Using Sparse Principal Component Methods for Approximating Restricted Isometry Constants of Complex-Valued Tight Frames

Igor KAPORIN

Dorodnicyn Computing Center RAS of the Federal Research Center
“Computer Science and Control” RAS, Vavilova 40, 119333 Moscow, Russia

Keywords: Tight frames, Compressed sensing, Quiasirandom matrices, Restricted isometry constants, Discrete combinatorial problems, L1-penalty relaxation, Power method

Abstract. For the estimation of constants in the standard restricted isometry condition for a complex-valued tight frame, a generalization of techniques related to Sparse Principal Component Analysis is developed and applied. We consider certain optimization reformulations of the problem and iterative algorithms for approximating sparse solutions. The efficiency of methods is verified by numerical results obtained for several important test examples of tight frames.

Introduction

The construction and analysis of the so-called low-coherence redundant bases is a hot topic in data retrieval and processing during the past few decades. Presented in the form of rectangular matrices, such objects are used, for instance, in compressed sensing technologies, see, e.g., recent book [5] and survey [3]. Considering a complex-valued rectangular matrix $A \in \mathbb{C}^{m \times n}$, where $m < n$, the key requirement is that any $m \times k$ submatrix of $A$ denoted by

$$A_J = [a_{j_1}, \ldots, a_{j_k}], \quad 1 \leq j_1 < \ldots < j_k \leq n,$$

is sufficiently well-conditioned. Here $k \leq m$ is as large as possible and $J = \{j_1, \ldots, j_k\}$ is an arbitrary ordered subset of pairwise distinct indices. This assumption of well-conditioning for each $A_J$ is often characterized by the presence of the restricted isometry property (hereafter RIP) of the $k$th order, requiring that the inequalities

$$1 - \delta_1(k; A) \leq \frac{\|Az\|^2}{\|z\|^2} \leq 1 + \delta_2(k; A)$$

hold with some $0 < \delta_1(k; A) < 1$ and $\delta_2(k; A) > 0$ for all sparse $n$-vectors $z \in \mathbb{C}^n$ such that $\|z\|_0 \leq k$ (hereafter, $\|z\|_0$ denotes the number of nonzero components of the vector $z$). We will consider the optimum (unimprovable) bounds in [2], which are defined via minimum and maximum eigenvalues of $k \times k$ matrices $A_J^H A_J$, where $k \leq m$, as follows:

$$1 - \delta_1(k; A) = \min_{|J| = k} \lambda_{\text{min}}(A_J^H A_J),$$

$$1 + \delta_2(k; A) = \max_{|J| = k} \lambda_{\text{max}}(A_J^H A_J).$$
Here and further on, $\|z\|^2 = z^H z$ is the squared Euclidean norm of $z$, and $B^H$ is the conjugate transpose of $B$.

Note that the problem under consideration can be readily reformulated in terms of the Gram matrix $G = A^H A$ as the problem of finding sparse maximum eigenpairs of Hermitian nonnegative definite matrices $G$ and $\xi I - \eta G$ (with properly chosen positive constants $\xi$ and $\eta$, see Section 3.2 below). Hence, as it was noticed in [4], the evaluation of the RIP constants is equivalent to the solution of a pair of sparse Principal Component Analysis (Sparse PCA) problems. A survey of sparse PCA applications and an efficient algorithm for the approximate solution of the problem (in the case of the real-valued matrix $G$) can be found in [8].

Recall that in [13], the NP-hardness of the exact evaluation of RIP constants was shown. Not surprisingly, it is still unknown how to quickly and precisely estimate a quantity like $\delta_1(k; A)$ for a given rectangular matrix $A$. Moreover, any deterministic designs for compressed sensing matrices possessing RIP with $k = O(m^{1-\epsilon})$, where $0 < \epsilon \ll 1$, are not found yet. The existing results in this direction are only of probabilistic nature and hold for random matrices of certain classes.

From viewpoint of applications, see, e. g., [5] it is reasonable to restrict the exposition to the case of “equal norm tight frames” $A$ satisfying

$$\text{Diag}(A^H A) = I_n \quad \text{and} \quad AA^H = \frac{n}{m} I_m. \tag{5}$$

The presence of such properties typically correlates with better RIP constants.

In [4], Section 6.2, it was noticed that finding the restricted isometry constants is equivalent to the solution of a pair of Sparse Principal Component problems. Therefore, one can try to adjust some of the numerous heuristic algorithms already developed for the latter problem, for numerical approximation of restricted isometry constants.

In the present paper, a reformulation of the Generalized Power method of [8] (both for L0- and L1-relaxations of Sparse Principal Component problem, whereas in [9] only L1 case was considered) is developed applicable to the analysis of complex-valued data matrices $A$. It should be stressed that the original Generalized Power method of [8] can only be applied to real-valued data. The proposed algorithm is targeted to an approximation of quantities (3) and (4) for all $k = 1, 2, \ldots, m$ rather than to solving a typical Sparse PCA problem (such as extracting a single sparse dominant principal component of the data matrix $A$ providing for a certain trade-off between better sparsity and larger value of target function). In addition to the results of [9], for both L0- and L1-relaxations of Sparse PCA problem we present upper bounds for the solution sparsity in terms of the penalty parameter and the sizes of the problem.

**Optimization Problem Setting**

According to the above discussion, the original sparse PCA problem is set as

$$z_k = \arg \max_{\|z\|_0 = k} \frac{\|Az\|}{\|z\|}.$$  

Next, we consider relaxations of this problem using L0- or L1-penalties as well as their reformulations in the form of nonlinear eigenvalue problems.
Parametrized L0- and L1-relaxations of Sparse PCA Problem

Following [8] and recalling that $\|z\|_1 = \sum_{j=1}^n |z_j|$, let us replace the original problem by

$$z_0(\gamma) = \arg \max_{\|z\|=1} (\|Az\|^2 - \gamma^2 \|z\|_0)$$

in the case of using L0-penalty, and by

$$z_1(\gamma) = \arg \max_{\|z\|=1} (\|Az\| - \gamma \|z\|_1)$$

in the case of using L1-penalty. Here $\gamma$ is a positive parameter which controls the sparsity degree $k = \|z\|_0$. As follows from the discussion in Section 2.1 of [8], if $A$ is a column normalized matrix (recall that $(A^HA)_{j,j} = 1$ for all $j$ by (5)), then it suffices to consider the variation of the parameter in the range $0 < \gamma < 1$; cf. (17) and (18) below. Note that smaller values of $\gamma$ typically correspond to larger values of $k$, see Section 4 below.

Reformulation of the Optimization Problem with L0-penalty

Using the representation $\|v\|^2 = \max_{\|x\|=1} |x^Hv|^2$ with $v = Az$, one has, similar to [8] and [4],

$$\max_{\|z\|=1} (\|Az\|^2 - \gamma^2 \|z\|_0)$$

$$= \max_{\|z\|=1} \left( \max_{\|x\|=1} |x^HAz|^2 - \gamma^2 \|z\|_0 \right) = \max_{\|x\|=1} \max_{\|z\|=1} (|x^HAz|^2 - \gamma^2 \|z\|_0)$$

$$= \max_{\|x\|=1} \max_{\|z\|=1} \sum_{j=1}^n \text{sign}(|z_j|) (|(A^Hx)_j|^2 - \gamma^2) = \max_{\|x\|=1} \sum_{j=1}^n (|(A^Hx)_j|^2 - \gamma^2)_+,$$

where $(\alpha)_+ = \max(0, \alpha)$; we also use $\text{sign}(0) = 0$. The last equality follows from the use of the closed-form solution of the maximization problem by $z$ (cf. formula (13) in [8]). The optimality conditions for the arising problem have the form

$$g_0(x) = \mu x,$$  

$$x^Hx = 1,$$  

where

$$g_0(x) = \sum_{j=1}^n a_j a_j^H x \left( \text{sign}(1 - \gamma/|a_j^Hx|) \right)_+.$$  

Here $a_j$ is the $j$th column of $A$, so that $(A^Hx)_j = a_j^Hx$, and $\mu$ is a real positive number.

Reformulation of the Optimization Problem with L1-penalty

Using the formula $\|v\| = \max_{\|x\|=1} |x^Hv|$ with $v = Az$ again, one has, as in [8],

$$\max_{\|z\|=1} (\|Az\| - \gamma \|z\|_1)$$

299
\[
\begin{align*}
= \max_{\|x\|=1} \left( \max_{\|z\|=1} |x^H A z| - \gamma \|z\|_1 \right) & = \max_{\|x\|=1} \max_{\|z\|=1} \left( |x^H A z| - \gamma \|z\|_1 \right) \\
= \max_{\|x\|=1} \max_{\|z\|=1} \sum_{j=1}^n |z_j| \left( |(A^H x)_j| - \gamma \right) & = \left( \max_{\|x\|=1} \sum_{j=1}^n \left( |(A^H x)_j| - \gamma \right)_+ \right)^{1/2}.
\end{align*}
\]

As earlier, the last equality follows from the use of the closed-form solution of the maximization problem by \( z \) (cf. formula (8) in [8]). The optimality conditions for the arising problem have the form
\[
g_1(x) = \mu x, \quad (9)
\]
\[
x^H x = 1, \quad (10)
\]
where
\[
g_1(x) = \sum_{j=1}^n a_j a_j^H x \left( 1 - \gamma/|a_j^H x| \right)_+. \quad (11)
\]

As before, we use \((A^H x)_j = a_j^H x\), and \(\mu\) denotes some real positive number.

Recall that our purpose is to find the spectral bounds (3) and (4) as functions of the discrete argument \( k \). Therefore, one needs a reasonable strategy to reduce the number of different problems (6)-(8) or (9)-(11) to solve.

**Generalized Power Method for RIP Spectral Bounds Problem**

Below we propose a construction of a discrete set of values for the penalty parameter \(\gamma\) allowing to estimate the required spectral bounds for all \( k = 2, \ldots, m \).

In accordance with the hardness result [13], the iterative approximation to the solution of equation (6) or (9) may not, in general, deliver the global optimum. Therefore, the use of many iteration restarts is necessary to obtain a satisfactory result. On the other hand, a high-performance implementation of such method on modern supercomputers is straightforward due to its inherently parallel structure.

**Estimating the Upper Spectral Bound**

The method is based on the variation of the penalty parameter using the formula
\[
\gamma_l = \gamma_1 + (\gamma_2 - \gamma_1) l/l_{\text{max}}, \quad l = 0, \ldots, l_{\text{max}}, \quad (12)
\]
where \( 0 < \gamma_1 < \gamma_2 < 1 \). The variation of quasirandom initial guess for the vector \( x \) is made by the use of the so-called logistic sequence [16, 17] with starting term equal to \( \gamma_l \):
\[
\xi_i = 1 - 2\xi_i^2, \quad i = 1, \ldots, n, \quad \xi_0 = \gamma_l, \quad (13)
\]
and the \( st \) initial guess is formed as
\[
y_i = \xi_i + i\xi_{i-1}, \quad x^{(0)} = y/\|y\|
\]
(here and elsewhere, we denote \( i = \sqrt{-1} \)).
For each \( l \) we perform a sufficient number of power iterations (quite similar to Algorithm 3 or Algorithm 2 of [8] except of our specialization (8) for \( g(\cdot) = g_0(\cdot) \) with L0-penalty or (11) for \( g(\cdot) = g_1(\cdot) \) with L1-penalty, respectively):

\[
x^{(t+1)} = g(x^{(t)})/\|g(x^{(t)})\|, \quad t = 0, 1, \ldots, t_{\text{max}}.
\]

The main result of converged iterations is the index set

\[
J^{(t)} = \{j_1, \ldots, j_k\}, \quad \text{where} \quad j_s \in J^{(t)} \quad \text{iff} \quad |a_{j_s}^H x^{(t)}| > \gamma, \quad s = 1, \ldots, k. \quad (14)
\]

Using \( J^{(t)} \) corresponding to a well-converged value of \( \mu^{(t)} = \|g(x^{(t)})\| \), one then finds the maximum eigenvalue of the matrix \( A_J^H A_J \). The latter is considered as an approximation from below to the quantity \( \max_{|J| = k} \lambda_{\text{max}}(A_J^H A_J) \).

**Estimating the Lower Spectral Bound**

In order to reduce the problem size (and, more important, to preserve the column normalization property (5)), we will use the standard Naimark complement techniques, see, e.g., [2]. Recall that the Naimark complement of an \( m \times n \) tight frame \( A \) is the \( (n - m) \times n \) tight frame \( B \) satisfying, in particular, the identity \( mA^H A + (n - m)B^H B = nI_n \). Thus, we first evaluate the Gram matrix of \( B \):

\[
G = \frac{n}{n - m}I_n - \frac{m}{n - m}A^H A.
\]

Next we obtain explicitly the Naimark complement \( B \) using, for instance, the first \( n - m \) steps of the Cholesky factorization (with complete diagonal pivoting) applied to \( G \):

\[
G = P^T U^H U P, \quad B = U P,
\]

where \( U \) is an upper trapezoidal \( (n - m) \times n \)-matrix and \( P \) is a permutation matrix. Finally, we apply the above described power method to the matrix \( B \) (of course, with replacement of \( m \) by \( n - m \) where necessary). The required result readily follows from the obvious relation

\[
A_J^H A_J = \frac{n}{m}I_k - \frac{n - m}{m}B_J^H B_J,
\]

holding for any index subset \( J \), which yields

\[
\min_{|J| = k} \lambda_{\min}(A_J^H A_J) = \frac{n}{m} - \frac{n - m}{m} \max_{|J| = k} \lambda_{\text{max}}(B_J^H B_J).
\]

In fact, this means that one should use the index subsets \( J \) arising in the generalized power algorithm applied to \( B \) in order to form matrices \( A_J^H A_J \) and to determine their minimum eigenvalues.

**Choosing \( \gamma_1 \) and \( \gamma_2 \) in Practical Computations**

One of the most important features of the power method revealed in the course of numerical tests was that the segment (12) of desired values for the penalty parameter is quite narrow (typically, 2% of the total length of the interval (0, 1), see Table 5 below). Therefore, the use of rather loose theoretical upper bounds for \( \gamma_2 \) presented below (see Corollary) may not essentially accelerate computations, especially for the upper restricted isometry constant.
In order to improve the precision and speed of computations, the following heuristic was used. First one performs a uniform computational sweep over \( \gamma \in (0, 1) \) in order to roughly localize the desirable range of values (e.g. resulting in \( 1 < k < m \)) for the penalty parameter \( \gamma \). Next one uses the results of the first rough sweep to determine a much more tight range of values \([\gamma_1, \gamma_2]\) for the penalty parameter, and then the uniform sweep is repeated over this contracted interval.

**Necessary Condition for Compatibility of Penalty Parameter \( \gamma \) with Sparsity \( k \)**

Next we consider the relationship between the given value of \( \gamma \) in (8) or (11) and the corresponding sparsity degree \( k \) associated with the solution of problem (6), (7) or (9), (10), respectively. The obtained estimates confirm the expected correspondence of smaller values of \( \gamma \) to the larger values of \( k \).

**Analysis of L0-penalty Case**

Assume that conditions (5) hold, i.e., \( A \) is an \( m \times n \) column normalized tight frame. Multiplying (6) by \( x^H \) from the left and using (7), one has

\[
0 < \mu = \sum_{j=1}^{n} x^H a_j a_j^H x \left( \sign(1 - \gamma/|a_j^H x|) \right)_+ = \sum_{j=1}^{k} |a_j^H x|^2, \tag{15}
\]

where the index set \( J = \{j_1, \ldots, j_k\} \) is defined similar to (14) and, by this definition,

\[
|a_j^H x| > \gamma, \quad s = 1, \ldots, k. \tag{17}
\]

One also has, by the column normalization condition for \( A \) and (10)

\[
|a_j^H x| \leq 1, \quad s = 1, \ldots, k. \tag{18}
\]

On the other hand, the following upper estimate holds:

\[
\sum_{s=1}^{k} |a_{j_s}^H x|^2 = x^H \left( \sum_{s=1}^{k} a_{j_s} a_{j_s}^H \right) x = x^H A_J A_J^H x \leq \|A_J A_J^H\| = \|A_J^H A_J\| = \frac{\|A_J v_J\|^2}{\|v_J\|^2} \leq 1 + \delta_2(k; A), \tag{19}
\]

where \( v_J \) is the right singular vector of \( A_J \) corresponding to its maximum singular value and \( \delta_2(k; A) \) is the upper RIP constant for \( A \) as defined in (2). We now apply the latter estimate and inequality (17) to (16), which gives

\[
\mu \leq 1 + \delta_2(k; A).
\]

On the other hand, by (16) and (17) one has \( \mu \geq k \gamma^2 \). Thus, the following result is shown:

**Proposition 1.** Let \( A \) be an \( m \times n \) unit norm tight frame, where \( m < n \). Then, in the L0-penalty setting, for a fixed \( \gamma \) the degree of sparsity \( k \) as defined in (14) must satisfy

\[
k \gamma^2 \leq 1 + \delta_2(k; A) \tag{20}
\]
where $\delta_2(k; A)$ and $\mu$ are defined by (2) and (9), respectively.

This result confirms the observation that sufficiently small values of $\gamma$ must be used in computations to obtain the approximations to RIP constants of larger degrees $k$.

Now using the inequalities

$$1 + \delta_2(k; A) \leq \frac{n}{m}, \quad 1 + \delta_2(k; B) \leq \frac{n}{n-m}$$

(holding by (5) written for $A$ and its Naimark complement $B$), one has

**Corollary.** Under assumptions of the above Proposition, the following simplified bounds

$$k\gamma^2 \leq \frac{n}{m} \quad \text{and} \quad k\gamma^2 \leq \frac{n}{n-m}$$

hold for the estimation of upper and lower RIP constants, respectively.

**Analysis of L1-penalty Case**

Assume again that conditions (5) hold. Multiplying (9) by $x^H$ from the left and using (10), one has

$$0 < \mu = \sum_{j=1}^{n} x^H a_j a_j^H x \left(1 - \gamma / |a_j^H x| \right)_+$$

$$= \sum_{s=1}^{k} \left(|a_{j_s}^H x|^2 - \gamma |a_{j_s}^H x|\right),$$

where the index set $J = \{j_1, \ldots, j_k\}$ is defined as in (14) and, by this definition, (17) holds. As before, estimate (18) is valid by the column normalization condition for $A$ and (10).

Using inequalities (19) and (17) with (23), one has $\mu \leq 1 + \delta_2(k; A) - k\gamma^2$. Thus, the following result is shown:

**Proposition 2.** Let $A$ be an $m \times n$ unit norm tight frame, where $m < n$. Then, in the L1-penalty setting, for a fixed $\gamma$ the degree of sparsity $k$ as defined in (14) must satisfy

$$k\gamma^2 \leq 1 + \delta_2(k; A) - \mu$$

(24)

where $\delta_2(k; A)$ and $\mu$ are defined by (2) and (9), respectively.

Finally, using $\mu > 0$ and (21), one can see that Corollary of the previous Subsection holds true. These results confirm the observation that sufficiently small values of $\gamma$ must be used in computations to obtain the approximations for RIP constants of larger degrees $k$.

**Theoretical Spectral Bounds for an Arbitrary Tight Frame**

In absence of fast algorithms for determining the exact values of RIP constants, it is useful to have at least “inner” bounds for these quantities. Such bounds (in general, not tight for $k \geq 4$) were found in [11].

**Theorem.** Let $A$ be an $m \times n$ tight frame, where $m < n$. Then, in the RIP condition of the form

$$\|z\|_0 \leq k \quad \rightarrow \quad 0 < 1 - \delta_1(k) \leq \|Az\|^2 / \|z\|^2 \leq 1 + \delta_2(k),$$

(25)

303
the best possible RIP constants

\[ \delta_1(k) = 1 - \min_{|J|=k} \lambda_{\min}(A_J^H A_J), \quad \delta_2(k) = \max_{|J|=k} \lambda_{\max}(A_J^H A_J) - 1 \]

must satisfy

\[ \delta_1(k) \geq 1 - \frac{n}{2m} \left( 1 - \rho_{\max}(k, m, n) \right), \quad \delta_2(k) \geq \frac{n}{2m} \left( 1 - \rho_{\min}(k, m, n) \right) - 1, \] (26)

where \( \rho_{\min}(k, m, n) \) and \( \rho_{\max}(k, m, n) \) are the smallest and the largest roots of the Jacobi polynomial \( P_k^{(m-k,n-m-k)}(x) \), respectively.

In other words, for a prescribed size parameters \( k, m, n \) the quantities \( \delta_1 \) and \( \delta_2 \) will be separated from zero below as indicated above for any choice of elements in \( A \).

**Description of Test Problems**

We will consider several tight frames of the sizes \( n \approx m^2 \), as well as another cases with \( n \approx 2m \), all of them having rather good conditioning. Recall that the coherence between the columns of normalized frame \( A \) is measured by the parameter

\[ \mu(A) = \max_{k \neq l} |a_k^H a_l| \geq \sqrt{\frac{n - m}{(n - 1)m}}, \] (27)

which is always positive if \( n > m \). Smaller values of \( \mu(A) \) typically correlate with better performance of compressed sensing algorithms. This inequality presents the well-known Welch bound [14]. The equality in the Welch bound is attained for certain \( m \) and \( n \) if \( A \) forms an optimum equiangular tight frame (further on ETF).

**Partial DFT Tight Frames**

The partial Discrete Fourier Transform (further on DFT) matrices are both well explored theoretically and important in applications. Generally, these are determined by

\[ (A)_{i,j} = \frac{1}{\sqrt{m}} \exp \left( \frac{2\pi i}{n} d(i) j \right), \] (28)

that is, \( A \) is formed by picking out of the \( n \times n \) DFT matrix a subset of its \( m \) pairwise different rows. It is known that such matrix is an optimum ETF with \( \mu(A) = \sqrt{m - 1/m} \) if \( n = m^2 - m + 1 \) and the sequence \( \{d(j)\} | j = 0, \ldots, m - 1 \) forms the so called difference set, see, e.g., [12].

We will consider the following equiangular tight frames based on different sets. Using in (28) the Singer difference set for \( m = 17 \) and \( n = 273 \) given by

\[ d(i) \in \{0, 1, 3, 7, 15, 31, 63, 90, 116, 127, 136, 181, 194, 204, 233, 238, 255\}, \]

see [7, 15, 6], we obtain test matrix denoted as SINGER17x273. Another test matrix is defined by the choice \( m = 31 \) and \( n = 273 \) given by

\[ d(i) \in \{0, 1, 2, 4, 5, 7, 8, 9, 10, 14, 15, 16, 17, 18, 20, 27, 28, 30, 32, 34, 35, 36, 39, 40, 45, 49, 51, 54, 56, 57, 60\}, \]

and we will refer it to as SINGER31x63.
Biangular Tight Frame Based on Chirp Functions.

Following the constructions described in [1, 12], consider normalized tight frame determined as

\[ (A)_{i,j} = \frac{1}{\sqrt{m}} \exp \left( 2\pi i \left( \frac{m}{m} (i j_1 + i j_2) \right) \right), \quad 0 \leq i, j_1, j_2 \leq m - 1, \quad j = j_1 + mj_2, \tag{29} \]

where \( m \) is a prime number not smaller than 5, and \( n = m^2 \). Here we use \( m = 17 \), and the resulting normalized tight frame is referred below as CHIRP17x289. Note that the pairwise correlation between columns of such matrices takes values 0 or 1/\( \sqrt{m} \), i.e. the frame is biangular.

Equivolume Tight Frame Based on Conference Matrices

The following construction for \( n = 2^m \), where \( m = 2^q \), and \( q = 1, 2, \ldots \), was proposed in [10]. These matrices have not only ETF property, but also demonstrate the coincidence of values of all principal minors of the third order in the Gram matrix \( A^H A \). (Recall that for column normalized matrices \( A \), the ETF property is equivalent to the coincidence of values of all second order principal minors of \( A^H A \)). Thus, according to the results of [10, 11], \( A \) has optimum RIP constants \( \delta_1(k) \) and \( \delta_2(k) \) not only for \( k = 2 \), but also for \( k = 3 \). Such matrix \( A \) is constructed as follows:

\[ A = [(\alpha + (m-1)\beta\gamma)I_m + i(\beta + \alpha\gamma)C_m] \quad | \quad ((m-1)\beta\gamma - i\alpha\gamma)I_m + (\beta\gamma + i\alpha\gamma)C_m], \]

where

\[ \alpha = \sqrt{\frac{2m-1 + \sqrt{2m^2 - m}}{2m}}, \quad \beta = -\sqrt{\frac{2m-1 - \sqrt{2m^2 - m}}{2m^2 - m}}, \quad \gamma = \frac{1}{\sqrt{2m-1}}, \]

and the skew-symmetric conference matrices \( C_m \) (see, e.g., [12]) are determined recursively as

\[ C_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad C_{2s} = \begin{bmatrix} C_s & C_s - I_s \\ C_s + I_s & -C_s \end{bmatrix}, \quad s = 2, 4, 8, \ldots, 2^q-1. \]

Simultaneous change of sign at \( \sqrt{2m^2 - m} \) in the expressions for \( \alpha \) and \( \beta \) gives another family of equivolume tight frames with the same properties. For testing, we used \( m = 32 \) and the corresponding matrix is referred to as CONFSK32x64.

Equal Norm Tight Frame Based on Logistic Sequence

We also used ENTF’s of arbitrary sizes obtained from quasirandom data. Further, we denote by ENTFLOGmxn an \( m \times n \) matrix \( A \) obtained from the following two-side normalization/orthogonalization procedure.

Initialization:

\[ (A)_{i,j} = \xi_{6l} + \xi_{6l-1}i, \quad l = i + (j-1)m, \quad 1 \leq i \leq m, \quad 1 \leq j \leq n; \]

Iterations:

\[ A := A(\text{Diag}(A^H A))^{-1/2}; \quad LL^H = AA^H; \quad A := L^{-1}A. \]

Here the same logistic sequence [13] was used with the initial value \( \xi_0 = 0.4 \). Normally, these iterations converge in less than 100 steps to a certain ENTF.
Test Results and Discussion

For general complex-valued tight frames, parallel algorithms for the exact evaluation of RIP constants (by the direct search over \(\binom{n}{k}\) submatrices) and for their iterative approximation (using the above-described power method corresponding to the L1-penalty reformulation) were developed. These procedures were implemented as a FORTRAN/MPI code and successfully tested on MVS-10P multiprocessor system at Joint Supercomputing Center of Russian Academy of Sciences. The timing results show close to the peak performance attained by the algorithms due to their very low communication costs and good computational task balancing.

It must be noted that using the L0-penalty reformulation have demonstrated performance considerably inferior to that obtained with L1-penalty one. Hence, in order to shorten the presentation, we decided to omit them from the discussion. Even worse performance was observed in experiments with the Truncated Power method [18], which may be explained by extremely high multiplicities of sparse maximum eigenvalues for the tight frames tested.

In Tables below, we present the following three types of data:

(i) the exact values of \(\lambda_{\text{exact}}^\text{min}\) and \(\lambda_{\text{exact}}^\text{max}\) representing the RIP bounds (3) and (4), respectively, obtained (if available) by the direct search over all possible index subsets \(J\) (shown in the 2nd and 7th columns of tables);

(ii) the approximations \(\lambda_{\text{iter}}^\text{min}\) and \(\lambda_{\text{iter}}^\text{max}\) for the RIP bounds (3) and (4), respectively, obtained by the use of the generalized power method based on L1-penalty with \(l_{\text{max}} = 10000\) and \(t_{\text{max}} = 200\) (shown in the 3rd and 6th columns of tables);

(iii) the values of theoretical estimates \(\lambda_{\text{est}}^\text{min}\) and \(\lambda_{\text{est}}^\text{max}\) for the RIP bounds (3) expressed via the extreme roots of the Jacobi polynomials as assert Theorem in Section 5 (shown in the 4th and 5th columns of tables).

By the definition, for every fixed \(k\) it holds \(\lambda_{\text{exact}}^\text{min} \leq \lambda_{\text{iter}}^\text{min} \leq \lambda_{\text{est}}^\text{min} < \lambda_{\text{est}}^\text{max} \leq \lambda_{\text{iter}}^\text{max} \leq \lambda_{\text{exact}}^\text{max}\).

In all cases, especially for the smallest eigenvalues, the approximations to the RIP spectral bounds obtained by the generalized power method are remarkably close to their exact values (whenever the latter are available).

Numerical results for SINGER17x273 and CHIRP17x289 tests are presented in Tables 1 and 2. It appears that the equiangularity property is not sufficient for the best conditioning of the tight frame (compare the data for \(k = 6\)). One can also notice that the theoretical bounds \(\lambda_{\text{est}}^\text{min}\) and \(\lambda_{\text{est}}^\text{max}\) are getting the more and more loose as the sparsity parameter \(k\) increases. Additionally, the four last lines in the Table clearly indicate that the RIP property cannot be adequately characterized by the single quantity \(\delta(k)\) since \(\lambda_{\text{max}}(A_J^H A_J)\) may simply be larger than 2.

Numerical results for SINGER31x63 and CONFSK32x64 tests are given in Tables 3 and 4. An analysis of the data shows that two equiangular tight frames of near the same size may differ considerably in conditioning. In general, it seems that better conditioning of frames with elements equal to complex roots of unity is correlated with the presence of lower degree roots. This conjecture may explain why the current frame-related research is biased towards various generalizations of equiangular frame constructions such as fusion frames, biangular frames, etc.

Compared to the case \(n \approx m^2\), the case \(n \approx 2m\) demonstrates a much better agreement of theoretical bounds \(\lambda_{\text{est}}^\text{min}\) and \(\lambda_{\text{est}}^\text{max}\) with the exact spectral bounds.

The range of values for the penalty parameter \(\gamma\) sufficient for the estimation of RIP spectral bounds for \(2 \leq k \leq m\) are shown in Table 5.

Numerical results for CONFSKEW128x256 and ENTFLOG128x256 tests are shown in Fig. [1].
Table 1. Results obtained for SINGER17x273 test matrix

| k  | $\lambda_{\text{exact}}$ | $\lambda_{\text{iter}}$ | $\lambda_{\text{est}}$ | $\lambda_{\text{max}}$ | $\lambda_{\text{iter}}$ | $\lambda_{\text{exact}}$ |
|----|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 2  | 0.764705                 | 0.764705                 | 0.764705                 | 1.235294                 | 1.235294                 | 1.235294                 |
| 3  | 0.529511                 | 0.529513                 | 0.610996                 | 1.423842                 | 1.470556                 | 1.470588                 |
| 4  | 0.295141                 | 0.295141                 | 0.496458                 | 1.589132                 | 1.705335                 | 1.705657                 |
| 5  | 0.142126                 | 0.142126                 | 0.405971                 | 1.739835                 | 1.940651                 | 1.941176                 |
| 6  | 0.0                      | 0.0                      | 0.332152                 | 1.880286                 | 2.175688                 | 2.176470                 |
| 7  | 0.0                      | 0.0                      | 0.270764                 | 2.013013                 | 2.411764                 | n/a                      |
| 8  | 0.0                      | 0.0                      | 0.219119                 | 2.139645                 | 2.614694                 | n/a                      |
| 9  | 0.0                      | 0.0                      | 0.175390                 | 2.261301                 | 2.840312                 | n/a                      |

Table 2. Results obtained for CHIRP17x289 test matrix

| k  | $\lambda_{\text{exact}}$ | $\lambda_{\text{iter}}$ | $\lambda_{\text{est}}$ | $\lambda_{\text{max}}$ | $\lambda_{\text{iter}}$ | $\lambda_{\text{exact}}$ |
|----|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 2  | 0.757464                 | 0.757464                 | 0.764297                 | 1.235702                 | 1.242535                 | 1.242535                 |
| 3  | 0.514928                 | 0.514928                 | 0.610436                 | 1.424666                 | 1.485071                 | 1.485071                 |
| 4  | 0.272393                 | 0.272393                 | 0.495848                 | 1.590380                 | 1.727606                 | 1.727606                 |
| 5  | 0.164500                 | 0.164500                 | 0.405363                 | 1.741519                 | 1.944834                 | 1.944834                 |
| 6  | 0.071470                 | 0.080396                 | 0.331575                 | 1.882415                 | 2.171241                 | 2.171241                 |
| 7  | n/a                      | 0.037318                 | 0.270234                 | 2.015598                 | 2.379614                 | n/a                      |
| 8  | n/a                      | 0.025543                 | 0.218646                 | 2.142695                 | 2.590712                 | n/a                      |
| 9  | n/a                      | 0.009153                 | 0.174979                 | 2.264826                 | 2.813357                 | n/a                      |

appears that for larger sparsities $k > 20$ the quasirandom design is (likely) much better conditioned compared to the deterministic construction (despite its nearly optimum RIP bounds for small $k$).

Estimating the Spark of Frame

The quantity $\text{spark}(A) = |J|$ is defined as the minimum number of columns in a singular submatrix $A_J$. If $\text{spark}(A) = m + 1$, such matrix is called a full spark frame. Surprisingly, for almost all examples of deterministic test matrices considered above, the developed procedure for approximation RIP spectral bounds was able to detect the presence of rank deficient of $m \times k$ submatrices $A_J$ with $k$ considerably smaller than $m$. For the matrices CHIRP17x 289, SINGER17x289, SINGER31x63, CONFSK32x64, CONFSK64x128, CONFSK128x256, CONFSK256x512, the corresponding upper estimates of $\text{spark}(A)$ were 13, 6, 26, 15, 23, 31, 47, respectively. At the same time, for the quasirandom ENTF, no singularities were found (though with no guarantee), see Fig. 1.

Conclusions

Based on analysis of exactly evaluated RIP constants in the standard restricted isometry condition, it was observed that the conditioning of equiangular tight frames based on partial DFT and difference sets may be inferior compared to other frame designs. This well agrees with the known fact that equiangularity alone does not guarantee good RIP constants in general. Also, for relatively large $k$, it was observed that the quasirandom frame designs seemingly outperform the
Table 3. Results obtained for SINGER31x63 test matrix

| k  | $\lambda_{\text{exact}}$ | $\lambda_{\text{iter}}$ | $\lambda_{\text{est}}$ | $\lambda_{\text{iter}}$ | $\lambda_{\text{max}}$ | $\lambda_{\text{iter}}$ | $\lambda_{\text{max}}$ |
|----|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 2  | 0.870967                 | 0.870967                 | 0.870967                 | 1.129032                 | 1.129032                 | 1.129032                 |                         |
| 3  | 0.742018                 | 0.742018                 | 0.776866                 | 1.223666                 | 1.258064                 | 1.258064                 |                         |
| 4  | 0.613090                 | 0.613090                 | 0.699788                 | 1.301082                 | 1.386262                 | 1.387096                 |                         |
| 5  | 0.506383                 | 0.508021                 | 0.633917                 | 1.367573                 | 1.512575                 | 1.516129                 |                         |
| 6  | 0.388512                 | 0.388512                 | 0.575911                 | 1.426268                 | 1.615928                 | 1.645161                 |                         |
| 7  | 0.332970                 | 0.332970                 | 0.523929                 | 1.478993                 | 1.705561                 | 1.774193                 |                         |
| 8  | 0.266829                 | 0.266829                 | 0.476788                 | 1.526919                 | 1.788357                 | 1.840001                 |                         |
| 9  | 0.217257                 | 0.217257                 | 0.436750                 | 1.570852                 | 1.832258                 | 1.905896                 |                         |
| 10 | n/a                      | 0.162697                 | 0.394005                 | 1.611372                 | 1.892932                 | n/a                      |                         |
| 11 | n/a                      | 0.116808                 | 0.357339                 | 1.648913                 | 1.942023                 | n/a                      |                         |
| 12 | n/a                      | 0.057246                 | 0.323339                 | 1.683812                 | 1.991718                 | n/a                      |                         |
| 13 | n/a                      | 0.042094                 | 0.291739                 | 1.716331                 | 2.009415                 | n/a                      |                         |
| 14 | n/a                      | 0.026014                 | 0.262325                 | 1.746683                 | 2.024966                 | n/a                      |                         |
| 15 | n/a                      | 0.021193                 | 0.234925                 | 1.775040                 | 2.032258                 | n/a                      |                         |
| 16 | n/a                      | 0.012760                 | 0.209394                 | 1.801545                 | 2.032258                 | n/a                      |                         |

Figure 1. RIP constant estimates for CONFSKEW128x256 matrix and ENTFlog128x256 quasirandom matrix

known deterministic ETF designs.

For the larger values of $k$ (when the direct search for the exact spectral bounds has prohibitively
Table 4. Results obtained for CONFSK32x64 test matrix

| k  | $\lambda_{\text{exact}}$ | $\lambda_{\text{iter}}$ | $\lambda_{\text{est}}$ | $\lambda_{\text{est}}$ | $\lambda_{\text{iter}}$ | $\lambda_{\text{exact}}$ |
|----|---------------------------|---------------------------|-------------------------|-------------------------|---------------------------|-------------------------|
| 2  | 0.874011                  | 0.874011                  | 1.125988                | 1.125988                |                            |                         |
| 3  | 0.781782                  | 0.781782                  | 1.218217                | 1.218217                |                            |                         |
| 4  | 0.695837                  | 0.706435                  | 1.293564                | 1.304162                |                            |                         |
| 5  | 0.612248                  | 0.641796                  | 1.358203                | 1.387751                |                            |                         |
| 6  | 0.529805                  | 0.584789                  | 1.415210                | 1.470194                |                            |                         |
| 7  | 0.448009                  | 0.533626                  | 1.466373                | 1.551990                |                            |                         |
| 8  | 0.366614                  | 0.487157                  | 1.512842                | 1.633385                |                            |                         |
| 9  | 0.285485                  | 0.444592                  | 1.555407                | 1.714514                |                            |                         |
| 10 | n/a                       | 0.204542                  | 1.594638                | 1.795457                |                            |                         |
| 11 | n/a                       | 0.147006                  | 1.630960                | 1.852993                |                            |                         |
| 12 | n/a                       | 0.069238                  | 1.664704                | 1.930761                |                            |                         |
| 13 | n/a                       | 0.043024                  | 1.696127                | 1.956975                |                            |                         |
| 14 | n/a                       | 0.008097                  | 1.725437                | 1.991902                |                            |                         |
| 15 | n/a                       | 0.000000                  | 1.752805                | 2.000000                |                            |                         |
| 16 | n/a                       | 0.000000                  | 1.778369                | 2.000000                |                            |                         |

Table 5. Sufficient range of values for $\gamma$

| Name of test matrix | $\gamma_1$ | $\gamma_2$ |
|---------------------|------------|------------|
| CONFSK32x64         | 0.15       | 0.17       |
| CONFSK64x128        | 0.10       | 0.13       |
| Chirp17x289(max)    | 0.36       | 0.38       |
| Chirp17x289(min)    | 0.01       | 0.03       |
| Singer17x273(max)   | 0.36       | 0.38       |
| Singer17x273(min)   | 0.01       | 0.03       |

large costs), a version of the generalized power method [8] adjusted for the processing of complex-valued data was developed, implemented and successively tested. The evaluated estimates for the RIP constants of ETF’s nearly coincide with their exact values for small k (if available). In particular, it was possible to detect the linear dependence of nontrivial subsets of columns of $A$. A future research may be directed towards the certification of optimality for the obtained solutions using SDP relaxation techniques similar to [4]. Also, certain more efficient strategies for the choice of penalty parameter $\gamma$ and initial guess $x^{(0)}$ may likely be found.

Acknowledgements

This work was supported by the Presidium of the Russian Academy of Sciences Program No.26 “Theory and methods for the solution of computational problems on super-high-performance computer systems”.

References

[1] Lorne Applebaum, Stephen D Howard, Stephen Searle, and Robert Calderbank. Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery. Applied and
[2] Peter G Casazza, Matthew Fickus, Dustin G Mixon, Jesse Peterson, and Ihar Smalyanau. Every hilbert space frame has a naimark complement. *Journal of Mathematical Analysis and Applications*, 406(1):111–119, 2013.

[3] Peter G Casazza and Gitta Kutyniok. *Finite Frames. Applied and Numerical Harmonic Analysis*. Birkhäuser/Springer, New York, 2013.

[4] Alexandre d’Aspremont, Francis Bach, and Laurent El Ghaoui. Optimal solutions for sparse principal component analysis. *Journal of Machine Learning Research*, 9(Jul):1269–1294, 2008.

[5] Simon Foucart and Holger Rauhut. *A mathematical introduction to compressive sensing*, volume 1. Birkhäuser Basel, 2013.

[6] D Gordon. La jolla difference set repository. *URL* = https://www.ccrwest.org/diffsets.html (Accessed 15 February 2017), 2017.

[7] Marshall Hall. A survey of difference sets. *Proceedings of the American Mathematical Society*, 7(6):975–986, 1956.

[8] Michel Journé, Yurii Nesterov, Peter Richtárik, and Rodolphe Sepulchre. Generalized power method for sparse principal component analysis. *Journal of Machine Learning Research*, 11(Feb):517–553, 2010.

[9] Igor Kaporin. Approximating the restricted isometry constants for a tight frame using parametrized l1-penalty formulation. In *VIII International Conference on Optimization and Applications (OPTIMA-2017)* (Yuri G. Evtushenko e.a., eds.) Petrovac, Montenegro, October 2-7, 2017, P.282-290. CEUR Workshop Proceedings, Vol.1987, ISSN 1613-0073. *URL* = http://ceur-ws.org/Vol-1987/paper41.pdf.

[10] Igor E Kaporin. Characterizing the restricted isometry property in compressed sensing matrices via the k-condition number. In: Optimization and Applications, Issue 4 (Collection of papers of the Applied Optimization Problems Dept. of CCRAS, V.G.Zhadan, ed.) - Moscow: Comput. Center R.A.S. Publ.: 131–146, 2015.

[11] Igor E Kaporin. Bounding the restricted isometry constants for a tight frame. *Sbornik: Mathematics*, 208(11):1646–1650, 2017.

[12] Thomas Strohmer and Robert W Heath Jr. Grassmannian frames with applications to coding and communication. *Applied and computational harmonic analysis*, 14(3):257–275, 2003.

[13] Andreas M Tillmann and Marc E Pfetsch. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing. *IEEE Transactions on Information Theory*, 60(2):1248–1259, 2014.

[14] Lloyd Welch. Lower bounds on the maximum cross correlation of signals (corresp.). *IEEE Transactions on Information theory*, 20(3):397–399, 1974.

[15] Pengfei Xia, Shengli Zhou, and Georgios B Giannakis. Achieving the welch bound with difference sets. *IEEE Transactions on Information Theory*, 51(5):1900–1907, 2005.

[16] Lei Yu, Jean Pierre Barbot, Gang Zheng, and Hong Sun. Compressive sensing with chaotic sequence. *IEEE Signal Processing Letters*, 17(8):731–734, 2010.

[17] Lei Yu, Jean Pierre Barbot, Gang Zheng, and Hong Sun. Toeplitz-structured chaotic sensing matrix for compressive sensing. In *Communication Systems Networks and Digital Signal Processing (CSNDSP), 2010 7th International Symposium on*, pages 229–233. IEEE, 2010.

[18] Xiao-Tong Yuan and Tong Zhang. Truncated power method for sparse eigenvalue problems. *Journal of Machine Learning Research*, 14(Apr):899–925, 2013.