FRAMES, ERASURES, AND SIGNAL ESTIMATION WITH STOCHASTIC MODELS

SOMANTIKA DATTA

ABSTRACT. Frame properties and conditions are determined that would minimize the error in signal reconstruction or estimation in the presence of noise and erasures. The special focus here is on stochastic models. These include estimating a random signal with zero mean and a general covariance matrix, minimizing the mean-squared error (MSE) when the frame coefficients are erased according to some a priori probability distribution in the presence of random noise, and also studying the use of stochastic frames in estimating a random signal. In estimating a random signal from noisy coefficients, when a frame coefficient is lost or erased, it is established that the MSE is minimized under certain geometric relationships between the frame vectors and the signal. When the coefficients are erased according to some a priori distribution, conditions are found for the norms of the frame vectors in terms of the probability distribution of the erasure so that the MSE is minimized. Results obtained here also show how using stochastic frames can lead to more flexibility in design and greater control on the MSE.

KEYWORDS: Erasures, estimation, frames, LMMSE estimation, random signal, stochastic frames
2010 MSC: 94A12; 42C15

1. Introduction and Background

Frame theory has gained significant attention in recent years due to applications in signal processing. Representation of signals in terms of frames are robust to transmission losses and resilient to noise. The purpose of this work is to determine conditions on frames so that signal reconstruction can be efficient in the presence of erasures. The study of optimal frames for erasures has been done by other authors, most notably in [5] and [4] where the latter also considers the presence of random noise in the frame measurements. Under this class of work, much of the focus in the literature has been on deterministic models only. In contrast, the work here is focused on stochastic models. These include considering a random signal, the frame coefficients being erased according to some a priori probability distribution in the presence of random noise, and also the use of stochastic frames. The analysis is almost entirely for the situation where there is a single erasure, i.e, one of the frame coefficients is lost. Where applicable, comments are made on the situation with more than one erasure. Under a stochastic model, the general idea is to minimize the mean-square error when the signal is reconstructed from frame measurements that are corrupted by noise and erasures. Some of the calculations involving random signals are inspired by the work in [7] where a random vector is reconstructed from its noisy projections onto low-dimensional subspaces and under subspace erasures, using fusion frames. However, some different aspects are considered here. Unlike in [7], the covariance matrix of the noise or the signal is not assumed to be diagonal. Additionally, the results here are for frame measurements, and not for fusion frame measurements. Also, as far as we are aware, assuming some a priori probability distribution on the erasure location and the use of stochastic frames have not been addressed elsewhere.

Some basic definitions and facts on frames are now given. For details, the reader is referred to [1]. In what follows, the setting will be a finite d-dimensional Hilbert space $\mathcal{H}$. Let $\Phi = \{f_i\}_{i=1}^m \subseteq \mathcal{H}$ denote a set of m vectors in $\mathcal{H}$. The set Φ is called a *frame* if there exist constants $0 < A \leq B < \infty$ such that for all $x \in \mathcal{H}$

$$A\|x\|^2 \leq \sum_{i=1}^{m} |\langle x, f_i\rangle|^2 \leq B\|x\|^2. \tag{1.1}$$

The constants A and B are called the upper and lower *frame bounds*, respectively. When $A = B$, the frame is said to be *tight*. A frame with all vectors having the same norm is called an *equal norm frame* or a *unit norm frame* (if the norms are all equal to one). For a finite dimensional Hilbert space, a frame is the same as a spanning set. Let F be the $m \times d$ matrix whose kth row is f_k^*. F is called the *analysis operator*. The $d \times d$ matrix $S = F^*F$ is called the *frame operator*. It is self-adjoint, positive definite, and invertible. For a tight frame, S is a multiple of the identity, i.e., S equals AI, where I is the identity. Under the above setting, the following result will be used [1].

Theorem 1.1. *(a) The smallest and largest eigenvalues of S are the optimal lower and upper frame bounds, respectively.*

(b) Let $\{\lambda_k\}_{k=1}^d$ denote the eigenvalues of S, including repetitions. Then

$$\sum_{k=1}^{d} \lambda_k = \sum_{i=1}^{m} \|f_i\|^2.$$

Due to Theorem 1.1 (b), if $\{f_i\}_{i=1}^m$ is a unit norm tight frame (UNTF) then its frame bound is $A = \frac{m}{d}$. The ratio of the frame size to the dimension, i.e., $\frac{m}{d}$ is called the *redundancy* of the frame. The $m \times m$ matrix of inner products, $FF^* =: G$, is called the *Gram matrix* of Φ. The rank of G is d, and the nonzero eigenvalues of G are the same as those of S. Let x be a signal that is a vector in $\mathcal{H}$. If S is the frame operator then

$$x = \sum_{i=1}^{m} \langle x, f_i\rangle S^{-1} f_i = \sum_{i=1}^{m} \langle x, S^{-1} f_i\rangle f_i. \tag{1.2}$$

The set $\{S^{-1}f_i\}_{i=1}^m$ is also a frame for $\mathcal{H}$. It is called the *canonical dual frame* of Φ and denoted by $\{\widetilde{f_i}\}_{i=1}^m$.

In the deterministic setting, for a deterministic signal with deterministic frames, it was shown in [5] that equiangular tight frames are optimal for multiple erasures. The presence of random noise in the frame coefficients $\{\langle x, f_i\rangle\}$ was considered in [4] where it was shown that unit norm tight frames minimize the average and maximum mean-squared error (MSE) over a single erasure. In the realm of fusion frames, estimating a random signal and determining properties of robust fusion frames, in the presence of random noise and subspace erasures, was done in [7].

As already stated, the purpose of this work is to study frames and signal reconstruction in the presence of erasures with a special focus on stochastic models for the erasure, signal, and also the frame. Suppose that one of the frame coefficients $\langle x, f_k \rangle$ is lost at random which is the same as one of the frame vectors f_k being deleted. Denote the remaining set by $\Phi_k := \Phi \setminus \{f_k\}$. If Φ_k is still a frame for $\mathcal{H}$ then exact reconstruction of x is possible from the remaining coefficients by using the frame operator S_k of Φ_k :

$$x = \sum_{i \neq k}^{m} \langle x, f_i \rangle S_k^{-1} f_i.$$

However, the frame properties of the random subframe Φ_k will be affected by the deletion, and this is studied in Section 2. When the deletion of a vector leaves behind a set that is no longer a frame then the standard choice is to set the lost coefficient to zero, however, this is not necessarily the best choice. In this case, conditions that determine the choice of a replacement for the erased coefficient so that the error can be minimized is discussed in Section 3. Estimating a random signal x from noisy frame coefficients when there is an erasure is studied in Section 4. Conditions on the frame that would minimize the mean-squared error (MSE) are established there. This was done in [7] using fusion frames when the covariance of x is of the form $\sigma_x^2 I$ whereas here a general covariance matrix is considered using traditional frames. For minimizing the MSE, this leads to some geometric connections between the signal and the frame vectors, see Theorem 4.3. Further, Section 4 also addresses the situation where the erasure is assumed to have some a priori distribution and minimizing the MSE in this case gives rise to conditions on the norms of the frame vectors based on the probability distribution of the location of the erasure, see Theorem 4.4. This was not done in [7] or elsewhere. Finally, in Section 5, the use of stochastic frames in estimating a random signal is studied, and it is shown how using stochastic frames can lead to more flexibility in design and greater control on the MSE, see Remark 5.2. The use of stochastic frames in signal reconstruction has been vastly untouched so far.

2. Changes in frame properties after deleting a frame vector

Consider the situation where a vector f_k is deleted from a frame Φ and the remaining set Φ_k is still a frame. This may cause some changes in the structure of the resulting frame, including changes in the frame bounds and the condition number of the frame. Such changes may effect the stability of the reconstruction in (1.2).

2.1. **Changes in frame bounds after deleting a frame vector.** The results below show how the frame bounds of Φ_k are related to those of Φ. All results are independent of the location k. The results can thus be interpreted as being for *random* subframes of Φ of size $m - 1$. The following is a slight generalization of Theorem 4.1 in [4].

Theorem 2.1. *(i) Let $\{f_i\}_{i=1}^m$ in $\mathcal{H}$, $m > d$, be a unit norm frame with bounds $1 < A \le B < \infty$. For any k, $\{f_i\}_{i \ne k}$ is a frame with a lower bound $A_1 = A - 1$ and an upper bound $B_1 = B$.*
(ii) If $\{f_i\}_{i=1}^m$ is a unit norm tight frame then the bounds A_1 and B_1 mentioned in part (i) are optimal*, and $\{f_i\}_{i \ne k}$ is non-tight.*

Proof. (i) By the Cauchy-Schwarz Inequality, using the fact that the frame vectors are unit norm,

$$|\langle x, f_k \rangle| \le \|x\| \|f_k\| = \|x\|.$$

Therefore, from (1.1),

$$(A-1)\|x\|^2 \le \sum_{i \ne k} |\langle x, f_i \rangle|^2 \le B\|x\|^2. \tag{2.1}$$

(ii) If the starting frame is unit norm and tight then $A = B = \frac{m}{d}$. For all $x \in \mathcal{H}$

$$\frac{m}{d}\|x\|^2 = \sum_{i=1}^m |\langle x, f_i \rangle|^2.$$

When $x = f_k$, equality is attained on the left inequality in (2.1). For x such that $\langle x, f_k \rangle = 0$, equality is attained on the right inequality in (2.1). Thus the optimal lower and upper bounds are $\frac{m}{d} - 1$ and $\frac{m}{d}$, respectively. Since these bounds are not equal to each other, $\{f_i\}_{i \ne k}$ is non-tight. □

When the number of erasures is more that one, the following is obtained in a similar manner.

Theorem 2.2 ([1], Proposition 1.5.6)**.** *For $e \ge 1$ erasures, the remaining vectors form a frame if the starting frame $\{f_i\}$ is unit norm with lower frame bound $A > e$. After e erasures, the lower frame bound of a unit norm frame can decrease by at most e.*

Theorem 2.3. *Let $\{f_i\}_{i=1}^m$ be a tight frame in $\mathcal{H}$, not necessarily unit norm, with bound equal to A. If the vector f_k is removed then the resulting set Φ_k is no longer a tight frame. Further, Φ_k is a frame if and only if $\|f_k\|^2 < A$.*

Proof. Due to tightness,

$$A\|x\|^2 = \sum_{i=1}^m |\langle x, f_i \rangle|^2, \quad \forall x \in \mathcal{H}.$$

Suppose that the kth vector is removed for some $1 \le k \le m$. By the Cauchy-Schwarz Inequality and (1.1),

$$(A - \|f_k\|^2)\|x\|^2 \le \sum_{i \ne k} |\langle x, f_i \rangle|^2 \le A\|x\|^2.$$

As done in the proof of Theorem 2.1 (ii), setting $x = f_k$ shows that $A - \|f_k\|^2$ is the *optimal* lower bound, and taking x such that $\langle x, f_k \rangle = 0$ shows that A is the *optimal* upper bound of Φ_k. Since f_k is not zero, the lower and upper bounds are not equal, and Φ_k is not tight. Further, since $A - \|f_k\|^2$ is the optimal lower bound, this implies that Φ_k is a frame if and only if $\|f_k\|^2 < A$. □

Theorem 2.4. *Let $\Phi = \{f_i\}_{i=1}^m$ be a frame for $\mathcal{H}$ that is not necessarily unit norm. It is not possible for every subset of $m-1$ vectors of Φ to be a tight frame.*[a]

Proof. Let F_k be the analysis operator of Φ_k. The frame operator of Φ_k can then be written as

$$F_k^* F_k = F^* F - f_k f_k^*.$$

Suppose that every subset of $m-1$ vectors of Φ is a tight frame. Then there exist constants $\lambda_1, \lambda_2, \ldots, \lambda_m$ such that for $k = 1, \ldots, m$,

$$F^* F - f_k f_k^* = \lambda_k I.$$

Adding over all k, gives

$$F^* F = \frac{\sum_{k=1}^m \lambda_k}{m-1} I$$

which implies that Φ is tight. However, in that case, by Theorem 2.3, if a vector is removed from Φ then the resulting frame is not tight. This contradicts the assumption that every subset of $m-1$ vectors is tight. □

2.2. **Change in the condition number of a frame after deleting a vector.** The condition number [b] of a tight frame is 1 making tight frames highly desirable as they offer numerical stability in the reconstruction formula (1.2) which involves inverting the frame operator. Thus, changes in the condition number with the deletion of a frame vector is of interest. This involves studying the eigenvalue properties of the principal submatrices of the Gram matrix of Φ.

Proposition 2.5. *Let Φ be a unit norm tight frame (UNTF) of m vectors in $\mathcal{H}$ with $m > d$. Then, regardless of the location, the condition number always increases after a single erasure.*

Proof. Recall that for a UNTF the frame bound is $A = \frac{m}{d}$, and the condition number is 1. By Theorem 2.1, the condition number of Φ_k is

$$c_k = \frac{A}{A-1} = 1 + \frac{1}{A-1} = 1 + \frac{1}{\frac{m}{d} - 1} > 1. \tag{2.2}$$

□

Remark 2.6. *From (2.2), it can be seen that higher the redundancy ($\frac{m}{d}$) of the starting UNTF Φ, the closer the condition number of Φ_k is to 1.*

[a]This result was proved in [4] with the assumption that the starting frame is unit norm.
[b]By condition number of a frame is meant the ratio of the maximum and the minimum eigenvalue of the frame operator, i.e., the ratio of the optimal upper and lower frame bounds.

Using the following Theorem 2.7 [6] one can construct a frame such that the deletion of a certain frame vector leaves the condition number unchanged (see Remark 2.8 and Example 2.9).

Theorem 2.7. [6] *Let m be a given positive integer, and let $\{\mu_i : i = 1, 2, \ldots, m\}$ and $\{\lambda_i : i = 1, 2, \ldots, m+1\}$ be two given sequences of real numbers such that*

$$\lambda_1 \le \mu_1 \le \lambda_2 \le \mu_2 \le \cdots \le \mu_{m-1} \le \lambda_m \le \mu_m \le \lambda_{m+1}.$$

Let $\Lambda = diag(\mu_1, \mu_2, \ldots, \mu_m)$. There exists a real number a and a real vector $y \in \mathbb{R}^m$ such that $\{\lambda_i : i = 1, 2, \ldots, m + 1\}$ is the set of eigenvalues of the real symmetric matrix

$$G \equiv \begin{bmatrix} \Lambda & y \\ y^T & a \end{bmatrix}.$$

Remark 2.8. *The proof of Theorem 2.7 [6] provides a way to construct the vector y and find the number a. G can be thought of as the Gram matrix of the starting frame Φ with $m+1$ frame vectors. When the last row and column of G are deleted then the diagonal matrix Λ is the Gram matrix of the frame Φ_{m+1}. With an appropriate choice of the μs and the λs, Φ_{m+1} has the same frame bounds as Φ thus leaving the condition number unchanged after erasure at location $m + 1$. The diagonal elements of G are the lengths of the frame vectors and so in this process one does not necessarily construct a unit norm frame. The sequence $\{\lambda_i\}$ must be strictly positive so that there are no zero vectors. Since Λ is diagonal, Φ_{m+1} is an orthogonal basis. It is thus implied that this is applicable for a starting frame of $d + 1$ vectors in a d-dimensional space such that removing the last vector leaves behind an orthogonal basis. An example demonstrating this is given next.*

Example 2.9. Suppose that $m = 4$, and $\{\lambda_i\} = \{0, 1, 3/2, 3/2\}$. Thus one is considering a starting frame Φ of four vectors in $\mathbb{R}^3$. Let $\{\mu_i\} = \{1, 5/4, 3/2\}$. Using Theorem 2.7, one can obtain the matrix G as

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 5/4 & 0 & \sqrt{5}/4 \\ 0 & 0 & 3/2 & 0 \\ 0 & \sqrt{5}/4 & 0 & 1/4 \end{bmatrix}.$$

The frame vectors of Φ have lengths 1, 5/4, 3/2, and 1/4, and so it is not unit norm. Φ has frame bounds $A = 1$ and $B = 3/2$, and so it is not tight. After deleting the last row and column of G one is left with the Gram matrix of an orthogonal basis of $\mathbb{R}^3$. The frame bounds and hence the condition number of this basis (frame) are the same as those of Φ. However, any other principal submatrix of G may not give the same result. For example, deleting the first row and column will give the Gram matrix of a tight frame for $\mathbb{R}^2$ (not $\mathbb{R}^3$), whereas, deleting the second row and column will give an orthogonal basis of $\mathbb{R}^3$ with frame bounds 1/4 and 3/2.

Let G_k denote the kth principal submatrix of G, i.e., G_k is the Gram matrix of Φ_k. Arrange the eigenvalues of G and G_k in increasing order. Let λ_j denote the jth eigenvalue of G and let $\mu_{k,j}$ denote the jth eigenvalue of G_k. The eigenvalues of G and those of the principal submatrices satisfy [6]

$$\max_{1\leq k\leq m} \mu_{k,j} \geq \frac{m-j}{m}\lambda_1 + \frac{j}{m}\lambda_{j+1}, \quad j=1,\dots,m, \tag{2.3}$$

and

$$\min_{1\leq k\leq m} \mu_{k,j} \leq \frac{m-j}{m}\lambda_j + \frac{j}{m}\lambda_{m}, \quad j=1,\dots,m. \tag{2.4}$$

If Φ_k is a frame for $\mathcal{H}$, then G_k has rank d. Define the best case condition number for any Φ_k as

$$\frac{\min\limits_{1\leq k\leq m} \mu_{k,m-1}}{\max\limits_{1\leq k\leq m} \mu_{k,m-d}},$$

and the worst case condition number as

$$\frac{\max\limits_{1\leq k\leq m} \mu_{k,m-1}}{\min\limits_{1\leq k\leq m} \mu_{k,m-d}}.$$

With these definitions, Proposition 2.10 below shows that frames with higher redundancy perform better.

Proposition 2.10. *Let c be the condition number of the starting frame Φ. For any k, if Φ_k is a frame for $\mathcal{H}$, then the best case condition number of Φ_k satisfies*

$$c \leq \frac{\min\limits_{1\leq k\leq m} \mu_{k,m-1}}{\max\limits_{1\leq k\leq m} \mu_{k,m-d}} \leq c\left(1+\frac{1}{\frac{m}{d}-1}\right)$$

and the worst case condition number satisfies

$$\frac{\max\limits_{1\leq k\leq m} \mu_{k,m-1}}{\min\limits_{1\leq k\leq m} \mu_{k,m-d}} \geq 1+\frac{1-\frac{1}{d}}{\frac{m}{d}-1}.$$

Proof. Since the rank of G is d, the eigenvalues of G can be written as

$$0=\lambda_1=\cdots=\lambda_{m-d}<\lambda_{m-d+1}\leq\cdots\leq\lambda_m.$$

The condition number of Φ is

$$c=\frac{\lambda_m}{\lambda_{m-d+1}}.$$

If Φ_k is a frame for $\mathcal{H}$, then G_k also has d nonzero eigenvalues. By the interlacing theorem [6],

$$0=\lambda_1=\mu_{k,1}=\lambda_2=\cdots=\mu_{k,m-d-1}=\lambda_{m-d}<\mu_{k,m-d}\leq\lambda_{m-d+1}\leq\cdots\leq\lambda_{m-1}\leq\mu_{k,m-1}\leq\lambda_m. \tag{2.5}$$

Substituting $m-1$ and $m-d$ for j in (2.4) and (2.3), respectively, gives

$$\min_{1\le k\le m}\mu_{k,m-1} \le \frac{1}{m}\lambda_{m-1} + \frac{m-1}{m}\lambda_m, \quad \text{and} \quad \max_{1\le k\le m}\mu_{k,m-d} \ge \frac{d}{m}\lambda_1 + \frac{m-d}{m}\lambda_{m-d+1}. \tag{2.6}$$

From (2.5) and (2.6)

$$c = \frac{\lambda_{m-1}}{\lambda_{m-d+1}} \le \frac{\min\limits_{1\le k\le m}\mu_{k,m-1}}{\max\limits_{1\le k\le m}\mu_{k,m-d}} \le \frac{\lambda_{m-1}+(m-1)\lambda_m}{(m-d)\lambda_{m-d+1}} \le \frac{m}{m-d}c = c\left(1+\frac{1}{\frac{m}{d}-1}\right)$$

which gives the bounds for the best case condition number for any Φ_k.

Next substitute $m-1$ and $m-d$ for j in (2.3) and (2.4) to get

$$\max_{1\le k\le m}\mu_{k,m-1} - \min_{1\le k\le m}\mu_{k,m-d} \ge \frac{\lambda_1}{m} + \frac{d-1}{m}\lambda_m - \frac{d}{m}\lambda_{m-d} = \frac{d-1}{m}\lambda_m \quad (\text{since } \lambda_1 = \lambda_{m-d} = 0),$$

or, $$\frac{\max\limits_{1\le k\le m}\mu_{k,m-1}}{\min\limits_{1\le k\le m}\mu_{k,m-d}} \ge 1 + \frac{d-1}{m}\frac{\lambda_m}{\min\limits_{1\le k\le m}\mu_{k,m-d}} \ge 1 + \frac{d-1}{m}\lambda_m\frac{1}{\frac{m-d}{m}\lambda_m} = 1 + \frac{1-\frac{1}{d}}{\frac{m}{d}-1}$$

where the last inequality is obtained by using $j = m-d$ in (2.4), and the fact that $\lambda_{m-d} = 0$. □

3. When erasure results in a set that is no longer a frame

If $\Phi_k = \{f_i\}_{i\neq k}$ is not a frame for $\mathcal{H}$ then exact reconstruction of x may not be possible when the kth frame coefficient is erased. The erased coefficient $\langle x, f_k\rangle$ can be replaced by some linear combination of the known coefficients at the reconstruction step. One wants to pick scalars $\{\alpha_j\}$ in such a way that when the linear combination $\sum_{i\neq k}\alpha_i\langle x, f_i\rangle$ is used in place of $\langle x, f_k\rangle$ then the reconstruction $\widehat{x}_k$ is closest to x in some sense. Here $\widehat{x}_k$ is calculated as

$$\widehat{x}_k = \sum_{i\neq k}\langle x, f_i\rangle S^{-1}f_i + \left(\sum_{i\neq k}\alpha_i\langle x, f_i\rangle\right)S^{-1}f_k = \sum_{i\neq k}\langle x, f_i\rangle\widetilde{f_i} + \langle x, \sum_{i\neq k}\overline{\alpha}_i f_i\rangle\widetilde{f_k} \tag{3.1}$$

Using (1.2) and Cauchy-Schwarz, the relative error can be bounded by

$$\frac{\|x-\widehat{x}_k\|}{\|x\|} \le \|f_k - \sum_{j\neq k}\overline{\alpha}_j f_j\|\|\widetilde{f_k}\| \tag{3.2}$$

To minimize the ℓ_2-norm of the upper bound in (3.2), one can choose the α_js such that $\sum_{j\neq k}\overline{\alpha}_j f_j$ is the orthogonal projection, $\mathrm{Proj}(f_k)$, of f_k on $W = \mathrm{span}\{f_j\}_{j\neq k}$. In other words, the deleted vector f_k is replaced by $\mathrm{Proj}(f_k)$.

Proposition 3.1. *Suppose that when the kth frame coefficient is erased or f_k is deleted, Proj(f_k) is used in place of f_k. Then the upper bound of the average relative error is minimized in ℓ_2-norm when the starting frame is tight. Further, the average relative error is lower for frames with higher redundancy.*

Proof. By the Pythagorean Identity

$$\|f_k\|^2 = \|f_k - \mathrm{Proj}(f_k)\|^2 + \|\mathrm{Proj}(f_k)\|^2 \geq \|f_k - \mathrm{Proj}(f_k)\|^2.$$

Using this in (3.2) gives

$$\frac{\|x - \widehat{x}_k\|}{\|x\|} \leq \|f_k - \mathrm{Proj}(f_k)\|\|\widetilde{f}_k\| \leq \|f_k\|\|\widetilde{f}_k\| \leq \frac{1}{A}\|f_k\|^2.$$

The average relative error is

$$\frac{1}{m}\sum_{k=1}^{m}\frac{\|x - \widehat{x}_k\|}{\|x\|} \leq \frac{1}{m}\frac{1}{A}\sum_{k=1}^{m}\|f_k\|^2 \leq \frac{1}{m}\frac{1}{A}\sum_{i=1}^{d}\lambda_i(S) \leq \frac{d}{m}\frac{1}{A}\lambda_{\max}(S) = \frac{d}{m}\frac{B}{A}.$$

The quantity $\frac{d}{m}\frac{B}{A}$ is inversely proportional to $\frac{m}{d}$, and is minimized when $B = A$, i.e., when the frame is tight. □

Often the deleted frame vector is replaced by zero in which case the reconstruction based on the remaining coefficients is

$$\widehat{x} = \sum_{i \neq k}\langle x, f_i\rangle\tilde{f}_i.$$

When the deleted frame vector is replaced by zero, the upper bound of the average relative error is minimized under the same conditions of Proposition 3.1.

Depending on the signal and the deleted vector, one has to choose a replacement for the erased coefficient. The following example shows that in some cases setting the lost coefficient to zero is better than setting it to $\langle x, \mathrm{Proj} f_k\rangle$ and vice versa, while at times, depending upon the signal, it makes no difference. At other times, neither is the best choice.

Example 3.2. Consider the frame for $\mathbb{R}^3$ given by

$$f_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, \quad f_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}, \quad f_3 = \begin{bmatrix}1\\1\\1\end{bmatrix}, \quad f_4 = \begin{bmatrix}1\\1\\0\end{bmatrix}.$$

The canonical dual frame is

$$\widetilde{f}_1 = \begin{bmatrix}2/3\\-1/3\\-1/3\end{bmatrix}, \quad \widetilde{f}_2 = \begin{bmatrix}-1/3\\2/3\\-1/3\end{bmatrix}, \quad \widetilde{f}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}, \quad \widetilde{f}_4 = \begin{bmatrix}1/3\\1/3\\-2/3\end{bmatrix}.$$

Remove f_3. The set $\{f_1, f_2, f_4\}$ is no longer a frame for $\mathbb{R}^3$. The orthogonal projection of f_3 onto $\mathrm{span}\{f_1, f_2, f_4\}$ is $\mathrm{Proj} f_3 = (1, 1, 0)$. The reconstructions of x obtained from replacing the erased coefficient $\langle x, f_3\rangle$ by zero and by $\langle x, \mathrm{Proj} f_3\rangle$ are compared below for several different choices of x.

(a) Take $x = (1, -1, 0)$. Then $\langle x, f_3\rangle = \langle x, \mathrm{Proj} f_3\rangle = 0$. Replacing $\langle x, f_3\rangle$ either by zero or by $\langle x, \mathrm{Proj} f_3\rangle$ gives *exact* reconstruction.

(b) Let $x = (1,0,2)$. First replace $\langle x, f_3\rangle$ by zero. The reconstruction is

$$\widehat{x} \quad = \quad \langle x, f_1\rangle \widetilde{f}_1 + \langle x, f_2\rangle \widetilde{f}_2 + \langle x, f_4\rangle \widetilde{f}_4 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \neq x.$$

The ℓ_2-norm of the error is $\|x - \widehat{x}\| = 3$. Next replace $\langle x, f_3\rangle$ by $\langle x, \mathrm{Proj} f_3\rangle$. Then

$$\widehat{x} = \langle x, f_1\rangle \widetilde{f}_1 + \langle x, f_2\rangle \widetilde{f}_2 + \langle x, f_4\rangle \widetilde{f}_4 + \langle x, \mathrm{Proj} f_3\rangle \widetilde{f}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \neq x.$$

The norm of the error is $\|x - \widehat{x}\| = 2$ which is lower than when the lost coefficient is set to zero. However, neither zero nor the orthogonal projection of the deleted vector is the best choice. Exact reconstruction can be obtained by replacing $\langle x, f_3\rangle$ by $\langle x, 2f_1 + cf_2 + f_4\rangle$, where c is any scalar.

(c) Lastly, take $x = (1, 0, -2)$. As before, first replace the lost coefficient $\langle x, f_3\rangle$ by zero. Then $\widehat{x} = (1, 0, -1)$, and $\|x - \widehat{x}\| = 1$. Replacing the lost coefficient by $\langle x, \mathrm{Proj} f_3\rangle$ gives $\widehat{x} = (1, 0, 0)$, and $\|x - \widehat{x}\| = 2$ which is higher than when the lost coefficient is set to zero.

Proposition 3.3 and Corollary 3.4 below give a structure on the frame with respect to the signal that guarantees exact reconstruction for both the zero replacement and when the deleted vector is replaced by its orthogonal projection $\mathrm{Proj} f_k$.

Proposition 3.3. *Let $\{e_i\}_{i=1}^d$ be an orthonormal basis (ONB) of $\mathcal{H}$. Suppose that $\{e_i\}_{i=1}^d \subseteq \{f_i\}_{i=1}^m$, the deleted frame vector is $f_k = e_n$ for some $1 \leq n \leq d$, and that $\{f_i\}_{i\neq k}$ is not a frame for $\mathcal{H}$. Then f_k belongs to span$\{f_i\}_{i\neq k}^{\perp}$ and span$\{e_i\}_{i\neq n}$ = span$\{f_j\}_{j\neq k}$.*

Proof. Since $\{f_i\}_{i\neq k}$ is not a frame for $\mathcal{H}$, there exists a nonzero $v \in \mathcal{H}$ such that

$$\sum_{i\neq k} |\langle v, f_i\rangle|^2 = 0$$

i.e., v is orthogonal to each f_i for $i \neq k$. In other words, v belongs to $\mathrm{span}\{f_i\}_{i\neq k}^{\perp}$. Since $\{e_i\}_{i=1}^d \subseteq \{f_i\}_{i=1}^m$, this further implies that v is orthogonal to each e_i, $i \neq n$, i.e., $v = ce_n$ for some nonzero scalar c. Thus $e_n = f_k$ belongs to $\mathrm{span}\{f_i\}_{i\neq k}^{\perp}$.

Since $\{e_i\}_{i=1}^d$ is an ONB of $\mathcal{H}$, the dimension of $\mathrm{span}\{e_i\}_{i\neq n}$ is $d - 1$. Also, since $\{f_i\}_{i\neq k}$ is no longer a frame but $\{f_i\}_{i=1}^m$ is, the dimension of $\mathrm{span}\{f_i\}_{i\neq k}$ is also $d - 1$. Since $\mathrm{span}\{e_i\}_{i\neq n} \subseteq \mathrm{span}\{f_i\}_{i\neq k}$ and both have the same dimension, therefore, $\mathrm{span}\{e_i\}_{i\neq n} = \mathrm{span}\{f_i\}_{i\neq k}$. □

Corollary 3.4. *Under the assumptions of Proposition 3.3, let $x \in$ span$\{e_j\}_{j\neq n}$. Then exact reconstruction is obtained both when the lost coefficient $\langle x, f_k\rangle$ is set to zero and when $\langle x, f_k\rangle$ is replaced by $\langle x, Projf_k\rangle$.*

Proof. $\langle x, e_n\rangle = \langle x, f_k\rangle = 0$, and note that $\{e_j\}_{j\neq n}$ is an ONB for span$\{e_j\}_{j\neq n}$, i.e., $\{e_j\}_{j\neq n}$ is an ONB for span$\{f_j\}_{j\neq k}$. Therefore, Proj$f_k = 0$, and $\langle x, \text{Proj}f_k\rangle = 0$. □

In what follows, when the lost coefficient $\langle x, f_k\rangle$ is replaced by zero denote the reconstruction by $\widehat{x}_0$. If $\langle x, f_k\rangle$ is replaced by $\sum_{j\neq k}\alpha_j\langle x, f_j\rangle$ denote the reconstruction by $\widehat{x}_k$ as given in (3.1).

Theorem 3.5. *Suppose that $\{f_i\}_{i=1}^m$ is a UNTF with bound A. When the coefficient $\langle x, f_k\rangle$ is lost, the average squared error is lower if it is replaced by $\sum_{j\neq k}\alpha_j\langle x, f_j\rangle$ than when replaced by zero, provided that the scalars α_js can be picked such that*

$$\sum_{k=1}^{m}\|f_k - \sum_{j\neq k}\overline{\alpha}_j f_j\|^2 \leq A. \tag{3.3}$$

Proof. If the lost coefficient is replaced by zero, the norm of the error is

$$\|x - \widehat{x}_0\| = \frac{1}{A}|\langle x, f_k\rangle|\|f_k\| = \frac{1}{A}|\langle x, f_k\rangle|.$$

Otherwise, the error is

$$\|x - \widehat{x}_k\| = \frac{1}{A}|\langle x, f_k - \sum_{j\neq k}\overline{\alpha}_j f_j\rangle|\|f_k\| = \frac{1}{A}|\langle x, f_k - \sum_{j\neq k}\overline{\alpha}_j f_j\rangle|$$

The average squared error over all k for $\widehat{x}_0$ is

$$\frac{1}{m}\frac{1}{A^2}\sum_{k=1}^{m}|\langle x, f_k\rangle|^2 = \frac{1}{m}\frac{1}{A^2}A\|x\|^2 = \frac{\|x\|^2}{mA}.$$

For $\widehat{x}_k$, the average squared error over all k is

$$\frac{1}{m}\frac{1}{A^2}\sum_{k=1}^{m}|\langle x, f_k - \sum_{j\neq k}\overline{\alpha}_j f_j\rangle|^2 \leq \frac{\|x\|^2}{mA^2}\sum_{k=1}^{m}\|f_k - \sum_{j\neq k}\overline{\alpha}_j f_j\|^2$$

where the bound is obtained by using the Cauchy Schwarz Inequality. Choosing the αs such that

$$\frac{\|x\|^2}{mA^2}\sum_{k=1}^{m}\|f_k - \sum_{j\neq k}\overline{\alpha}_j f_j\|^2 \leq \frac{\|x\|^2}{mA}$$

simplifies to (3.3). □

4. Reconstructing random signals from noisy measurements with erasure

Let x be a random vector in $\mathcal{H}$ with mean zero and covariance matrix $E[xx^T] = R_{xx}$. In this section, the goal is to reconstruct x from noisy measurements with a real frame $\{f_i\}_{i=1}^m$. This has been studied in the realm of fusion frames [7] where the authors have shown that by using a linear minimum mean-squared error (LMMSE) estimator the mean-squared error (MSE) is minimum when the fusion frame is tight. In addition, the authors have shown that in the presence of a subspace erasure, maximum robustness is achieved with a tight fusion frame having subspaces with the same dimension. In [7], the analysis is limited to the case where the signal covariance is of the form $R_{xx} = \sigma_x^2 I$, and a white noise vector with covariance $\sigma_n^2 I$. Here a general R_{xx} as well as noise vector is considered. Further, when considering a single erasure, a priori distribution is assumed on the location of the erasure, and this was not considered in [7]. The notation used below is the same as in [7].

Suppose that the frame coefficients of x are corrupted by noise, i.e., one has to reconstruct x from

$$z_i = \langle x, f_i \rangle + n_i, \quad i = 1, \ldots, m,$$

where each n_i has mean zero and the noise vector $n = (n_1, n_2, \ldots, n_m)^T$ has covariance matrix Σ. Suppose that x and n are uncorrelated. If $z = (z_1, z_2, \ldots, z_m)^T$ then

$$z = Fx + n.$$

Denote the covariance matrix of z by R_{zz} and that between x and z by R_{xz}. The goal is to estimate x from z. The LMMSE filter or the Wiener filter is $K = R_{xz} R_{zz}^{-1}$. Following [8, 7], the LMMSE estimator of x is

$$\widehat{x} = Kz.$$

Using the fact that the error $x - Kz$ is orthogonal to Kz [8], the covariance of the error is

$$\begin{aligned} R_{ee} = E[(x-Kz)(x-Kz)^T] &= R_{xx} - R_{xz} R_{zz}^{-1} R_{zx} = R_{xx} - R_{xx} F^T (F R_{xx} F^T + \Sigma)^{-1} F R_{xx} \\ &= (R_{xx}^{-1} + F^T \Sigma^{-1} F)^{-1} \end{aligned} \tag{4.1}$$

where the last step uses the Sherman-Morrison-Woodbury matrix inversion formula [3, p. 50].

Theorem 4.1. *When estimating a zero mean random signal x from noisy measurements with the LMMSE estimator:*
(i) Among all frames with m vectors, the MSE is minimized when the frame is tight and when the covariance of the noise is of the form $\sigma_n^2 I$.
(ii) When $\Sigma = \sigma_n^2 I$, among unit norm tight frames (UNTFs), higher the redundancy of the frame, lower the MSE.

Proof. (i) Since Σ is a Hermitian matrix, there is a unitary matrix U and a diagonal matrix D such that

$$\Sigma = U^T D U.$$

Then from (4.1)

$$R_{ee} = E[(x - Kz)(x - Kz)^T] = (R_{xx}^{-1} + F^T U^T D^{-1/2} D^{-1/2} UF)^{-1}.$$

Let $P = D^{-1/2}UF$. Then

$$R_{ee} = (R_{xx}^{-1} + P^T P)^{-1}.$$

The fact that U is unitary leads to

$$A\|x\|^2 \leq \|UFx\|^2 = \|Fx\|^2 = \sum_{i=1}^{m} |\langle x, f_i\rangle|^2 \leq B\|x\|^2.$$

Let $\lambda_{\max}$ and $\lambda_{\min}$ denote the maximum and minimum eigenvalues of D or Σ. Then

$$\frac{1}{\lambda_{\max}} A\|x\|^2 \leq \|D^{-1/2}UFx\|^2 = \|Px\|^2 \leq \frac{1}{\lambda_{\min}} B\|x\|^2.$$

Note that the rows of P form a frame with bounds $\frac{A}{\lambda_{\max}}$ and $\frac{B}{\lambda_{\min}}$. Thus the eigenvalues of its frame operator $P^T P$ lie in $[\frac{A}{\lambda_{\max}}, \frac{B}{\lambda_{\min}}]$. Denote the eigenvalues of R_{xx} by $0 < \mu_1 \leq \mu_2 \leq \ldots \leq \mu_d$ and those of $R_{xx}^{-1} + P^T P$ by $\phi_1 \geq \phi_2 \geq \ldots \geq \phi_d > 0$. Therefore,

$$\sum_i \frac{1}{\frac{1}{\mu_i} + \frac{B}{\lambda_{\min}}} \leq \sum_i \frac{1}{\phi_i} \leq \sum_i \frac{1}{\frac{1}{\mu_i} + \frac{A}{\lambda_{\max}}}. \tag{4.2}$$

The MSE equals the trace of R_{ee} which is $\sum_i \frac{1}{\phi_i}$, and attains the lower bound when

$$\frac{B}{\lambda_{\min}} = \frac{A}{\lambda_{\max}}$$

$$\text{or,} \quad \frac{B}{A} = \frac{\lambda_{\min}}{\lambda_{\max}}. \tag{4.3}$$

The left side of (4.3) is the condition number of the frame operator whereas the right side is the inverse of the condition number of Σ. The two can be equal if and only if they are both equal to 1 which means that the frame is tight and the covariance of the noise is a constant multiple of the identity, say $\sigma_n^2 I$.

(ii) For a UNTF

$$A = \frac{1}{d}\sum_{i=1}^{m} \|f_i\|^2 = \frac{m}{d}$$

Additionally, if $\Sigma = \sigma_n^2 I$, then from (4.2)

$$MSE = \sum_{i=1}^{d} \frac{1}{\frac{1}{\mu_i} + \frac{A}{\sigma_n^2}} = \sum_{i=1}^{d} \frac{1}{\frac{1}{\mu_i} + \frac{1}{\sigma_n^2}\frac{m}{d}} \tag{4.4}$$

The expression in (4.4) suggests that higher the redundancy $\frac{m}{d}$, the lower the MSE. □

4.1. **Erasures.** Suppose that the kth frame coefficient is erased. Let E_k be the $m \times m$ diagonal matrix whose kth diagonal entry is 1 and every other entry is zero. Where the location of the erasure is irrelevant, the subscript k is dropped and just E is written. The measurements in the presence of noise would be

$$\widetilde{z} = (I - E)z.$$

As done in [7], the estimate of a random signal x is taken to be

$$\widetilde{x} = K\widetilde{z}$$

where $K = R_{xz}R_{zz}^{-1}$ is the no erasure LMMSE filter used above. From [7], the covariance of the error is

$$\begin{aligned}\widetilde{R}_{ee} &= E[(x - K\widetilde{z})(x - K\widetilde{z})^{\mathrm{T}}] = R_{xx} - R_{xz}R_{zz}^{-1}R_{zx} + R_{xz}R_{zz}^{-1}ER_{zz}E^{\mathrm{T}}R_{zz}^{-1}R_{zx} \\ &= R_{ee} + \overline{R}_{ee}\end{aligned}$$

where $R_{ee} = R_{xx} - R_{xz}R_{zz}^{-1}R_{zx}$ is the no erasure covariance and

$$\overline{R}_{ee} = R_{xz}R_{zz}^{-1}ER_{zz}E^{\mathrm{T}}R_{zz}^{-1}R_{zx} \tag{4.5}$$

is the extra covariance due to the erasure. The mean-squared error is

$$\mathrm{MSE} = \mathrm{tr}[\widetilde{R}_{ee}] = \mathrm{tr}[R_{ee}] + \mathrm{tr}[\overline{R}_{ee}]. \tag{4.6}$$

Remark 4.2. *The use of the full covariance matrix R_{zz} and the no erasure LMMSE filter K in the case of an erasure was justified in* [7] *as being done to avoid recalculating the LMMSE filter for every erasure and hence reduce computation costs. If one wishes for better accuracy, then, for erasure at location k, one would instead use $\widehat{K} = R_{xz_k}R_{z_kz_k}^{-1}$ where $z_k = (I - E_k)z$ and R_{xz_k}, $R_{z_kz_k}$ are the respective covariance matrices. However, due to the nature of the matrix E_k, $R_{z_kz_k}$ is not invertible. Therefore, to calculate $\widehat{K}$, one can use the pseudo inverse of $R_{z_kz_k}$.*

In the following, tight frames are considered because it is known from Theorem 4.1 that these minimize the no erasure MSE, $\mathrm{tr}[R_{ee}]$, in (4.6). In the same spirit, it is also assumed that the noise has covariance matrix of the form $\sigma_n^2 I$.

Theorem 4.3. *(i) Let x be a d-dimensional random signal with mean zero and covariance R_{xx}, and suppose that the kth coefficient is erased. Consider a noise vector with covariance $\sigma_n^2 I$. The MSE is minimized when the frame is tight and the frame vector corresponding to the location of loss, i.e. f_k, is an eigenvector of R_{xx} corresponding to its minimum eigenvalue.*
(ii) Let the covariance of x be $R_{xx} = \sigma_x^2 I$. Among all frames with a given frame bound A, the average MSE taken over all erasures is minimized for an equal norm tight frame. Further, the redundancy of the frame must be greater than 2, and the common norm of the frame vectors must be less than or equal to $\frac{\sigma_n^2}{\sigma_x^2(\frac{m}{d}-2)}$.

Proof. (i) From Theorem 4.1 it is known that tight frames minimize the no erasure MSE, $\mathrm{tr}[R_{ee}]$, in (4.6). Assume a tight frame with bound equal to A. Since R_{xx} is a Hermitian matrix, there exists a unitary matrix Q, and a diagonal matrix $D = \mathrm{diag}(d_1, \ldots, d_m)$ consisting of the eigenvalues of R_{xx} such that

$$R_{xx} = QDQ^{\mathrm{T}}.$$

Then, in this case,

$$R_{xz} = R_{xx}F^{\mathrm{T}} = QDQ^{\mathrm{T}}F^{\mathrm{T}}, \tag{4.7}$$

$$R_{zz} = FR_{xx}F^{\mathrm{T}} + \sigma_n^2 I = BB^{\mathrm{T}} + \sigma_n^2 I, \quad \text{where } B := FQD^{1/2}, \tag{4.8}$$

$$R_{zz}^{-1} = \frac{1}{\sigma_n^2}I - \frac{1}{\sigma_n^4}B\left(I + \frac{1}{\sigma_n^2}B^{\mathrm{T}}B\right)^{-1}B^{\mathrm{T}} = \frac{1}{\sigma_n^2}I - \frac{1}{\sigma_n^4}B\left(I + \frac{A}{\sigma_n^2}D\right)^{-1}B^{\mathrm{T}} \tag{4.9}$$

where (4.9) is obtained by using the definition of B in (4.8) and the fact that the frame is tight. The additional MSE due to erasure, $\overline{R}_{ee}$, can be calculated from (4.5) by using the matrix inversion formula [3, p. 50].

$$\begin{aligned}\overline{R}_{ee} &= \frac{1}{\sigma_n^4}QDQ^{\mathrm{T}}F^{\mathrm{T}}\left(I - \frac{1}{\sigma_n^2}B\left(I + \frac{A}{\sigma_n^2}D\right)^{-1}B^{\mathrm{T}}\right)E_k\left(BB^{\mathrm{T}} + \sigma_n^2 I\right)E_k\left(I - \frac{1}{\sigma_n^2}B\left(I + \frac{A}{\sigma_n^2}D\right)^{-1}B^{\mathrm{T}}\right)FQDQ^{\mathrm{T}}\\ &= \frac{1}{\sigma_n^4}QD\left(I - \mathrm{diag}\left(\frac{Ad_i}{\sigma_n^2 + Ad_i}\right)\right)Q^{\mathrm{T}}F^{\mathrm{T}}E_k\left(FR_{xx}F^{\mathrm{T}} + \sigma_n^2 I\right)E_kFQ\left(I - \mathrm{diag}\left(\frac{Ad_i}{\sigma_n^2 + Ad_i}\right)\right)DQ^{\mathrm{T}}.\end{aligned}$$

The above can be simplified by noting that

$$I - \mathrm{diag}\left(\frac{Ad_i}{\sigma_n^2 + Ad_i}\right) = \mathrm{diag}\left(\frac{\sigma_n^2}{\sigma_n^2 + Ad_i}\right).$$

Denoting $\Delta := \mathrm{diag}\left(\frac{d_i}{\sigma_n^2 + Ad_i}\right)$, and using the cyclic property of the trace

$$\begin{aligned}\mathrm{tr}[\overline{R}_{ee}] &= \mathrm{tr}[Q^{\mathrm{T}}F^{\mathrm{T}}E_k(FR_{xx}F^{\mathrm{T}} + \sigma_n^2 I)E_kFQ\Delta^2]\\ &= \mathrm{tr}[\Delta Q^{\mathrm{T}}f_kf_k^*QDQ^{\mathrm{T}}f_kf_k^*Q\Delta + \sigma_n^2\Delta Q^{\mathrm{T}}f_kf_k^*Q\Delta].\end{aligned} \tag{4.10}$$

Let $u_1, \ldots, u_m$ be a set of orthonormal eigenvectors of R_{xx} corresponding to the eigenvalues $d_1, \ldots, d_m$, respectively. These eigenvectors form the columns of the matrix Q. For convenience take $A = 1$. Then the second term in (4.10) is

(4.11)

$$\mathrm{tr}[\sigma_n^2\Delta Q^{\mathrm{T}}f_kf_k^*Q\Delta] = \sigma_n^2\sum_{j=1}^{m}\frac{1}{\left(\frac{\sigma_n^2}{d_j} + 1\right)^2}|\langle u_j, f_k\rangle|^2 \geq \frac{\sigma_n^2}{\left(\frac{\sigma_n^2}{d_{\min}} + 1\right)^2}\sum_{j=1}^{m}|\langle u_j, f_k\rangle|^2 = \frac{\sigma_n^2}{\left(\frac{\sigma_n^2}{d_{\min}} + 1\right)^2}\|f_k\|^2.$$

Suppose that f_k is an eigenvector of R_{xx} corresponding to $d_{\min}$. Let $S \subseteq \{1, 2, \ldots, m\}$ be such that if $i \in S$ then u_i is an eigenvector of $d_{\min}$. There exist scalars $\{\alpha_i\}_{i\in S}$ such that

$$f_k = \sum_{i\in S} \alpha_i u_i, \quad \text{or, } \|f_k\|^2 = \sum_{i\in S} |\alpha_i|^2.$$

Then

$$\sigma_n^2 \sum_{j=1}^{m} \frac{1}{\left(\frac{\sigma_n^2}{d_j}+1\right)^2} |\langle u_j, f_k\rangle|^2 = \sigma_n^2 \frac{1}{\left(\frac{\sigma_n^2}{d_{\min}}+1\right)^2} \sum_{i\in S} |\alpha_i|^2 = \sigma_n^2 \frac{1}{\left(\frac{\sigma_n^2}{d_{\min}}+1\right)^2} \|f_k\|^2.$$

This implies that the lower bound in (4.11) is attained or in other words, the second term in (4.10) is minimized, if f_k is an eigenvector of R_{xx} corresponding to its minimum eigenvalue.

The first term in (4.10) is

$$\mathrm{tr}[\Delta Q^{\mathrm{T}} f_k f_k^* Q D Q^{\mathrm{T}} f_k f_k^* Q \Delta] = \sum_{i=1}^{m} d_i |\langle u_i, f_k\rangle|^2 \sum_{j=1}^{m} \frac{d_j^2}{(\sigma_n^2 + d_j)^2} |\langle u_j, f_k\rangle|^2 \geq \frac{d_{\min}^3}{(\sigma_n^2 + d_{\min})^2} \|f_k\|^4.$$

Once again, the lower bound is attained if f_k is an eigenvector of R_{xx} corresponding to its minimum eigenvalue.

(ii) Due to Theorem 4.1, a tight frame minimizes $\mathrm{tr}[R_{ee}]$, and it is assumed that the noise n has covariance $\sigma_n^2 I$. If $R_{xx} = \sigma_x^2 I$, then

$$R_{xz} = \sigma_x^2 F^{\mathrm{T}}, \quad R_{zz} = \sigma_x^2 F F^{\mathrm{T}} + \sigma_n^2 I,$$

and from (4.5)

$$\overline{R}_{ee} = \sigma_x^4 F^{\mathrm{T}} (\sigma_x^2 F F^{\mathrm{T}} + \sigma_n^2 I)^{-1} E (\sigma_x^2 F F^{\mathrm{T}} + \sigma_n^2 I) E^{\mathrm{T}} (\sigma_x^2 F F^{\mathrm{T}} + \sigma_n^2 I)^{-1} F.$$

By using the matrix inversion formula [3, p. 50],

$$\begin{aligned}
\overline{R}_{ee} &= \sigma_x^4 F^{\mathrm{T}} \left(\frac{1}{\sigma_n^2} I - \frac{\sigma_x^2}{\sigma_n^2(\sigma_n^2 + A\sigma_x^2)} F F^{\mathrm{T}}\right) E (\sigma_x^2 F F^{\mathrm{T}} + \sigma_n^2 I) E^{\mathrm{T}} \left(\frac{1}{\sigma_n^2} I - \frac{\sigma_x^2}{\sigma_n^2(\sigma_n^2 + A\sigma_x^2)} F F^{\mathrm{T}}\right) F, \\
\mathrm{tr}[\overline{R}_{ee}] &= \frac{\sigma_x^4}{(\sigma_n^2 + A\sigma_x^2)^2} \mathrm{tr}[F^{\mathrm{T}} E (\sigma_x^2 F F^{\mathrm{T}} + \sigma_n^2 I) E^{\mathrm{T}} F] = \frac{\sigma_x^4}{(\sigma_n^2 + A\sigma_x^2)^2} \mathrm{tr}[\sigma_x^2 (F^{\mathrm{T}} E F)^2 + \sigma_n^2 F^{\mathrm{T}} E F] \\
&= \frac{\sigma_x^4}{(\sigma_n^2 + A\sigma_x^2)^2} \mathrm{tr}[\sigma_x^2 (f_k f_k^*)^2 + \sigma_n^2 f_k f_k^*] = \frac{\sigma_x^4}{(\sigma_n^2 + A\sigma_x^2)^2} (\sigma_x^2 \|f_k\|^4 + \sigma_n^2 \|f_k\|^2).
\end{aligned}$$

Let $\alpha = \frac{\sigma_x^2}{\sigma_n^2 + A\sigma_x^2}$. Averaging over all erasures, and recalling that for a tight frame $\sum_{i=1}^{m} \|f_i\|^2 = Ad$ gives

$$\mathrm{avg}(\mathrm{tr}[\overline{R}_{ee}]) = \frac{1}{m} \alpha^2 \left(\sigma_x^2 \sum_{k=1}^{m} \|f_k\|^4 + \sigma_n^2 A d\right) \geq \alpha^2 \left(\sigma_x^2 \min_{1\leq k\leq m} \|f_k\|^4 + \sigma_n^2 A \frac{d}{m}\right).$$

Equality is attained if every frame vector f_k has the same norm. Thus, putting the conditions for $\mathrm{tr}[R_{ee}]$ and $\mathrm{tr}[\overline{R}_{ee}]$ together, it is implied that the average MSE is minimized for an equal norm tight frame.

Let $\{f_k\}_{k=1}^m$ be an equal norm tight frame such that for all k, $\|f_k\|^2 = L$, where L is some constant. From Theorem 1.1 (b), the frame bound A is

$$A = mL/d.$$

For a fixed L, the minimum value of $\mathrm{tr}[\overline{R}_{ee}]$ is

$$\mu(L) := \mathrm{tr}[\overline{R}_{ee}] = \frac{\sigma_x^4}{(\sigma_n^2 + \frac{mL\sigma_x^2}{d})^2}\left(\sigma_x^2 L^2 + \sigma_n^2 L\right).$$

As a function of L, $\mu(L)$ attains maxima at $L = \frac{\sigma_n^2}{\sigma_x^2(\frac{m}{d}-2)}$. This means that the common norm of the frame vectors must be less than or equal to $\frac{\sigma_n^2}{\sigma_x^2(\frac{m}{d}-2)}$, and that the redundancy $\frac{m}{d}$ must be greater than 2. □

4.2. **A priori distribution of the erasure.** Suppose that the kth frame coefficient is erased according to some known probability distribution $\{p_i\}_{i=1}^m$. That is, if a single erasure occurs at location k then

$$\text{Probability}[k = i] = p_i, \quad i = 1, \ldots, m,$$

and this distribution is independent of the signal x. The LMMSE estimator when the kth coefficient is lost is denoted by $\widehat{x}_k$. When the location of the erasure is unknown, denote the estimator of x by $\widehat{x}$ dropping the subscript k. The mean-squared error with $\widehat{x}$ is

$$\mathrm{MSE} = E[(x-\widehat{x})^{\mathrm{T}}(x-\widehat{x})] = \sum_{k=1}^m E[(x-\widehat{x}_k)^{\mathrm{T}}(x-\widehat{x}_k)]\; p_k = \sum_{k=1}^m \mathrm{MSE}_k\; p_k$$

where MSE_k is the mean-squared error with $\widehat{x}_k$. As in (4.6)

$$\mathrm{MSE}_k = \mathrm{tr}[R_{ee} + (\overline{R}_{ee})_k]$$

where R_{ee} denotes the no erasure covariance, and $(\overline{R}_{ee})_k = R_{xz}R_{zz}^{-1}E_kR_{zz}E_kR_{zz}^{-1}R_{zx}$ denotes the extra covariance due to erasure at location k. Thus

$$\mathrm{MSE} = \sum_{k=1}^m \mathrm{tr}\left[R_{ee} + (\overline{R}_{ee})_k\right] p_k = \mathrm{tr}[R_{ee}] + \sum_{k=1}^m \mathrm{tr}\left[(\overline{R}_{ee})_k\right]\; p_k.$$

Theorem 4.4. *Let x be a zero mean random signal of dimension d. With a frame $\{f_k\}_{k=1}^m$, suppose that the coefficients are erased according to some probability distribution $\{p_k\}_{k=1}^m$.*

(i) Suppose that the covariance of x is $\sigma_x^2 I$. Then:

(a) The MSE with a single erasure is minimized with an equal norm tight frame when the erasure follows a uniform distribution.

(b) For a given probability distribution, considering tight frames, the MSE is minimized when for each $j = 1, \ldots, m$, $\|f_j\|^2 p_j$ is a constant. This implies that a frame vector having a higher probability of getting erased must have smaller norm.

(ii) When the covariance of x is a general matrix R_{xx} then the MSE is minimized when each frame vector is an eigenvector of R_{xx} corresponding to its minimum eigenvalue.

Proof. (i) (a) Recall

$$\text{MSE} = \text{tr}[R_{ee}] + \sum_{k=1}^{m} \text{tr}\left[(\overline{R}_{ee})_k\right]\ p_k.$$

It is known that a tight frame will minimize $\text{tr}[R_{ee}]$. Let the frame bound be $A = 1$. From calculations done in the proof of Theorem 4.3 (ii)

$$\text{tr}[(\overline{R}_{ee})_k] = \frac{\sigma_x^4}{(\sigma_n^2 + \sigma_x^2)^2}\left(\sigma_x^2\|f_k\|^4 + \sigma_n^2\|f_k\|^2\right),$$

(4.12) $$\text{and,}\quad \text{MSE} = \text{tr}[R_{ee}] + \frac{\sigma_x^4}{(\sigma_n^2 + \sigma_x^2)^2}\left[\sum_{k=1}^{m} \sigma_x^2\|f_k\|^4 p_k + \sum_{k=1}^{m} \sigma_n^2\|f_k\|^2\ p_k\right].$$

Since the frame bound is taken to be 1, $\sum_{i=1}^{m} \|f_i\|^2 = d$. To minimize the first term inside the brackets in (4.12) first consider the minimization problem

$$\text{minimize } \mu(\|f_1\|^2, \ldots, \|f_m\|^2, p_1, \ldots, p_m) := \sum_{k=1}^{m} \|f_k\|^4\ p_k$$

subject to the constraints

$$g := \sum_{k=1}^{m} p_k = 1 \quad \textit{and} \quad h := \sum_{k=1}^{m} \|f_k\|^2 = d.$$

Using the method of Lagrange multipliers, it can be shown that μ is minimized with an equal norm tight frame when the erasures follow a uniform distribution, i.e., $p_k = \frac{1}{m}$ for all k. Using Lagrange multipliers to minimize the other term $\sum_{k=1}^{m} \|f_k\|^2\ p_k$ in (4.12) gives the same conditions on the frame and the distribution.

(b) For convenience, consider tight frames with bound $A = 1$. Then $\sum_{i=1}^{m} \|f_i\|^2 = d$. If $\{p_j\}$ is given and fixed, minimization of the first term within brackets in (4.12) can be stated as

$$\text{minimize } \mu(\|f_1\|^2, \ldots, \|f_m\|^2) := \sum_{k=1}^{m} \|f_k\|^4\ p_k$$

subject to

$$h := \sum_{k=1}^{m} \|f_k\|^2 = d.$$

By the method of Lagrange multipliers, it can be shown that the MSE is minimized if $\|f_j\|^2 p_j$ is constant for all $j = 1, \ldots, m$ given by

$$\|f_j\|^2 p_j = \frac{d}{\sum_{k=1}^{m} \frac{1}{p_k}}.$$

This means that in order to minimize the MSE, for a smaller p_j the corresponding frame vector f_j must have a higher norm.

Minimizing the second term inside the brackets in (4.12) results in a uniform distribution and hence gives no additional conditions on the frame.

(ii) Using the same notation as above, and following the calculations done in the proof of Theorem 4.3 (i)

$$\text{MSE} = \text{tr}[R_{ee}] + \sum_{k=1}^{m} \text{tr}[(\overline{R}_{ee})_k]\ p_k \geq \text{tr}[R_{ee}] + \frac{d_{\min}^2}{(\sigma_n^2 + d_{\min})^2} \sum_{k=1}^{m} \left[d_{\min}\|f_k\|^4 + \sigma_n^2 \|f_k\|^2\right]\ p_k$$

The lower bound is attained when each frame vector is an eigenvector of R_{xx} corresponding to its minimum eigenvalue.

□

Remark 4.5. *In Theorem 4.4, setting $R_{xx} = \sigma_x^2 I$ in part (ii) gives the same result as in part (i). Besides, since $\{f_k\}_{k=1}^m$ is a frame for a d-dimensional space, the condition obtained on the frame in part (ii) implies that the dimension of the eigenspace of the minimum eigenvalue of R_{xx} must be at least d.*

5. Stochastic frames and erasures

Representing a signal by stochastic frames may make it harder for a third party to intercept the signal. Therefore, stochastic frames may be an appealing choice in many situations. A way to construct stochastic frames from low autocorrelation stochastic sequences was done in [2]. The sketch of such a construction is as follows. Let $\{Y_{rs}\}_{r,s\in\mathbb{Z}}$ be independent identically distributed (i.i.d.) random variables. One can then define the following two dimensional sequence.[c] For $r, s \in \mathbb{Z}$,

$$X_{rs} = e^{\frac{2\pi}{\epsilon} i Y_{rs}}. \tag{5.1}$$

Consider the mapping $u : \mathbb{Z} \to \mathbb{C}^d$ given by

$$u(\ell) = \frac{1}{\sqrt{d}} \begin{pmatrix} X_{1\ell} \\ X_{2\ell} \\ \vdots \\ X_{d\ell} \end{pmatrix}. \tag{5.2}$$

[c]As a one dimensional sequence, if $\{Y_k\}_{k\in\mathbb{Z}}$ are i.i.d. random variables following a Gaussian distribution with mean zero and variance σ^2, then $X : \mathbb{Z} \to \mathbb{C}$ with $X(k) = e^{\frac{2\pi}{\epsilon} i Y_k}$ is a sequence whose autocorrelation can be made arbitrarily small, depending on ϵ, everywhere except at the origin [2].

Let $u_\ell := u(\ell)$. Consider the set of m unit vectors $V = \{u_1, u_2, \ldots, u_m\}$ in $\mathbb{C}^d$. Let

$$F^* = \frac{1}{\sqrt{d}} \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1m} \\ X_{21} & X_{22} & \cdots & X_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ X_{d1} & X_{d2} & \cdots & X_{dm} \end{bmatrix}$$

so that the frame operator of V is $S = F^*F$. Recall that for a tight frame the eigenvalues of the frame operator are all equal to each other. It has been shown in [2] that, by taking ϵ to be small in (5.1) and a suitable shift, the eigenvalues of S can be made arbitrarily close to each other with a high probability.

Assume that the Ys in (5.1) follow a Gaussian distribution [d] with mean zero and variance σ^2. It can be shown [2] that

$$E[X_{rs}] = E[\overline{X}_{rs}] = e^{-\frac{\sigma^2}{2}\left(\frac{2\pi}{\epsilon}\right)^2}. \tag{5.3}$$

Theorem 5.1. *Let x be a d-dimensional random signal with mean zero and covariance $\sigma_x^2 I$. Assume a noise vector n with zero mean and covariance $\sigma_n^2 I$. Consider estimating x from noisy measurements $z = Fx + n$ using the stochastic frame $\{u_\ell\}_{\ell=1}^m$ constructed in (5.2) and (5.1) for a fixed ϵ. Assume that x, n, and the frame vectors u_ℓs are uncorrelated. Then:*

(1) The MSE is
$$d\sigma_x^2 \left[1 - \sigma_x^2 \frac{m}{d} \frac{a-b}{\Delta} e^{-\sigma^2\left(\frac{2\pi}{\epsilon}\right)^2}\right]$$
where $a = \sigma_x^2 + \sigma_n^2$, $b = \sigma_x^2 e^{-\sigma^2\left(\frac{2\pi}{\epsilon}\right)^2}$, and $\Delta = a^2 + (m-2)ab - (m-1)b^2$.

(2) When there is a single erasure at location k, the extra MSE in (4.6) due to erasure is given by

$$tr[\overline{R}_{ee}] = d\sigma_x^4 t^2 \frac{(a-b)^2 a}{\Delta^2}$$

where a, b, and Δ are the same as given above in (1).

Proof. (1) Let $J_{d\times m}$ be the matrix of size $d \times m$ whose entries are all ones, and let $t = \frac{1}{\sqrt{d}} e^{-\frac{\sigma^2}{2}\left(\frac{2\pi}{\epsilon}\right)^2}$. Then

$$E[F] = tJ_{d\times m},$$

$$R_{xz} = E[xx^*]E[F^*] = t\sigma_x^2 J_{d\times m},$$

$$R_{zz} = E[Fxx^*F^*] + \sigma_n^2 I.$$

To compute R_{zz}, note that Fxx^*F^* is the matrix of inner products

$$[Fxx^*F^*]_{k,\ell} = \langle x, u_k\rangle \langle \overline{x, u_\ell}\rangle.$$

[d] Even though the calculations here are with the Gaussian distribution, some other distribution may be assumed.

Using (5.3) and the i.i.d. property of the Ys

$$E[Fxx^*F^*]_{k,\ell} = E[\langle x, u_k\rangle \overline{\langle x, u_\ell\rangle}] = \begin{cases} \sigma_x^2 & \text{when } k = \ell \\ \sigma_x^2 e^{-\sigma^2(\frac{2\pi}{\epsilon})^2} & \text{when } k \neq \ell. \end{cases}$$

Let $a = \sigma_x^2 + \sigma_n^2$, $b = \sigma_x^2 e^{-\sigma^2(\frac{2\pi}{\epsilon})^2}$, and $\Delta = a^2 + (m-2)ab - (m-1)b^2$. Then

$$R_{zz} = \begin{bmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix},$$

and

$$R_{zz}^{-1} = \begin{bmatrix} \frac{a+(m-2)b}{\Delta} & \frac{-b}{\Delta} & \cdots & \frac{-b}{\Delta} \\ \frac{-b}{\Delta} & \frac{a+(m-2)b}{\Delta} & \cdots & \frac{-b}{\Delta} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{-b}{\Delta} & \frac{-b}{\Delta} & \cdots & \frac{a+(m-2)b}{\Delta} \end{bmatrix}.$$

From (4.1), the covariance of the error is

$$R_{ee} = R_{xx} - R_{xz}R_{zz}^{-1}R_{zx} = \sigma_x^2 I - \sigma_x^4 t^2 J_{d\times m} R_{zz}^{-1} J_{m\times d} = \sigma_x^2 I - \sigma_x^4 t^2 \frac{a-b}{\Delta} m J_{d\times d}.$$

Thus, R_{ee} is of size $d \times d$, and has constant diagonal which gives

$$\text{MSE} = \text{tr}[R_{ee}] = d\sigma_x^2 \left[1 - \sigma_x^2 \frac{m}{d}\frac{a-b}{\Delta} e^{-\sigma^2(\frac{2\pi}{\epsilon})^2}\right].$$

(2) From (4.6), with a single erasure at location k, the extra MSE comes from the trace of $\overline{R}_{ee}$.

$$\begin{aligned} \overline{R}_{ee} &= R_{xz}R_{zz}^{-1}ER_{zz}E^{\mathrm{T}}R_{zz}^{-1}R_{zx} = \sigma_x^4 t^2 J_{d\times m} R_{zz}^{-1} E R_{zz} E R_{zz}^{-1} J_{m\times d} = \sigma_x^4 t^2 \frac{(a-b)^2}{\Delta^2} J_{d\times m} E R_{zz} E J_{m\times d} \\ &= \sigma_x^4 t^2 \frac{a(a-b)^2}{\Delta^2} J_{d\times d}, \\ \text{tr}[\overline{R}_{ee}] &= d\sigma_x^4 t^2 \frac{a(a-b)^2}{\Delta^2}. \end{aligned}$$

□

Remark 5.2. *In Theorem 5.1, part (1), by taking the frame parameter ϵ tending to zero and infinity, respectively, the upper and lower bounds for the MSE can be obtained as*

$$d\sigma_x^2 \left[1 - \frac{1}{d}\frac{m\sigma_x^2}{\sigma_n^2 + m\sigma_x^2}\right] \leq MSE \leq d\sigma_x^2.$$

Acknowledgement

The author would like to thank Doug Cochran for many useful discussions on the topic of this work. The author was supported by the National Science Foundation under Award No. CCF-1422252.

References

[1] O. Christensen. *An Introduction to Frames and Riesz Bases*. Birkhäuser, second edition, 2016.

[2] S. Datta. Frame properties of low autocorrelation random sequences. *Advances in Pure and Applied Mathematics*, 5(4):243 – 253, 2014.

[3] G. H. Golub and C. F. Van Loan. *Matrix Computations (3rd Ed.)*. Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[4] V. K. Goyal, J. Kovacevic, and J. A. Kelner. Quantized frame expansions with erasures. *Applied and Computational Harmonic Analysis*, 10:203 – 233, 2001.

[5] R. B. Holmes and V. I. Paulsen. Optimal frames for erasures. *Linear Algebra Appl.*, 377:31–51, 2004.

[6] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1990.

[7] G. Kutyniok, A. Pezeshki, R. Calderbank, and T. Liu. Robust dimension reduction, fusion frames, and Grassmannian packings. *Applied and Computational Harmonic Analysis*, 26(1):64 – 76, 2009.

[8] L. Scharf. *Statistical Signal Processing*. Addison-Wesley, 1991.

Department of Mathematics, University of Idaho, ID 83844-1103, USA

Email address: `sdatta@uidaho.edu`