

# Construction of Incoherent Unit Norm Tight Frames With Application to Compressed Sensing

Evaggelia V. Tsiligiani, *Student Member, IEEE*, Lisimachos P. Kondi, *Senior Member, IEEE*,  
and Aggelos K. Katsaggelos, *Fellow, IEEE*

**Abstract**—Despite the important properties of unit norm tight frames (UNTFs) and equiangular tight frames (ETFs), their construction has been proven extremely difficult. The few known techniques produce only a small number of such frames while imposing certain restrictions on frame dimensions. Motivated by the application of incoherent tight frames in compressed sensing (CS), we propose a methodology to construct incoherent UNTFs. When frame redundancy is not very high, the achieved maximal column correlation becomes close to the lowest possible bound. The proposed methodology may construct frames of any dimensions. The obtained frames are employed in CS to produce optimized projection matrices. Experimental results show that the proposed optimization technique improves CS signal recovery, increasing the reconstruction accuracy. Considering that the UNTFs and ETFs are important in sparse representations, channel coding, and communications, we expect that the proposed construction will be useful in other applications, besides the CS.

**Index Terms**—Unit norm tight frames, Grassmannian frames, compressed sensing.

## I. INTRODUCTION

WHEN abandoning orthonormal bases for overcomplete spanning systems, we are led to *frames* [1]. In signal processing, frames are a decomposition tool that adds more flexibility to signal expansions, facilitating various signal processing tasks [2], [3]. Although their existence has been known for over half a century [4], frames have been introduced in the signal processing community only in the recent decades, offering the advantage of redundancy in signal representations and providing numerical stability of reconstruction, resilience to additive noise and resilience to quantization. Frames have mainly become popular due to *wavelets* [5]; however, many other frame families have been employed in numerous applications including source coding, robust transmission, code division multiple access (CDMA) systems, operator theory, coding theory, quantum theory and quantum computing [2], [6]. Certain frame categories such as Grassmannian frames have connections to Grassmannian packings, spherical codes and

graph theory [7]. Therefore, frame theory and its applications have experienced a growing interest among mathematicians, engineers, computer scientists, and others.

To introduce some notation, a finite frame  $F_m^N$  in a real or complex  $m$ -dimensional Hilbert space  $\mathbb{H}^m$  is a sequence of  $N \geq m$  vectors  $\{f_k\}_{k=1}^N$ ,  $f_k \in \mathbb{H}^m$ , satisfying the following condition

$$\alpha \|f\|_2^2 \leq \sum_{k=1}^N |\langle f, f_k \rangle|^2 \leq \beta \|f\|_2^2, \quad \forall f \in \mathbb{H}^m, \quad (1)$$

where  $\alpha, \beta$  are positive constants, called the lower and upper frame bounds respectively [1]. Viewing  $\mathbb{H}^m$  as  $\mathbb{R}^m$  or  $\mathbb{C}^m$ , the  $m \times N$  matrix  $F = [f_1 \ f_2 \ \dots \ f_N]$  with columns the frame vectors  $f_k$ , is known as the frame synthesis operator. We usually identify the synthesis operator with the frame itself. The redundancy of the frame is defined by  $\rho = N/m$  and is a “measure of overcompleteness” of the frame [7].

When designing a frame for a specific application, certain requirements are imposed. The orthogonality of frame rows is a common one; frames exhibiting equal-norm orthogonal rows are known as *tight frames* or *Welch-bound sequences* [8], [9] and have been employed in sparse approximation [10], [11]. Equality of column norms is also important. *Unit norm tight frames (UNTFs)*, that is, tight frames with unit norm column vectors, have been used in the construction of signature sequences in CDMA systems [8], [12], [13]. Moreover, they are robust against additive noise and erasures, and allow for stable reconstruction in communications [14]–[18]. *Equiangularity* is a property related to the dependency between frame columns; column vectors forming equal angles exhibit minimal dependency. *Equiangular tight frames (ETFs)* have been popular due to their use in sparse approximation [19], robust transmission [17], [18] and quantum computing [20].

Frames are employed in signal processing when there is a need for redundancy. Redundancy provides representations resilient to coding noise, enabling signal recovery even when some coefficients are lost. Moreover, a redundant dictionary can be chosen to fit its content to the data, yielding highly sparse representations that would not be easily achieved using an orthonormal basis. However, a significant drawback when working with frames is that the frame elements may be linearly dependent. Therefore, the advantages provided by the frame redundancy come at the cost that the signal representation may not be unique.

Manuscript received April 15, 2013; revised November 8, 2013; accepted February 8, 2014. Date of publication February 25, 2014; date of current version March 13, 2014.

E. V. Tsiligiani and L. P. Kondi are with the Department of Computer Science and Engineering, University of Ioannina, Ioannina 45110, Greece (e-mail: etsiligia@cs.uoi.gr; lkoni@cs.uoi.gr).

A. K. Katsaggelos is with the Department of Electrical and Engineering Computer Sciences, Northwestern University, Evanston, IL 60208 USA (e-mail: aggk@eecs.northwestern.edu).

Communicated by M. Elad, Associate Editor for Signal Processing.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2014.2308171

*Mutual coherence* is a simple numerical way to characterize the degree of similarity between the columns of a frame and is defined as the largest absolute normalized inner product between different frame columns [10], [21],

$$\mu\{F\} = \max_{\substack{1 \leq i, j \leq N \\ i \neq j}} \frac{|\langle f_i, f_j \rangle|}{\|f_i\|_2 \|f_j\|_2}. \quad (2)$$

Frames with small mutual coherence are known as *incoherent*. Equiangular tight frames are the unit norm ensembles that achieve equality in the Welch bound [see eq. (12)]. However, the construction of equiangular tight frames has been proven extremely difficult.

In this paper, we rely on frame theory to construct *incoherent unit norm tight frames*. Based on recent theoretical results, we employ these frames in compressed sensing to improve reconstruction of sparse signals. Sparse signal recovery was introduced in signal processing in the context of sparse and redundant representations as the problem of finding a signal representation under an overcomplete basis or redundant dictionary. An overcomplete representation is described by an underdetermined linear system of the form

$$x = D\alpha, \quad (3)$$

where  $x \in \mathbb{R}^K$  is the treated signal,  $D \in \mathbb{R}^{K \times N}$ ,  $K \leq N$ , is a redundant dictionary, and  $\alpha \in \mathbb{R}^N$  is the vector of the unknown coefficients. Assuming that  $D$  is full rank, we need additional criteria to find a unique signal satisfying (3). Thus, we employ a penalty function  $J(\alpha)$ , defining the general optimization problem

$$\min_{\alpha} J(\alpha) \quad \text{subject to} \quad x = D\alpha. \quad (4)$$

In sparse representations, we are interested in a solution of (4) with a few nonzero coefficients, that is,  $\|\alpha\|_0 = T$ ,  $T \ll K$ , where  $\|\cdot\|_0$  is the so-called  $\ell_0$  norm (which is actually not a norm) counting the nonzero coefficients of the respective signal. According to well-known results [10], uniqueness of a sparse solution is guaranteed if the sparsity level  $\|\alpha\|_0$  satisfies

$$\|\alpha\|_0 < \frac{1}{2} \left( 1 + \frac{1}{\mu\{D\}} \right). \quad (5)$$

It is obvious that in order to obtain a unique sparse representation, we need a sufficiently incoherent dictionary. In addition, mutual coherence plays an important role in the performance of the algorithms deployed to find sparse solutions in problems of the form (4) [11]. More recent results [22] concerning the algorithms' performance highlight the role of tightness, requiring  $D$  to be an *incoherent tight frame* (see also Section IV).

Compressed sensing (CS) is a novel theory [23], [24] that merges compression and acquisition, exploiting sparsity to recover signals that have been sampled at a drastically smaller rate than the conventional Shannon/Nyquist theorem imposes. Undersampling implies that the number of measurements,  $m$ , is much smaller than the dimension  $K$  of the signal. The sensing mechanism employed by CS leads to an underdetermined linear system, described by the following equation

$$y = Px, \quad (6)$$

with  $y \in \mathbb{R}^m$  and  $P \in \mathbb{R}^{m \times K}$ ,  $m \ll K$ , a proper *sensing or projection matrix*. Considering a sparse representation of  $x$ , we obtain

$$y = PD\alpha. \quad (7)$$

Setting  $F = PD$ ,  $F \in \mathbb{R}^{m \times N}$ , which is referred to as the *effective dictionary*, we rewrite (7) in the form

$$y = F\alpha. \quad (8)$$

Following the above discussion about finding sparse representations satisfying underdetermined linear systems, we require  $F$  to be an incoherent tight frame. Considering the optimization of the sampling process, we note that, given the dictionary  $D$ , in order to obtain a nearly optimal effective dictionary  $F$  with respect to mutual coherence and tightness, we need to find a projection matrix  $P$  such that  $F$  is as close as possible to an incoherent UNTF.

The idea of optimizing the projection matrix such that it leads to an effective dictionary with small mutual coherence was introduced by Elad in [25]. Based on frame theory, we proposed in [26] the construction of an effective dictionary that forms an incoherent unit norm tight frame. In this paper, we extend our technique to obtain unit norm tight frames that exhibit a significantly improved incoherence level compared to [25], [26], resulting in accurate signal reconstruction when employed in CS. Moreover, for certain frame dimensions, i.e., when the frame redundancy is not very high, the achieved mutual coherence becomes very close to the lowest possible bound. Considering that the construction of unit norm tight frames and equiangular tight frames has been proven notoriously difficult, whereas the few known design techniques impose certain restrictions on the frame dimensions, we expect that the proposed methodology will be useful to other applications besides CS.

The rest of the paper is organized as follows: In Section II we review basic definitions and concepts from frame theory, while highlighting challenges and difficulties in frame design. In Section III we present two algorithms for the construction of incoherent UNTFs and discuss their convergence. Section IV includes the application of the proposed construction to CS. We also present previous work on projection matrix optimization and recent theoretical results justifying our optimization strategy. Experimental results can be found in Section V. Finally, conclusions are drawn in Section VI.

## II. BACKGROUND

Based on the discussion in the previous section about the requirements in frame design, we conclude that even though there are considerable reasons to abandon orthonormal bases for frames, in most applications, we still want to use frames that preserve as many properties of orthonormal bases as possible [7].

When designing frames close to orthonormal bases, row orthogonality is a potential such desirable property. Therefore, we are led to tight frames. Let  $F_m^N = \{f_k\}_{k=1}^N$  be a finite redundant frame in  $\mathbb{H}^m$ . Then, if we set  $\alpha = \beta$  in (1),

we have

$$f = \frac{1}{\alpha} \sum_{k=1}^N \langle f, f_k \rangle f_k, \quad \forall f \in \mathbb{H}^m, \quad (9)$$

thus obtaining an  $\alpha$ -tight frame. The rows of  $\alpha^{-1/2} F_m^N$  form an orthonormal family [1]. Constructing a tight frame is straightforward; we take an orthonormal basis and select the desired number of rows. For example,  $m \times N$  harmonic tight frames are obtained by deleting  $(N - m)$  rows of an  $N \times N$  DFT matrix.

However, in many problems tightness is not sufficient. Most applications require some additional structure, such as, specific column norms or small column correlation. Given an  $\alpha$ -tight frame, we obtain a *unit norm tight frame (UNTF)*, if,  $\|f_k\|_2 = 1$ , for all  $k$ . For  $\alpha$ -tight frames, the following relation holds

$$\sum_{k=1}^N \|f_k\|_2^2 = \alpha m. \quad (10)$$

It is clear that we cannot design a UNTF with an arbitrary tightness parameter; a UNTF  $F_m^N$  exists only for  $\alpha = N/m$ , or  $\alpha = \rho$ , the redundancy of the frame.

Orthonormal bases provide unique representations as their elements are perfectly uncorrelated. Requiring unique sparse representations in overcomplete dictionaries, it would be convenient to employ frames whose columns give identical inner products, that is,

$$|\langle f_k, f_\ell \rangle| = c_\mu, \quad \text{for } k \neq \ell, \quad (11)$$

where  $c_\mu$  is a constant. For a unit norm frame, the absolute value of the inner product between two frame vectors equals the cosine of the acute angle between the lines spanned by the two vectors. For this reason, frames satisfying (11) are called equiangular.

The maximal correlation between frame elements depends on the frame dimensions  $m, N$ . The lowest bound on the minimal achievable correlation for equiangular frames, also known as *Welch bound* is given by [7]

$$\mu\{F\} \geq \sqrt{\frac{N - m}{m(N - 1)}}. \quad (12)$$

Among all unit norm frames with the same redundancy, the equiangular ones that are characterized by the property of minimal correlation between their elements are called *Grassmannian* frames [7]. If minimal correlation is the lowest achievable, as (12) implies, then we obtain an *optimal Grassmannian* frame. According to [7], an equiangular unit norm tight frame is an optimal Grassmannian frame. As unit norm tight frames with dimensions  $m, N$  exist for a specific tightness parameter ( $\alpha = N/m$ ), an optimal Grassmannian frame is an equiangular  $N/m$ -tight frame (ETF).

Despite their important properties, neither UNTFs nor Grassmannian frames are easy to construct. Tightness implies certain restrictions on singular values and singular vectors; the  $m$  nonzero singular values of an  $m \times N$   $\alpha$ -tight frame equal  $\sqrt{\alpha}$ . This property combats either column normalization or the requirement for constant inner products between

columns [27]. Two techniques are known to provide general UNTF constructions; the work of [28], where the authors start from a tight frame and by solving a differential equation they approach a UNTF, and the work of [29], where the authors start from a unit norm frame and increase the degree of tightness using a gradient-descent-based algorithm. Relative primeness of  $m$  and  $N$  is a condition assumed by both techniques, though in [29] in a weaker sense. Regarding the construction of equiangular tight frames, it is known that these frames exist for certain frame dimensions. When  $\mathbb{H}^m = \mathbb{R}^m$  the frame dimensions should satisfy  $m \leq N(N + 1)/2$ , while for  $\mathbb{H}^m = \mathbb{C}^m$ , there must hold  $m \leq N^2$ . Moreover, even if we know the existence of such frames, there is no explicit way of constructing them. The only general construction techniques are reported in [30]–[32] and they impose additional restrictions on the frame dimensions.

### III. CONSTRUCTING INCOHERENT UNIT NORM TIGHT FRAMES

The role of incoherence in sparse signal recovery, both in redundant representations and compressed sensing, makes optimal Grassmannian frames ideal candidates for these problems. Considering the design difficulties discussed in Section II, we aim at the construction of frames that are as close as possible to optimal Grassmannian frames. Based on the observation that optimal Grassmannian frames not only exhibit minimal mutual coherence, but  $N/m$ -tightness as well, we propose the following design methodology: Suppose we compute a matrix with small mutual coherence. Then, the problem of approximating a Grassmannian frame reduces to finding a UNTF that is nearest to the computed incoherent matrix in Frobenius norm. This is a matrix nearness problem, which can be solved algebraically by employing the following theorem [27], [33].

**Theorem 1.** *Given a matrix  $F \in \mathbb{R}^{m \times N}$ ,  $N \geq m$ , suppose  $F$  has singular value decomposition (SVD)  $U \Sigma V^*$ . With respect to the Frobenius norm, a nearest  $\alpha$ -tight frame  $F'$  to  $F$  is given by  $\sqrt{\alpha} \cdot UV^*$ . Assume in addition that  $F$  has full row-rank. Then  $\sqrt{\alpha} \cdot UV^*$  is the unique  $\alpha$ -tight frame closest to  $F$ . Moreover, one may compute  $UV^*$  using the formula  $(FF^*)^{-1/2} F$ .*

Therefore, the main idea of the proposed design methodology is alternating between tightness and incoherence. Next, we present two algorithms implementing this construction strategy.

#### A. Algorithm 1

The first algorithm starts from an arbitrary  $m \times N$  matrix that has full rank and sequentially applies a “shrinkage” process and Theorem 1. The “shrinkage” process reduces the matrix mutual coherence, while Theorem 1 finds an  $N/m$ -tight frame that is nearest to the incoherent matrix. In order to minimize the correlation between the columns of a given matrix, it is a common strategy to work with the Gram matrix [25], [27]. Given a matrix  $F \in \mathbb{R}^{m \times N}$ , formed by the frame vectors  $\{f_k\}_{k=1}^N$  as its columns, the Gram matrix is the Hermitian matrix of the column inner products, that is  $G = F^*F$ .

For unit norm frame vectors, the maximal correlation is obtained as the largest absolute value of the off-diagonal entries of  $G$ .

In [25], Elad proposes a “shrinkage” operation on the off-diagonal Gram matrix entries (see eq. (19), Section IV). In this work, we bound the off-diagonal entries according to

$$\hat{g}_{ij} = \begin{cases} \text{sgn}(g_{ij}) \cdot (1/\sqrt{m}), & \text{if } 1/\sqrt{m} < |g_{ij}| < 1, \\ g_{ij}, & \text{otherwise,} \end{cases} \quad (13)$$

where  $g_{ij}$  is the  $(i, j)$  entry of the Gram matrix. The selected bound  $1/\sqrt{m}$  is approximately equal to the lowest bound [see eq. (12)] for large values of  $N$ . Other choices of the bound might work better depending on the frame dimensions. Combined with Theorem 1, the proposed Gram matrix processing yields highly incoherent UNTFs.

The algorithm we propose is iterative. Our initial matrix  $F_0$  is a tight frame nearest to a random Gaussian matrix. As the process that reduces the mutual coherence involves “shrinkage” operations on the Gram matrix  $G_q$ , a column normalization step precedes the main steps of our method. After applying (13), the modified Gram matrix  $\tilde{G}_q$  may have rank larger than  $m$ ; thus, we use SVD to reduce the rank back to  $m$ . Decomposing the new Gram matrix  $\check{G}_q$ , we obtain the incoherent matrix  $S_q$  such that  $S_q^* S_q = \check{G}_q$ . Next, Theorem 1 is applied to  $S_q$  to obtain an incoherent tight frame. Therefore, the  $q$ -th iteration of Algorithm 1 involves the following.

- 1) Obtain the matrix  $\hat{F}_q$ , after column normalization of  $F_q$ .
- 2) Calculate the Gram matrix  $\hat{G}_q = \hat{F}_q^* \hat{F}_q$  and apply (13) to bound the absolute values of the off-diagonal entries, producing  $\tilde{G}_q$ .
- 3) Apply SVD to  $\tilde{G}_q$  to force the matrix rank to be equal to  $m$ , obtaining  $\check{G}_q$ .
- 4) A matrix  $S_q \in \mathbb{R}^{m \times N}$  is obtained such that  $S_q^* S_q = \check{G}_q$ .
- 5) Find  $S'_q$ , the nearest  $N/m$ -tight frame to  $S_q$ , according to  $S'_q = \sqrt{N/m} \cdot (S_q S_q^*)^{-1/2} S_q$ . Set  $F_{q+1} = S'_q$ .

### B. Convergence of Algorithm 1

The proposed algorithm is actually an alternating projections algorithm. In alternating projections we find a point in the intersection of two or more sets by iteratively projecting a point sequentially onto every set [34]. More particularly, the proposed algorithm projects onto the following sets.

- i. The set  $Y$  of  $N \times N$  Gram matrices of  $m \times N$  unit norm frames,

$$Y = \left\{ G \in \mathbb{R}^{N \times N} : G = G^*, g_{ii} = 1, i = 1, \dots, N \right\}.$$

- ii. The set  $Z$  of  $N \times N$  symmetric matrices with bounded off diagonal entries,

$$Z = \left\{ G \in \mathbb{R}^{N \times N} : G = G^*, |g_{ij}| \leq 1/\sqrt{m}, i \neq j, i, j = 1, \dots, N \right\}.$$

- iii. The set  $W$  of rank- $m$ ,  $N \times N$  symmetric matrices,

$$W = \left\{ G \in \mathbb{R}^{N \times N} : G = G^*, \text{rank}(G) = m \right\}.$$

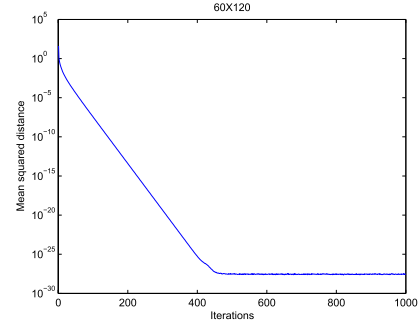


Fig. 1. Convergence of Algorithm 1 for a  $60 \times 120$  matrix. The mean squared distance between the current iteration and the sets we project on reduces in a linear rate.

- iv. The set  $S$  of  $N \times N$  Gram matrices of  $m \times N$   $\alpha$ -tight frames,

$$S = \left\{ G \in \mathbb{R}^{N \times N} : G = G^*, \text{ with only } m \text{ nonzero eigenvalues, all equal to } \alpha \right\}.$$

Alternating projections is a popular method and has been well-studied for closed convex sets. However, from the above sets only  $Y$  and  $Z$  are convex, whereas  $W$  and  $S$  are smooth manifolds [35]. Therefore, well known convergence results for alternating projections on convex sets [34] cannot be applied to the proposed method.

Only a few recent extensions of alternating projections consider the case of nonconvex sets [35], [36]. In [35] the authors study alternating projections on manifolds and prove convergence when two smooth manifolds intersect transversally. A more recent publication [36] considers alternating projections on two nonconvex sets, one of which is assumed to be suitably “regular”; the term refers to convex sets, smooth manifolds or feasible regions satisfying the Mangasarian-Fromovitz constraint qualification. The authors show that the method converges locally to a point in the intersection at a linear rate. The convergence of alternating projections on more than two sets some of which are nonconvex is still an open problem.

Therefore, our discussion regarding convergence of Algorithm 1 is mainly based on numerical results. To illustrate convergence, we need to define the mean squared distance of the current iteration from the sets involved in the projections, that is

$$D(q) = \frac{1}{8} (d^2(G_q, Y) + d^2(G_q, Z) + d^2(G_q, W) + d^2(G_q, S)),$$

where the distance  $d(G_q, F)$  between the current iteration  $G_q$  and the set  $F$  we project on is defined as  $d(G_q, F) = d_F = \inf \{ \|G_q - X\|_{\mathcal{F}} : X \in F \}$ , denoting by  $\|\cdot\|_{\mathcal{F}}$  the Frobenius norm.

Figs. 1 and 2 show  $\log_{10} D(q)$  when Algorithm 1 is applied to a  $60 \times 120$  and a  $25 \times 120$  matrix, respectively. Fig. 1 shows that the proposed algorithm converges at a linear rate, constructing a frame that belongs to the intersection of the involved sets. The zeroing of the mean squared distance implies that the produced frame is indeed an incoherent UNTF. When the frame redundancy increases, the numerical results become a little different. Fig. 2(a) shows that the convergence

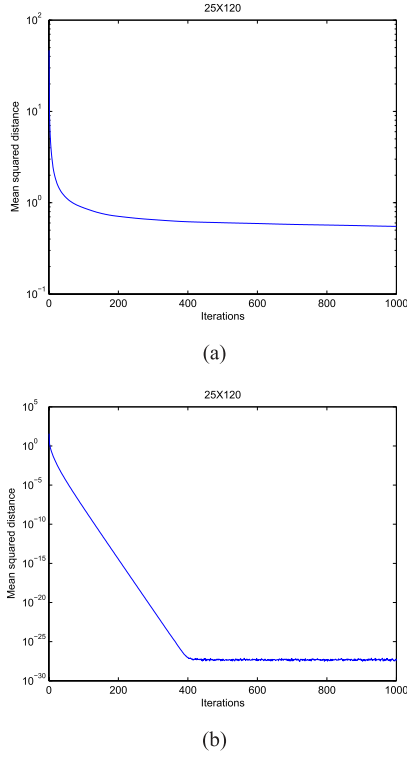


Fig. 2. Convergence of Algorithm 1 (alternating projections) for a  $25 \times 120$  matrix. The convergence rate depends on the bound used in eq. (13). In (a) we observe a sub-linear convergence rate when the bound equals  $1/\sqrt{m}$ . In (b) the convergence rate becomes linear as the bound is relaxed to  $3/2\sqrt{m}$ .

rate for a  $25 \times 120$  frame is sub-linear and the produced frame does not belong to the intersection of the involved sets. Considering the increased difficulties of constructing incoherent frames of high redundancy, this result is not surprising; it is possible that either the intersection is empty or it has properties that bring on difficulties to the proposed algorithm. Experiments performed with a relaxed incoherence level, which is determined by the bound  $1/\sqrt{m}$  in eq. (13) confirm our conjecture. A relaxed bound yields a broader set  $Z$  and increases the probability that the intersection has good properties. Fig. 2(b) illustrates convergence of Algorithm 1 when the bound  $1/\sqrt{m}$  in eq. (13) is replaced by  $3/2\sqrt{m}$ . We can see that the convergence rate becomes linear and the produced matrix belongs to the intersection of the involved sets.

### C. Algorithm 2

Similar to alternating projections is the method of *averaged projections*. At every step of averaged projections, we project the current iteration onto every set and average the results to obtain the value for the next iteration. Considering our problem, if  $G_q$  is the Gram matrix calculated in the  $q$ -th iteration and  $\mathcal{P}_Y(G_q)$  represents the projection of  $G_q$  on  $Y$ , then a modified version of Algorithm 1 would consider as input in the  $(q+1)$ -st iteration the average

$$G_{q+1} = \frac{1}{4}(\mathcal{P}_Y(G_q) + \mathcal{P}_Z(G_q) + \mathcal{P}_W(G_q) + \mathcal{P}_S(G_q)). \quad (14)$$

The projections can be calculated in the same way as in the algorithm presented at the beginning of this section, assuming

an additional calculation of the Gram matrix in the first and last steps.

Again we start from a random Gaussian matrix and apply Theorem 1 to obtain a nearest tight frame  $F_0$ ; then we calculate the Gram matrix  $G_0 = F_0^* F_0$ . In the  $q$ -th iteration we execute the following steps.

1. Obtain the matrix  $\hat{F}_q$ , after column normalization of  $F_q$  and calculate the Gram matrix  $\hat{F}_q^* \hat{F}_q$ , which is the projection  $\mathcal{P}_Y(G_q)$ .
2. Apply (13) on  $G_q$  to bound the absolute values of the off-diagonal entries, producing  $\mathcal{P}_Z(G_q)$ .
3. Apply SVD to  $G_q$  to force the matrix rank to be equal to  $m$ , obtaining  $\mathcal{P}_W(G_q)$ .
4. Find  $F'_q$ , the nearest  $N/m$ -tight frame to  $F_q$ , according to  $F'_q = \sqrt{N/m} \cdot (F_q F_q^*)^{-1/2} F_q$ . Obtain the Gram matrix  $S_q^* S_q$ , which is the projection  $\mathcal{P}_S(G_q)$ .
5. Calculate the average Gram matrix according to (14).
6. Find a matrix  $F_{q+1}$  s.t.  $F_{q+1}^* F_{q+1} = G_{q+1}$ .

### D. Convergence of Algorithm 2

According to recent results [36] concerning averaged projections, when several *prox-regular* sets have *strongly regular intersection* at some point, the method converges locally at a linear rate to a point in the intersection. Let us provide some definitions before proceeding. Prox-regular sets is a large class of sets that admits unique projections locally. It is known in [35] that convex sets and smooth manifolds belong to this category. Strongly regular intersection is important to prevent the algorithm from projecting near a *locally extremal point*. The notion of a locally extremal point in the intersection of some sets is the following: if we restrict to a neighborhood of such a point and then translate the sets by small distances, their intersection may render empty. Therefore, not choosing a locally extremal point as initial point in a projections algorithm is a critical hypothesis for convergence. In order to make clear that strong regularity implies local extremality, we cite here the related definitions for the case of two sets. For more details the reader is referred to [36].

**Definition 1** (Locally extremal point). Denoting by  $\mathbb{E}$  the Euclidean space, consider the sets  $F, G \subset \mathbb{E}$ . A point  $\bar{x} \in F \cap G$  is locally extremal for this pair of sets if there exists  $\rho > 0$  and a sequence of vectors  $z_r \rightarrow 0$  in  $\mathbb{E}$  such that

$$(F + z_r) \cap G \cap B_\rho(\bar{x}) = \emptyset, \text{ for all } r = 1, 2, \dots$$

where  $B_\rho(\bar{x})$  is the closed ball of radius  $\rho$  centered at  $\bar{x}$ . Clearly  $\bar{x}$  is not locally extremal if and only if

$$0 \in \text{int}(((F - \bar{x}) \cap \rho B) - ((G - \bar{x}) \cap \rho B)), \text{ for all } \rho > 0,$$

where  $B$  is the closed unit ball in  $\mathbb{E}$ .

**Definition 2** (Strongly regular intersection). Two sets  $F, G \subset \mathbb{E}$  have strongly regular intersection at a point  $\bar{x} \in F \cap G$  if there exists a constant  $\alpha > 0$  such that

$$\alpha \rho B \subset ((F - x) \cap \rho B) - ((G - z) \cap \rho B)$$

for all  $x \in F$  near  $\bar{x}$  and  $z \in G$  near  $\bar{x}$ .

By considering the case  $x = z = \bar{x}$ , we see that strongly regular intersection at a point  $\bar{x}$  implies that  $\bar{x}$  is not

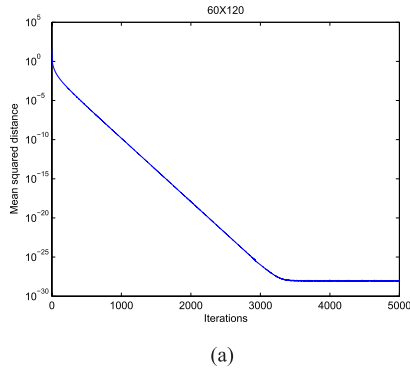


Fig. 3. Convergence of Algorithm 2 (averaged projections) for a  $60 \times 120$  matrix. The mean squared distance between the current iteration and the sets we project on reduces in a linear rate.

locally extremal. Conversely, finding a point in the intersection of the involved sets that is not locally extremal, implies that the sets have strongly regular intersection at this point.

Now, we can summarize the results of [36] regarding averaged projections.

**Theorem 2.** Consider prox-regular sets  $F_1, F_2, \dots, F_L \subset \mathbb{E}$  having strongly regular intersection at a point  $\bar{x} \in \cap F_i$ , and any constant  $k > \text{cond}(F_1, F_2, \dots, F_L | \bar{x})$ . Then, starting from any point near  $\bar{x}$ , one iteration of the method of averaged projections reduces the mean squared distance

$$D = \frac{1}{2L} \sum_{i=1}^L d_{F_i}^2$$

by a factor of at least  $1 - \frac{1}{k^2 L}$ .

The condition modulus  $\text{cond}(F_1, F_2, \dots, F_L | \bar{x})$  is a positive constant that quantifies strong regularity [36].

The sets  $Y, Z, W$  and  $S$  involved in Algorithm 2 are prox-regular ( $Y, Z$  are convex and  $W, S$  are smooth manifolds) and their intersection is very likely to be strongly regular; the fact that our initial matrix is a random Gaussian matrix minimizes the probability of choosing an initial point that is near to a locally extremal point. Though we cannot guarantee strong regularity for the above sets, randomness seems to prevent us from irregular solutions. Therefore, we expect that the averaged projections algorithm converges linearly to a point in the intersection of the above sets.

Let us see what experimental results show. Figs. 3 and 4 show mean squared distance for the averaged projections algorithm. Indeed, in Fig. 3 the results for a matrix of redundancy equal to 2 confirm a linear convergence rate and are in agreement with our theoretical expectations. Moreover, the zero mean squared distance implies that the obtained frame belongs to the intersection of the involved sets, that is, it is an incoherent UNTF. The results are a little different for a matrix with higher redundancy. As we can see in Fig. 4(a), the rate of convergence becomes sub-linear, indicating that the intersection of the involved sets is either empty or does not have the desired properties. Relaxing the imposed incoherence level, i.e., using a larger bound than  $1/\sqrt{m}$  in eq. (13), we obtain a broader set  $Z$ , increasing the probability that the intersection of the involved sets satisfies the necessary conditions formulated in Theorem 2. The experiments performed with the new set

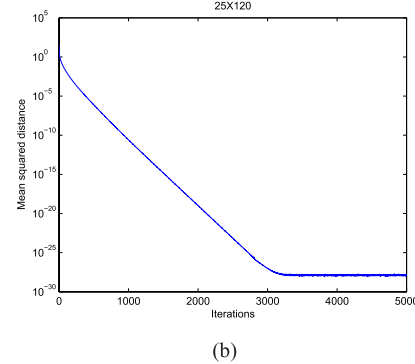
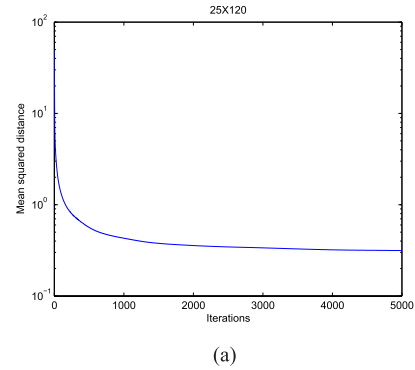
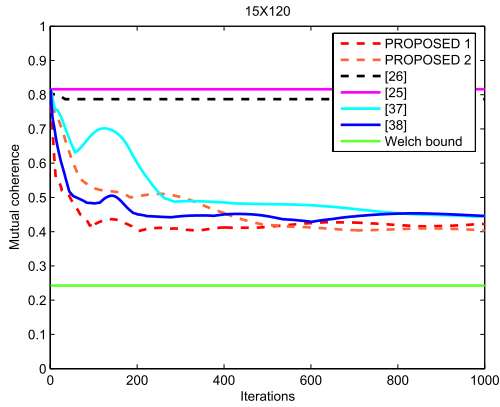


Fig. 4. Convergence of Algorithm 2 (averaged projections) for a  $25 \times 120$  matrix. The convergence rate depends on the bound used in eq. (13). In (a) we observe a sub-linear convergence rate when the bound equals  $1/\sqrt{m}$ . In (b) the convergence rate becomes linear as the bound is relaxed to  $3/2\sqrt{m}$ .

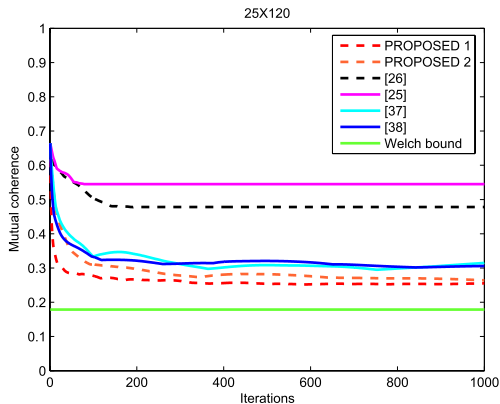
$Z$  yield a linear convergence rate [Fig. 4(b)], confirming our conjecture.

Comparing the convergence of the two proposed algorithms, an important note is that the presented experiments show that the results of the proposed averaged projections algorithm are similar to the alternating projections. Of course, there is a significant difference regarding the slope of the convergence curve; alternating projections is faster than averaged projections. However, the shapes of the curves are identical in all the examples employed in our experiments. Therefore, even though the theoretical justification of the proposed alternating projections needs further investigation, the experimental results encourage its use for the proposed constructions. In the next subsection, we present some experiments demonstrating the desired properties of the obtained frames, showing that both algorithms give similar results.

Before proceeding to more experiments and applications, we would like to clarify a point concerning the incoherence level constraint. One might wonder what is the effect of the imposed incoherence level on the proposed construction. Do we obtain frames with similar properties, regardless of the bound used in eq. (13)? The answer is that the frame properties are similar but not identical. Depending on the frame redundancy, there is a lower incoherence bound that should not be exceeded; otherwise, the smaller the incoherence bound we impose, the worse the incoherence level we finally obtain is. Thus, the selected bound needs fine tuning. However, the proposed bound  $1/\sqrt{m}$  works well for the constructions considered in this paper.



(a)



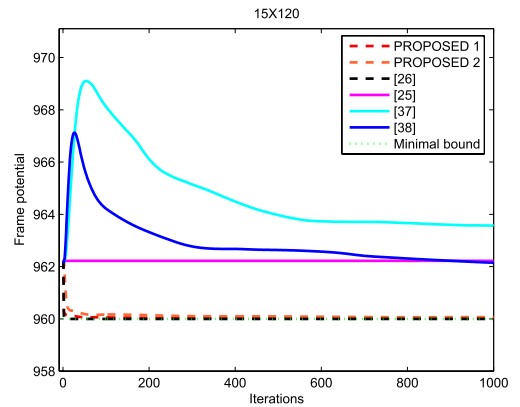
(b)

Fig. 5. Mutual coherence as a function of the number of iterations. The experiments involve (a) a  $15 \times 120$  frame and (b) a  $25 \times 120$  frame.

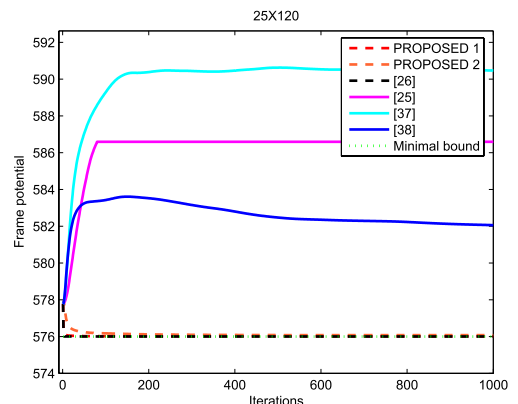
### E. Preliminary Experimental Results

Before we see the application of the proposed methodology in CS, we discuss some experimental results regarding the main properties of our constructions. As our goal is to use the proposed frame in CS, we compare the proposed algorithms to [25], [26], [37], [38], which also produce incoherent frames for CS. We briefly present these methods in Section IV.

Fig. 5 illustrates two snapshots of execution including 1000 iterations, depicting the achieved mutual coherence at every iteration. The examples involve a  $15 \times 120$  and a  $25 \times 120$  matrix showing that the proposed algorithms behave well as the mutual coherence improves significantly after a few iterations. Even though the averaged projections algorithm is slower, both algorithms finally converge to similar values regarding the mutual coherence. Moreover, the improvement is smooth even when applied to matrices with high redundancy. Regarding the other methods under testing, Fig. 5 shows that, in general, all of them have a stable behaviour and converge to low mutual coherence values. Only [37] seems to be inaccurate for very redundant matrices. Comparing the proposed algorithms to the other methods presented here, it is obvious that the incoherence level achieved by our methodology outperforms all the existing methods. Average results for mutual coherence confirm these observations (see Section V).



(a)



(b)

Fig. 6. Frame potential as a function of the number of iterations. The experiments involve (a) a  $15 \times 120$  frame and (b) a  $25 \times 120$  frame. The curves corresponding to the proposed algorithms and [26] coincide, meeting the minimal bound after a few iterations.

A metric used to evaluate how close the obtained frame is to a UNTF is the *frame potential*; it was defined in [39] as

$$FP(F) = \sum_{\substack{1 \leq i, j \leq N \\ i \neq j}} |\langle f_i, f_j \rangle|^2. \quad (15)$$

Benedetto and Fickus [39] discovered that the frame potential is bounded below by  $N^2/m$ , with equality if and only if  $F$  is a UNTF for  $\mathbb{R}^m$ . The frames obtained with the proposed methodology, employed in the present algorithms and in [26], exhibit frame potentials that coincide with this bound after very few iterations. Therefore, the proposed methodology produces UNTFs. This is an important advantage over existing techniques, as is demonstrated experimentally by the results shown in Fig. 6.

## IV. OPTIMIZED COMPRESSED SENSING

In sparse representations and compressed sensing we seek a solution of a sparse representation problem of the form

$$\min_{\alpha} J(\alpha) \text{ subject to } y = F\alpha, \quad (16)$$

where  $J(\alpha)$  is a function that imposes sparsity constraints on  $\alpha$ . One way to promote a sparse solution is the  $\ell_0$ -norm.

Choosing  $J(\alpha) \equiv \|\alpha\|_0$ , we are led to the following  $\ell_0$ -minimization problem,

$$\min_{\alpha \in \mathbb{R}^N} \|\alpha\|_0 \quad \text{subject to} \quad y = F\alpha. \quad (17)$$

As this problem is intractable [40], requiring combinatorial search, approximate numerical methods have been adopted for its solution. The main techniques include greedy methods and convex relaxation. The former generate a sequence of locally optimal choices in hope of determining a globally optimal solution. Orthogonal Matching Pursuit (OMP) [41] and its variations belong to this category. Instead of (17), convex relaxation methods solve a related convex program in hope that solutions coincide. Smoothing the penalty function with the  $\ell_1$ -norm, we obtain the problem

$$\min_{\alpha \in \mathbb{R}^N} \|\alpha\|_1 \quad \text{subject to} \quad y = F\alpha. \quad (18)$$

Many algorithms have been proposed to solve (18); Basis Pursuit (BP) [42] is a typical one.

Performance guarantees for the aforementioned algorithms underline that  $F$  should have certain properties. According to well known results, OMP and BP can find the solution of the  $\ell_0$  and  $\ell_1$  minimization problems with high probability provided that the system matrix  $F$  has small enough mutual coherence [11]. Besides  $F$  having small enough mutual coherence, the true solution must also be sufficiently sparse.

In compressed sensing,  $F$  comes from the product of a given representation dictionary  $D$  and the projection matrix  $P$ . CS theory asserts that the projection matrix and the representation dictionary should be incoherent, that is, highly uncorrelated.  $D$  is designated by the characteristics of the treated signals, as its role is to provide a high level of sparsity. As far as the choice of  $P$  is concerned, a matrix satisfying the restricted isometry property (RIP) is known to be incoherent with any orthonormal basis [43]. We can loosely say that when a matrix obeys RIP of order  $s$ , then all subsets of  $s$  columns are nearly orthogonal. Unfortunately, constructing RIP matrices has been proven extremely difficult. Most theoretical results in CS are based on the assumption that the projection matrix is drawn at random; random Gaussian or Bernoulli matrices are known to satisfy RIP with high probability, provided that the number of measurements  $m$  is in the order of  $T \log(N/T)$  [44]. Recall that  $N$  is the dimension of the  $T$ -sparse representation of the treated signal.

The improvement of performance guarantees on one hand, and the need to design sensing operators corresponding to the application of interest on the other, have produced several new theoretical and practical results regarding projection matrices, while research still goes on [45]. Deterministic constructions is one working direction [46], [47], providing the convenience to verify RIP without checking up all  $s$ -column submatrices. However, when employing deterministic matrices for CS, the sparsity level must be in the order of  $\sqrt{m}$ , a limitation known as the ‘‘square root bottleneck.’’ A construction that managed to go beyond this bottleneck [48] provided only a slight improvement. A more optimistic result concerning a specific deterministic construction can be found in [49]; the authors

conjecture that ETFs corresponding to Paley graphs of prime order are RIP in a manner similar to random matrices.

The approach adopted in this paper toward the design of the projection matrix involves its optimization. Recent publications [25], [26], [37], [38], [50], [51] have shown that reconstruction accuracy may be improved if the projection matrix is in a sense optimized for a certain signal class as this is determined by the representation dictionary. Following conditions set in sparse representations, the optimization process introduced in [25] involves the construction of an effective dictionary that exhibits small mutual coherence. Given the representation dictionary  $D$ , the optimization of the projection matrix reduces to finding a matrix  $P$  that yields an incoherent effective dictionary  $F = PD$ .

Aiming at the minimization of the average correlation between the columns of  $F$ , Elad in [25] proposes an iterative ‘‘shrinkage’’ operation on the off-diagonal Gram matrix entries according to

$$\hat{g}_{ij} = \begin{cases} \gamma g_{ij}, & |g_{ij}| \geq t, \\ \gamma t \cdot \text{sgn}(g_{ij}), & t > |g_{ij}| \geq \gamma t, \\ g_{ij}, & \gamma t > |g_{ij}|, \end{cases} \quad (19)$$

where  $\gamma$ ,  $t$  are appropriate parameters ( $0 < \gamma$ ,  $t < 1$ ) determining the convergence speed. In the same spirit, similar ‘‘shrinkage’’ operations are proposed in [37] and [38].

In [50], the authors’ goal is to produce a Gram matrix that is close to the identity matrix, by introducing the minimization problem

$$\min_F \|F^*F - I\|_{\mathcal{F}}. \quad (20)$$

Their solution, based on SVD, can work for either the single optimization of the projection matrix given the dictionary or the joint design and optimization of the dictionary and the projection matrix, from a set of training images. In the latter case the authors combine their method with K-SVD [52]. The problem of (20) is also treated in [51], proposing a solution based on gradient descent.

While working with the mutual coherence is simpler than working with the complex restricted isometry property, the analysis from the point of view of the mutual coherence leads to pessimistic results regarding CS measurements. Recovering  $T$  components from a sparse signal requires an order  $T^2$  measurements [25]. However, it has been demonstrated that the mutual coherence expresses worst case results for signal recovery. In addition, recent theoretical results provide a more optimistic perspective for incoherent matrices provided that they also form tight frames. A theorem that relates an incoherent tight frame with performance guarantees for the algorithms deployed to solve (17) and (18) has been formulated by Tropp [22].

**Theorem 3.** *Let  $F$  be an  $m \times N$  incoherent tight frame, and  $a$  a sparse vector observed by  $y = Fa$ . If  $a$  has  $T \leq cm / \log N$  nonzero entries drawn at random ( $c$  is some positive constant), then it is the unique solution for  $\ell_0$  and  $\ell_1$  minimization problems with probability greater than 99.44%. The author characterizes as *incoherent* the frames with mutual coherence equal to or smaller than  $1/\sqrt{m}$ . According to*



Theorem 3, using an incoherent tight frame, the number of necessary measurements to recover a  $T$ -sparse signal of length  $N$ , is in the order of  $T \log N$ .

The goal of the optimization technique proposed in this paper is the construction of an effective dictionary that satisfies the constraints introduced by Theorem 3, that is, small mutual coherence and tightness. Starting with a random Gaussian projection matrix  $P$  and a matrix  $D$  that yields sparse signal representations, we obtain an initial effective dictionary  $F = PD$ . Based on the algorithm presented in Section III, we produce an effective dictionary  $F_{CS}$  that forms an incoherent UNTF. Then solving a least squares problem, we compute a matrix  $P_{opt}$  satisfying  $F_{CS} = P_{opt}D$  to obtain an optimized projection matrix for CS.

Following this optimization strategy, we present experimental results for CS signal recovery in the next section.

## V. EXPERIMENTAL RESULTS

The algorithms proposed in this paper build incoherent UNTFs, which can be used for compressed sensing and other applications as well. As we have already seen in Section III, both algorithms yield similar constructions, therefore, we have decided to employ only one of the proposed algorithms in the experiments presented in this section. We choose the proposed alternating projections, as it exhibits higher convergence speed. We first present significant properties of the obtained frames, that is, their mutual coherence and frame potential. Then, we present reconstruction results when the proposed construction is applied to CS.

In CS, we may consider either naturally sparse signals or signals that are sparse with respect to a representation matrix  $D$ . For sparse signals, we use an incoherent UNTF,  $F$ , as a projection matrix; the proposed methodology is applied to an initial random Gaussian matrix and the resulting frame is used to take measurements according to  $y = Fa$ . Otherwise, we consider the product  $F = PD$ , where  $D$  is a fixed dictionary, take an initial random Gaussian  $P$ , and optimize  $F$  over  $P$ . Here, we follow the latter consideration.

The proposed method is compared to [25], our previous work [26] and the methods of [37] and [38]. Although our experiments included the methods of [50] and [51] as well, we only report results with the methods of [37] and [38] since they seem to perform better.

### A. Incoherent UNTFs

Our first experiments investigate the properties of the obtained incoherent UNTFs. In the following experiments, we take the initial  $F$  to be an  $m \times N$  random Gaussian matrix with  $m \in [15 : 5 : 60]$  and  $N = 120$ . For every value of  $m$ , we carry out 10000 experiments and compute average results. As our work aims at improving CS reconstruction accuracy, we compare the proposed construction to incoherent matrices produced by other methods used for optimized CS.

The mutual coherence is presented in Fig. 7. We can see that the proposed method leads to a significant reduction of the mutual coherence of the initial random matrix starting by a factor of approximately 45% for very redundant frames

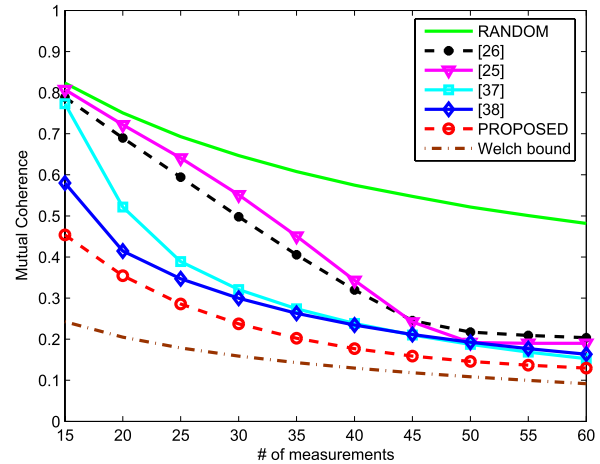


Fig. 7. Mutual coherence as a function of the number of measurements. The bottom brown dash-dotted line represents the lowest possible bound [see eq. (12)].

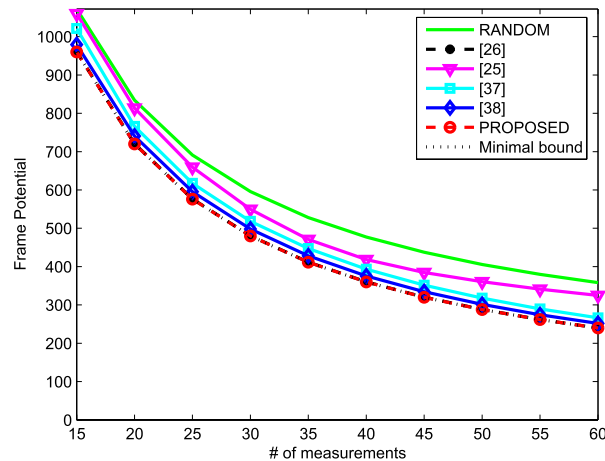


Fig. 8. Frame potential as a function of the number of measurements. The black and red dotted lines corresponding to our methodology coincide with the lowest possible bound  $N^2/m$ .

and becoming closer to the lowest possible bound when redundancy ( $\rho = N/m$ ) decreases (the brown dash-dotted line, in Fig. 7, stands for the lowest possible bound [see eq. (12)]. This is a very significant improvement compared to the results of our work in [26] and the other methods presented here. The fact that the proposed method performs well even for very redundant frames is an important advantage over the other competing methods. In Fig. 8 we demonstrate the frame potential [see eq. (15)] of the frames under testing, answering the question “how close are the obtained constructions to UNTFs?” The measurements corresponding to the proposed method and [26] coincide with the lowest bound  $N^2/m$ , confirming that the proposed methodology leads to UNTFs.

Another way to evaluate the obtained frames is to consider the distribution of the inner products between distinct columns. Fig. 9 illustrates a representative example of a  $25 \times 120$  frame. The histogram depicts the distribution of the absolute values of the corresponding Gram matrix entries. The results concern

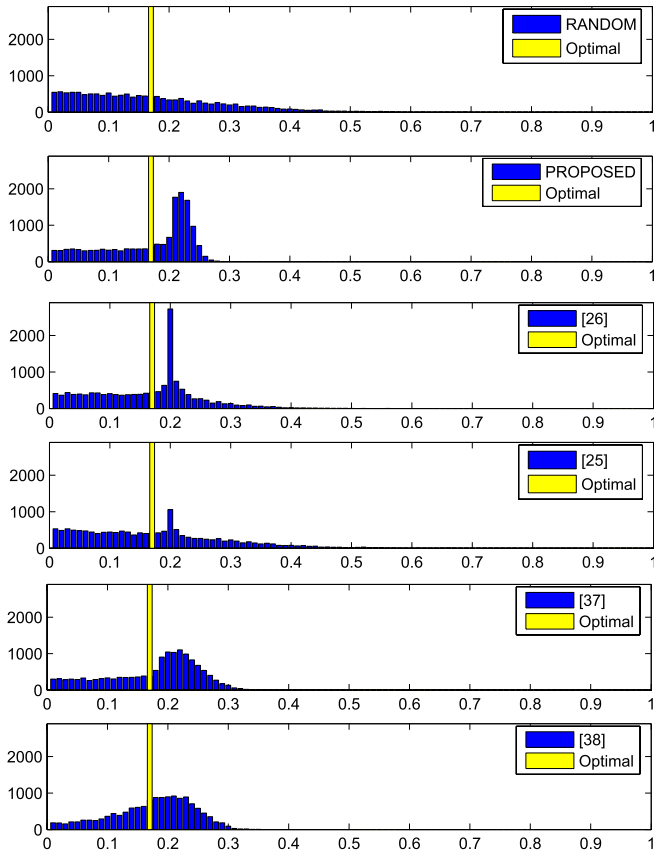


Fig. 9. Changes in the distribution of the column correlation of a  $25 \times 120$  frame.

the initial random matrix and all the matrices produced by the employed iterative algorithms, after 50 iterations. The yellow bar rises at the critical interval that includes the minimal achievable correlation, corresponding to the distribution of an optimal Grassmannian frame (the bar's actual height is constrained for clear demonstration of the methods under testing). The proposed method exhibits a significant concentration near the critical interval, combined with a short tail after it, showing that the number of the Gram entries that are closer to the ideal Welch bound is larger than in any other method presented here. Such a result is in agreement with the small mutual coherence values depicted in Fig. 7.

### B. CS Performance

The second group of experiments concerns the application of the obtained incoherent UNTFs to compressed sensing. The effective dictionary built with the proposed methodology is used to acquire sparse synthetic signals. A recovery algorithm is employed to reconstruct the original signal from the obtained measurements.

For each experiment, we generate a  $T$ -sparse vector of length  $N$ ,  $a \in \mathbb{R}^N$ , which constitutes a sparse representation of the  $K$ -length synthetic signal  $x = Da$ ,  $x \in \mathbb{R}^K$ ,  $K \leq N$ . We choose the dictionary  $D \in \mathbb{R}^{K \times N}$  to be a random Gaussian matrix. Experiments with DCT dictionaries lead to similar results. The locations of the nonzero coefficients in the sparse vector are chosen at random. Besides the effectiveness

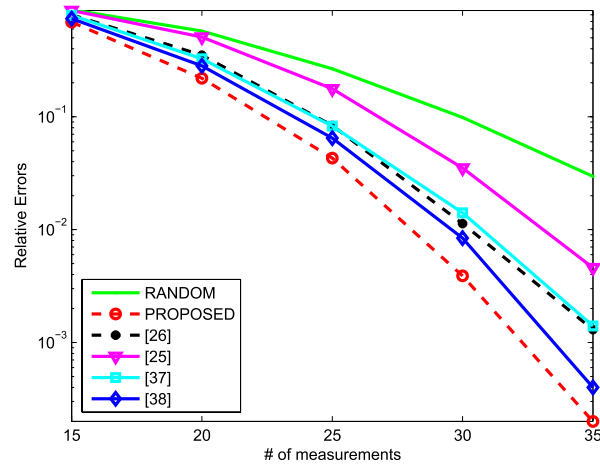


Fig. 10. CS performance for random and optimized projection matrices. Keeping the sparsity level fixed, reconstruction experiments are performed for varying number of measurements. The proposed method clearly outperforms the existing methods even when the number of measurements is small.

of the projection matrix  $P$ , the reconstruction results also depend on the number of measurements  $m$  and the sparsity level of the representation  $T$ . Thus, our experiments include varying values of these two parameters. For a specified number of measurements  $m \ll K$ , we create a random projection matrix  $P \in \mathbb{R}^{m \times K}$ . After the optimization process, we obtain  $m$  projections of the original signal according to (7). We reconstruct the original sparse signal with OMP.

In all experiments presented here, the synthetic signals are of length  $K = 80$  and the respective sparse representations, under the dictionary  $D$ , of length  $N = 120$ . The execution of the optimization algorithm included up to 50 iterations. Two sets of experiments have been considered; the first one includes varying values of the number of measurements  $m$  and the second one includes varying values of the sparsity level  $T$  of the treated signals. For every value of the aforementioned parameters we perform 10000 experiments and calculate the relative error rate; if the mean squared error of a reconstruction exceeds a threshold of order  $\mathcal{O}(10^{-4})$ , the reconstruction is considered to be a failure.

Fig. 10 presents the relative errors as a function of the number of measurements  $m$ , for a fixed sparsity level ( $T = 4$ ) of the treated signal. Fig. 11 presents the relative errors for a fixed number of measurements ( $m = 25$ ) and varying values of the sparsity level of the signal. It is clear that the effective dictionary obtained by the proposed algorithm leads to better reconstruction results compared to random matrices and to matrices produced by the other methods. This is due to the improvement in the effective dictionary properties.

An important observation regarding CS performance, is that although we achieved a high quality signal recovery, the fact that for some values of measurements (e.g. 15) this improvement is not of the same order as the improvement in the mutual coherence, indicates that additional properties should be taken into consideration to decide about the appropriateness of the effective dictionary. This has been pointed out by other authors [25], [50] as well and should be explored both theoretically and experimentally.

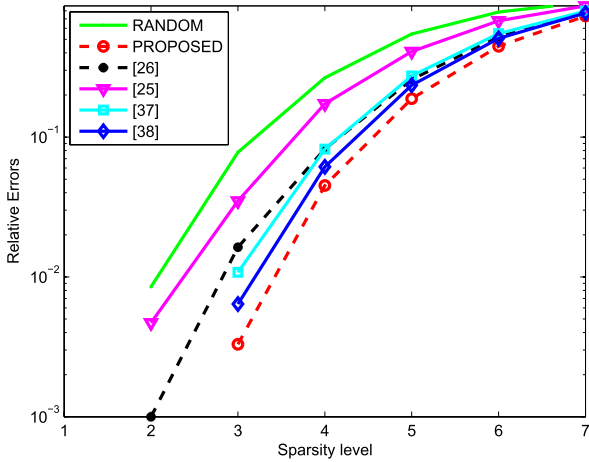


Fig. 11. CS performance for random and optimized projection matrices. Keeping the number of measurements fixed, reconstruction experiments are performed for varying sparsity levels. The proposed method clearly outperforms the existing methods even for high sparsity levels. A vanishing graph implies a zero error rate.

### VI. CONCLUSION AND FUTURE WORK

Based on new concepts of frame theory and recent results in sparse representations, we developed a methodology for optimizing the projection matrix that yields an incoherent UNTF as effective dictionary. Employing the obtained projection matrix in CS, we recovered sparse signals with high accuracy from a small number of measurements.

Considering that previous work on projections’ optimization involves only the minimization of the mutual coherence, the proposed method introduces a new parameter in the optimization process, namely tightness. Requiring the effective dictionary to be a tight frame results in additional reduction of the mutual coherence and improves other properties such as frame potential. The proposed methodology, combined with more efficient techniques of coherence minimization, could further improve the characteristics of the obtained dictionary.

Concluding, we expect that the construction of incoherent UNTFs with the proposed methodology will be useful for other applications as well, besides CS.

### REFERENCES

- [1] O. Christensen, *An Introduction to Frames and Riesz Bases*. Boston, MA, USA: Birkhäuser, 2003.
- [2] J. Kovačević and A. Chebira, “Life beyond bases: The advent of frames (Part I),” *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 86–104, Jul. 2007.
- [3] A. M. Bruckstein, D. L. Donoho, and M. Elad, “From sparse solutions of systems of equations to sparse modeling of signals and images,” *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, 2009.
- [4] R. J. Duffin and A. C. Schaeffer, “A class of nonharmonic fourier series,” *Trans. Amer. Math. Soc.*, vol. 72, no. 2, pp. 341–366, 1952.
- [5] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA, USA: SIAM, 1992.
- [6] J. Kovačević and A. Chebira, “Life beyond bases: The advent of frames (Part II),” *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 115–125, Sep. 2007.
- [7] T. Strohmer and R. W. Heath, “Grassmannian frames with applications to coding and communication,” *Appl. Comput. Harmon. Anal.*, vol. 14, no. 3, pp. 257–275, 2003.

- [8] J. L. Massey and T. Mittelholzer, “Welch’s bound and sequence sets for code-division multiple-access systems,” in *Sequences II: Methods in Communication, Security and Computer Science*. Berlin, Germany: Springer-Verlag, 1991, pp. 63–78.
- [9] S. Waldron, “Generalized welch bound equality sequences are tight frames,” *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2307–2309, Sep. 2003.
- [10] R. Gribonval and M. Nielsen, “Sparse representations in unions of bases,” *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3320–3325, Dec. 2003.
- [11] J. A. Tropp, “Greed is good: Algorithmic results for sparse approximation,” *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.
- [12] P. Viswanath and V. Anantharam, “Optimal sequences and sum capacity of synchronous CDMA systems,” *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 1984–1991, Sep. 1999.
- [13] P. Viswanath and V. Anantharam, “Optimal sequences for CDMA under colored noise: A Schur–Saddle function property,” *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1295–1318, Jun. 2002.
- [14] V. K. Goyal, M. Vetterli, and N. T. Thao, “Quantized overcomplete expansions in  $\mathbb{R}^N$ : Analysis, synthesis, and algorithms,” *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 16–31, Jan. 1998.
- [15] V. K. Goyal, J. Kovačević, and J. A. Kelner, “Quantized frame expansions with erasures,” *Appl. Comput. Harmon. Anal.*, vol. 10, no. 3, pp. 203–233, 2001.
- [16] P. G. Casazza and J. Kovačević, “Equal-norm tight frames with erasures,” *Adv. Comput. Math.*, vol. 18, nos. 2–4, pp. 387–430, 2003.
- [17] R. B. Holmes and V. I. Paulsen, “Optimal frames for erasures,” *Linear Algebra Appl.*, vol. 377, no. 1, pp. 31–51, 2004.
- [18] M. Fickus and D. G. Mixon, “Numerically erasure-robust frames,” *Linear Algebra Appl.*, vol. 437, no. 6, pp. 1394–1407, 2012.
- [19] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York, NY, USA: Springer-Verlag, 2010.
- [20] Y. C. Eldar and G. D. Forney, “Optimal tight frames and quantum measurement,” *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 599–610, Mar. 2002.
- [21] D. L. Donoho and M. Elad, “Optimally sparse representation in general (nonorthogonal) dictionaries via  $l^1$ -minimization,” *Proc. Nat. Acad. Sci. USA*, vol. 100, no. 5, pp. 2197–2202, 2003.
- [22] J. A. Tropp, “On the conditioning of random subdictionaries,” *Appl. Comput. Harmon. Anal.*, vol. 25, no. 1, pp. 1–24, 2008.
- [23] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [24] E. J. Candès, J. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.
- [25] M. Elad, “Optimized projections for compressed sensing,” *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5695–5702, Dec. 2007.
- [26] E. Tsiligianni, L. P. Kondi, and A. K. Katsagelos, “Use of tight frames for optimized compressed sensing,” in *Proc. EUSIPCO*, Bucharest, Romania, Aug. 2012, pp. 1439–1443.
- [27] J. A. Tropp, I. S. Dhillon, R. W. Heath, and T. Strohmer, “Designing structured tight frames via an alternating projection method,” *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 188–209, Jan. 2005.
- [28] B. G. Bodmann and P. G. Casazza, “The road to equal-norm Parseval frames,” *J. Funct. Anal.*, vol. 258, no. 2, pp. 397–420, 2010.
- [29] P. G. Casazza, M. Fickus, and D. G. Mixon, “Auto-tuning unit norm frames,” *Appl. Comput. Harmon. Anal.*, vol. 32, no. 1, pp. 1–15, 2012.
- [30] P. Xia, S. Zhou, and G. B. Giannakis, “Achieving the Welch bound with difference sets,” *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1900–1907, May 2005.
- [31] S. Waldron, “On the construction of equiangular frames from graphs,” *Linear Algebra Appl.*, vol. 431, no. 11, pp. 2228–2242, 2009.
- [32] M. Fickus, D. G. Mixon, and J. C. Tremain, “Steiner equiangular tight frames,” *Linear Algebra Appl.*, vol. 436, no. 5, pp. 1014–1027, 2012.
- [33] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ Press, 1985.
- [34] H. H. Bauschke and J. M. Borwein, “On projection algorithms for solving convex feasibility problems,” *SIAM Rev.*, vol. 38, no. 3, pp. 367–426, 1996.
- [35] A. S. Lewis and J. Malick, “Alternating projections on manifolds,” *Math. Oper. Res.*, vol. 33, no. 1, pp. 216–234, 2008.
- [36] A. S. Lewis, D. R. Luke, and J. Malick, “Local convergence for alternating and averaged nonconvex projections,” *Found. Comput. Math.*, vol. 9, no. 4, pp. 485–513, Aug. 2009.

- [37] J. Xu, Y. Pi, and Z. Cao, "Optimized projection matrix for compressive sensing," *EURASIP J. Adv. Signal Process.*, vol. 2010, pp. 43:1–43:8, Feb. 2010.
- [38] H. Huang and A. Makur, "Optimized measurement matrix for compressive sensing," *Sampl. Theory Signal Image Process.*, 2013, to be published.
- [39] J. J. Benedetto and M. Fickus, "Finite normalized tight frames," *Adv. Comput. Math.*, vol. 18, nos. 2–4, pp. 357–385, 2003.
- [40] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.
- [41] G. Davis, S. Mallat, and M. Avellaneda, "Greedy adaptive approximation," *Construct. Approx.*, vol. 13, no. 1, pp. 57–58, 1997.
- [42] S. S. Chen, D. L. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1999.
- [43] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [44] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Construct. Approx.*, vol. 28, no. 3, pp. 253–263, 2008.
- [45] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4053–4085, Sep. 2011.
- [46] R. A. DeVore, "Deterministic constructions of compressed sensing matrices," *J. Complex.*, vol. 23, nos. 4–6, pp. 918–925, 2007.
- [47] L. Applebaum, S. D. Howard, S. Searle, and R. Calderbank, "Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 2, pp. 283–290, 2009.
- [48] J. Bourgain, S. Dilworth, K. Ford, S. Konyagin, and D. Kutzarova, "Explicit constructions of RIP matrices and related problems," *Duke Math. J.*, vol. 159, no. 1, pp. 145–185, 2011.
- [49] A. S. Bandeira, M. Fickus, D. G. Mixon, and P. Wong, "The road to deterministic matrices with the restricted isometry property," *J. Fourier Anal., Appl.*, vol. 19, pp. 1123–1149, 2013.
- [50] J. M. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1395–1408, Jul. 2009.
- [51] V. Abolghasemi, D. Jarchi, and S. Sanei, "A robust approach for optimization of the measurement matrix in compressed sensing," in *Proc. 2nd Int. Workshop CIP*, Jun. 2010, pp. 388–392.
- [52] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

**Evaggelia V. Tsiligiani** (S'10) received the Diploma degree in electrical and computer engineering from the National Technical University of Athens, Greece, in 2001 and the M.S. degree in Computer Science from University of Ioannina, Greece, in 2010. She is currently pursuing the Ph.D. degree in the Department of Computer Science and Engineering, University of Ioannina, Greece.

Her research interests include video coding, sparse representations and compressive sensing.

**Lisimachos P. Kondi** (SM'11) received the Diploma in electrical engineering from the Aristotle University of Thessaloniki, Greece, in 1994 and the M.S. and Ph.D. degrees in electrical and computer engineering from Northwestern University, Evanston, IL, in 1996 and 1999, respectively.

During the 1999–2000 academic year, he was a Post-Doctoral Research Associate at Northwestern University. He is currently an Associate Professor with the Department of Computer Science and Engineering, University of Ioannina, Greece. He was previously with the faculty of the University at Buffalo, The State University of New York, and has also held summer appointments at the Naval Research Laboratory, Washington, DC, and the Air Force Research Laboratory, Rome, NY. His research interests are in the general areas of signal and image processing and communications, including image and video compression and transmission over wireless channels and the Internet, sparse representations and compressive sensing, super-resolution of video sequences, and shape coding.

Dr. Kondi has been an Associate Editor of the *IEEE SIGNAL PROCESSING LETTERS* (2008–2012), the *EURASIP Journal on Advances in Signal Processing* (2005–present), and the *International Journal of Distributed Sensor Networks* (2013–present). He was also Technical Program Chair of the International Conference on Digital Signal Processing (Santorini, Greece, July 2013).

**Aggelos K. Katsaggelos** (F'98) received the Diploma degree in electrical and mechanical engineering from the Aristotelian University of Thessaloniki, Greece, in 1979, and the M.S. and Ph.D. degrees in Electrical Engineering from the Georgia Institute of Technology, in 1981 and 1985, respectively.

In 1985, he joined the Department of Electrical Engineering and Computer Science at Northwestern University, where he is currently a Professor holder of the AT&T chair. He was previously the holder of the Ameritech Chair of Information Technology (1997–2003). He is also the Director of the Motorola Center for Seamless Communications, a member of the Academic Staff, NorthShore University Health System, an affiliated faculty at the Department of Linguistics and he has an appointment with the Argonne National Laboratory.

He has published extensively in the areas of multimedia signal processing and communications (over 200 journal papers, 500 conference papers and 40 book chapters) and he is the holder of 21 international patents. He is the coauthor of *Rate-Distortion Based Video Compression* (Kluwer, 1997), *Super-Resolution for Images and Video* (Claypool, 2007) and *Joint Source-Channel Video Transmission* (Claypool, 2007).

Among his many professional activities Prof. Katsaggelos was Editor-in-Chief of the *IEEE SIGNAL PROCESSING MAGAZINE* (1997–2002), a BOG Member of the IEEE Signal Processing Society (1999–2001), and a member of the Publication Board of the IEEE Proceedings (2003–2007). He is a Fellow of the IEEE (1998) and SPIE (2009) and the recipient of the IEEE Third Millennium Medal (2000), the IEEE Signal Processing Society Meritorious Service Award (2001), the IEEE Signal Processing Society Technical Achievement Award (2010), an IEEE Signal Processing Society Best Paper Award (2001), an IEEE ICME Paper Award (2006), an IEEE ICIP Paper Award (2007), an ISPA Paper Award (2009), and a EUSIPCO paper award (2013). He was a Distinguished Lecturer of the IEEE Signal Processing Society (2007–2008).