

On the Benefit of using Tight Frames for Robust Data Transmission and Compressive Data Gathering in Wireless Sensor Networks

Wei Chen*, Miguel R. D. Rodrigues[†] and Ian J. Wassell*

* Computer Laboratory, University of Cambridge, UK

[†] Instituto de Telecomunicações, Departamento de Ciência de Computadores, Universidade do Porto, Portugal

Abstract—Compressive sensing (CS), a new sampling paradigm, has recently found several applications in wireless sensor networks (WSNs). In this paper, we investigate the design of novel sensing matrices which lead to good expected-case performance - a typical performance indicator in practice - rather than the conventional worst-case performance that is usually employed when assessing CS applications. In particular, we show that tight frames perform much better than the common CS Gaussian matrices in terms of the reconstruction average mean squared error (MSE). We also showcase the benefits of tight frames in two WSN applications, which involve: i) robustness to data sample losses; and ii) reduction of the communication cost.

I. INTRODUCTION

WIRELESS sensor networks (WSNs) have attracted world-wide attention in recent years owing to the reduced cost of sensor nodes and the increasing demand for gathering various sorts of physical information from the environment. These sensor nodes have low power and limited processing resources compared to the fusion center (FC), which collects and analyzes all the sensed data in the field. The communications cost in data gathering is crucial for a WSN as it determines the lifetime of the network and the amount of wireless spectrum resources required. Another problem in WSNs is data loss, which is inevitable due to the characteristic of the wireless medium. While a number of schemes have been proposed to deal with these issues [1], [2], the success of a new sampling paradigm, known as compressive sensing (CS) [3], [4], motivates novel designs for WSNs.

CS leverages the compressibility of signals in nature to reduce the number of samples required for perfect or nearly-perfect reconstruction. Therefore, this is a promising technique for applications where the data gathering and communication is expensive. In addition, CS also trades-off sampling rate for reconstruction complexity, a property which suits the fact that a FC has much more computational power than the sensor nodes in WSNs.

Several CS-based schemes have been proposed in the literature for data gathering in WSNs. For example, Bajwa et al. [5] proposed a distributed communication architecture using an analog coherent transmission scheme. Charbiwala et al. [6] exploited CS as an application layer erasure coding strategy for

recovering missing data. Luo et al. [7] proposed a compressive data gathering scheme to reduce communication cost for large scale WSNs. A comprehensive survey of CS-based network data gathering is given in [8].

The key novel feature of CS is using incoherent sampling instead of the traditional Nyquist sampling method [3], [4]. The majority of the work in the literature considers the use of Gaussian sensing matrices to perform the sampling operation. These matrices obey the so-called restricted isometry property (RIP), which represents a sufficient condition for successful reconstruction of the signals [4].

In this paper, we propose to use specialized random matrices, named tight frames, for CS sampling owing to its superior expected-case reconstruction performance compared to random Gaussian matrices. We also propose the use of tight frames based sensing matrices in a variety of WSN applications. In particular, we show that tight frames are robust against sample losses in WSN applications, and that tight frames can reduce the transmission cost in WSN applications.

The following notation is used. Upper-case letters denote numbers, boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors and calligraphic upper-case letters denote sets. $(\cdot)^T$, $\text{Tr}(\cdot)$ and $\text{rank}(\cdot)$ denote the transpose, trace and rank of a matrix, respectively. $X_{i,j}$ denotes the element of the i th row and j th column of the matrix \mathbf{X} . $\mathcal{N}(\mu, \sigma^2)$ denotes the Gaussian distribution with a mean μ and a variance σ^2 . $\|\mathbf{x}\|_0$ denotes the ℓ_0 norm, i.e., the number of nonzero elements of the vector \mathbf{x} . $\|\mathbf{x}\|_1$ denotes the ℓ_1 norm, i.e., the summation of the absolute value of elements of the vector \mathbf{x} . $\|\mathbf{x}\|_2 = \sqrt{\sum_i x_i^2}$ denotes the ℓ_2 norm of the vector \mathbf{x} . $\|\mathbf{X}\|_F$ denotes the Frobenius norm of the matrix \mathbf{X} .

II. COMPRESSIVE SENSING AND TIGHT FRAMES: THE BASICS

A. Compressive Sensing

Signals in nature are often compressible due to inherent redundancy. Unlike the traditional Nyquist sampling method, CS directly acquires a reduced number of signal samples on the premise of the compressibility of signals thereby obviating the need for an extra compression process after the sampling process. We assume that a discrete signal $\mathbf{f} \in \mathbb{R}^N$ can be represented by a sparse signal $\mathbf{x} \in \mathbb{R}^N$ in some orthonormal

The work of M. R. D. Rodrigues was supported by Fundação para a Ciência e a Tecnologia through the research project CMU-PT/SIA/0026/2009.

basis, i.e.,

$$\mathbf{f} = \mathbf{D}\mathbf{x}, \quad (1)$$

where $\mathbf{D} \in \mathbb{R}^{N \times N}$ denotes the orthonormal basis. In particular, the signal \mathbf{f} is S -sparse if it is a linear combination of S basis vectors only, i.e., the vector \mathbf{x} contains only S nonzero elements.

The conventional CS sensing model can be written as follows:

$$\mathbf{y} = \Phi \mathbf{D}\mathbf{x} + \mathbf{n}, \quad (2)$$

where $\Phi \in \mathbb{R}^{M \times N}$ ($M < N$) denotes a sensing matrix and $\mathbf{n} \in \mathbb{R}^M$ denotes a noise vector with i.i.d. random elements drawn according to $\mathcal{N}(0, \sigma^2)$.

In general, solving (2) with a given Φ , \mathbf{D} and \mathbf{y} is impossible even in the absence of noise. Fortunately, with a sparse constraint, i.e., $S \leq \frac{M}{2}$ [4], it is possible to estimate the sparse vector \mathbf{x} by solving the following optimization problem:

$$\begin{aligned} \min_{\hat{\mathbf{x}}} \quad & \|\hat{\mathbf{x}}\|_0 \\ \text{s.t.} \quad & \|\Phi \mathbf{D}\hat{\mathbf{x}} - \mathbf{y}\|_2 \leq \epsilon, \end{aligned} \quad (3)$$

where $\epsilon \geq \sqrt{M\sigma^2}$. In the view of the fact that such an optimization problem is NP-complete, the following convex relaxation of the original optimization problem is applied:

$$\begin{aligned} \min_{\hat{\mathbf{x}}} \quad & \|\hat{\mathbf{x}}\|_1 \\ \text{s.t.} \quad & \|\Phi \mathbf{D}\hat{\mathbf{x}} - \mathbf{y}\|_2 \leq \epsilon, \end{aligned} \quad (4)$$

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

The reconstruction performance of (4) depends upon the ‘‘quality’’ of a sensing matrix Φ . A common metric to assess the quality of a sensing matrix is the mutual coherence [9] given by:

$$\mu = \max_{1 \leq i, j \leq N, i \neq j} |\Psi_{i,j}|, \quad (5)$$

where $\Psi = \mathbf{D}^T \Phi^T \Phi \mathbf{D}$ denotes the coherence matrix of $\Phi \mathbf{D}$. In general, the smaller the mutual coherence, the better quality the sensing matrix. Another common metric to assess the quality of the sensing matrix is the Babel function, also called cumulative coherence [10], [11], given by:

$$\mu_p(S) = \max_{\|\mathcal{L}\|_0 = S, j \notin \mathcal{L}} \sum_{i \in \mathcal{L}} \|\Psi_{i,j}\|_p, \quad (6)$$

where \mathcal{L} is a subset of the column indexes of Ψ , and $p = 1$ and $p = 2$ denote the 1-Babel function and the 2-Babel function respectively. The Babel function evaluates the matrix quality in a vector scale, while mutual coherence considers the entry-wise incoherence. Currently, only coherence based techniques such as mutual coherence and its refined variants, i.e., Babel functions, are computationally tractable thereby providing the means to evaluate the quality of the sensing matrix and reconstruction performance.

B. Tight Frames

For finite-dimensional real spaces, a frame can be seen as a matrix $\Phi \in \mathbb{R}^{M \times N}$ such that for any vector $\mathbf{z} \in \mathbb{R}^N$,

$$A \|\mathbf{z}\|_2^2 \leq \|\Phi^T \mathbf{z}\|_2^2 \leq B \|\mathbf{z}\|_2^2, \quad (7)$$

where $A > 0$ and $B > 0$ are called frame bounds. Tight frames are a class of frames with equal frame bounds, i.e., $A = B$.

Proposition 1:

- 1) Let $M = N$ and $A = B = 1$. Then the matrix Φ is a tight frame if and only if it is an orthogonal matrix.
- 2) Let $M \leq N$. Then the matrix $\Phi \in \mathbb{R}^{M \times N}$ is a tight frame if and only if it is a solution to $\min_{\Phi} \|\Phi^T \Phi - A \mathbf{I}_N\|_F^2$, where \mathbf{I}_N is the $N \times N$ identity matrix.
- 3) Let $M \leq N$. Then the matrix $\Phi \in \mathbb{R}^{M \times N}$ is a tight frame if and only if $\Phi \Phi^T = A \mathbf{I}_M$, where \mathbf{I}_M is the $M \times M$ identity matrix.

Proof: See Appendix. ■

Proposition 1 illustrates the connection between tight frames and orthogonal matrices. In particular, a tight frame represents a matrix whose coherence matrix is as close as possible to an orthogonal matrix according the Frobenius norm. Frames have been widely used in many applications such as denoising, CDMA systems and multiantenna code design [12].

III. THE SUPERIOR EXPECTED-CASE RECONSTRUCTION PERFORMANCE BY USING TIGHT FRAMES

We are interested in the expected-case performance of CS reconstruction. We wish to minimize the mean squared error (MSE) in estimating \mathbf{x} from \mathbf{y} , according to the sensing model in (2), given by

$$\text{MSE}(\Phi) = \mathbb{E}_{\mathbf{x}, \mathbf{n}} (\|\mathcal{F}(\Phi \mathbf{D}\mathbf{x} + \mathbf{n}) - \mathbf{x}\|_2^2), \quad (8)$$

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} . We assume a random signal model where the signal is exactly S -sparse. We also assume a random signal model where the positions of the S non-zero elements follows an equiprobable distribution.

We consider the well-known oracle estimator, which uses least square estimation based on the prior knowledge of the positions of the S non-zero elements of the sparse signal, in order to bound the MSE in (8). The reason for using the oracle estimator is based on the fact that the oracle MSE is equal to the unbiased Cramér-Rao bound (CBD) for exactly S sparse deterministic vectors [13]. The oracle MSE - for a fixed sparse vector \mathbf{x} - is given by:

$$\begin{aligned} \text{MSE}_{\mathbf{n}}^{\text{oracle}}(\Phi, \mathbf{x}) &= \mathbb{E}_{\mathbf{n}} (\|\mathcal{F}^{\text{oracle}}(\Phi \mathbf{D}\mathbf{x} + \mathbf{n}) - \mathbf{x}\|_2^2) \\ &= \sigma^2 \text{Tr} \left((\mathbf{E}_{\mathcal{J}}^T \mathbf{D}^T \Phi^T \Phi \mathbf{D} \mathbf{E}_{\mathcal{J}})^{-1} \right), \end{aligned} \quad (9)$$

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

Consequently, the average value of the oracle MSE - which acts as a lower bound to the MSE in (8) - is given by:

$$\begin{aligned} \text{MSE}^{\text{oracle}}(\Phi) &= \mathbb{E}_{\mathbf{x}} \left(\text{MSE}_{\mathbf{n}}^{\text{oracle}}(\Phi, \mathbf{x}) \right) \\ &= \mathbb{E}_{\mathbf{x}} \left(\sigma^2 \text{Tr} \left((\mathbf{E}_{\mathcal{J}}^T \mathbf{D}^T \Phi^T \Phi \mathbf{D} \mathbf{E}_{\mathcal{J}})^{-1} \right) \right) \\ &= \sigma^2 \mathbb{E}_{\mathcal{J}} \left(\text{Tr} \left((\mathbf{E}_{\mathcal{J}}^T \mathbf{D}^T \Phi^T \Phi \mathbf{D} \mathbf{E}_{\mathcal{J}})^{-1} \right) \right), \end{aligned} \quad (10)$$

where $\mathbb{E}_{\mathbf{x}}(\cdot)$ and $\mathbb{E}_{\mathcal{J}}(\cdot)$ denote expectation with respect to the distribution of the random vector \mathbf{x} and the random set \mathcal{J} , respectively. In (10), 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 positions of the nonzero elements of the random vector \mathbf{x} , owing to the use of the oracle.

We define the coherence matrix of the sensing matrix as $\mathbf{Q} = \Phi^T \Phi$. We now look for the coherence matrix, which up to a rotation leads to the sensing matrix, that minimizes the average value of the oracle MSE subject to an energy constraint by posing the following optimization problem:

$$\begin{aligned} \min_{\mathbf{Q}} \quad & \mathbb{E}_{\mathcal{J}} \left(\text{Tr} \left((\mathbf{E}_{\mathcal{J}}^T \mathbf{D}^T \mathbf{Q} \mathbf{D} \mathbf{E}_{\mathcal{J}})^{-1} \right) \right) \\ \text{s.t.} \quad & \mathbf{Q} \succeq 0, \\ & \text{Tr}(\mathbf{Q}) = N, \\ & \text{rank}(\mathbf{Q}) \leq M, \end{aligned} \quad (11)$$

where $\mathbf{Q} \succeq 0$ denotes that the matrix \mathbf{Q} is positive semi-definite. However, (11) is non-convex due to the rank constraint. Therefore, we first consider a convex relaxation of (11) by ignoring the rank constraint, and then look for a feasible solution that is the closest to the solution to the relaxed problem and also satisfies the rank constraint.

Proposition 2: The solution of the optimization problem:

$$\begin{aligned} \min_{\mathbf{Q}} \quad & \mathbb{E}_{\mathcal{J}} \left(\text{Tr} \left((\mathbf{E}_{\mathcal{J}}^T \mathbf{D}^T \mathbf{Q} \mathbf{D} \mathbf{E}_{\mathcal{J}})^{-1} \right) \right) \\ \text{s.t.} \quad & \mathbf{Q} \succeq 0, \\ & \text{Tr}(\mathbf{Q}) = N, \end{aligned} \quad (12)$$

which represents a convex relaxation of the original optimization problem in (11), is the $N \times N$ identity matrix \mathbf{I}_N .

Proof: In the interest of space, the proof is omitted, but it can be found in [14]. ■

Proposition 2, in conjunction with Proposition 1, leads us to conclude that a tight frame is the closest design - in the Frobenius norm sense - to the solution of the convex relaxation of the original optimization problem in (11). Fig. 1 illustrates the average MSE performance of the oracle estimator and the conventional ℓ_1 minimization estimator. The orthonormal basis used here is a fixed orthogonal matrix generated randomly and the MSE is calculated by averaging over 1000 trials, where in each trial we generate randomly a sparse vector with S randomly placed ± 1 spikes. The reconstruction performance of tight frames based sensing matrices clearly outperforms the reconstruction performance of other sensing matrix designs including Gaussian matrices, Elad's design [15] and Sapiro's design [16].

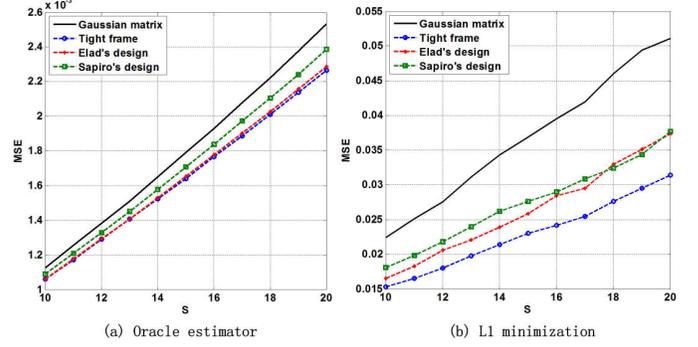


Fig. 1. Comparison of the MSE of different sensing matrices ($M = 100$, $N = 200$ and $\sigma^2 = 10^{-4}$).

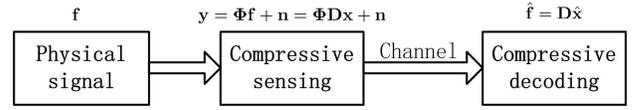


Fig. 2. Compressive sensing erasure coding.

IV. APPLICATIONS IN WSNs

We have seen improved expected-case CS reconstruction performance by using tight frames based sensing matrices rather than Gaussian sensing matrices. In this section, we further show two beneficial characteristics of tight frames, and their usage in WSNs.

A. Robustness of Tight Frames Against Data Loss

In [6], Charbiwala et al. use CS as an application layer erasure coding strategy for recovering missing data. In particular, as shown in Fig. 2, physical signals are sampled using the CS approach rather than traditional sampling used in conjunction with source and channel coding. The data, which is transmitted through a memoryless erasure channel, is reconstructed at the FC. The objective is to explore CS to shift the computation burden from the sensor nodes to the FC. The approach is robust to data loss at the physical layer, because the CS decoder can reconstruct the signal at the application layer.

We assume that K out of the M samples are erased or lost. Then, CS sensing model in (2) can be re-written as follows:

$$\mathbf{y}^{\mathcal{K}} = \mathbf{E}^{\mathcal{K}} \mathbf{y} = \mathbf{E}^{\mathcal{K}} \Phi \mathbf{D} \mathbf{x} + \mathbf{E}^{\mathcal{K}} \mathbf{n}, \quad (13)$$

where where \mathcal{K} represents a set with containing the positions of the K erased samples, $\mathbf{y}^{\mathcal{K}}$ denotes the vector that results from the vector \mathbf{y} by deleting the elements in the positions indexed by the elements of the set \mathcal{K} , and $\mathbf{E}^{\mathcal{K}} \in \mathbb{R}^{K \times M}$ denotes the matrix that results from the identity matrix by deleting the set of rows indexed by the elements of the set \mathcal{K} . Then the average reconstruction MSE becomes

$$\begin{aligned} \text{MSE}(\Phi) &= \mathbb{E}_{\mathbf{x}, \mathbf{n}, \mathcal{K}} \left(\|\mathcal{F}(\mathbf{E}^{\mathcal{K}} \Phi \mathbf{D} \mathbf{x} + \mathbf{E}^{\mathcal{K}} \mathbf{n}) - \mathbf{x}\|_2^2 \right) \\ &= \mathbb{E}_{\mathcal{K}} \left(\text{MSE}(\mathbf{E}^{\mathcal{K}} \Phi) \right), \end{aligned} \quad (14)$$

where $\mathbb{E}_{\mathbf{x}, \mathbf{n}, \mathcal{K}}(\cdot)$ denotes expectation with respect to the joint distribution of the random vector \mathbf{x} , \mathbf{n} and the random set \mathcal{K} ,

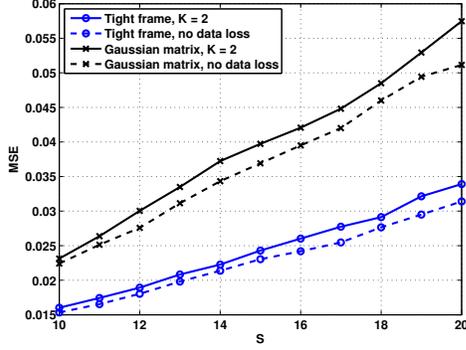


Fig. 3. Comparison of the MSE of Gaussian matrix and tight frames by using ℓ_1 minimization reconstruction in the case of data loss. ($M = 100$, $N = 200$ and $\sigma^2 = 10^{-4}$).

and $\mathbb{E}_{\mathcal{K}}(\cdot)$ denotes expectation with respect to the distribution of the random set \mathcal{K} .

The following Proposition shows that if Φ is a tight frame then $\mathbb{E}^{\mathcal{K}}\Phi$ is also a tight frame, which suggests that the expected case performance of a tight frame is also better than that of other sensing matrix designs in this application.

Proposition 3: Let $\mathcal{K} \subset \{1, \dots, M\}$ denote a set of integers. Let also $\mathbb{E}^{\mathcal{K}} \in \mathbb{R}^{K \times M}$ ($K < M$) denote the matrix that results from the identity matrix by deleting the set of rows indexed by the elements of the set \mathcal{K} . It follows that if $\Phi \in \mathbb{R}^{M \times N}$ ($M \leq N$) is a tight frame, then $\mathbb{E}^{\mathcal{K}}\Phi$ is also a tight frame.

Proof: Proposition 1 shows that a matrix is a tight frame if and only if its rows are orthogonal. Thus, any matrix that results from deleting a certain number of rows from a tight frame is still a tight frame. ■

Fig. 3 demonstrates that indeed tight frames based sensing matrices are more robust to data loss than Gaussian sensing matrices. The orthonormal basis used here is a fixed orthogonal matrix generated randomly and the MSE is calculated by averaging over 1000 trials, where in each trial we generate randomly a sparse vector with S randomly placed ± 1 spikes. The positions of the K erased samples are also randomly generated.

B. Efficient Compressive Data Gathering

In [7], Luo et al. proposed a compressive data gathering (CDG) scheme to reduce the communication cost of WSNs. In particular, they considered a WSN composed of one FC and N sensor nodes arranged in a chain-type topology. Let us denote by f_i the measurement of sensor node i ($i = 1, \dots, N$). The baseline data gathering scheme shown in Fig. 4 (a) requires $\mathcal{O}(N^2)$ total message transmissions in the network and $\mathcal{O}(N)$ maximum message transmissions for any single sensor node. On the other hand, in the CDG scheme shown in Fig. 4 (b), each sensor node transmits M ($M < N$) messages that are the sum of the received message vector $\sum_{i=1}^{j-1} \phi_i f_i$ and its own message vector $\phi_j f_j$ generated by multiplying its message f_j with a spreading code $\phi_j \in \mathbb{R}^M$. By exploiting the CS principle, the CDG scheme only requires $\mathcal{O}(MN)$ total

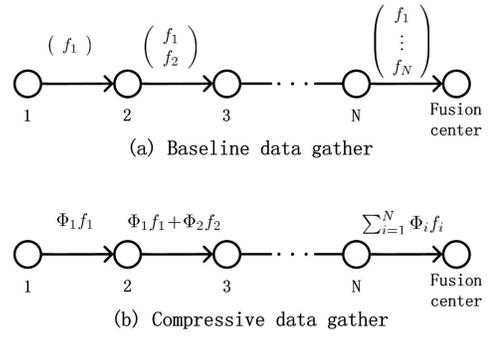


Fig. 4. Comparing compressive data gathering and baseline data gathering in a chain-type topology.

message transmissions in the network and $\mathcal{O}(M)$ maximum message transmissions for any single node.

The success of the CDG scheme relies on the assumption that the signal sensed by those sensor nodes has spatial correlation, which has been observed in many applications of WSNs.¹

The fact that the mathematical model of the CDG scheme is equivalent to the general CS model in (2), suggests the use of a tight frame, rather than the traditional Gaussian matrices, to define the spreading codes in the CDG application will lead to better average MSE reconstruction performance. We consider a more efficient CDG scheme in terms of transmission cost where the (sensing) matrix that defines the spreading codes exhibits the structure:

$$\Phi = \begin{bmatrix} 1 & 0 & \cdots & 0 & \Phi_{1,M+1} & \cdots & \Phi_{1,N} \\ 0 & 1 & \cdots & 0 & \Phi_{2,M+1} & \cdots & \Phi_{2,N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \Phi_{M,M+1} & \cdots & \Phi_{M,N} \end{bmatrix}.$$

This matrix can be seen as the concatenation of an identity matrix and another matrix. Therefore, the first M sensor nodes in the chain only need to transmit one message per sensor, while M messages per sensor are required to be transmitted if the traditional spreading codes (having M non-zero entries) are used in the first M columns of Φ . That is, the total number of message transmissions in the network can be further reduced to $M(N - M + 1)$.

The following Proposition shows that if the matrix $\Phi \in \mathbb{R}^{(N-M) \times N}$ is a tight frame then the matrix $[\mathbf{I}_M \Phi]$, where \mathbf{I}_M is an $M \times M$ identity matrix, is also a tight frame, which suggests that the expected case performance of the spreading codes that correspond to such a design is better than that of the traditional spreading codes in this application.

Proposition 4: Let \mathbf{I}_M as an $M \times M$ identity matrix. It follows that if $\Phi \in \mathbb{R}^{(N-M) \times N}$ ($M < N$) is also a tight frame, then the concatenated matrix $[\mathbf{I}_M \Phi]$ is a tight frame.

¹This CDG scheme can also be applied to a tree-type topology. Although the transmission cost for the tree-type topology is different to the chain-type topology, the superiority of the CDG scheme against the baseline scheme does not change [7]. Here we only consider the chain-type topology due to the limited space, but our results can also be applied to a tree-type topology.

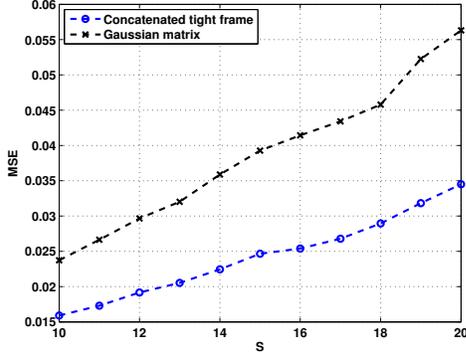


Fig. 5. Comparison of the MSE of Gaussian matrix and concatenated tight frames by using ℓ_1 minimization reconstruction. ($M = 100$, $N = 200$ and $\sigma^2 = 10^{-4}$).

Proof: Proposition 1 shows that a matrix is a tight frame if and only if its rows are orthogonal. Then, the matrix $[\mathbf{I}_M \ \Phi]$ is a tight frame because its rows are orthogonal. ■

Fig. 5 demonstrates that indeed tight frames based design leads to superior reconstruction than a Gaussian matrix based design. The orthonormal basis used here is a fixed orthogonal matrix generated randomly and the MSE is calculated by averaging over 1000 trials, where in each trial we generate randomly a sparse vector with S randomly placed ± 1 spikes.

V. CONCLUSIONS

In this paper, it is shown that tight frames based sensing matrices provide better expected case performance than other sensing matrix designs, most notably, the widely used Gaussian sensing matrices. It is also shown that tight frames based sensing matrices exhibit properties suitable for WSN applications, such as robustness to data sample losses.

APPENDIX A PROOF OF PROPOSITION 1

Let $\lambda_1 \geq \dots \geq \lambda_N \geq 0$ be the eigenvalues of the positive semi-definite matrix $\Phi^T \Phi$. According to the definition of frame in (7), we have the maximum eigenvalue of $\Phi^T \Phi$ given by

$$\lambda_1 = \max_{\mathbf{z} \neq 0} \frac{\|\Phi^T \mathbf{z}\|_2}{\|\mathbf{z}\|_2} = B. \quad (15)$$

As $M \leq N$, the matrix $\Phi^T \Phi$ has at most M nonzero eigenvalues. Then, the minimum non-zero eigenvalue of the matrix $\Phi^T \Phi$ is given by

$$\lambda_M = \min_{\mathbf{z} \neq 0} \frac{\|\Phi^T \mathbf{z}\|_2}{\|\mathbf{z}\|_2} = A. \quad (16)$$

Assume that $M = N$. If $A = B = 1$ then $\lambda_1 = \dots = \lambda_M = 1$ and so the matrix Φ is orthogonal. Conversely, if the matrix Φ is orthogonal then $\lambda_1 = \dots = \lambda_M = 1$ and so $A = B = 1$.

Assume now that $M \leq N$. The minimum of

$$\|\Phi^T \Phi - \mathbf{A} \mathbf{I}_N\|_F^2 = \sum_{i=1}^N (\lambda_i - A)^2 = \sum_{i=1}^M (\lambda_i - A)^2 + (N - M)A^2, \quad (17)$$

is achieved when $\lambda_i = A$ ($i = 1, \dots, M$), i.e., Φ is a tight frame. Conversely, if a matrix Φ is a tight frame then it minimizes $\|\Phi^T \Phi - \mathbf{A} \mathbf{I}_N\|_F^2$.

Consider the singular value decomposition $\Phi = \mathbf{U} \mathbf{R} \mathbf{V}^T$, where $\mathbf{U} \in \mathbb{R}^{M \times M}$ and $\mathbf{V} \in \mathbb{R}^{N \times N}$ are orthogonal matrices, and $\mathbf{R} \in \mathbb{R}^{M \times N}$ is a matrix whose main diagonal entries are the singular values of Φ . Thus, we have

$$\Phi \Phi^T = \mathbf{U} \mathbf{R} \mathbf{R}^T \mathbf{U}^T, \quad (18)$$

where the diagonal entries of the matrix $\mathbf{R} \mathbf{R}^T$ are λ_i ($i = 1, \dots, M$). Therefore, if a matrix Φ is a tight frame, i.e., $\lambda_1 = \dots = \lambda_M = A$, then $\Phi \Phi^T = \mathbf{A} \mathbf{I}_M$. Conversely, if $\Phi \Phi^T = \mathbf{A} \mathbf{I}_M$ then $\lambda_1 = \dots = \lambda_M = A$, i.e., the matrix Φ is a tight frame.

REFERENCES

- [1] "Energy conservation in wireless sensor networks: A survey," *Ad Hoc Networks*, vol. 7, no. 3, pp. 537 – 568, 2009.
- [2] I. Akyildiz and M. Vuran, *Wireless sensor networks*. John Wiley & Sons Inc, 2010.
- [3] D. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289 – 1306, April 2006.
- [4] 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.
- [5] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Joint source-channel communication for distributed estimation in sensor networks," *Information Theory, IEEE Transactions on*, vol. 53, no. 10, pp. 3629 – 3653, Oct. 2007.
- [6] Z. Charbiwala, S. Chakraborty, S. Zahedi, Y. Kim, M. Srivastava, T. He, and C. Bisdikian, "Compressive oversampling for robust data transmission in sensor networks," in *INFOCOM, 2010 Proceedings IEEE*, March 2010, pp. 1 – 9.
- [7] C. Luo, F. Wu, J. Sun, and C. W. Chen, "Efficient measurement generation and pervasive sparsity for compressive data gathering," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 12, pp. 3728 – 3738, December 2010.
- [8] J. Haupt, W. Bajwa, M. Rabbat, and R. Nowak, "Compressed sensing for networked data," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 92 – 101, March 2008.
- [9] D. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *Information Theory, IEEE Transactions on*, vol. 52, no. 1, pp. 6 – 18, 2006.
- [10] K. Schnass and P. Vandergheynst, "Dictionary preconditioning for greedy algorithms," *Signal Processing, IEEE Transactions on*, vol. 56, no. 5, pp. 1994 – 2002, May 2008.
- [11] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *Information Theory, IEEE Transactions on*, vol. 54, no. 5, pp. 2210 – 2219, May 2008.
- [12] J. Kovačević and A. Chebira, "An introduction to frames," *Signal Processing*, vol. 2, no. 1, pp. 1–94, 2008.
- [13] Z. Ben-Haim and Y. Eldar, "The cramer-rao bound for estimating a sparse parameter vector," *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 3384 – 3389, June 2010.
- [14] W. Chen, M. Rodrigues, and I. Wassell, "On the use of unit-norm tight frames to improve the average mse performance in compressive sensing applications," *Signal Processing Letters, IEEE*, vol. 19, no. 1, pp. 8 – 11, Jan. 2012.
- [15] M. Elad, "Optimized projections for compressed sensing," *Signal Processing, IEEE Transactions on*, vol. 55, no. 12, pp. 5695 – 5702, Dec. 2007.
- [16] 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.