira, L. Lovisolo, E. A. B. Da Silva, and M. L. R. De Campos, "On the design of maximally incoherent sensing

matrices for compressed sensing using orthogonal bases and its extension for biorthogonal bases case," *Digital Signal Processing*, vol. 27, no. 1, pp. 12–22, 2014.

[13] E. V. Tsiligianni, L. P. Kondi, and A. K. Katsaggelos, "Construction of incoherent unit norm tight frames with application to compressed sensing," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2319–2330, 2014.

[14] C. Rusu and N. González-Prelcic, "Designing incoherent frames through convex techniques for optimized compressed sensing," *IEEE Transactions on Signal Processing*, vol. 64, no. 9, pp. 2334–2344, 2016.

[15] M. Yaghoobi, L. Daudet, and M. E. Davies, "Parametric dictionary design for sparse coding," *IEEE Transactions on Signal Processing*, vol. 57, no. 12, pp. 4800–4810, 2009.

[16] M. Yaghoobi, L. Daudet, and M. E. Davies, "Structured and incoherent parametric dictionary design," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '10)*, pp. 5486–5489, IEEE, Dallas, Tex, USA, March 2010.

[17] I. Ramirez, F. Lecumberry, and G. Sapiro, "Sparse modeling with universal priors and learned incoherent dictionaries," in *Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP '09)*, pp. 197–200, IEEE Signal Processing Society, December 2009.

[18] C. D. Sigg, T. Dikk, and J. M. Buhmann, "Learning dictionaries with bounded self-coherence," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 861–864, 2012.

[19] V. Abolghasemi, S. Ferdowsi, and S. Sanei, "Fast and incoherent dictionary learning algorithms with application to fMRI," *Signal, Image and Video Processing*, vol. 9, no. 1, pp. 147–158, 2015.

[20] C. Bao, Y. Quan, and H. Ji, "A convergent incoherent dictionary learning algorithm for sparse coding," in *Computer Vision—ECCV 2014*, vol. 8694 of *Lecture Notes in Computer Science*, pp. 302–316, Springer, New York, NY, USA, 2014, Proceedings of the IEEE Conference on Computer Vision—ECCV.

[21] B. Mailhé, D. Barchiesi, and M. D. Plumbley, "INK-SVD: learning incoherent dictionaries for sparse representations," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '12)*, pp. 3573–3576, Kyoto, Japan, March 2012.

[22] D. Barchiesi and M. D. Plumbley, "Learning incoherent dictionaries for sparse approximation using iterative projections and rotations," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 2055–2065, 2013.

[23] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2845–2862, 2001.

[24] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

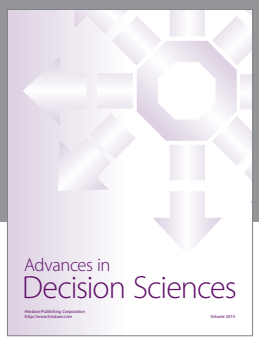
[25] L. R. Welch, "Lower bounds on the maximum cross correlation of signals," *IEEE Transactions on Information Theory*, vol. 20, no. 3, pp. 397–399, 1974.

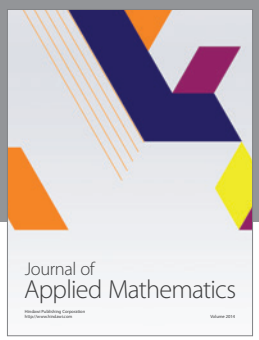
[26] M. A. Sustik, J. A. Tropp, I. S. Dhillon, and R. W. Heath Jr., "On the existence of equiangular tight frames," *Linear Algebra and Its Applications*, vol. 426, no. 2-3, pp. 619–635, 2007.

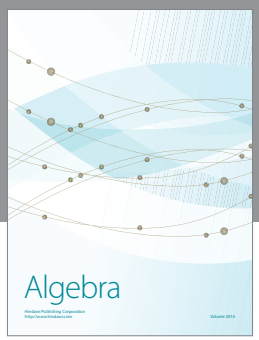
[27] J. A. Tropp, I. S. Dhillon, J. Heath, and T. Strohmer, "Designing structured tight frames via an alternating projection method," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 188–209, 2005.

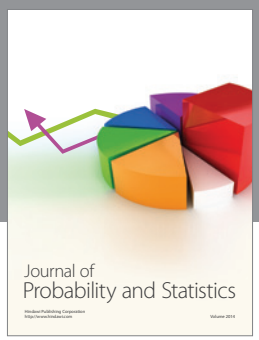
[28] N. Boumal, *Optimization and estimation on manifolds [Ph.D. thesis]*, Princeton University, 2014.

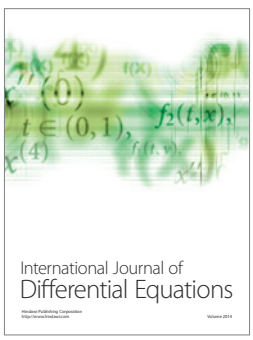
[29] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, "Manopt, a matlab toolbox for optimization on manifolds," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1455–1459, 2014.

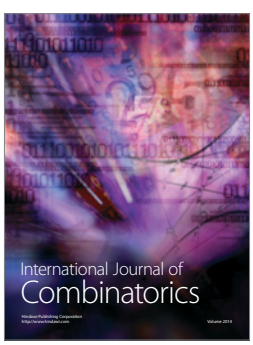


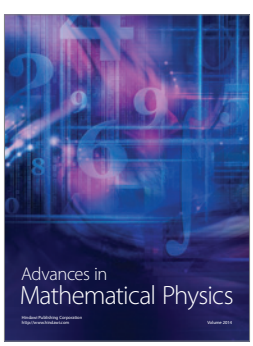


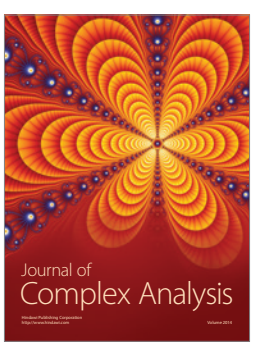


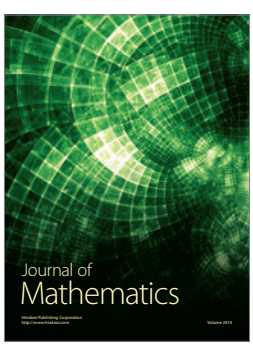


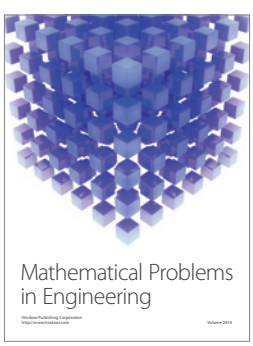


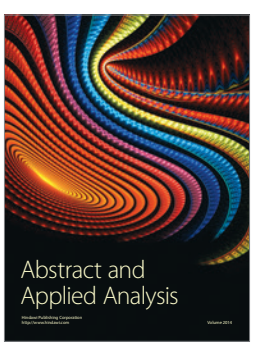


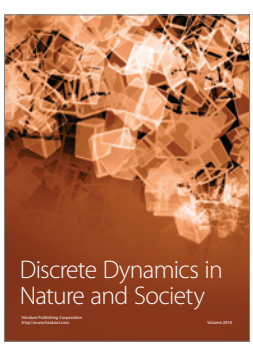


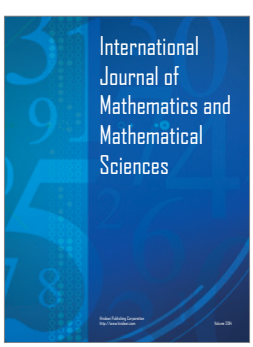


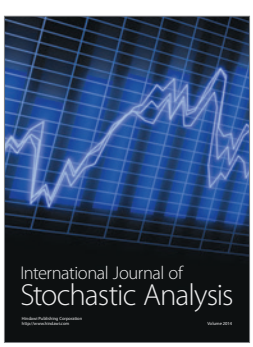












Hindawi

Submit your manuscripts at
http://www.hindawi.com