

Robust Measurement Design for Detecting Sparse Signals: Equiangular Uniform Tight Frames and Grassmannian Packings

Ramin Zahedi, Ali Pezeshki, and Edwin K. P. Chong

Abstract—Detecting a sparse signal in noise is fundamentally different from reconstructing a sparse signal, as the objective is to optimize a detection performance criterion rather than to find the sparsest signal that satisfies a linear observation equation. In this paper, we consider the design of low-dimensional (compressive) measurement matrices for detecting sparse signals in white Gaussian noise. We use a *lexicographic optimization* approach to maximize the worst-case signal-to-noise ratio (SNR). More specifically, we find an optimal solution for a k -sparse signal among optimal solutions subject to sparsity level $k - 1$. We show that for all sparse signals, columns of the optimal measurement matrix must form a *uniform tight frame*. For 2-sparse signals, the smallest angle among angles between element pairs of this frame must be maximized. In this case, the optimal solution matrix is an optimal *Grassmannian packing*. For k -sparse signals where $k > 2$, the largest angle among such angles must be as close to the maximum smallest angle as possible. We show that under certain conditions, columns of the optimal measurement matrix form an *equiangular uniform tight frame*. For this case, we derive an expression for the maximal SNR in the worst-case scenario, as a function of the signal dimension and the number of measurements.

I. INTRODUCTION

Over the past few years, considerable progress has been made towards developing a mathematical framework for *reconstructing* sparse or compressible signals. In particular, the advent of compressed sensing (see, e.g., [1]–[3]) has created a great deal of enthusiasm among the signal processing community, as it suggests that a high dimensional signal can be accurately reconstructed from a small number of measurements, using linear programming, provided that the signal is sparse in a known basis. However, little attention has been paid to statistical inference based on compressive measurements from sparse signals, which is the main objective in many sensing applications.

Detecting a sparse signal in noise is fundamentally different from reconstructing a sparse signal, as the objective in detection is to maximize the probability of detection or to minimize Bayes risk, rather than to find the sparsest signal that satisfies a linear observation equation. Therefore, sufficient conditions required in compressive sensing for signal recovery may not apply to signal detection. For instance, a sufficient condition for the so-called basis pursuit principle for sparse signal recovery is that the compressive measurement matrix must satisfy a restricted isometry

This work was supported in part by ONR Grant N00014-08-1-110 and NSF Grants ECCS-0700559 and CCF-0916314.

The authors are with the Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523-1373, USA. Emails: {ramin.zahedi, ali.pezeshki, edwin.chong}@colostate.edu.

property (RIP), or, equivalently, it must be incoherent with the sparsity basis for the signal [3]–[5]. However, it is not clear whether or not this condition is in any sense optimal for detecting sparse signals. The literature on sparse signal detection (see, e.g., [6]–[8]) is mainly focused on deriving bounds on the performance of Neyman-Pearson or Bayesian detectors when the compressive measurements are made with a random matrix, and not on the design of measurement matrices that optimize the detection performance.

In this paper, we consider the design of compressive measurement matrices for detecting sparse signals in white Gaussian noise. We consider the following binary hypothesis test:

$$\begin{cases} \mathcal{H}_0 : \mathbf{x} = \mathbf{n}, \\ \mathcal{H}_1 : \mathbf{x} = \mathbf{s} + \mathbf{n}, \end{cases} \quad (1)$$

where \mathbf{x} is an $(N \times 1)$ vector that describes the state of a physical phenomenon. Under the null hypothesis \mathcal{H}_0 , \mathbf{x} is a white Gaussian noise vector with covariance matrix $E[\mathbf{nn}^H] = \sigma_n^2 \mathbf{I}$. Under the alternative hypothesis \mathcal{H}_1 , $\mathbf{x} = \mathbf{s} + \mathbf{n}$ consists of a deterministic signal \mathbf{s} distorted by additive white Gaussian noise \mathbf{n} . We assume that the signal of interest \mathbf{s} is composed as

$$\mathbf{s} = \Psi \boldsymbol{\theta}, \quad (2)$$

where $\Psi \in \mathbb{R}^{N \times N}$ is a known matrix, whose columns form an orthonormal basis for \mathbb{R}^N , and $\boldsymbol{\theta}$ is a k -sparse ($k \ll N$) vector, which means that it has at most k nonzero elements (but at least one). In this case, we say that \mathbf{s} is sparse in the basis $\Psi = [\psi_1, \dots, \psi_N]$.

We wish to decide between the two hypotheses based on a limited number $m \leq N$ in the vector $\mathbf{y} = \Phi^H \mathbf{x}$ from \mathbf{x} , where $\Phi^H \in \mathbb{R}^{m \times N}$ is a compressive measurement matrix that we will design, and the superscript H is the Hermitian transpose. The observation vector $\mathbf{y} = \Phi^H \mathbf{x}$ belongs to one of the following hypothesized models:

$$\begin{cases} \mathcal{H}_0 : \mathbf{y} = \Phi^H \mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_n^2 \Phi^H \Phi), \\ \mathcal{H}_1 : \mathbf{y} = \Phi^H (\mathbf{s} + \mathbf{n}) \sim \mathcal{N}(\Phi^H \mathbf{s}, \sigma_n^2 \Phi^H \Phi). \end{cases} \quad (3)$$

We consider a log-likelihood linear detector (e.g., the Neyman-Pearson detector, which yields the maximum detection probability for a given SNR and false alarm rate). Since the detection performance for this detector is a monotonically increasing function of the SNR, we consider optimizing an SNR criterion for designing the matrix Φ . To avoid coloring the noise vector \mathbf{n} , we constraint the compressive measurement matrix Φ^H to be left orthogonal, that is we force $\Phi^H \Phi = \mathbf{I}$.

We use a *lexicographic optimization* (see, e.g., [9], [10], and [11]) approach to design the matrix Φ that maximizes the worst-case detection SNR, where the worst-case is with respect to the location of nonzero entries of θ and their values. This is a design for robustness with respect to the worst sparse signal that can be produced in the basis Ψ . We show that the worst-case detection SNR is maximized when the columns of the product $\Phi^H \Psi$ between the compressive measurement matrix Φ^H and the sparsity basis Ψ form a *uniform tight frame*. A uniform tight frame is a frame system, in which the frame operator is a scalar multiple of the identity operator and every frame element has the same norm. We also show that when the signal is 2-sparse, the smallest angle among angles between frame element pairs must be maximized. This means that the frame in this case is an optimal *Grassmannian packing* (see, e.g., [12], [13], and [14]). For the case where the sparsity level of the signal is greater than 2, we provide a lower bound on the worst-case performance. To maximize this lower bound, vector pairs of the optimal Grassmannian packing must have the minimum largest angle possible. Under certain conditions, the minimum and maximum angle between frame element pairs become equal, and we will have an *equiangular uniform tight frame* (see, e.g., [15]–[13]). For this case, we derive an expression for the maximal worst-case SNR in terms of the number of measurements m and the signal dimension N .

II. DETECTOR PERFORMANCE AND LEXICOGRAPHIC OPTIMIZATION

The log-likelihood ratio function (see, e.g., [20]) for (3) is given by

$$t(\mathbf{y}) = \mathbf{y}^H \Phi^H \mathbf{s},$$

and is distributed as

$$\begin{cases} \mathcal{H}_0 : t(\mathbf{y}) \sim \mathcal{N}(0, \sigma_n^4 \text{SNR}), \\ \mathcal{H}_1 : t(\mathbf{y}) \sim \mathcal{N}(\sigma_n^2 \text{SNR}, \sigma_n^4 \text{SNR}), \end{cases}$$

where the detection SNR is

$$\text{SNR} = (\mathbf{s}^H \Phi \Phi^H \mathbf{s}) / \sigma_n^2. \quad (4)$$

Our aim is to design Φ to maximize the worst-case SNR with respect to the location and values of the nonzero entries in θ . As pointed out earlier, the rationale for maximizing SNR is that detection performance is monotonically related to the SNR. For example, consider the Neyman-Pearson test of size γ , which is the log-likelihood ratio test

$$t(\mathbf{y}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \eta,$$

where the threshold η is chosen to maintain a constant false alarm probability γ . The detection probability is given by

$$P_D = Q\left(Q^{-1}(\gamma) - \sqrt{\text{SNR}}\right),$$

where $Q(\cdot)$ is the Q -function. This is the maximum detection probability for a given SNR under a constant false alarm rate γ constraint. The SNR, however, depends on the choice of the compressive measurement matrix Φ^H as in (4).

In practice, we might have signals that are not strictly sparse. However, we can often find a basis (e.g., Fourier or wavelet) in which the signal has k large coefficients. In this case, we consider the other coefficients to be negligible and treat them as zeros. For the detection problem, one approach is to assume a known value for k and design the measurement matrix Φ based on this assumption. This approach, however, runs the risk that the true sparsity level might be different.

An alternative approach is not to assume any specific sparsity level. Instead, when designing the measurement matrix Φ , we prioritize the level of importance of different values of sparsity k . In other words, we first find a set of solutions that are optimal for a k_1 -sparse signal. Then, within this set, we find a subset of solutions that are also optimal for k_2 -sparse signals. We follow this procedure until we find a subset that contains a family of optimal solutions for sparsity levels k_1, k_2, k_3, \dots . This approach is known as the *lexicographic optimization* method (see, e.g., [10] and [11]).

III. THE WORST-CASE PROBLEM STATEMENT

As mentioned above, we will use a lexicographic optimization approach to maximize the worst-case SNR. Since all sparse signals share the fact that they might only have one nonzero entry, it seems natural to start with finding an optimal measurement matrix for parameter vectors θ with one nonzero entry. Next, among the set of optimal solutions for this case, we find matrices that are optimal for vectors θ with two nonzero entries. This procedure is continued for vectors with more nonzero entries at each step.

Consider the k th step of the lexicographic approach. In this step, the vector θ has up to k nonzero entries. We do not impose any prior constraints on the locations and the values of the nonzero entries of θ . Without loss of generality, we assume that $\|\mathbf{s}\|^2 = \|\theta\|^2 = 1$. We wish to maximize the minimum (worst-case) SNR, produced by assigning the worst possible locations and values to the nonzero entries of the k -sparse vector θ . Referring to (4), this is a worst-case design for maximizing the signal energy $\mathbf{s}^H \Phi \Phi^H \mathbf{s}$ inside the subspace $\langle \Phi \rangle$ spanned by the columns of Φ , since $\Phi \Phi^H$ is the orthogonal projection operator onto $\langle \Phi \rangle$.

To define the k th step of the optimization procedure more precisely, we need some additional notation. Let \mathcal{A}_0 be the set containing all $(N \times m)$ left orthogonal matrices Φ . Then, we recursively define the set \mathcal{A}_k , $k = 1, 2, \dots$, as the set of solutions to the following optimization problem:

$$\begin{aligned} \max_{\Phi} \min_{\mathbf{s}} \quad & \|\Phi^H \mathbf{s}\|^2, \\ \text{s.t.} \quad & \Phi^H \Phi = \mathbf{I}, \Phi \in \mathcal{A}_{k-1}, \\ & \|\mathbf{s}\| = 1. \end{aligned} \quad (5)$$

In our lexicographic formulation, the optimization problem for the k th problem (5) involves a worst-case objective restricted to the set of solutions \mathcal{A}_{k-1} from the $(k-1)$ th problem. So, $\mathcal{A}_k \subset \mathcal{A}_{k-1} \subset \dots \subset \mathcal{A}_0$.

Before we present a complete solution to these problems, we first simplify them in three steps. First, since the matrix

Ψ is known, the matrix Φ can be written as

$$\Phi = \Psi \mathbf{C},$$

where \mathbf{C} is an $(N \times m)$ matrix. Then, $\Phi^H \Psi = \mathbf{C}^H \Psi^H \Psi = \mathbf{C}^H$, and also $\Phi^H \Phi = \mathbf{C}^H \Psi^H \Psi \mathbf{C} = \mathbf{C}^H \mathbf{C} = \mathbf{I}$. Using (2), the max-min problems (5) become

$$\begin{aligned} \max_{\mathbf{C}} \min_{\boldsymbol{\theta}} \quad & \|\mathbf{C}^H \boldsymbol{\theta}\|^2, \\ \text{s.t.} \quad & \mathbf{C}^H \mathbf{C} = \mathbf{I}, \mathbf{C} \in \mathcal{B}_{k-1}, \\ & \|\boldsymbol{\theta}\| = 1, \end{aligned} \quad (6)$$

where, similar to the sets \mathcal{A}_k , the sets \mathcal{B}_k ($k = 1, 2, \dots$) are recursively defined to contain all the optimal solutions of (6). It is easy to see that $\mathcal{B}_k = \{\mathbf{C} : \Psi \mathbf{C} \in \mathcal{A}_k\}$.

Let Ω be the set $\Omega = \{1, 2, \dots, N\}$. Consider a nonempty subset T of Ω with cardinality $|T| = k$. Given a vector $\boldsymbol{\theta}$, let $\boldsymbol{\theta}_T$ be the subvector of size $(k \times 1)$ that contains all the components of $\boldsymbol{\theta}$ corresponding to indices in T . Similarly, given a matrix \mathbf{C} , let \mathbf{C}_T^H be the $(m \times k)$ submatrix consisting of all columns of \mathbf{C}^H whose indices are in T . Now, suppose that $\boldsymbol{\theta}$ has at most k nonzero elements. Then, $\mathbf{C}^H \boldsymbol{\theta}$ can be written as $\mathbf{C}_T^H \boldsymbol{\theta}_T$ for some T . Here, the elements of T include the location of the nonzero elements of $\boldsymbol{\theta}$. If we replace $\mathbf{C}^H \boldsymbol{\theta}$ with $\mathbf{C}_T^H \boldsymbol{\theta}_T$ in the max-min problem, then besides considering the worst $\boldsymbol{\theta}_T$ that minimizes $\|\mathbf{C}_T^H \boldsymbol{\theta}_T\|^2$, we also have to take into account the case where the set T consists of locations in $\boldsymbol{\theta}$ that cause $\|\mathbf{C}_T^H \boldsymbol{\theta}_T\|^2$ to be minimum. Thus, the max-min problem becomes

$$\begin{aligned} \max_{\mathbf{C}} \min_T \min_{\boldsymbol{\theta}_T} \quad & \|\mathbf{C}_T^H \boldsymbol{\theta}_T\|^2, \\ \text{s.t.} \quad & \mathbf{C}^H \mathbf{C} = \mathbf{I}, \mathbf{C} \in \mathcal{B}_{k-1}, \\ & \|\boldsymbol{\theta}_T\| = 1, |T| = k. \end{aligned} \quad (7)$$

The solution to (7) is the most robust design with respect to the locations and values of the nonzero entries of the parameter vector $\boldsymbol{\theta}$.

The solution to the minimization subproblem

$$\begin{aligned} \min_{\boldsymbol{\theta}_T} \quad & \|\mathbf{C}_T^H \boldsymbol{\theta}_T\|^2, \\ \text{s.t.} \quad & \|\boldsymbol{\theta}_T\| = 1, \end{aligned}$$

is well known; see, e.g., [21]. The optimal objective function is $\lambda_{\min}(\mathbf{C}_T \mathbf{C}_T^H)$, the smallest eigenvalue of the matrix $\mathbf{C}_T \mathbf{C}_T^H$. Therefore, the max-min-min problem (7) simplifies to

$$(P_k) \quad \begin{cases} \max_{\mathbf{C}} \min_T \quad & \lambda_{\min}(\mathbf{C}_T \mathbf{C}_T^H), \\ \text{s.t.} \quad & \mathbf{C}^H \mathbf{C} = \mathbf{I}, \mathbf{C} \in \mathcal{B}_{k-1}, \\ & |T| = k. \end{cases} \quad (8)$$

At each step k , the optimal compressive measurement matrix, denoted by Φ^{*H} , is determined from the optimizer \mathbf{C}^* of (8) as $\Phi^{*H} = \mathbf{C}^{*H} \Psi^H$. Next, we describe how to solve the max-min problem (P_k) in (8).

IV. SOLUTION TO THE WORST-CASE PROBLEM

Let \mathbf{c}_i be the i th column of the matrix \mathbf{C}^H . As mentioned earlier, we first find the solution set \mathcal{A}_1 for problem (P_1) . Then, we find a subset $\mathcal{A}_2 \subset \mathcal{A}_1$ as the solution for (P_2) . We continue this procedure for general sparsity level k .

A. Sparsity Level $k = 1$

If $k = 1$, then any T such that $|T| = 1$ can be written as $T = \{i\}$ with $i \in \Omega$, and $\mathbf{C}_T^H = \mathbf{c}_i$ consists of only the i th column of \mathbf{C}^H . Therefore,

$$\mathbf{C}_T \mathbf{C}_T^H = \mathbf{c}_i^H \mathbf{c}_i = \|\mathbf{c}_i\|^2,$$

and the max-min problem becomes

$$\begin{aligned} \max_{\mathbf{C}} \min_i \quad & \|\mathbf{c}_i\|^2, \\ \text{s.t.} \quad & \mathbf{C}^H \mathbf{C} = \mathbf{I}, \mathbf{C} \in \mathcal{B}_0, \\ & i \in \Omega. \end{aligned} \quad (9)$$

Because \mathcal{B}_0 is the set of $(N \times m)$ matrices \mathbf{C} with the property that $\mathbf{C}^H \mathbf{C} = \mathbf{I}$, the constraint $\mathbf{C} \in \mathcal{B}_0$ can be ignored.

Theorem 1: The optimal value of the objective function of the max-min problem (9) is m/N . A necessary and sufficient condition for a matrix \mathbf{C}^* to be in the solution set \mathcal{B}_1 is that the columns $\{\mathbf{c}_i^*\}_{i=1}^N$ of \mathbf{C}^{*H} form a *uniform tight frame* with norm values equal to $\sqrt{m/N}$.

Proof: We first prove the claim about the optimal value. Assume false, i.e., assume there exists an optimal matrix $\mathbf{C}^* \in \mathcal{B}_1$ for which the value of the cost function is either less than or greater than m/N . Suppose the former is true. Let \mathbf{C}_1^H be an $(m \times N)$ matrix, satisfying $\mathbf{C}_1^H \mathbf{C}_1 = \mathbf{I}$, whose columns have equal norm $\sqrt{m/N}$. Then, the value of the objective function in (9) for $\mathbf{C} = \mathbf{C}_1$ is m/N . This means that our proposed matrix \mathbf{C}_1 achieves a higher SNR than \mathbf{C}^* which is a contradiction. Now, assume the latter is correct, that is the value of the objective function for \mathbf{C}^* is greater than m/N . This means that

$$\min_{i \in \Omega} \|\mathbf{c}_i^*\|^2 = \|\mathbf{c}_j^*\|^2 > m/N.$$

Knowing this, we write

$$\begin{aligned} \text{tr}(\mathbf{C}^{*H} \mathbf{C}^*) &= \text{tr}(\mathbf{C}^* \mathbf{C}^{*H}) = \sum_{i=1}^N \|\mathbf{c}_i^*\|^2 \\ &> \sum_{i=1}^N m/N = m. \end{aligned}$$

However, from the constraint in (9) we know that $\mathbf{C}^{*H} \mathbf{C}^* = \mathbf{I}$, and $\text{tr}(\mathbf{C}^{*H} \mathbf{C}^*) = m$. This is also a contradiction. Thus, the assumption is false and the optimal value for the objective function of (9) is m/N .

We now prove the claim about the optimizer \mathbf{C}^* . From the preceding part of the proof, it is easy to see that all columns of \mathbf{C}^{*H} must have equal norm $\sqrt{m/N}$. If not, since none of them can be less than $\sqrt{m/N}$, then the sum of all column norms will be greater than m , which is a contradiction. Moreover, we write

$$\mathbf{C}^{*H} \mathbf{C}^* = \sum_{i=1}^N \mathbf{c}_i^* \mathbf{c}_i^{*H} = \mathbf{I}. \quad (10)$$

Multiplying both sides of (10) by an arbitrary $(m \times 1)$ vector \mathbf{x} from the right side and \mathbf{x}^H from the left side, we get

$$\sum_{i=1}^N \|\mathbf{c}_i^{*H} \mathbf{x}\|^2 = \|\mathbf{x}\|^2.$$

This equation represents a tight frame with frame elements $\{\mathbf{c}_i^*\}$ and frame bound 1. In other words, it represents a Parseval frame. Since the frame elements have equal norms, the frame is also uniform. Therefore, for a matrix \mathbf{C}^* to be in \mathcal{B}_1 , the columns of \mathbf{C}^{*H} must form a *uniform tight frame*. ■

This completes the $k = 1$ case.

Remark 1: The reader is referred to [15]–[14], and the references therein, for examples of constructions of uniform tight frames.

B. Sparsity Level $k = 2$

The next step is to solve (P₂). Since our solution for this case should lie among the family of optimal solutions for $k = 1$, results concluded in the previous part should also be taken into account, i.e., the columns of the optimal matrix \mathbf{C}^{*H} must form a uniform tight frame, where the frame elements \mathbf{c}_i^* have norm $\sqrt{m/N}$.

For $T \subset \Omega$ such that $|T| = 2$, the matrix \mathbf{C}_T^H consists of two columns, e.g., \mathbf{c}_i and \mathbf{c}_j . So, the matrix $\mathbf{C}_T \mathbf{C}_T^H$ in the max-min problem (8) is a (2×2) matrix:

$$\mathbf{C}_T \mathbf{C}_T^H = \begin{bmatrix} \langle \mathbf{c}_i, \mathbf{c}_i \rangle & \langle \mathbf{c}_i, \mathbf{c}_j \rangle \\ \langle \mathbf{c}_i, \mathbf{c}_j \rangle & \langle \mathbf{c}_j, \mathbf{c}_j \rangle \end{bmatrix}.$$

From the $k = 1$ case, we have $\|\mathbf{c}_i\|^2 = \|\mathbf{c}_j\|^2 = m/N$. Therefore,

$$\mathbf{C}_T \mathbf{C}_T^H = (m/N) \begin{bmatrix} 1 & \cos \alpha_{ij} \\ \cos \alpha_{ij} & 1 \end{bmatrix},$$

where α_{ij} is the angle between vectors \mathbf{c}_i and \mathbf{c}_j . The minimum eigenvalue of this matrix is

$$\lambda_{\min}(\mathbf{C}_T \mathbf{C}_T^H) = (m/N)(1 - |\cos \alpha_{ij}|).$$

We assume for simplicity that $\alpha_{ij} \leq \pi/2$ (justified for the case where $m \ll N$), so that

$$\lambda_{\min}(\mathbf{C}_T \mathbf{C}_T^H) = (m/N)(1 - \cos \alpha_{ij}). \quad (11)$$

Let α_{kl} be the minimum angle among angles of all possible vector pairs \mathbf{c}_i and \mathbf{c}_j satisfying the constraint of (P₂). Let α be the maximum possible value of α_{kl} . So,

$$\alpha \leq \alpha_{ij}, \quad i, j \in \Omega, i \neq j. \quad (12)$$

Theorem 2: The optimal value of the objective function of the max-min problem (P₂) is $(m/N)(1 - \cos \alpha)$. A matrix \mathbf{C}^* is in \mathcal{B}_2 if and only if the columns of \mathbf{C}^{*H} form a *uniform tight frame* with norm values $\sqrt{m/N}$ and the minimum angle among angles between column pairs is α .

Proof: Since our solution must be chosen from the family of uniform tight frames with frame elements of equal norm $\sqrt{m/N}$, the objective function of (P₂) is only a function of the angle α_{ij} . Using (11), it is easy to see that the minimum $\lambda_{\min}(\mathbf{C}_T \mathbf{C}_T^H)$ is $(m/N)(1 - \cos \alpha_{kl})$. Using (12), we conclude that the largest possible value of the objective function of (P₂) is $(m/N)(1 - \cos \alpha)$. Note that if we consider any other uniform tight frame with elements having norms equal to $\sqrt{m/N}$ and a minimum angle β among angles of all possible pairs of frame elements, then because

$\beta \leq \alpha$, the value of the corresponding objective function is less than $(m/N)(1 - \cos \alpha)$. ■

This completes the $k = 2$ case.

Remark 2: Methods for constructing uniform tight frames with frame elements that have a maximum smallest angle among angles of frame element pairs is equivalent to optimal *Grassmannian packings* of one-dimensional subspaces (see, e.g., [15], [13], and [12]). We will say more about this point later in the paper.

C. Sparsity Level $k > 2$

We now consider cases where $k > 2$. In this case, $T \subset \Omega$ with $|T| = k$ can be written as $T = \{i_1, i_2, \dots, i_k\}$ where $i_h \in \Omega$ for $h \in \{1, \dots, k\}$. From the previous results, we know that an optimal matrix $\mathbf{C}^* \in \mathcal{B}_k$ must already satisfy two properties, in addition to $\mathbf{C}^{*H} \mathbf{C}^* = \mathbf{I}$:

- Columns of \mathbf{C}^{*H} must build a uniform tight frame with equal norms $\sqrt{m/N}$ (to be in the set \mathcal{B}_1),
- The minimum angle among angles of all possible column pairs of \mathbf{C}^{*H} must be equal to the maximum possible such angle α (to be in the set \mathcal{B}_2).

Taking the above properties into account for \mathbf{C}^* , the matrix $\mathbf{C}_T^* \mathbf{C}_T^{*H}$ will be a $(k \times k)$ symmetric matrix that can be written as $\mathbf{C}_T^* \mathbf{C}_T^{*H} = (m/N)[\mathbf{I} + \mathbf{A}_T]$ where \mathbf{A}_T is

$$\mathbf{A}_T = \begin{bmatrix} 0 & \cos \alpha_{i_1 i_2} & \dots & \cos \alpha_{i_1 i_k} \\ \cos \alpha_{i_1 i_2} & 0 & \dots & \cos \alpha_{i_2 i_k} \\ \vdots & \vdots & \ddots & \vdots \\ \cos \alpha_{i_1 i_k} & \cos \alpha_{i_2 i_k} & \dots & 0 \end{bmatrix},$$

where $i_h \neq i_f \in T$ for the entry $\cos \alpha_{i_h i_f}$ in the (i_h, i_f) th location. Then,

$$\lambda_{\min}(\mathbf{C}_T^* \mathbf{C}_T^{*H}) = (m/N)(1 + \lambda_{\min}(\mathbf{A}_T)). \quad (13)$$

Let $\{\alpha_{i_h i_f}\}$ be the collection of k largest angles among angles between column pairs of the matrix \mathbf{C}^{*H} that satisfy the constraint in (P_k). Also, let T_1 be the set of indexes of these angles. Thus,

$$\alpha \leq \alpha_{i_h i_j} \leq \alpha_{i_l i_f}, \quad i_h \neq i_f \in T_1, i_l \neq i_j \in T \neq T_1.$$

Moreover, let $\delta_{i_l i_j}$ be defined as

$$\delta_{i_l i_j} = \cos \alpha - \cos \alpha_{i_l i_j}, \quad i_l \neq i_j \in T.$$

It is easy to see that

$$\delta_{i_h i_f} \geq \delta_{i_l i_j}, \quad i_h \neq i_f \in T_1, i_l \neq i_j \in T \neq T_1.$$

The following theorem holds.

Theorem 3: The optimal value of the objective function of the max-min problem (P_k) for $k > 2$ lies between $(m/N)(1 - \cos \alpha - \sum_{i_h \neq i_f \in T_1} \delta_{i_h i_f})$ and $(m/N)(1 - \cos \alpha)$.

Proof: Let \mathbf{x}_{ij} be a $(k \times 1)$ vector that contains values $(1/\sqrt{2})$ and $(-1/\sqrt{2})$ in the i th and j th locations ($i \neq j$) and zeros elsewhere. Then, by using Raleigh's inequality for the matrix \mathbf{A}_T defined above and the family of vectors $\{\mathbf{x}_{ij}\}$ defined by i and j , we conclude that

$$\lambda_{\min}(\mathbf{A}_T) \leq -\cos \alpha_{i_l i_j}, \quad i_l \neq i_j \in T.$$

Thus,

$$\min_T \lambda_{\min}(\mathbf{A}_T) \leq \min_T (-\cos \alpha_{i_l i_j}) = -\cos \alpha. \quad (14)$$

On the other hand, the matrix \mathbf{A}_T can be written as $\mathbf{A}_T = \cos \alpha \mathbf{B} + \mathbf{F}_T$ where \mathbf{B} is a matrix with zeros on the diagonal and ones elsewhere, and \mathbf{F}_T is a symmetric matrix with zeros on the diagonal and the value $-\delta_{i_l i_j}$ in the (l, j) th location for $i_l \neq i_j \in T$. Then,

$$\begin{aligned} \lambda_{\min}(\mathbf{A}_T) &\geq \cos \alpha \lambda_{\min}(\mathbf{B}) + \lambda_{\min}(\mathbf{F}_T) \\ &= -\cos \alpha + \lambda_{\min}(\mathbf{F}_T). \end{aligned}$$

The matrix \mathbf{F}_T can be written as $\mathbf{F}_T = \sum_{i_l \neq i_j \in T} \mathbf{F}_{i_l i_j}$ where $\mathbf{F}_{i_l i_j}$ is a symmetric matrix with the value $-\delta_{i_l i_j}$ in the (l, j) th location and zeros elsewhere. Using matrix properties (see, e.g., [22]), we can write

$$\lambda_{\min}(\mathbf{F}_T) \geq \sum_{i_l \neq i_j \in T} \lambda_{\min}(\mathbf{F}_{i_l i_j}) = -\sum_{i_l \neq i_j \in T} \delta_{i_l i_j}.$$

Thus,

$$\lambda_{\min}(\mathbf{A}_T) \geq -\cos \alpha - \sum_{i_l \neq i_j \in T} \delta_{i_l i_j}.$$

It is easy to conclude that

$$\min_T \lambda_{\min}(\mathbf{A}_T) \geq -\cos \alpha - \sum_{i_h \neq i_f \in T_1} \delta_{i_h i_f}. \quad (15)$$

Using (13), (14), and (15) we get

$$\begin{aligned} (m/N)(1 - \cos \alpha - \sum_{i_h \neq i_f \in T_1} \delta_{i_h i_f}) &\leq \min_T \lambda_{\min}(\mathbf{C}_T^* \mathbf{C}_T^{*H}) \\ &\leq (m/N)(1 - \cos \alpha). \end{aligned} \quad (16)$$

D. Equiangular Uniform Tight Frames and Grassmannian Packings

The inequality (16) in Theorem (3) suggests that if the largest and smallest angles among angles between column pairs are equal, then the optimal value of the objective function of (\mathbf{P}_k) for $k > 2$ will reach its upper bound. In this case, the columns of \mathbf{C}^{*H} (where $\mathbf{C}^* \in \mathcal{B}_k$) in fact form an *equiangular uniform tight frame*. Equiangular uniform tight frames are optimal Grassmannian packings, where a collection of N one-dimensional subspaces are packed in \mathbb{R}^m so that the chordal distance between each pair of subspaces is maximal (see, e.g., [12], [15], and [13]). Each one-dimensional subspace is the span of one of the frame element vectors \mathbf{c}_i . The chordal distance between the i th subspace $\langle \mathbf{c}_i \rangle$ and the j th subspace $\langle \mathbf{c}_j \rangle$ is given by

$$d_c(i, j) = \sqrt{\sin^2 \alpha_{ij}}, \quad (17)$$

where α_{ij} is the angle between \mathbf{c}_i and \mathbf{c}_j . When all the α_{ij} , $i \neq j$, are equal and the frame is tight, the chordal distances between all pairs of subspaces become equal, i.e., $d_c(i, j) = d_c$ for all $i \neq j$, and they take their maximum value. This maximum value is the simplex bound given by

$$d_c = \sqrt{(N(m-1))/(m(N-1))}. \quad (18)$$

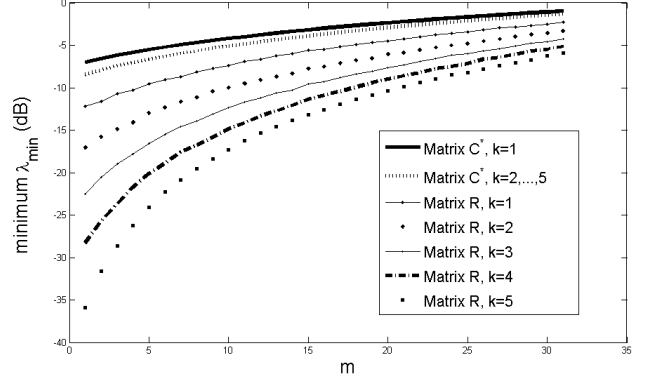


Fig. 1. Performance comparison between matrices \mathbf{C}^* and \mathbf{R} for some equiangular cases

This bound, however, can only be reached for some values of m and N . It is shown in [18] that vectors \mathbf{c}_i could be equiangular only when $1 < m < N - 1$ and

$$N \leq \min\{m(m+1)/2, (N-m)(N-m+1)/2\} \quad (19)$$

for frames with real elements, and

$$N \leq \min\{m^2, (N-m)^2\} \quad (20)$$

for frames with complex elements. If the above conditions hold, then the optimal solution for (\mathbf{P}_k) for $k > 2$ is a matrix \mathbf{C}^{*H} such that its columns form an equiangular uniform tight frame with frame elements of equal norm $\sqrt{m/N}$ and angle α defined as

$$\alpha = \arcsin \left(\sqrt{\left(\frac{m-1}{m}\right) \left(\frac{N}{N-1}\right)} \right). \quad (21)$$

The optimal value of the objective function of (\mathbf{P}_k) in this case is $(m/N)(1 - \cos \alpha)$.

In other cases where N and m do not satisfy the condition (19) or (20), the bound (16) suggests that we use an optimal Grassmannian packing where the k largest angles among angles between column pairs of the matrix \mathbf{C}^{*H} are as close to the angle α as possible. This is, however, a very difficult problem since even finding optimal Grassmannian packings for different values of N and m is still an open problem. The reader is referred to [12] and [15] for the state of the art in this field.

We have thus considered a worst-case design criterion in which we assume nothing about the vector $\boldsymbol{\theta}$, and our design is robust against arbitrary possibilities of this unknown.

V. SIMULATION RESULTS

We have compared the performance of our robust (worst-case) designed matrix \mathbf{C}^* with that of a random matrix \mathbf{R} with i.i.d Gaussian $\mathcal{N}(0, (1/m))$ entries, which is typically used in compressive sensing for signal recovery (see, e.g., [3]). To satisfy the constraint in problem (8), we make \mathbf{R} to be left orthogonal. We have run two sets of simulations. In both cases, the value of the objective function in (8) when

TABLE I

PERFORMANCE COMPARISON BETWEEN MATRICES \mathbf{C}^* AND \mathbf{R} FOR SOME NON-EQUIANGULAR CASES

m	N	minimum λ_{\min} (dB)			
		Matrix \mathbf{C}^*		Matrix \mathbf{R}	
		$k = 1$	$k = 2$	$k = 1$	$k = 2$
4	40	-10	-16.517	-19.393	-33.979
6	36	-7.78	-10.793	-14.609	-22.596
9	48	-7.27	-9.03	-12.815	-15.297

the matrix \mathbf{R} is used is an average taken over objective functions of 100 realizations of the matrix \mathbf{R} . In the first case, the value m increases from 10 to 40 and $N = 50$. For such values, the condition (19) is satisfied and columns of the optimal matrix \mathbf{C}^{*H} form an equiangular uniform tight frame. Figure 1 shows the comparison between the performance of our designed matrix \mathbf{C}^* with the matrix \mathbf{R} for $k = 1, 2, \dots, 5$.

For cases where N and m do not satisfy the condition (19), we were unable to find an optimal matrix \mathbf{C}^* for $k > 2$. But, for cases $k = 1$ and 2, we found three optimal matrices from the website [23]. Table I shows the comparison between the performance of these matrices and the matrix \mathbf{R} .

Note that values of the objective functions in Figure 1 and Table I are in dB. As can be seen, in all scenarios, the performance of the optimal matrix \mathbf{C}^* is better than the matrix \mathbf{R} .

VI. CONCLUSIONS

In this paper, we have considered the design of low-dimensional (compressive) measurement matrices for detecting sparse signals in white Gaussian noise. The detector could be any log-likelihood detector (e.g., the Neyman-Pearson detector) since for all such detectors, the detection performance is an increasing function of the SNR. We have found optimal solutions for the problem of maximizing the worst-case detection SNR, and consequently the worst-case detection probability for 1- and 2-sparse signals. When the signal's sparsity level is larger than 2, we have found lower and upper bounds on the performance of the optimizer, which meet under certain conditions. We have given an expression for the maximal SNR in the worst-case scenario, as a function of the signal dimension and the number of measurements, by utilizing the equivalence between equiangular uniform tight frames and optimal Grassmannian packings of one-dimensional subspaces.

REFERENCES

- [1] E. Candés, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 489–509, 2006.
- [2] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [3] R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, 2007.
- [4] E. Candés, "Compressive sampling," in *Int. Congress of Mathematics*, Madrid, Spain, Aug. 22–30 2006, vol. 3, pp. 1433–1452.
- [5] E. Candés, "The restricted isometry property and its implications for compressed sensing," *Compte Rendus de l'Academie des Sciences*, vol. 346, pp. 589–592, 2008.
- [6] J. Haupt and R. Nowak, "Compressive sampling for signal detection," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2007, vol. 3, pp. III–1509–III–1512.
- [7] M. F. Duarte, M. A. Davenport, M. B. Wakin, and R. G. Baraniuk, "Sparse signal detection from incoherent projections," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2006, vol. 3, pp. III–305–III–308.
- [8] M. F. Davenport, M. B. Wakin, and R. G. Baraniuk, "Detection and estimation with compressive measurements," Tech. Rep. TREE 0610, ECE Department, Rice University, 2006.
- [9] M. Ehrgott, *Multicriteria Optimization*, Springer, 2nd edition, June 2005.
- [10] H. Isermann, "Linear lexicographic optimization," *OR Spectrum*, vol. 4, no. 4, pp. 223–228, 1982.
- [11] B. Hajek and P. Seri, "Lex-optimal online multiclass scheduling with hard deadlines," *Mathematics of Operations Research*, vol. 30, no. 3, pp. 562–596, 2005.
- [12] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Experimental Mathematics*, vol. 5, no. 2, pp. 139–159, 1996.
- [13] G. Kutyniok, A. Pezeshki, R. Calderbank, and T. Liu, "Robust dimension reduction, fusion frames, and Grassmannian packings," *Applied and Computational Harmonic Analysis*, vol. 26, no. 1, pp. 64–76, 2009.
- [14] T. Strohmer, "A note on equiangular tight frames," *Linear Algebra and its Applications*, vol. 429, no. 1, pp. 326–330, 2008.
- [15] T. Strohmer and R. W. Heath Jr., "Grassmannian frames with applications to coding and communication," *Appl. Comput. Harmon. Anal.*, vol. 14, no. 3, pp. 257–275, 2003.
- [16] P. G. Casazza and J. Kovacevic, "Equal-norm tight frames with erasures," *Appl. Comput. Harmon. Anal.*, vol. 18, no. 2–4, pp. 387–430, 2003.
- [17] J. Renes, "Equiangular tight frames from Paley tournaments," *Linear Algebra Appl.*, vol. 426, no. 2–3, pp. 497–501, 2007.
- [18] M. Sustik, J. A. Tropp, I. S. Dhillon, and R. W. Heath Jr., "On the existence of equiangular tight frames," *Linear Algebra Appl.*, vol. 426, no. 2–3, pp. 619–635, 2007.
- [19] V. N. Malozemov and A. B. Pevnyi, "Equiangular tight frames," *Journal of Mathematical Sciences*, vol. 157, no. 6, pp. 789–815, 2009.
- [20] L. L. Scharf, *Statistical Signal Processing*, Addison-Wesley, Cambridge, MA, 1991.
- [21] E. K. P. Chong and S. H. Zak, *An Introduction to Optimization*, John Wiley and Sons, Inc., New York, NY, 3rd edition, Feb. 2008.
- [22] H. Lütkepohl, *Handbook of Matrices*, John Wiley and Sons, Inc., 1st edition, February 1997.
- [23] N. J. A. Sloane, "How to pack lines, planes, 3-spaces, etc.," <http://www2.research.att.com/njas/grass/index.html>.