On the Use of Unit-Norm Tight Frames to Improve the Average MSE Performance in Compressive Sensing Applications

Wei Chen*, *Student Member, IEEE,* Miguel R. D. Rodrigues, *Member, IEEE,* and Ian J. Wassell

***Abstract*—This letter considers the design of sensing matrices with good expected-case performance for compressive sensing applications. By capitalizing on the mean squared error (MSE) of the oracle estimator, whose performance has been shown to act as a benchmark to the performance of standard sparse recovery algorithms, we demonstrate that a unit-norm tight frame is the closest design - in the Frobenius norm sense - to the solution of a convex relaxation of the optimization problem that relates to the minimization of the MSE of the oracle estimator with respect to the sensing matrix. Simulation results reveal that the MSE performance of a unit-norm tight frame based sensing matrix surpasses that of other standard sensing matrix designs in various scenarios, which include sparse recovery with basis pursuit denoise (BPDN), the Dantzig selector and orthogonal matching pursuit (OMP). This also has important practical implications because a unit-norm tight frame based sensing matrix can be designed very efficiently.**

I. Introduction

COMPRESSIVE sensing deals with the recovery of a high-dimensional sparse vector from a lower-dimensional linear measurement vector [1], [2]. We consider the standard measurement model given by:

$$\mathbf{y} = \mathbf{\Phi}\mathbf{f} + \mathbf{n}, \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^M$ is the measurement signal vector, $\mathbf{f} \in \mathbb{R}^N$ is the original signal vector, $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}) \in \mathbb{R}^N$ is a white Gaussian noise vector, and $\mathbf{\Phi} \in \mathbb{R}^{M\times N}$ (with $M < N$) is a measurement matrix with unit norm columns, i.e., $\|\boldsymbol{\phi}_j\|_{\ell_2} = 1$. We assume that the original signal is sparse in some orthonormal basis, i.e.,

$$\mathbf{f} = \mathbf{D}\mathbf{x}, \tag{2}$$

where $\mathbf{D} \in \mathbb{R}^{N\times N}$ is a matrix that represents the orthonormal basis and $\mathbf{x} \in \mathbb{R}^N$ is a sparse representation of $\mathbf{f} \in \mathbb{R}^N$, i.e., $\|\mathbf{x}\|_{\ell_0} \leq S$. We also assume a random signal model where the distinct support patterns of the same sparsity level in the sparse representation of the original signal occur with equal probability.

The prevailing method used to reconstruct the sparse vector $\mathbf{x}$ from the measurement vector $\mathbf{y}$ involves solving a ℓ_1 minimization problem, which acts as a convex relaxation of a ℓ_0 minimization problem, given by [1], [2]:

$$\begin{aligned} &\min_{\hat{\mathbf{x}}} && \|\hat{\mathbf{x}}\|_{\ell_1} \\ &\text{subject to:} && \|\mathbf{\Phi}\mathbf{D}\hat{\mathbf{x}} - \mathbf{y}\|_{\ell_2} \leq \epsilon, \end{aligned} \tag{3}$$

where $\epsilon \geq \sqrt{M\sigma^2}$.

A number of conditions have been put forth in order to establish successful reconstruction in this class of underdetermined sparse recovery problems as well as study the quality of the sensing matrices and recovery algorithms. For example, the null space property represents a necessary and sufficient condition for sparse recovery [1]. The restricted isometry property (RIP) also represents a sufficient condition for sparse recovery [2]. Another common method to measure the quality of a sensing matrix, which is not computationally intractable as the null space property and the RIP, is the mutual coherence given by:

$$\mu = \max_{1\leq i,j\leq N, i\neq j} |\psi_{i,j}|, \tag{4}$$

where $\mathbf{\Psi} = \mathbf{D}^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{D}$ denotes the coherence matrix of $\mathbf{\Phi}\mathbf{D}$ and $\psi_{i,j}$ denotes the element in ith row and jth column of $\mathbf{\Psi}$.

These conditions are mainly used to address the worst-case performance of sparse recovery [3]–[5]. However, the actual reconstruction performance in practice is often much better than the worst-case performance, so that this viewpoint can be too conservative. In addition, the worst-case performance is a less typical indicator of quality in signal processing applications than the expected-case performance. This motivates the design of sensing matrices with adequate expected-case performance. In [6], Elad proposes an iterative algorithm to minimize the t-averaged mutual coherence. In [7], Duarte-Carvajalino and Sapiro propose an algorithm to iteratively optimize both the sensing matrix and the basis simultaneously. Xu et al. consider the equiangular tight frame (ETF) as their target design and proposed an iterative algorithm to make the sensing matrix approach that design [8]. The sensing dictionaries for optimal expected-case performance of the thresholding algorithm are characterized in [9], where probability plays a critical role.

In this letter, we investigate the quality of a sensing matrix with respect to the mean squared error (MSE) performance of the oracle estimator [3], whose performance has been shown to act as a benchmark to the performance of various common sparse recovery algorithms. Since the minimization of the oracle estimator MSE with respect to the sensing

W. Chen and I. J. Wassell are with the Digital Technology Group, Computer Laboratory, University of Cambridge, Cambridge CB3 0FD, United Kingdom (e-mail: wc253, ijw24@cam.ac.uk).

Miguel R. D. Rodrigues is with the Instituto de Telecomunicações and the Department of Computer Science, University of Porto, Porto 4169-007, Portugal (e-mail: mrodrigues@dcc.fc.up.pt).

matrix is a non-convex optimization problem, we adopt convex relaxation techniques that enable us to show that a unit-norm tight frame is the closest design in the Frobenius norm sense to the solution of the relaxed problem. The superior reconstruction quality of a unit-norm tight frame based sensing matrix, which can be very efficiently constructed within at most $N-1$ iterations each requiring $\mathcal{O}(MN)$ operations [10], with respect to other sensing matrix designs is illustrated in various practical scenarios.

We use the notation: Italic upper-case letters denote numbers, boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, and calligraphic upper-case letters denote supports. The superscripts $(\cdot)^T$ and $(\cdot)^{-1}$ denote matrix transpose and matrix inverse, respectively. $\|\cdot\|_{\ell_0}$, $\|\cdot\|_{\ell_1}$, $\|\cdot\|_{\ell_2}$ and $\|\cdot\|_F$ denote the ℓ_0 norm, the ℓ_1 norm, the ℓ_2 norm and the Frobenius norm of matrices, respectively. The rank and trace of a matrix are denoted by $\mathtt{rank}(\cdot)$ and $\mathtt{Tr}(\cdot)$, respectively. $q_{i,j}$ denotes the element of the *i*th row and *j*th column of the matrix $\mathbf{Q}$, and $\mathbf{q}_i$ denotes the *i*th column of the matrix $\mathbf{Q}$. $\mathbf{Q} \succeq 0$ denotes that the matrix $\mathbf{Q}$ is positive semi-definite. Finally, $\mathbf{I}_N$ denotes the $N \times N$ identity matrix.

II. Design of The Sensing Matrix

We define the MSE in estimating a sparse random vector $\mathbf{x}$ corrupted by a random Gaussian noise vector $\mathbf{n}$ as follows:

$$\mathtt{MSE}(\mathbf{\Phi}) = \mathbb{E}_{\mathbf{x},\mathbf{n}}\left(\|\mathcal{F}(\mathbf{\Phi}\mathbf{D}\mathbf{x}+\mathbf{n})-\mathbf{x}\|_{\ell_2}^2\right), \qquad (5)$$

where $\mathcal{F}(\cdot)$ denotes an estimator, and $\mathbb{E}_{\mathbf{x},\mathbf{n}}(\cdot)$ denotes expectation with respect to the joint distribution of the random vectors $\mathbf{x}$ and $\mathbf{n}$.

The calculation of the average MSE in (5) depends upon the actual estimator. Consequently, to avoid the analysis of a single or several practical sparse recovery algorithms such as the basis pursuit de-noise (BPDN), the Dantzig selector, and orthogonal matching pursuit (OMP), we capitalize on the well-known oracle estimator that performs ideal least square (LS) estimation based on prior knowledge of the sparse vector support $\mathcal{J} \subset \{1,\ldots,N\}$ [3]. The oracle estimator MSE incurred in the estimation of a sparse deterministic vector $\mathbf{x}$ in the presence of a standard Gaussian noise vector $\mathbf{n}$, according to the model in (1), is given by [3]:

$$\begin{aligned}\mathtt{MSE}_{\mathbf{n}}^{\text{oracle}}(\mathbf{\Phi},\mathbf{x}) &= \mathbb{E}_{\mathbf{n}}\left(\|\mathcal{F}^{\text{oracle}}(\mathbf{\Phi}\mathbf{D}\mathbf{x}+\mathbf{n})-\mathbf{x}\|_{\ell_2}^2\right)\\ &= \sigma^2 \mathrm{Tr}\left(\left(\mathbf{E}_{\mathcal{J}}^T\mathbf{D}^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{D}\mathbf{E}_{\mathcal{J}}\right)^{-1}\right),\end{aligned} \qquad (6)$$

where $\mathbb{E}_{\mathbf{n}}(\cdot)$ denotes expectation with respect to the distribution of the random vector $\mathbf{n}$, and $\mathbf{E}_{\mathcal{J}}$ denotes the matrix that results from the identity matrix by deleting the set of columns out of the support $\mathcal{J}$.

The rationale of this approach is also supported by the fact that the oracle MSE coincides with the unbiased Cramér-Rao bound (CBD) for exactly *S*-sparse deterministic vectors [11], so that it represents the best achievable performance for any unbiased estimator. Equally important, the oracle MSE performance acts as a performance benchmark for the key sparse recovery algorithms. For example, Ben-Haim, Eldar and Elad [12] demonstrate both theoretically and numerically that the BPDN, the Dantzig selector, the OMP and thresholding algorithms all achieve performances that are proportional to the oracle MSE.

Consequently, the average value of the oracle MSE is given by:

$$\begin{aligned}\mathtt{MSE}^{\text{oracle}}(\mathbf{\Phi}) &= \mathbb{E}_{\mathbf{x}}\left(\mathtt{MSE}_{\mathbf{n}}^{\text{oracle}}(\mathbf{\Phi},\mathbf{x})\right)\\ &= \mathbb{E}_{\mathbf{x}}\left(\sigma^2\mathrm{Tr}\left(\left(\mathbf{E}_{\mathcal{J}}^T\mathbf{D}^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{D}\mathbf{E}_{\mathcal{J}}\right)^{-1}\right)\right)\\ &= \sigma^2\mathbb{E}_{\mathcal{J}}\left(\mathrm{Tr}\left(\left(\mathbf{E}_{\mathcal{J}}^T\mathbf{D}^T\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{D}\mathbf{E}_{\mathcal{J}}\right)^{-1}\right)\right),\end{aligned} \qquad (7)$$

where $\mathbb{E}_{\mathbf{x}}(\cdot)$ denote expectation with respect to the distribution of the random vector $\mathbf{x}$, and we have used the fact that the expectation with respect to the distribution of the random vector $\mathbf{x}$ is equal to the expectation with respect to the distribution of the support of the random vector $\mathbf{x}$, due to the use of the oracle. Note that the $\mathtt{MSE}_{\mathbf{n}}(\mathbf{\Phi},\mathbf{x})$ and $\mathtt{MSE}(\mathbf{\Phi})$ emphasize that the MSEs depend upon the sensing matrix and/or the deterministic parameters.

By defining $\mathbf{Q} = \mathbf{\Phi}^T\mathbf{\Phi}$, which represents the coherence matrix of the sensing matrix, we now pose the optimization problem:

$$\begin{aligned}\min_{\mathbf{Q}} \quad & \mathbb{E}_{\mathcal{J}}\left(\mathrm{Tr}\left(\left(\mathbf{E}_{\mathcal{J}}^T\mathbf{D}^T\mathbf{Q}\mathbf{D}\mathbf{E}_{\mathcal{J}}\right)^{-1}\right)\right)\\ \text{s.t.} \quad & \mathbf{Q} \succeq 0,\\ & q_{j,j} = 1 \quad (j=1,\ldots,N),\\ & rank(\mathbf{Q}) \le M.\end{aligned} \qquad (8)$$

This optimization problem, which up to a rotation leads to the sensing matrix that optimizes the average MSE of the oracle estimator, is non-convex due to the rank constraint, and so is very difficult to solve. Therefore, we will adopt a two-step procedure that involves: i) the determination of the solution to a convex relaxation of the original optimization problem, which ignores the rank constraint, that will be given in Proposition 1; ii) the determination of the feasible solution that is closest in the Frobenius norm sense to the solution to the convex relaxed problem, that will be given in Proposition 2. This procedure does not necessarily produce the optimal sensing matrix, but extensive simulation results demonstrate that this design outperforms other designs in the literature.

Proposition 1: The solution of the optimization problem:

$$\begin{aligned}\min_{\mathbf{Q}} \quad & \mathbb{E}_{\mathcal{J}}\left(\mathrm{Tr}\left(\left(\mathbf{E}_{\mathcal{J}}^T\mathbf{D}^T\mathbf{Q}\mathbf{D}\mathbf{E}_{\mathcal{J}}\right)^{-1}\right)\right)\\ \text{s.t.} \quad & \mathbf{Q} \succeq 0,\\ & q_{j,j} = 1 \quad (j=1,\ldots,N),\end{aligned} \qquad (9)$$

which represents a convex relaxation of the original optimization problem in (8), is the $N \times N$ identity matrix $\mathbf{I}_N$.

Proof: Let $s \le S$ be a positive integer. Let also $\mathcal{J}_s^t \subset \{1,2,\ldots,N\}$ $(t=1,\ldots,T_s)$ denote a support set with cardinality s, where $T_s = \binom{N}{s} = \frac{N!}{s!(N-s)!}$. We let $\mathbf{\Psi}_{\mathcal{J}_s^t} = \mathbf{E}_{\mathcal{J}_s^t}^T\mathbf{D}^T\mathbf{Q}\mathbf{D}\mathbf{E}_{\mathcal{J}_s^t}$. We also let $\lambda_s^{\mathcal{J}_s^t} \ge \ldots \ge \lambda_1^{\mathcal{J}_s^t}$ be the eigenvalues of $\mathbf{\Psi}_{\mathcal{J}_s^t}$. Let $p(\mathcal{J}_s)$ denote the probability that the support size is s.

We now note that

$$\sum_{t=1}^{T_s} \text{Tr}\left(\mathbf{\Psi}_{\mathcal{J}_s^t}\right) = \text{Tr}\left(\left(\sum_{t=1}^{T_s} \mathbf{E}_{\mathcal{J}_s^t}\mathbf{E}_{\mathcal{J}_s^t}^T\right)\mathbf{D}^T\mathbf{Q}\mathbf{D}\right)$$
$$= \text{Tr}\left(\frac{sT_s}{N}\mathbf{I}_N\mathbf{D}^T\mathbf{Q}\mathbf{D}\right) = sT_s. \tag{10}$$

By the arithmetic mean - harmonic mean inequality, it follows that:

$$\sum_{t=1}^{T_s}\sum_{k=1}^{s}\frac{1}{\lambda_k^{\mathcal{J}_s^t}} \geq \frac{(sT_s)^2}{\sum_{t=1}^{T_s}\sum_{k=1}^{s}\lambda_k^{\mathcal{J}_s^t}} = \frac{(sT_s)^2}{\sum_{t=1}^{T_s}\text{Tr}\left(\mathbf{\Psi}_{\mathcal{J}_s^t}\right)} = sT_s, \tag{11}$$

where one achieves the lower bound with $\lambda_k^{\mathcal{J}_s^t} = 1$ ($k = 1,\ldots,s$; $t = 1,\ldots,T_s$). This implies immediately that the matrix $\mathbf{Q} = \mathbf{I}_N$, which is consistent with the constraints, minimizes:

$$\mathbb{E}_{\mathcal{J}_s}\left(\text{Tr}\left((\mathbf{\Psi}_{\mathcal{J}_s})^{-1}\right)\right) = \frac{1}{T_s}\sum_{t=1}^{T_s}\sum_{k=1}^{s}\frac{1}{\lambda_k^{\mathcal{J}_s^t}}, \tag{12}$$

and hence also minimizes:

$$\mathbb{E}_{\mathcal{J}}\Big(\text{Tr}\Big((\mathbf{\Psi}_{\mathcal{J}})^{-1}\Big)\Big) = \sum_{s=1}^{S} p(\mathcal{J}_s)\mathbb{E}_{\mathcal{J}_s}\Big(\text{Tr}\left((\mathbf{\Psi}_{\mathcal{J}_s})^{-1}\right)\Big). \tag{13}$$

■

It is evident that the solution to the convex relaxation of the original optimization problem is not feasible, because $\texttt{rank}(\mathbf{I}_N) = N > M$. Therefore, we now propose to determine the $M \times N$ sensing matrix $\mathbf{\Phi}$ whose coherence matrix $\mathbf{Q} = \mathbf{\Phi}^T\mathbf{\Phi}$ is closest to the $N \times N$ identity matrix $\mathbf{I}_N$.

Proposition 2: Any $M \times N$ unit-norm tight frame is a solution to the optimization problem

$$\min_{\mathbf{\Phi}} \; \left\|\mathbf{\Phi}^T\mathbf{\Phi} - \mathbf{I}_N\right\|_F^2 \quad \text{s.t. } \|\boldsymbol{\phi}_j\|_{\ell_2} = 1 \; (j = 1,\ldots,N). \tag{14}$$

Proof: Let us note, in view of the fact that the set $\left\{\mathbf{\Phi} \in \mathbb{R}^{M\times N}: \; \texttt{Tr}\left(\mathbf{\Phi}^T\mathbf{\Phi}\right) = N\right\}$ contains the set $\left\{\mathbf{\Phi} \in \mathbb{R}^{M\times N}: \; \|\boldsymbol{\phi}_j\|_{\ell_2} = 1, \; j = 1,\ldots,N\right\}$, that the solution $\mathbf{\Phi}^*$ to the optimization problem

$$\min_{\mathbf{\Phi}} \; \left\|\mathbf{\Phi}^T\mathbf{\Phi} - \mathbf{I}_N\right\|_F^2 \quad \text{s.t. } \texttt{Tr}\left(\mathbf{\Phi}^T\mathbf{\Phi}\right) = N, \tag{15}$$

is also a solution to the optimization problem (14) if and only if $\|\boldsymbol{\phi}_j\|_{\ell_2} = 1$ ($j = 1,\ldots,N$). By using the singular value decomposition $\mathbf{\Phi} = \mathbf{U}_{\mathbf{\Phi}}\mathbf{R}_{\mathbf{\Phi}}\mathbf{V}_{\mathbf{\Phi}}^T$, where $\mathbf{U}_{\mathbf{\Phi}} \in \mathbb{R}^{M\times M}$ and $\mathbf{V}_{\mathbf{\Phi}} \in \mathbb{R}^{N\times N}$ are orthonormal matrices, and $\mathbf{R}_{\mathbf{\Phi}} \in \mathbb{R}^{M\times N}$ is a matrix whose main diagonal entries are the singular values of $\mathbf{\Phi}$, and the off-diagonal entries are zeros, we pose the convex optimization problem:

$$\min_{r_{1,1},\ldots,r_{M,M}} \; \sum_{k=1}^{M}(r_{k,k}^2 - 1)^2$$
$$\text{s.t. } \sum_{k=1}^{M} r_{k,k}^2 = N, \; r_{k,k} \geq 0 \; (k = 1,2,\ldots,M). \tag{16}$$

which, in view of the fact that $\left\|\mathbf{\Phi}^T\mathbf{\Phi} - \mathbf{I}_N\right\|_F^2 = \sum_{k=1}^{M}(r_{k,k}^2 - 1)^2$ and $\texttt{Tr}\left(\mathbf{\Phi}^T\mathbf{\Phi}\right) = \sum_{k=1}^{M} r_{k,k}^2$, leads to the solution of (15). Since the solution of (16) is $r_{k,k}^2 = \frac{N}{M}$ ($k = 1,2,\ldots,M$), it follows that any tight frame is the solution of (15). By proper choice of $\mathbf{V}_{\mathbf{\Phi}}$, i.e., $\sum_{j=1}^{M} v_{i,j}^2 = \frac{M}{N}$ for $\forall i \in \{1,\ldots,N\}$, any unit-norm (i.e., $\|\boldsymbol{\phi}_j\|_{\ell_2} = 1$ [13]) tight frame is also the solution of (15) and thus the solution of (14). ■

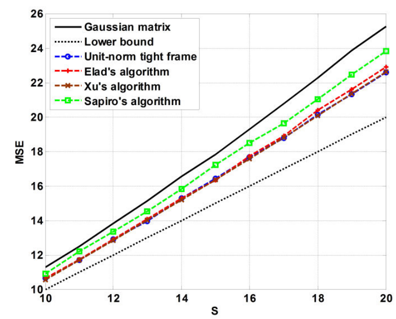


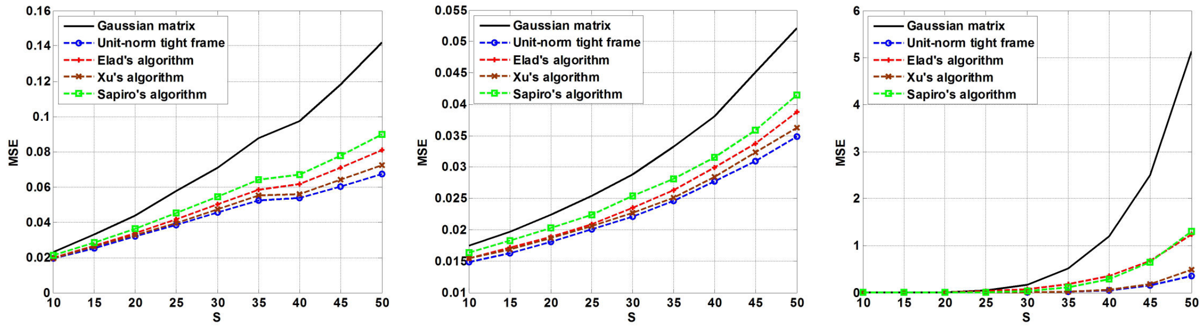
Fig. 1. Comparison of the MSE of different sensing matrices using the oracle estimator (N=200, M=100 and σ^2=1).

This proposition, which also follows from the Welch bound equality (WBE) sequences [14], establishes that a unit-norm tight frame is the closest design to the solution of the convex relaxation of the original optimization problem. Note that unit-norm tight frames can be constructed within a finite number of steps. For example, Davies and Higham's algorithm [10] generates a random unit-norm tight frame by performing $N-1$ rotations of a striped orthogonal matrix in at most $\mathcal{O}(MN)$ operations. This represents a clear computational advantage and consequently, provides an incentive to use unit-norm tight frames based sensing matrices rather than sensing matrix based on complex optimization approaches [6]–[8]. Note also that an alternative way to prove Proposition 2, which has been motivated by the optimization problem put forth by Duarte-Carvajalino and Sapiro [7], is also provided in [15]. The current problem differs from the problems in [7], [15] in various important aspects: i) our optimization approach is based on a metric with operational significance, the MSE, whereas the optimization approach in [7], [15] is based on mutual coherence; ii) our optimization approach applies to a very general sparse model whereas the optimization approach in [15] applies only to block-sparse models; and iii) our model is applicable only to orthonormal dictionaries but the models in [7], [15] are applicable to more general overcomplete dictionaries.

III. Performance Evaluation

We now compare the MSE performance of a random unit-norm tight frame based sensing matrix design to other sensing matrix designs including a Gaussian random matrix with normalized columns and matrices generated by three optimization methods, namely, Elad's algorithm [6], Sapiro's algorithm [7] and Xu's algorithm [8]. We consider throughout the sensing model in (1) with a fixed orthonormal basis consisting of Gaussian random vectors.

Fig 1 illustrates the MSE performance of the oracle estimator for the various sensing matrices. We calculate the MSE of the oracle estimator by averaging over 1000 different realizations of the sparse vector support. We also plot a lower



(a) BPDN (b) The Dantzig selector (c) OMP

Fig. 2. Comparison of the MSE of different sensing matrices ($M = 256$, $N = 512$ and $\sigma^2 = 10^{-4}$).

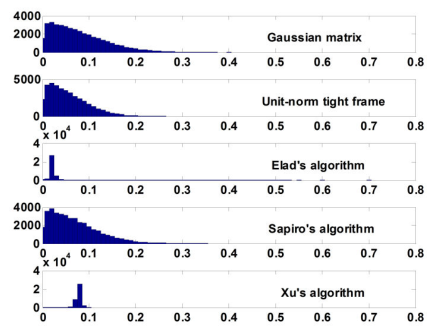


Fig. 3. Histogram of the off-diagonal absolute values of the coherence matrix (N=200 and M=100).

bound to the oracle MSE which, in view of Proposition 1, is given by $\sigma^2\mathbb{E}_{\mathcal{J}}\mathrm{Tr}\left(\left(\mathbf{E}_{\mathcal{J}}^T\mathbf{D}^T\mathbf{I}_N\mathbf{D}\mathbf{E}_{\mathcal{J}}\right)^{-1}\right) = \sigma^2 S$. It is evident that the unit-norm tight frame based sensing matrix outperforms all the other candidates, particularly, the random Gaussian matrix, for all sparsity levels.

Figs 2 illustrates the MSE performance of different practical estimators, namely, BPDN, Dantzig selector and OMP, for the various sensing matrices. We calculate the MSE between the original vector and the reconstructed vector by averaging over 1000 experiment trials, where in each trial we generate randomly a sparse vector with S randomly placed ± 1 spikes as well as a Gaussian noise vector. Once again it is evident that the unit-norm tight frame based sensing matrix outperforms the other sensing matrices for all these reconstruction algorithms. This superiority also justifies the use of the oracle MSE to enable sensing matrix design.

Finally, Fig 3 presents the distribution of the absolute values of the off-diagonal elements of the coherence matrices. Note that the mutual coherence of the unit-norm tight frame is not smaller than the mutual coherence of other sensing matrix designs. This shows that mutual coherence is not the most appropriate metric for the optimization of the average performance in compressive sensing applications, as already suggested in [6]. In contrast, we put forth a sensing matrix design method that is based on a metric with direct operational significance, the MSE, which leads naturally to better performance.

IV. Conclusion

In this letter, we propose the use of unit-norm tight frames for compressive sensing applications. This has been justified from optimization considerations, and has been shown to lead to MSE performance gains when used in conjunction with standard sparse recovery algorithms.

References

[1] D. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289 –1306, April 2006.

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

[3] E. Candès and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," *The Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, 2007.

[4] L. Yu, J. Barbot, G. Zheng, and H. Sun, "Compressive sensing with chaotic sequence," *Signal Processing Letters, IEEE*, vol. 17, no. 8, pp. 731 –734, Aug. 2010.

[5] S. Sarvotham and R. Baraniuk, "Deterministic bounds for restricted isometry of compressed sensing matrices," *Information Theory, IEEE Transactions on*, submitted for publication.

[6] M. Elad, "Optimized projections for compressed sensing," *Signal Processing, IEEE Transactions on*, vol. 55, no. 12, pp. 5695 –5702, Dec. 2007.

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

[8] J. Xu, Y. Pi, and Z. Cao, "Optimized projection matrix for compressive sensing," *EURASIP J. Adv. Signal Process*, vol. 43, pp. 1–8, Feb. 2010.

[9] K. Schnass and P. Vandergheynst, "Average performance analysis for thresholding," *Signal Processing Letters, IEEE*, vol. 14, no. 11, pp. 828 –831, Nov. 2007.

[10] P. Davies and N. Higham, "Numerically stable generation of correlation matrices and their factors," *BIT Numerical Mathematics*, vol. 40, no. 4, pp. 640–651, 2000.

[11] Z. Ben-Haim and Y. Eldar, "The cramèr-rao bound for estimating a sparse parameter vector," *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 3384 –3389, June 2010.

[12] Z. Ben-Haim, Y. Eldar, and M. Elad, "Coherence-based performance guarantees for estimating a sparse vector under random noise," *Signal Processing, IEEE Transactions on*, vol. 58, no. 10, pp. 5030 –5043, Oct. 2010.

[13] P. G. Casazza and J. Kovaevi, "Equal-norm tight frames with erasures," *Advances in Computational Mathematics*, vol. 18, pp. 387–430, 2003.

[14] S. Waldron, "Generalized welch bound equality sequences are tight frames," *Information Theory, IEEE Transactions on*, vol. 49, no. 9, pp. 2307 – 2309, Sep. 2003.

[15] L. Zelnik-Manor, K. Rosenblum, and Y. Eldar, "Sensing matrix optimization for block-sparse decoding," *Signal Processing, IEEE Transactions on*, vol. 59, no. 9, pp. 4300 –4312, Sep. 2011.