FRAMES AS CODES
FOR STRUCTURED ERASURES

A Dissertation

Presented to

the Faculty of the Department of Mathematics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Pankaj K. Singh

December 2012

FRAMES AS CODES FOR STRUCTURED ERASURES

Pankaj K. Singh

APPROVED:

Dr. Bernhard G. Bodmann (Committee Chair)
Department of Mathematics, University of Houston

Dr. Peter G. Casazza
Department of Mathematics, University of Missouri

Dr. Demetrio Labate
Department of Mathematics, University of Houston

Dr. Vern Paulsen
Department of Mathematics, University of Houston

Dean, College of Natural Sciences and Mathematics

Acknowledgements

First and foremost I would like to thank my advisor Dr. Bernhard Bodmann for his constant support, motivation, time, and effort. He guided me with a balanced combination of motivation, encouragement, and firmness, making sure that I was never out of my depth. His enormous experience in mathematics and innovative ideas were crucial to the success of the projects we worked on. Despite his busy schedule he was always ready for discussions and approachable for any problem I had. He gave me the freedom to work independently while at the same time kept track that I didn't lose focus. He was very supportive and was never short of new ideas when I hit a road block in the project.

I would like to thank the committee members Dr. Pete Casazza, Dr. Vern Paulsen, and Dr. Demetrio Labate for serving on my thesis defense committee, and for their insightful suggestions for improving the presentations of this dissertation.

I am indebted to Dr. Dinesh Singh for being a source of inspiration, a mentor, and a teacher. He not only encouraged me to pursue higher studies in mathematics, but also bestowed me with such an opportunity.

I greatly acknowledge the financial, academic, and technical support from the Department of Mathematics at the University of Houston. I would like to extend sincere thanks to Dr. Jeff Morgan for being a constant source of encouragement. I would also like to thank all the faculty and staff of the department.

I owe many thanks to my teachers and mentors including Mr. Brahmdeo N. Singh, Dr. Surya N. Singh, Mr. Mahendra P. Singh, Dr. Ajit I. Singh, Dr. Amber

Habib, Dr. Sanjeev Agrawal, and Dr. Geetha Venkatraman for training me in various foundational areas of mathematics and guiding me through my mathematical journey. I truly appreciate the financial support provided by Mathematical Sciences Foundation(MSF).

My friends have always been there for me whenever I need them. My batchmates- Natasha, Anando, Ankita, Joseph, Aanchal, and Charu, who started this graduate journey with me in India deserve a special mention. Natasha has been one of the motivators and believers in my abilities right from my early days at the DU South Campus. I would like to thank Manisha for her support, encouragement, work behind the scenes, and her belief in me. It is a pleasure to thank all my friends, seniors and colleagues for being there by my side.

This journey would not have been possible without the constant support, unparalleled love and care of my grandparents, my father Ram Sevak Singh, my mother Kamta Devi, my sister Bibha Singh, my brothers Shailendra Singh, Praveen Singh, Rajkapoor Singh, my niece Pallavi, my nephews Ankit, Rahul, Krishna, Kanhaiya, Mayank, Utkarsh, and all the close relatives.

Finally, I would like to thank everyone who has been instrumental in shaping my career and me as an individual. I ask forgiveness from those who have been omitted unintentionally.

To my mother Kamta Devi

FRAMES AS CODES
FOR STRUCTURED ERASURES

An Abstract of a Dissertation

Presented to

the Faculty of the Department of Mathematics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Pankaj K. Singh

December 2012

Abstract

This dissertation studies the role of frames as codes. Frames are families of vectors that give rise to embeddings of Hilbert spaces. These embeddings can be interpreted as codes, because possible linear dependencies among frame vectors can be used to recover missing components in the embedded data, so-called erasures. This dissertation is dedicated to structured erasures. One type of structured erasure occurs when consecutive frame coefficients are lost due to the occurrence of random burst errors. Assuming that the distribution of bursts is invariant under cyclic shifts and that the burst-length statistics are known, we wish to find frames of a given size, which minimize the mean-square reconstruction error for the encoding of vectors in a complex finite-dimensional Hilbert space. We derive statistical error bounds for a given Parseval frame and relate them to its generalized frame potential. In the case of cyclic Parseval frames, we find a family of frames which minimizes the upper bound. Under certain conditions, these minimizers are identical to complex Bose-Chaudhuri-Hocquenghem(BCH) codes discussed in the literature. The accuracy of our upper bounds for the mean-square error is substantiated by complementary lower bounds.

Another part of the dissertation concerns the transmission of digital media, typically following a protocol that splits data into a number of packets having a fixed size. When such packets are sent over a network such as the Internet, there is in principle no guarantee of reliability, that is, the contents of each packet may become corrupted in the course of transmission or entire packets may be lost due to buffer overflows. We assume that during the transmission, only a few of these packets are

corrupted or lost. In this part of the dissertation we adapt ideas by Candes and Tao in order to construct frames as codes for such erasures. The frames are associated with consistency checks for the data that are obtained from random matrices whose entries are independent realizations of a Gaussian random variable. In addition to the random Gaussian matrices, we use random projections to achieve recovery based on a low dimensional check-sum measurement. We use a generalized technique of l_1 minimization to reconstruct the error vector from these measurements.

Contents

List of Figures

CHAPTER 1

Introduction

Redundancy is common and useful in our daily lives. When we leave for work, we double and triple-check that we turned off appliances and lights, that we took our keys, money, etc. The same possibility of a consistency check can be incorporated in signal representations. A signal is understood to be a vector of a Hilbert space and a representation is a map to another Hilbert space. Typically, we would choose a basis and map the vector to its inner products with the basis vectors, the transform coefficients. One of the reasons for choosing a new basis is that often the signal characteristics are more readily apparent in the transform coefficients. Such a representation is non-redundant, and thus corruption or loss of any transform coefficient is irrecoverable. However, instead of a basis, we can choose a linearly dependent family of vectors that introduces redundancy, when the inner products are computed. This way we try to build a safety net into our representation so that we can retain recoverability, stable linear reconstruction, from part of the inner products. Allowing the vectors to be linearly dependent generalizes the concept of a basis to that of a frame.

Frames for a Hilbert space were formally defined by Duffin and Schaeffer [32] in 1952 to study some deep problems in non-harmonic Fourier series. Despite being over half a century old, frames gained popularity only in the last two decades after the landmark paper of Daubechies, Grossmann and Meyer [30] was published in 1986. Traditionally, frames have been used in signal processing, image processing, data compression, and sampling theory.

1.1 Encoding via frames

In the first part of this dissertation, we study frames for redundant encoding of an analog signal and evaluate their performance when the signal is sent through a channel and burst erasure occurs, that is, consecutive frame coefficients are lost. There are two general possibilities, so-called blind reconstruction which replaces the lost coefficients by zeros and active error correction, which aims at perfect reconstruction provided enough coefficients are known. Previous results were concerned with the worst-case scenario and the most generic error models for data loss [38, 37, 55, 5], and many investigated the performance of blind reconstruction [24, 62, 40, 12, 43]. The underlying assumption was that the erasures happen so infrequently that one wants to successively narrow the choice of frames by demanding perfect reconstruction for no loss, then optimal error suppression for one lost coefficient, and at each step select from the remaining frames the ones which are optimal for the next higher number of erasures. Another series of works evaluated the performance of active error correction with low-pass cyclic frames in the presence of quantization errors and burst erasures [56, 57] and with other cyclic frames when more general erasures occur [58]. The performance measure in this case is the mean-square error remaining after the erasure correction is applied. In the absence of erasures, it is known that the error due to quantization of frame coefficients is minimized by equal-norm Parseval frames [38, 37, 57, 58].

Next, we discuss specifics of frames used for encoding signal and resulting error corrction capabilities. For a signal x in a Hilbert space $\mathcal{H}$ of dimension k, we encode this signal using a frame $F = \{f_i : 1 \leq i \leq n\}$ as the frame coefficients $\{\langle x, f_i \rangle : 1 \leq i \leq n\}$. These coefficients are sent through a channel and the receiver does not receive all the coefficients as they are. Some of the coefficients are lost or corrupted. The problem is to linearly reconstruct the vector from these remaining coefficients. We evaluate the performance of frames in the case of burst erasures when we linearly reconstruct the vector. We measure the performance in terms of the mean-square error (MSE), which is the average of the square of the Euclidean norm of the difference between the input and the recovered vector over all unit-norm input vectors. We derive bounds for the MSE and investigate frames which have optimality properties in this setting, with special emphasis on cyclic frames. These frames are a counterpart to cyclic block codes in digital error correction, for example

Bose-Chaudhuri-Hocquenghem (BCH) codes [39, 15] or Reed Solomon codes [60], see also [7, 61]. Burst error correction by Reed-Solomon codes has found a wide range of applications, from deep space transmissions or other wireless communications to storing audio on compact discs or even bar-codes [66]. In addition to their suitability for bursty environments, such codes can correct a maximum number of corrupted or missing symbols in a block due to their maximum distance separability. However, in many cases the transmitted string of symbols is a digitized analog signal, in which case the number of correctable symbols may not be the ultimate measure for performance. For this reason, we consider a redundant linear encoding over the real or complex number field and evaluate the performance by the mean-square error.

Averaging the entries of the Grammian cyclically lowers the mean-square reconstruction error. This is the reason we primarily study cyclic frames, which have Grammians that are invariant under cyclic averaging. For cyclic frames, we estimate the mean-square reconstruction error in the presence of bursty data loss. The mean-square error is defined by averaging over all cyclic bursts with given burst-length probabilities and over the set of all uniformly distributed unit-norm input vectors. The uniform distribution among all unit-norm input vectors is justified by assuming that the data has been compressed before transmission which would result in seemingly independent, identically Gaussian distributed entries for the vectors, and then if the dimension is large enough, a concentration of measure argument would allow replacing them by vectors of a constant norm. The optimal design of frames for this type of performance measure has applications for streaming media and wireless communications [41, 42, 48].

Our first main result is an identity between the mean-square error for Parseval frames resulting from burst erasures and subsequent blind reconstruction, and a weighted squared Hilbert-Schmidt norm of the Grammian. Then we derive upper and lower bounds for the mean-square error. In the case of cyclic frames, we find minimizers for the upper bound. These minimizers can be described by the frequency support of the encoded vectors, or alternatively, by its complement, the syndrome frequencies. If the number of frame vectors n and the dimension of the Hilbert space k are relatively prime, if k is odd, and if the redundancy ratio is sufficiently large, then the optimizers for the upper bound are complex BCH codes that have been introduced in [49] and whose error correcting capabilities have been investigated in previous works [56, 57, 58]. In this case, the syndrome frequencies can be arranged in a uniformly spaced sequence. If the redundancy ratio is sufficiently small, then the optimizers have syndrome frequencies whose complement is uniformly spaced. We also investigate a specific burst-length distribution motivated by a channel model introduced by Gilbert and Elliott [36, 33].

1.2 Packet-based encoding and fusion frame

Signal transmission through a noisy digital channel typically uses the following strategy: a generic signal is decomposed (encoded) into a sequence of coefficients which are then grouped into a number of packets of the same size. When such packets are sent over a network such as the Internet, there is in principle no guarantee of reliability, that is, the contents of each packet may become corrupted in the course of transmission or entire packets may be lost due to buffer overflows[9]. The integrity of the data in each packet is typically protected by some error correction scheme, so for practical purposes one may assume that packets arrive either intact or not at all. The noise of the channel may cause the loss of some packets so that the reconstruction of the signal is done possibly without the whole set of packets. Hence, we search for encoding-decoding schemes that minimize, with respect to some measure, the worst case error between the original signal and the reconstructed signal for a fixed number of packet losses. In practice, it is commonly assumed that losing one packet in the transmission process is rare, and that the occurrence of two lost packets is much less likely. A similar hierarchy of probabilities usually holds for a higher number of lost packets. When the coefficients are sent individually, these types of problems have been considered recently in [24, 40, 14, 62, 12], where they describe the structure of optimal encoding-decoding schemes based on a particular choice to measure the worst case reconstruction error. Casazza and Kovačević [24] find the optimal scheme in case of one erasure using equal-norm tight frames, Holmes and Paulsen [40] in the case of two erasures using equiangular uniform frames, Bodmann and Paulsen [12], and Strohmer and Heath [62] provide conditions for higher number of erasures using graph theoretic apporach, and Bodmann, Paulsen and Kribs [14] generalize the idea in the quantum setting. For related problems in packet encoding, see the work by Bodmann [9], the most recent overview by Bodmann [10], Oswald on stable space splittings in Hilbert spaces [53] which are equivalent to Sun's notion of g-frames [63], and the concept of frames for subspaces introduced by Casazza and Kutyniok [27]. This concept was applied under the name of fusion frames to distributed processing [28]. To study a similar problem, Massey [50] uses so-called uniform (N, m, D)-reconstruction systems in terms of N weighted rank-m operators satisfying certain conditions on the D-dimensional Hilbert space containing the vectors to be transmitted.

The second part of this dissertation deals with recovering the transmitted vector assuming that there are only a few corrupt packets received at the receiver end. Instead of setting the corrupt packets to zero, we try to recover them using a generalization of l_1 minimization techniques by Candes and Tao[22] and Candes [21] in the recovery of sparse vectors. The signal under consideration is a vector v in the D-dimensional Hilbert space $\mathbb{R}^D$. An encoding scheme will map this vector v to N packets of size m, thus to a vector in $\mathbb{R}^{Nm}$. We send this encoded vector through a noisy channel and at the receiving end we receive k corrupted packets, k being small compared to N. We wish to recover these packets in order to get the transmitted vector. We observe that the range of an ecoding map is a D-dimensional suspace of $\mathbb{R}^{Nm}$ and the kernel of its adjoint is also a subspace of $\mathbb{R}^{Nm}$ of dimension $M = Nm - D$. For purposes of designing a recovery algorithm, we use a random matrix $T$of size $M \times Nm$ with entries $t_{i,j}$, which are independent realizations of a Gaussian random variable with mean 0 and variance $\frac{1}{M}$. We denote the range of the map $T^* : \mathbb{R}^M \to \mathbb{R}^{Nm}$ by R and $Ker(T)$ by K, which is an D-dimensional subspace of $\mathbb{R}^{Nm}$ satisfying $K = R^{\perp}$. From an orthonormal basis of K, we obtain a Parseval frame F for $\mathbb{R}^D$ and use its analysis operator Q as the encoding map. Thus, for any $v \in \mathbb{R}^D$, the encoded vector is $Qv \in K \subset \mathbb{R}^{Nm}$. After the transmission, we receive $Qv + x$, where x is the error during the transmission. If we know x, then we will have the encoded vector Qv. As $K = Ker(T)$, applying T to the received vector gives us $T(Qv + x) = 0 + Tx = Tx \in \mathbb{R}^M$. We then apply T^* and a multiple of a random projection P of rank p to get the measurements. From these measurements, we try to recover x and once we have x, we get the transmitted vector as $(Qv+x)-x$ at the receiver end. The reason behind taking random projections P of rank p on $\mathbb{R}^{Nm}$ and random T described above is that the range of the encoding map Q is a D-dimensional random subspace of $\mathbb{R}^{Nm}$ so that the vector $PT^*T(Qv + x) = 0 + PT^*Tx = PT^*Tx$ is the projection of x on a random p dimensional subspace of $\mathbb{R}^{Nm}$. Under certain conditions on k, N, m, M, and p, we show that with overwhelming probability, the error vector x can be recovered with good accuracy even in the presence of noise of small magnitude when we allow M, m, and p to grow to infinity. We use the structure of the erasures to achieve better probability bounds for the probability of success compared to a setting in which unstructured erasure of km coefficients occur.

CHAPTER 2

Burst Erasures and the Mean-Square Error for Cyclic Parseval Frames

This chapter is concerned with the concepts of burst erasure and mean-square reconstruction error for a given frame corresponding to a given burst statistics. We compute the statistical error bounds for a given Parseval frame and focus on cyclic Parseval frames. We also discuss some numerical experiments to validate the theoretical results towards the end of the chapter.

2.1 Preliminaries

First we fix our notations for frames, define concepts related to frames and discuss some properties of frames.

Definition 2.1.1. *A sequence* $\{f_j\}_j$ *of elements of a Hilbert space* $\mathcal{H}$ *is called a frame if there are constants* $A, B > 0$ *such that*

$$A\|x\|^2 \leq \sum_j |\langle x, f_j\rangle|^2 \leq B\|x\|^2 \tag{2.1}$$

for each $x \in \mathcal{H}$*. If* $A = B$*, it is called tight and if* $A = B = 1$*, it is called a Parseval frame.*

In a finite dimensional Hilbert space $\mathcal{H}$ of dimension k, any spanning family of n vectors $F = \{f_j\}_{j=1}^n \subset \mathcal{H}$ is a frame for $\mathcal{H}$. We will work with finite dimensional Hilbert space $\mathcal{H}$ over $\mathbb{R}$ or $\mathbb{C}$.

Definition 2.1.2. *For a frame* $F = \{f_j : 1 \leq j \leq n\}$ *for* $\mathcal{H}$*, we define the* ***analysis oparator*** *for* F *as a linear map* $V : \mathcal{H} \to \mathbb{F}^n$ *via*

$$(Vx)_j = \langle x, f_j\rangle$$

for all $x \in \mathcal{H}$*.*

The map $V^* : \mathbb{F}^n \to \mathcal{H}$ *is called the* ***synthesis operator*** *for* F *and the map* $T = V^*V : \mathcal{H} \to \mathcal{H}$ *is the frame operator.*

Definition 2.1.3. *We call family* F *an* $(\mathbf{n}, \mathbf{k})$***-frame*** *if the Parseval-type equality*

$$\|x\|^2 = \sum_{j=1}^n |\langle x, f_j\rangle|^2$$

holds for every $x \in \mathcal{H}$. If the Grammian G with entries $G_{i,j} = \langle f_j, f_i \rangle$ satisfies $G_{i+l,j+l} = G_{i,j}$ for all $i, j, l \in \{1, 2, \ldots, n\}$, where the summation of indices is taken modulo n, then we call F a ***cyclic*** *(n, k)-frame.*

Given a frame F for $\mathcal{H}$, we say that a frame $\tilde{F}$ for a Hilbert space $\tilde{\mathcal{H}}$ of dimension $(n-k)$ is complementary to F if its Grammian $\tilde{G}$ is related to that of F by $\tilde{G} = I - G$.

Proposition 1. *If V and G are the analysis operator and the Grammian respectively for a frame $F = \{f_j : 1 \leq j \leq n\}$, then we have*

(i) $G = VV^*$.

(ii) *G is an orthogonal projection if and only if the frame is Parseval.*

(iii) *If the frame F is a cyclic (n, k) frame, then G is of rank k and $\|f_j\|^2 = k/n$ for all $j \in \{1, 2, \ldots n\}$. Moreover, $G = q(S)$ with a complex polynomial q which assumes the value one at k of the n-th roots of unity and the value zero at the others, where S is the cyclic shift matrix. Hence, G commutes with S.*

Proof. (i) Let $\{e_i : 1 \leq i \leq n\}$ be the standard orthonormal basis for $\mathbb{F}^n$. For any fixed $1 \leq i \leq n$, we have

$$
\begin{aligned}
VV^* e_i &= \sum_{j=1}^{n} \langle V^* e_i, f_j \rangle e_j = \sum_{j=1}^{n} \langle e_i, V f_j \rangle e_j \\
&= \sum_{j=1}^{n} \langle e_i, \sum_{l=1}^{n} \langle f_j, f_l \rangle e_l \rangle e_j \\
&= \sum_{j=1}^{n} \overline{\langle f_j, f_i \rangle} e_j = \sum_{j=1}^{n} \langle f_i, f_j \rangle e_j \\
&= G e_i.
\end{aligned}
$$

This shows that $G = VV^*$.

(ii) By definition, $G = G^*$.

For a Parseval frame, we have

$$\|Vx\|^2 = \sum_{j=1}^{n} |\langle x, f_j\rangle|^2 = \|x\|^2.$$

Also, we observe that $\|Vx\|^2 = \langle Vx, Vx\rangle = \langle V^*Vx, x\rangle$ and $\|x\|^2 = \langle Ix, x\rangle$. Since both sides coincide as quadratic forms, the operators coincide, that is, $V^*V = I$. Therefore, G is an orthogonal projection as

$$G^2 = VV^*VV^* = V(V^*V)V^* = VIV^* = VV^* = G\,.$$

Conversely, let us assmume that G is an orthogonal projection. As V is one to one and V^* is onto, we have

$$\begin{aligned} &G^2 = G \\ \Rightarrow &VV^*VV^* = VV^* \\ \Rightarrow &V(V^*V - I)V^* = 0 \\ \Rightarrow &(V^*V - I) = 0\,. \end{aligned}$$

Therefore, for any $x \in \mathcal{H}$, we obtain

$$\|x\|^2 = \langle Ix, x\rangle = \langle V^*Vx, x\rangle = \langle Vx, Vx\rangle = \|Vx\|^2 = \sum_{j=1}^{n} |\langle x, f_j\rangle|^2\,.$$

Therefore, F is a Parseval frame.

(iii) As $F = \{f_j : 1 \leq j \leq n\}$ is aS cyclic (n, k)-frame, $\langle f_1, f_1\rangle = \langle f_{1+l}, f_{1+l}\rangle$ for $1 \leq l \leq (n-1)$. Therefore, all the diagonal entries of G are the same. Now, we have

$$\operatorname{tr}(G) = \operatorname{tr}(VV^*) = \operatorname{tr}(V^*V) = k.$$

Also, we have $\operatorname{tr}(G) = \sum_{j=1}^{n} \|f_j\|^2$. Therefore, we conclude that

$$\|f_j\|^2 = \frac{k}{n} \qquad \text{for all } 1 \le j \le n.$$

Using the cyclic property $\langle f_i, f_j \rangle = \langle f_{i+l}, f_{j+l} \rangle$ for all $1 \le i, j, l \le n$ and identifying f_j and f_{j+n}, we obtain

$$G = \begin{pmatrix} \langle f_1, f_1 \rangle & \langle f_n, f_1 \rangle & \cdots & \langle f_3, f_1 \rangle & \langle f_2, f_1 \rangle \\ \langle f_2, f_1 \rangle & \langle f_1, f_1 \rangle & \cdots & \langle f_4, f_1 \rangle & \langle f_3, f_1 \rangle \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \langle f_{(n-1)}, f_1 \rangle & \langle f_{(n-2)}, f_1 \rangle & \cdots & \langle f_1, f_1 \rangle & \langle f_n, f_1 \rangle \\ \langle f_n, f_1 \rangle & \langle f_{(n-1)}, f_1 \rangle & \cdots & \langle f_2, f_1 \rangle & \langle f_1, f_1 \rangle \end{pmatrix}$$

$$= \langle f_1, f_1 \rangle I + \langle f_2, f_1 \rangle S + \langle f_3, f_1 \rangle S^2 + \cdots + \langle f_n, f_1 \rangle S^{(n-1)},$$

$$\text{where } S = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Thus, we conclude that $G = q(S) = \sum_{j=0}^{(n-1)} q_j S^j$ with $q_j = \langle f_{j+1}, f_1 \rangle$. Hence, G commutes with S. As S is the companion matrix for the polynomial $x^n - 1$, the characteristic polynomial for S is $x^n - 1$. Therefore, the eignevalues of S are the nth roots of unity $e^{\frac{2\pi il}{n}}$, $0 \leq l \leq (n-1)$. For $\lambda \in \{e^{\frac{2\pi il}{n}} : 0 \leq l \leq (n-1)\}$, if f_λ is an eigenvector of S, then we have

$$Gf_\lambda = q(S)f_\lambda = q(\lambda)f_\lambda$$

Thus, $q(\lambda)$ is an eigenvalue of G with the same eigenvector f_λ. Also, G is an orthogonal projection onto a k-dimensional space, eigenvalues of G belong to $\{0, 1\}$. Therefore, $q(\lambda) \in \{0, 1\}$ and $q(\lambda) = 1$ for k of the nth roots of unity and zero for others.

□

For a fixed input vector $x \in \mathcal{H}$, we send the encoded coefficients through a noisy channel and loose some of the coefficients. In the case of blind reconstruction, we need to set these coefficients to zero and recover the vector from a subset of coefficients. The norm of the difference of the vector and recovered vector is the reconstruction error. Formally, we define the reconstruction error for one input vector as follows:

Definition 2.1.4. *Let F be an (n, k)-frame for a real or complex Hilbert space $\mathcal{H}$. The* ***reconstruction error*** *for an input vector $x \in \mathcal{H}$ and an erasure of frame coefficients with indices $\mathbb{J} = \{j_1, j_2, \ldots j_m\}$, $m \leq n$, is given by*

$$\|V^*EVx - x\| = \|(V^*EV - I)x\| = \|V^*DVx\|$$

where E is the diagonal $n \times n$ matrix with $E_{j,j} = 0$ if $j \in \mathbb{J}$ and $E_{j,j} = 1$ otherwise, and $D = I - E$.

Since the error is proportional to the norm of the input vector, we could use the operator norm $\|V^*DV\|$ as a measure for the worst case error among all inputs. This has been investigated elsewhere [40, 12, 43]. In the following, we will define a measure for average performance when the probabilities for erasures are known. We average the square of the reconstruction error with the distribution of erasures and input vectors. As a simple model, we have chosen a uniform measure on the unit sphere in $\mathcal{H}$ for the input. We focus on a particular type of errors, the burst erasures.

We use the following notations, in order to define the mean-square reconstruction error by taking the average of the difference of norms of each input vector and recovered vector, where the average is taken over all the unit vectors:

Definition 2.1.5. *Let $\mathcal{B}_m$ denote the set of diagonal $n \times n$ matrices with a block (indices modulo n) of exactly $m \leq n$ diagonal entries equal to 0 and a block of $n-m$ diagonal entries equal to 1.*

We also denote $\mathcal{D}_m = \{D : D = I - E, E \in \mathcal{B}_m\}$. To simplify expressions, we write $D^{(m)}$ for the $n \times n$ matrix with $D^{(m)}_{j,j} = 1$ for $1 \leq j \leq m$ and zero elsewhere. Then $\mathcal{D}_m = \{(S^*)^j D^{(m)} S^j\}_{j=1}^n$.

Definition 2.1.6. *Given an (n,k)-frame F for a Hilbert space $\mathcal{H}$ with analysis operator V and a probability vector $p = (p_0, p_1, \ldots, p_n)$, we define the* **mean-square reconstruction error** *for burst erasures with length statistics p as*

$$\mathcal{E}(V,p) = p_n + \sum_{m=1}^{n-1} \frac{p_m}{n} \sum_{D \in \mathcal{D}_m} \int_{\|\Omega\|=1} \|V^*DV\Omega\|^2 d\mu(\Omega)$$

whereby μ denotes the uniform probability measure on the unit sphere in $\mathcal{H}$. We will later use the same definition and notation for the mean-square reconstruction error *corresponding to some vector* $p = (p_0, p_1, \dots, p_n)$ *even if it is not a probability vector.*

2.2 Computing the mean-square reconstruction error

First, we show that the mean-square error is the square of a weighted Frobenius norm of the Grammian. We also estimate the mean-square reconstruction error.

As preparation, we review a fact from representation theory.

Lemma 2.2.1. *For any self-adjoint operator* A *on a real or complex Hilbert space* $\mathcal{H}$ *of dimension* k*,*

$$\int_{\|\Omega\|=1} \langle A\Omega, \Omega\rangle d\mu(\Omega) = \frac{1}{k} \operatorname{tr} A \,.$$

Proof. It is enough to show the claimed identity if A is a self-adjoint rank-one projection, because both sides of the equation are linear in A and any self-adjoint A can be written as a sum of real multiples of mutually orthogonal rank-one projections. To verify the claim for the rank one projection $P_{\Omega'}$ onto the space generated by a fixed vector $\Omega' \in \mathcal{H}$, we compute

$$\int_{\|\Omega\|=1} \|P_{\Omega'}\Omega\|^2 d\mu(\Omega) = \int_{\|\Omega\|=1} |\langle \Omega', \Omega\rangle|^2 d\mu(\Omega) \tag{2.2}$$

in cylindrical coordinates. We first consider the case of a real Hilbert space. when Ω is projected onto $x = \langle \Omega', \Omega\rangle$, the measure μ induces a measure $d\mu_\lambda(x) = A_\lambda(1-x^2)^{\lambda-\frac{1}{2}}dx$

with $\lambda = \frac{k-2}{2}$ on the interval $[-1, 1]$ [52]. Thus, we have

$$\begin{aligned} 1 = \int d\mu_\lambda(x) &= \int_{-1}^{1} A_\lambda (1-x^2)^{\lambda - \frac{1}{2}} dx \\ &= 2A_\lambda \int_0^1 (1-x^2)^{\lambda-\frac{1}{2}} dx \\ &= A_\lambda \int_0^1 (1-u)^{\lambda - \frac{1}{2}} u^{\frac{-1}{2}} du \\ &= A_\lambda \int_0^1 u^{\frac{1}{2}-1} (1-u)^{(\lambda+\frac{1}{2})-1} du \\ &= A_\lambda B\left(\frac{1}{2}, \lambda + \frac{1}{2}\right) \end{aligned}$$

By definition of beta function and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, we obtain

$$\begin{aligned} A_\lambda \frac{\Gamma(\frac{1}{2})\Gamma(\lambda+\frac{1}{2})}{\Gamma(\lambda + \frac{1}{2} + \frac{1}{2})} &= 1 \\ \Rightarrow A_\lambda = \frac{\Gamma(\lambda+1)}{\sqrt{\pi}\Gamma(\lambda+\frac{1}{2})}. \end{aligned}$$

Now, the integral in 2.2 becomes

$$
\begin{aligned}
&\int_{\|\Omega\|=1} |\langle \Omega', \Omega \rangle|^2 d\mu(\Omega) \\
&= \int_{-1}^{1} x^2 A_\lambda (1-x^2)^{\lambda - \frac{1}{2}} dx \\
&= \frac{\Gamma(\frac{k}{2})}{\sqrt{\pi}\Gamma(\frac{k-1}{2})} \int_{-1}^{1} x^2 (1-x^2)^{\frac{k-3}{2}} dx \\
&= \frac{\Gamma(\frac{k}{2})}{\sqrt{\pi}\Gamma(\frac{k-1}{2})} 2 \int_{0}^{1} x^2 (1-x^2)^{\frac{k-3}{2}} dx \\
&= \frac{\Gamma(\frac{k}{2})}{\sqrt{\pi}\Gamma(\frac{k-1}{2})} \int_{0}^{1} u^{\frac{1}{2}} (1-u)^{\frac{k-3}{2}} du \\
&= \frac{\Gamma(\frac{k}{2})}{\sqrt{\pi}\Gamma(\frac{k-1}{2})} \int_{0}^{1} u^{\frac{3}{2}-1} (1-u)^{\frac{k-1}{2}-1} du \\
&= \frac{\Gamma(\frac{k}{2})}{\sqrt{\pi}\Gamma(\frac{k-1}{2})} B\left(\frac{3}{2}, \frac{k-1}{2}\right) \\
&= \frac{\Gamma(\frac{k}{2})}{\sqrt{\pi}\Gamma(\frac{k-1}{2})} \frac{\Gamma(\frac{3}{2})\Gamma(\frac{k-1}{2})}{\Gamma(\frac{k-1}{2}+\frac{3}{2})} \\
&= \frac{\Gamma(\frac{k}{2})}{\sqrt{\pi}} \frac{\Gamma(\frac{1}{2}+1)}{\Gamma(\frac{k}{2}+1)} \\
&= \frac{\Gamma(\frac{k}{2})}{\sqrt{\pi}} \frac{\frac{1}{2}\Gamma(\frac{1}{2})}{\frac{k}{2}\Gamma(\frac{k}{2})} = \frac{\sqrt{\pi}}{\sqrt{\pi}k} = \frac{1}{k} \\
&= \frac{\operatorname{tr}(P_{\Omega'})}{k}.
\end{aligned}
$$

The integral for the case of a complex Hilbert space gives the same result because $\mathcal{H}$ can be viewed as a real Hilbert space of twice the dimension and $P_{\Omega'}$ as the sum of the two projections onto the real subspaces spanned by Ω' and $i\Omega'$. □

In the following result, we express $\mathcal{E}(V, p)$ in terms of Frobenius norm of compression of the Grammian using the above lemma and its definition.

Theorem 2.2.2. *Let $\mathcal{H}$ be a real or complex Hilbert space of dimension k and F an (n,k)-frame for $\mathcal{H}$ with analysis operator V. Given a vector $p = (p_0, p_1, \ldots, p_n)$ with $p_m \geq 0$ for $m \geq 1$, then*

$$\mathcal{E}(V,p) = p_n + \frac{1}{k}\sum_{m=1}^{n-1}\frac{p_m}{n}\sum_{D\in\mathcal{D}_m}\mathrm{tr}[(DGD)^2]\,.$$

Proof. We observe that

$$\|V^*DV\Omega\|^2 = \langle (V^*DV)\Omega, (V^*DV)\Omega\rangle = \langle (V^*DV)^2\Omega, \Omega\rangle$$

and $(V^*DV)^2$ is self adjoint. From definition of mean-square reconstruction error 2.1.6 and applying the preceding lemma 2.2.1, we get

$$\mathcal{E}(V,p) = p_n + \sum_{m=1}^{n-1}\frac{p_m}{n}\sum_{D\in\mathcal{D}_m}\frac{1}{k}\,\mathrm{tr}\left[(V^*DV)^2\right]\,.$$

Since $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, $G = VV^*$ and $D^2 = D$, we have

$$\begin{aligned}
\mathrm{tr}\left[(V^*DV)^2\right] &= \mathrm{tr}\left[V^*DVV^*DV\right] \\
&= \mathrm{tr}\left[DVV^*DVV^*\right] \\
&= \mathrm{tr}\left[DD(VV^*)DD(VV^*)\right] \\
&= \mathrm{tr}\left[(DGD)(DGD)\right] \\
&= \mathrm{tr}\left[(DGD)^2\right]
\end{aligned}$$

Therefore, we have

$$\mathcal{E}(V,p) = p_n + \frac{1}{k}\sum_{m=1}^{n-1}\frac{p_m}{n}\sum_{D\in\mathcal{D}_m}\mathrm{tr}\left[(DGD)^2\right]\,.$$

□

Since $(\mathrm{tr}[(DGD)^2])^{1/2}$ is the Frobenius norm of the compression DGD of the Grammian G, $\mathcal{E}(V,p)$ can be viewed as a generalized frame potential [23]. In the following, we define the weights and express $\mathcal{E}(V,p)$ in terms of the weighted Frobenius norm of the Grammian.

Definition 2.2.3. *Let $p=(p_0,p_1,\ldots,p_n)$ be a probability vector. The sequence of weights associated to p is defined to be the sequence of an even, n-periodic sequence $\{w_j\}_{j\in\mathbb{Z}}$ specified by its values for $0\le j\le n/2$,*

$$w_j=\sum_{l=j+1}^{n-j}(l-j)p_l+\sum_{l=n-j+1}^{n}(2l-n)p_l\,.$$

We will use the same definition and notation for the sequence of weights associated to a vector $p\in\mathbb{R}^n$ even if it is not a probability vector.

Theorem 2.2.4. *Given an (n,k)-frame F for a real or complex Hilbert space $\mathcal{H}$ and a vector $p=(p_0,p_1,\ldots,p_n)$, the resulting mean-square reconstruction error is given by the square of the weighted Frobenius norm*

$$\mathcal{E}(V,p)=\frac{1}{kn}\sum_{j,l=1}^{n}w_{j-l}|G_{j,l}|^2$$

of the Grammian G, where the sequence of weights $\{w_i\}$ are as in definition 2.2.3.

Proof. We recall that $\mathcal{D}_m=\{(S^*)^jD^{(m)}S^j\}_{j=1}^n$ and the cyclic shift matrix S satisfies

$SS^* = S^*S = I$, where S is given by

$$S = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Thus, using $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, we have

$$\begin{aligned} \frac{1}{n}\sum_{D\in\mathcal{D}_m} \mathrm{tr}[(DGD)^2] &= \frac{1}{n}\sum_{j=1}^{n} \mathrm{tr}[((S^*)^j D^{(m)} S^j G (S^*)^j D^{(m)} S^j)^2] \\ &= \frac{1}{n}\sum_{j=1}^{n} \mathrm{tr}[(S^*)^j D^{(m)} S^j G (S^*)^j D^{(m)} S^j (S^*)^j D^{(m)} S^j G (S^*)^j D^{(m)} S^j] \\ &= \frac{1}{n}\sum_{j=1}^{n} \mathrm{tr}[S^j (S^*)^j D^{(m)} S^j G (S^*)^j D^{(m)} S^j (S^*)^j D^{(m)} S^j G (S^*)^j D^{(m)}] \\ &= \frac{1}{n}\sum_{j=1}^{n} \mathrm{tr}[D^{(m)} S^j G (S^*)^j D^{(m)} D^{(m)} S^j G (S^*)^j D^{(m)}] \\ &= \frac{1}{n}\sum_{j=1}^{n} \mathrm{tr}[(D^{(m)} S^j G (S^*)^j D^{(m)})^2]\,. \end{aligned} \tag{2.3}$$

We observe that if the vector p has only one non-zero entry p_m with $1 \leq m < n$, then the sum in the resulting formula for $\mathcal{E}(V,p)$ is of the form

$$\frac{1}{kn}\sum_{j=1}^{n} p_m \,\mathrm{tr}[(D^{(m)} S^j G (S^*)^j D^{(m)})^2] = \frac{1}{kn}\sum_{l,j=1}^{n} q_{j,l} |G_{j,l}|^2$$

with certain weights $\{q_{j,l}\}_{j,l=1}^{n}$. This form is preserved under linear combinations of p_m's. So, we need to find the weights applying to each $|G_{j,l}|^2$. To compute them, we evaluate the square of the Frobenius norm or the Hilbert Schmidt norm for a matrix unit $E_{l,l+d}$ instead of G, for fixed $l, d \in \{1, 2, \dots, n\}$.

$$\begin{aligned} w_d &= \sum_m p_m \sum_{j=1}^{n} \|(D^{(m)} S^j E_{l,l+d} (S^*)^j D^{(m)})\|_{HS}^2 \\ &= \sum_m p_m \sum_{j=1}^{n} \sum_{p,q=1}^{m} (S^j E_{l,l+d} (S^*)^j)_{p,q}^2 \\ &= \sum_m p_m \sum_{j=1}^{n} \sum_{p,q=1}^{m} (E_{l+j,l+d+j})_{p,q}^2 \, . \end{aligned}$$

Now using that the entries of the matrix unit are zero or one, $(E_{j,l})_{p,q}^2 = (E_{j,l})_{p,q}$ for any $j, l \in \{1, 2, \dots, n\}$, and that $\sum_{j=1}^{n} S^j E_{l,l+d} (S^*)^j = S^{-d}$, produces

$$\begin{aligned} w_d &= \sum_m p_m \sum_{j=1}^{n} \sum_{p,q=1}^{m} (E_{l+j,l+d+j})_{p,q} \\ &= \sum_m p_m \sum_{p,q=1}^{m} \sum_{j=1}^{n} (S^j E_{l,l+d} (S^*)^j)_{p,q} \\ &= \sum_m p_m \sum_{p,q=1}^{m} (S^{-d})_{p,q} \\ &= \sum_m p_m \sum_{p,q=1}^{m} (S^{-d})_{p,q}^2 \\ &= \sum_m p_m \|(D^{(m)} S^{-d} D^{(m)})\|_{HS}^2 \, . \end{aligned}$$

For $1 \le m \le n$, we thus have

$$\begin{aligned} \|(D^{(m)} S^{-d} D^{(m)})\|_{HS}^2 &= \sum_{p,q=1}^{m} (S^{-d})_{p,q} \\ &= \begin{cases} 0 & \text{if } m \le d \\ (m-d) & \text{if } d+1 \le m \le n-d \\ (2m-n) & \text{if } n-d+1 \le m \le n \end{cases} . \end{aligned}$$

Therefore, we get

$$w_d = \sum_{m=d+1}^{n-d} (m-d)p_m + \sum_{m=n-d+1}^{n} (2m-n)p_m\,.$$

This concludes the proof. □

Theorem 2.2.5. *Let F and $\hat{F}$ be (n,k) and $(n,n-k)$-frames for $\mathcal{H}$ and $\hat{\mathcal{H}}$ respectively and $\hat{F}$ be complementary to F. If $p=(p_0,p_1,\cdots,p_n)$ is a given vector, then the resulting mean-square reconstruction error for $\hat{F}$ is given by*

$$\mathcal{E}(\hat{V},p) = \frac{n-2k}{kn} w_0 + \mathcal{E}(V,p), \tag{2.4}$$

where $\mathcal{E}(V,p)$ is as in Theorem 2.2.4.

Proof. Using Theorem 2.2.4, for $(n,n-k)$-frame $\hat{F}$ the mean-square reconstruction error is given by

$$\mathcal{E}(\hat{V},p) = \frac{1}{kn}\sum_{j,l=1}^{n} w_{j-l}|\hat{G}_{j,l}|^2\,,$$

where $\hat{G}$ is the Grammian for $\hat{F}$.

As $\hat{G}=I-G$, we have $|\hat{G}_{j,l}|=|G_{j,l}|$ for all $j\neq l$ and $\hat{G}_{j,j}=1-G_{j,j}=1-\dfrac{k}{n}$. Thus, $\mathcal{E}(\hat{V},p)$ is equal to

$$\begin{aligned}
&\frac{1}{kn}\sum_{j=1}^{n} w_0\left(1-\frac{k}{n}\right)^2 + \frac{1}{kn}\sum_{j\neq l} w_{j-l}|G_{j,l}|^2 \\
&= \frac{1}{kn}\sum_{j=1}^{n} w_0\left(1-\frac{2k}{n}+\frac{k^2}{n^2}\right) + \frac{1}{kn}\sum_{j\neq l} w_{j-l}|G_{j,l}|^2 \\
&= \frac{1}{kn}\sum_{j=1}^{n} w_0\left(1-\frac{2k}{n}\right) + \frac{1}{kn}\sum_{j,l=1}^{n} w_{j-l}|G_{j,l}|^2\,.
\end{aligned}$$

From the definition of $\mathcal{E}(V,p)$, the above expression reduces to

$$\mathcal{E}(\hat{V},p) = \frac{n-2k}{kn} w_0 + \mathcal{E}(V,p)\,.$$

□

The Frobenius norm of DGD is also the Euclidean norm of the vector formed by its eigenvalues. Therefore, it is equivalent to the operator norm and we could estimate the worst-case reconstruction error when the frame coefficients are subjected to some erasure in terms of the Frobenius norm. Furthermore, we could then derive an estimate for the average of all worst-case reconstruction errors for given erasure statistics. We focus exclusively on the mean-square reconstruction error.

2.3 Estimating the mean-square reconstruction error

Now, we try to establish a lower bound and an upper bound for the mean-square reconstruction error for an (n,k)-frame F. Using relation (2.4), we will then automatically get these bounds for an $(n,n-k)$-frame $\hat{F}$, which is complementary to F. We first estimate them in general and then focus on particular cases of interest.

2.3.1 A lower bound for the mean-square reconstruction error

In order to get a lower bound of $\mathcal{E}(V,p)$, we cyclically average the entries of the Grammian. We show that by doing so, we only lower the mean-square reconstruction error.

Proposition 2.3.1. *For an (n,k)-frame with an analysis operator V and a probability vector $p=(p_0,p_1,\ldots,p_n)$, the mean-square reconstruction error has the lower bound*

$$\mathcal{E}(V,p) \geq p_n + \frac{1}{k}\sum_{m=1}^{n-1} p_m \mathrm{tr}[(D^{(m)}\widetilde{G}D^{(m)})^2] \ ,$$

where the diagonal matrix $D^{(m)}$ has entries $D^{(m)}_{j,j}=1$ for $1\leq j\leq m$ and zero elsewhere, and $\widetilde{G}=\frac{1}{n}\sum_{j=1}^{n} S^jG(S^*)^j$. Assuming all $p_m>0$ for $1\leq m\leq n-1$, then this inequality is saturated if and only if $|\langle f_i,f_j\rangle| = |\langle f_{i+l},f_{j+l}\rangle|$ for all $i,j,l\in\mathbb{Z}$ (modulo n).

Proof. The inequality follows from an application of Jensen's inequality to the convex function $g: A\mapsto \mathrm{tr}[A^2]$ defined on the real vector space of all self-adjoint operators $\{A\in B(\mathcal{H}) : A=A^*\}$ (See [54]) and theorem 2.2.2. Using the relation 2.3 and averaging gives

$$\begin{aligned}\frac{1}{n}\sum_{D\in\mathcal{D}_m}\mathrm{tr}[(DGD)^2] &= \frac{1}{n}\sum_{j=1}^{n}\mathrm{tr}[(D^{(m)}S^jG(S^*)^jD^{(m)})^2] \\ &\geq \mathrm{tr}\left[\Big(\frac{1}{n}\sum_{j=1}^{n} D^{(m)}S^jG(S^*)^jD^{(m)}\Big)^2\right] \\ &= \mathrm{tr}[(D^{(m)}\widetilde{G}D^{(m)})^2]\,.\end{aligned}$$

To characterize cases of equality, we observe that Jensen's inequality is strict unless $\mathrm{tr}[(DGD)^2]$ depends only on $m = \mathrm{tr}\, D$, not the particular choice of $D \in \mathcal{D}_m$.

We proceed by induction. For $m = 1$, the trace of all DGD is identical if and only if there is c_1 such that all $\|f_j\| = c_1$. Now suppose we have shown this for all entries up to the $(m-1)$-th super and subdiagonal. Comparing the Frobenius norms, we get

$$\begin{aligned}&\mathrm{tr}[(D^{(m)}GD^{(m)})^2 + (D^{(m-2)}SGS^*D^{(m-2)})^2]\\&= 2|\langle f_m, f_1\rangle|^2 + \mathrm{tr}[(D^{(m-1)}GD^{(m-1)})^2 + (D^{(m-1)}SGS^*D^{(m-1)})^2]\end{aligned}$$

Repeating this for the cyclically shifted G and using the assumption shows that there is a constant c_m such that all $|\langle f_j, f_{j+m-1}\rangle| = c_m$. □

2.3.2 A lower bound for cyclic frames

Although cyclic frames saturate this lower bound for the mean-square reconstruction error, it is unclear whether there is always a cyclic frame among the minimizers. Nevertheless, further below we focus on designing the best cyclic frames for given erasure statistics. Even in this special case, there is no simple, explicit way known to us to characterize the best frame. As a first step, we investigate how well a cyclic frame could possibly perform. We prepare by examining convexity properties of the weights.

Lemma 2.3.2. *Let $(p_0, p_1, \ldots, p_n)$ be a probability vector, and let the weights $\{w_j\}_{j\in\mathbb{Z}}$ be as in the definition 2.2.3. Then the restriction of the sequence $\{w_j\}$ to indices*

$0 \le j \le n/2$ is convex and decreasing, that is, $w_j \le \frac{1}{2}(w_{j-1} + w_{j+1})$ and $w_j \le w_{j-1}$ for each $1 \le j \le \frac{n}{2} - 1$.

Proof. Let $p_l = 1$ for a fixed $l \in \{1, 2, \dots n\}$. If $l \le n/2$, we have

$$w_j = \begin{cases} l - j, & j \le l - 1 \\ 0, & \text{else} \end{cases}$$

and if $l \le n/2$, we have

$$w_j = \begin{cases} l - j, & j \le n - l \\ 2l - n, & \text{else} \end{cases} .$$

In both cases, w_j is convex and decreasing in j. Therefore, any convex combination of such sequences, obtained by allowing more than one index l for which $p_l \neq 0$, is convex and decreasing. □

Proposition 2. *Let n be even, p be a probability vector and $\{w_j\}$ be the weights associated to p. If $\mathcal{E}(V_{min}, p)$ is the minimum of $\mathcal{E}(V, p)$, where the minimum is taken over all cyclic (n, k)-frames, then*

$$\mathcal{E}(V_{min}, p) \ge \frac{1}{n} \left[\frac{k}{n} w_0 + \left(1 - \frac{k}{n}\right) w_{n/2} \right] .$$

Proof. First we note that for a cyclic (n, k)- frame F with Grammian G

$$k = \sum_{l,j=1}^{n} |G_{j,l}|^2 . \tag{2.5}$$

As $G_{j,j} = k/n$ for each $1 \le j \le n$, we have

$$\sum_{\substack{l,j=1 \\ l \neq j}}^{n} |G_{j,l}|^2 = k - n\frac{k^2}{n^2} = \frac{k}{n}(n - k) . \tag{2.6}$$

We employ a variational argument. To this end, we define the set

$$\mathcal{G} = \{G : G = G^*G, \operatorname{tr} G = k, GS = SG\},$$

containing the Grammians of all cyclic (n,k)-frames, and the larger set of $n \times n$ matrices

$$\mathcal{G}_1 = \{G = (G_{l,j}) : G_{j,j} = k/n, \sum_{l,j=1}^{n} |G_{j,l}|^2 = k \text{ and } GS = SG\}.$$

Define $f : \mathcal{G}_1 \to \mathbb{R}$ via

$$f(G) = \frac{1}{kn} \sum_{j,l=1}^{n} w_{j-l} |G_{j,l}|^2 .$$

Using the fact that $\mathcal{G} \subset \mathcal{G}_1$ and from definitions of f and $\mathcal{E}(V_{min}, p)$, we conclude that

$$\mathcal{E}(V_{min}, p) = \min_{G \in \mathcal{G}} f(G) \geq \min_{G \in \mathcal{G}_1} f(G) .$$

First we observe that relations (2.5) and (2.6) hold true for all $G \in \mathcal{G}_1$ as well. As $\{w_j\}$ is convex and decreasing for $0 \leq j \leq n/2$, the minimum value of $f(G)$ is attained at $\overline{G} \in \mathcal{G}_1$ satisfying $\overline{G}_{(n/2+j)(mod\ n),j} = \alpha$ for some fixed α and $1 \leq j \leq n$, and all other off-diagonal entries zero. Using the above relations (2.5) and (2.6) for this $\overline{G}$, we get

$$k = n\frac{k^2}{n^2} + n|\alpha|^2 .$$

This gives us

$$n|\alpha|^2 = \frac{k}{n}(n-k)$$

and the value of $f(\overline{G})$ is given by

$$\begin{aligned} f(\overline{G}) &= \frac{1}{kn}\left[n\frac{k^2}{n^2} w_0 + n|\alpha|^2 w_{n/2}\right] \\ &= \frac{1}{n}\left[\frac{k}{n} w_0 + \left(1 - \frac{k}{n}\right) w_{n/2}\right] . \end{aligned}$$

Hence, we obtain

$$\mathcal{E}(V_{min}, p) \geq \frac{1}{n}\left[\frac{k}{n}w_0 + \left(1 - \frac{k}{n}\right) w_{n/2}\right] .$$

□

2.3.3 An upper bound for the mean-square reconstruction error

A different convexity argument gives an upper bound for $\mathcal{E}(V, p)$. The idea is to replace the probability vector for the burst lengths by another vector $\tilde{p}$ such that the weights can only increase. In the special case of a transmission that erases no coefficients, approximately half, or all of the coefficients, this estimate is saturated under certain conditions.

Theorem 2.3.3. *Let F be an (n, k)-frame with analysis operator V on a real or complex Hilbert space. Let $(p_0, p_1, \ldots, p_n)$ be a probability vector, and let the weights $\{w_j\}_{j\in\mathbb{Z}}$ be defined as in Definition 2.2.3. If n is even, then there exists a probability vector $\tilde{p}$ with $\tilde{p}_j = 0$ for all $j \notin \{0, n/2, n\}$ such that $\tilde{w}_0 = w_0, \tilde{w}_{n/2} = w_{n/2}$. If n is odd, then there exists a vector $\tilde{p}$ with $\tilde{p}_j = 0$ for all $j \notin \{0, (n-1)/2, n\}$ and $\sum_{j=0}^{n} \tilde{p}_j = 1$ such that $\tilde{w}_0 = w_0$ and $\tilde{w}_{(n-1)/2} = w_{(n-1)/2}$, where $\{\tilde{w}_j\}_{j\in\mathbb{Z}}$ is the sequence of weights associated to $\tilde{p}$. Moreover, in both cases,*

$$\mathcal{E}(V, p) \leq \mathcal{E}(V, \tilde{p}) \ . \tag{2.7}$$

Proof. We first consider even n. Define $\tilde{p}$ by $\tilde{p}_{n/2} = \frac{2}{n}(w_0 - w_{n/2}), \tilde{p}_n = \frac{1}{n}w_{n/2}, \tilde{p}_0 = 1 - \tilde{p}_{n/2} - \tilde{p}_n$ and $\tilde{p}_j = 0$ for all $j \notin \{0, n/2, n\}$. As $w_0 \geq w_{n/2}$ and $w_{n/2} \geq 0$, we conclude that $\tilde{p}_{n/2} \geq 0$ and $\tilde{p}_n \geq 0$. From the definition of $\tilde{p}$, we verify that the weights associated with $\tilde{p}$ satisfy $\tilde{w}_0 = \frac{n}{2}\tilde{p}_{n/2} + n\tilde{p}_n = w_0$ and $\tilde{w}_{n/2} = n\tilde{p}_n = w_{n/2}$. Now, we need to show that vector $\tilde{p}$ is a probability vector. From the definition, we have

$$\begin{aligned}
\tilde{p}_{n/2} + \tilde{p}_n &= 2\frac{w_0}{n} - \frac{1}{n}w_{n/2} \\
&= \frac{1}{n}\left[2p_1 + 4p_2 + \cdots + np_{\frac{n}{2}} + np_{(\frac{n}{2}+1)} + \cdots + np_n\right] \\
&\leq \frac{1}{n}\left[np_1 + np_2 + \cdots + np_n\right] \\
&= \sum_{l=1}^{n} p_l \leq 1.
\end{aligned}$$

Thus $\tilde{p}$ is a probability vector. As $\tilde{w}_0 = w_0, \tilde{w}_{n/2} = w_{n/2}$ and the sequence $\{\tilde{w}_j\}_{j\in\mathbb{Z}}$ restricted to $0 \leq j \leq n/2$ is linear in j and is a convex combination of $\tilde{w}_0$ and $\tilde{w}_{n/2}$, so by the convexity from Lemma 2.3.2, $w_j \leq \tilde{w}_j$ for all $j \in \mathbb{Z}$ Therefore, by definition, we obtain

$$\mathcal{E}(V,p) \leq \mathcal{E}(V,\tilde{p})\,.$$

For n being odd, we define $\tilde{p}$ by $\tilde{p}_{(n-1)/2} = \dfrac{2}{n-1}(w_0 - w_{(n-1)/2}), \tilde{p}_n = \dfrac{1}{n}w_{(n-1)/2}$, $\tilde{p}_0 = 1 - \tilde{p}_{(n-1)/2} - \tilde{p}_n$ and $\tilde{p}_j = 0$ for all $j \notin \{0, (n-1)/2, n\}$. From the definition of $\tilde{p}$, we have $\tilde{w}_0 = \dfrac{n-1}{2}\tilde{p}_{(n-1)/2} + n\tilde{p}_n = w_0$ and $\tilde{w}_{(n-1)/2} = n\tilde{p}_n = w_{(n-1)/2}$.

Again, in this case, the sequence $\{\tilde{w}_j\}_{j\in\mathbb{Z}}$ restricted to $0 \leq j \leq (n-1)/2$ is linear in j and is a convex combination of $\tilde{w}_0$ and $\tilde{w}_{(n-1)/2}$, so by the convexity from the Lemma 2.3.2, $w_j \leq \tilde{w}_j$ for all $j \in \mathbb{Z}$ and hence

$$\mathcal{E}(V,p) \leq \mathcal{E}(V,\tilde{p})\,.$$

□

Now, we examine the conditions under which equality in (2.7) holds in the above theorem. The following theorem states a necessary and sufficient condition for the equality in (2.7).

Theorem 2.3.4. *Let n be a positive integer and p and $\tilde{p}$ be as in Theorem 2.3.3. A necessary and sufficient condition for $\mathcal{E}(V,p) = \mathcal{E}(V,\tilde{p})$ for all (n,k)-frames with analysis operator V is $p_i = \tilde{p}_i$ for $1 \leq i \leq n$. Moreover, when n is odd, $\tilde{p}$ is a probability vector if*

$$\sum_{l=0}^{(n-1)/2} (n-1-2l)p_l \geq \sum_{l=(n+1)/2}^{n} p_l \; .$$

Proof. If $p_i = \tilde{p}_i$ for $1 \leq i \leq n$, then $w_i = \tilde{w}_i$. Therefore, by definition, for any (n,k)-frame with analysis operator V, we have

$$\mathcal{E}(V,p) = \mathcal{E}(V,\tilde{p}) \, .$$

Conversely, if $\mathcal{E}(V,p) = \mathcal{E}(V,\tilde{p})$, then using Theorem 2.2.4, we have

$$\frac{1}{kn} \sum_{j,l=1}^{n} \left(\tilde{w}_{j-l} - w_{j-l}\right) |G_{j,l}|^2 = 0 \, . \tag{2.8}$$

From Theorem 2.3.3, we know that $\tilde{w}_i \geq w_i$ for each i and therefore, from relation (2.8), we conclude that $\tilde{w}_{j-l} = w_{j-l}$ whenever $|G_{j,l}| \neq 0$. As $\mathcal{E}(V,p) = \mathcal{E}(V,\tilde{p})$ for any (n,k)-frame with analysis operator V and corresponding Grammian G, we can always find one such G such that for fixed $1 \leq j,l \leq n$, $|G_{j,l}| \neq 0$ and therefore, we have $\tilde{w}_i = w_i$ for each $i \in \mathbb{Z}$. Hence, we conclude that $p_j = \tilde{p}_j$ for all $1 \leq j \leq n$.

From definition of $\tilde{p}_n$ and $\tilde{p}_{(n-1)/2}$ as in the proof of Theorem 2.3.3 both are non-negative. For $\tilde{p}$ to be a probability vector, it is enough to have $\tilde{p}_n + \tilde{p}_{(n-1)/2} \leq 1$.

We have

$$\begin{aligned}
\tilde{p}_n + \tilde{p}_{(n-1)/2} &= \frac{1}{n-1}\left[2w_0 - \frac{n+1}{n}w_{(n-1)/2}\right] \\
&\leq \frac{1}{n-1}\left[2w_0 - w_{(n-1)/2}\right] \\
&= \frac{1}{n-1}\left[2\sum_{l=1}^{n} lp_l - \sum_{l=(n+1)/2}^{n}(2l-n)p_l\right] \\
&= \frac{1}{n-1}\left[\sum_{l=1}^{(n-1)/2} 2lp_l + \sum_{l=(n+1)/2}^{n}(2l-2l+n)p_l\right] \\
&= \frac{1}{n-1}\left[\sum_{l=1}^{(n-1)/2} 2lp_l + \sum_{l=(n+1)/2}^{n} np_l\right] \\
&= \sum_{l=0}^{n} p_l + \frac{1}{n-1}\left[\sum_{l=(n+1)/2}^{n} p_l - \sum_{l=0}^{(n-1)/2}(n-1-2l)p_l\right]
\end{aligned}$$

As $\sum_{l=0}^{n} p_l = 1$, we conclude that if

$$\sum_{l=0}^{(n-1)/2}(n-1-2l)p_l \geq \sum_{l=(n+1)/2}^{n} p_l\,,$$

then $\tilde{p}_n + \tilde{p}_{(n-1)/2} \leq 1$. □

2.3.4 An upper bound for cyclic frames

The next result focuses on cyclic (n, k)-frames. From proposition 1, we know that if F is a cyclic frame for $\mathcal{H}$, then the Grammian is a polynomial of the cyclic shift, $G = q(S)$, where $q(z) = q_0 + q_1 z + \cdots + q_{n-1}z^{n-1}$, $z \in \mathbb{C}$ with $q_j = \langle f_{j+1}, f_1\rangle = G_{1,(j+1)}$ and q assumes the value one at k of the n-th roots of unity and the value zero at the others. Similarly, we define $w(z) = w_0 + w_1 z + \ldots w_{n-1} z^{n-1}$. We show that $\mathcal{E}(V, p)$ can be expressed in terms of sum of evaluations of the polynomial $w(z)$ for finite number of z on the unit circle.

Proposition 2.3.5. *Let F be a cyclic (n, k)-frame, with Grammian $G = q(S)$ and $q(e^{2\pi ij/n}) = 1$ for $j \in \mathbb{K} \subset \{1, 2, \ldots n\}$, $|\mathbb{K}| = k$. Let $p = (p_0, p_1, \ldots, p_n)$ be a vector with $p_j \geq 0$ for $1 \leq j \leq n$ and $\{w_j\}$ the sequence of associated weights as in definition 2.2.3, then*

$$\mathcal{E}(V, p) = \frac{1}{kn^2} \sum_{j,l \in \mathbb{K}} w(e^{2\pi i(j-l)/n}) \,.$$

Proof. This follows from the Plancherel theorem and the convolution theorem for the Discrete Fourier Transform.

For cyclic frames, we know that $G_{i,j} = G_{i+l,j+l}$ for all $1 \leq i, j, l \leq n$ (where sum is taken *modulo* n). Thus, we have

$$\begin{aligned}\mathcal{E}(V, p) &= \frac{1}{kn} \sum_{t,l=1}^{n} w_{t-l} |G_{t,l}|^2 = \frac{1}{k} \sum_{j=0}^{n-1} w_j |G_{(j+1),1}|^2 \\ &= \frac{1}{k} \sum_{j=0}^{n-1} w_j |q_j|^2 = \frac{1}{k} \sum_{j=0}^{n-1} w_j q_j \bar{q}_j \,.\end{aligned}$$

Let $\alpha = (w_1 q_1, \cdots, w_n q_n)^T, \beta = (q_1, \cdots, q_n)^T, \gamma = (w_1, \cdots, w_n)^T$ and $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$ be their Fourier transforms respectively. Using the Plancherel Theorem and convolution theorem, the above relation reduces to

$$\begin{aligned}\mathcal{E}(V,p) &= \frac{1}{k}\langle \alpha, \beta\rangle \\ &= \frac{1}{kn}\langle \hat{\alpha}, \hat{\beta}\rangle \\ &= \frac{1}{kn}\langle \frac{1}{n}\hat{\gamma} * \hat{\beta}, \hat{\beta}\rangle \\ &= \frac{1}{kn^2}\sum_{j=1}^{n}(w * q)(e^{2\pi i j/n})q(e^{2\pi i j/n})\end{aligned}$$

As $q(e^{2\pi i j/n}) = 1$ whenever $j \in \mathbb{K}$ and 0 otherwise, we have

$$\begin{aligned}\mathcal{E}(V,p) &= \frac{1}{kn^2}\sum_{j\in\mathbb{K}}(w * q)(e^{2\pi i j/n}) \\ &= \frac{1}{kn^2}\sum_{j\in\mathbb{K}}\sum_{l=1}^{n} w(e^{2\pi i(j-l)/n})q(e^{2\pi i l/n}) \\ &= \frac{1}{kn^2}\sum_{j,l\in\mathbb{K}} w(e^{2\pi i(j-l)/n}) .\end{aligned}$$

□

Consequently, to design an optimal cyclic frame, we want to find the set $\mathbb{K}$ containing indices for which the average of the pair potential $w(e^{2\pi i(j-l)/n})$ over all $j, l \in \mathbb{K}$, is minimized. In principle, this can be done by an exhaustive search. In order to obtain an analytic result, we could consider special cases for the probability vector. Another possibility is to estimate the associated weights $\{w_j\}$ and minimize the bound for the resulting error. We consider using the previously obtained upper bound $\mathcal{E}(V,p) \le \mathcal{E}(V,\tilde{p})$. In preparation to finding out what choice of $\mathbb{K}$ will give rise to the minimum value of $\mathcal{E}(V,\tilde{p})$ in the case of cyclic frame, we prove the following lemma.

Lemma 2.3.6. *Let $k \le n/2$ and $\mathbb{L} = \{\beta_1, \beta_2, \cdots, \beta_k\} \subset \mathbb{S} = \{1, 2, \cdots, 2n\}$ with $\beta_{l+1} = \beta_l + 2$ for all $1 \le l \le k-1$. Let $\mathbb{M} = \{\alpha_1, \alpha_2, \cdots, \alpha_k\}$ be any subset of $\mathbb{S}$ with $\alpha_l - \alpha_j \in 2\mathbb{Z}$ for all $1 \le l, j \le k$.*

Let $f : 2\mathbb{Z} \to \mathbb{R}$ be a symmetric increasing function, then,

$$\sum_{l,j=1}^{k} f(\beta_l - \beta_j) \le \sum_{l,j=1}^{k} f(\alpha_l - \alpha_j).$$

Proof. Without loss of generality, we may assume that $\alpha_{m+1} \ge \alpha_m$ for $1 \le m \le k-1$. As $\alpha_{m+1} - \alpha_m \in 2\mathbb{Z}$, there exists $c_m \in \mathbb{N}$ such that $\alpha_{m+1} = \alpha_m + 2c_m$. Thus, for $1 \le m \le k$, $\alpha_{m+1} = \alpha_1 + 2(c_1 + c_2 + \cdots + c_{m-1})$. Therefore, for $l > j$, we have

$$\begin{aligned}\alpha_l - \alpha_j &= [\alpha_1 + 2(c_1 + c_2 + \cdots + c_{l-1})] - [\alpha_1 + 2(c_1 + c_2 + \cdots + c_{j-1})] \\ &= 2(c_j + c_{j+1} + \cdots + c_{l-1}) \\ &\le 2(l-j) = \beta_l - \beta_j.\end{aligned}$$

As f is symmetric increasing, we have

$$\sum_{l,j=1}^{k} f(\beta_l - \beta_j) \le \sum_{l,j=1}^{k} f(\alpha_l - \alpha_j)\,.$$

□

We now describe how to choose $\mathbb{K} \subset \{1, 2, \cdots, n\}$ with $|\mathbb{K}| = k$ such that $\mathcal{E}(V, \tilde{p})$ is minimized and give the minimum values.

Theorem 2.3.7. *Let $p = (p_0, p_1, p_2, \ldots, p_n)$ be a probability vector. For n even and $k \le n/2$, a cyclic (n, k)-frame F with Grammian $G = q(S)$ and $q(e^{2\pi i j/n}) = 1$ for $j \in \mathbb{K} \subset \{1, 2, \ldots, n\}$ with $|\mathbb{K}| = k$, minimizes the upper bound $\mathcal{E}(V, \tilde{p})$ for the mean-square error given by Theorem 2.3.3 if and only if any two elements of $\mathbb{K}$ have an* *even difference. For n even and $k \geq n/2$, the frame minimizes the upper bound if and only if any two elements of the complement $\{1, 2, \ldots, n\} \setminus \mathbb{K}$ have an even difference. For odd n and $k < n/2$, a cyclic (n, k)-frame F minimizes the upper bound $\mathcal{E}(V, \tilde{p})$ if and only if $\mathbb{K}$ is formed by a sequence of evenly spaced numbers, with a difference of two modulo n for each consecutive pair. For odd n and $k > n/2$, the frame minimizes the upper bound if and only if the complement $\{1, 2, \ldots, n\} \setminus \mathbb{K}$ is formed by a sequence of evenly spaced numbers, with a difference of two modulo n for each consecutive pair.*

Proof. To evaluate $\mathcal{E}(V, \tilde{p})$, we use the fact that the sequence $\{\tilde{w}_j\}$ is piecewise linear.

In preparation for the following, we denote the l-th Dirichlet kernel by

$$D(\theta, l) = \frac{\sin(l\theta/2)}{\sin(\theta/2)}, \theta \in [0, 2\pi) \,.$$

We recall the Bartlett window sequence. If n is even, we define the sequence $\{b_j : -n/2 \leq j \leq n/2\}$ by

$$b_j = 1 - |\frac{2j}{n}|,$$

which gives the polynomial

$$b(e^{i\theta}) = \sum_{j=-n/2}^{n/2} b_j e^{ij\theta} = \frac{2}{n} D^2 \left(\theta, \frac{n}{2}\right) \,.$$

If n is odd, we let

$$b_j = 1 - |\frac{2j}{n-1}| \text{ for } -(n-1)/2 \leq j \leq (n-1)/2,$$

which gives

$$b(e^{i\theta}) = \sum_{j=-(n-1)/2}^{(n-1)/2} b_j e^{ij\theta} = \frac{2}{n-1} D^2 \left(\theta, \frac{n-1}{2}\right) \,.$$

Then for even n and $0 \leq j \leq n/2$, we have

$$\begin{aligned}\tilde{w}_j &= \left(\frac{n}{2} - j\right)\tilde{p}_{n/2} + n\tilde{p}_n \\ &= \left(1 - \frac{2j}{n}\right)(w_0 - w_{n/2}) + w_{n/2} \\ &= w_{n/2} + b_j(w_0 - w_{n/2})\,.\end{aligned}$$

For $l \in \{0, 1, \cdots, n-1\}$, using the definition of $\tilde{w}(z)$ and $\tilde{w}_j$, we have

$$\begin{aligned}\tilde{w}(e^{2\pi il/n}) &= \sum_{j=0}^{n} \tilde{w}_j e^{2\pi ilj/n} \\ &= \sum_{j=-n/2+1}^{n/2-1} \tilde{w}_j e^{2\pi ilj/n} + \tilde{w}_{n/2}e^{\pi il} \\ &= \sum_{j=-n/2+1}^{n/2-1} \left(w_{n/2} + b_j(w_0 - w_{n/2})\right) e^{2\pi ilj/n} + w_{n/2}e^{\pi il} \\ &= (w_0 - w_{n/2}) \sum_{j=-n/2}^{n/2} b_j e^{2\pi ilj/n} + w_{n/2}\left[e^{\pi il} + \sum_{j=-n/2+1}^{n/2-1} e^{2\pi ilj/n}\right] \\ &= (w_0 - w_{n/2})\frac{2}{n}D^2\left(2\pi l/n, \frac{n}{2}\right) + w_{n/2}D(2\pi l/n, n)\,.\end{aligned}$$

In order to minimize $\mathcal{E}(V, \tilde{p}) = \frac{1}{kn^2}\sum_{j,l\in\mathbb{K}} \tilde{w}(e^{2\pi i(j-l)/n})$, we must suppress the contribution from summing over all $j \neq l$ because that from $j = l$ does not depend on the particular choice of $\mathbb{K}$ with $|\mathbb{K}| = k$.

If n is even and $k \leq n/2$, we observe that $D^2(2\pi j/n, n/2) = 0$ if $j \in 2\mathbb{Z} \setminus \{0\}$ and is otherwise strictly positive. Therefore, if all pairs of indices $j, l \in \mathbb{K}$ have even differences, then we achieve the minimum and the minimum is given by

$$\begin{aligned}\mathcal{E}(\tilde{V}_{min}, \tilde{p}) &= \frac{1}{kn^2}\left[knw_{n/2} + \frac{2}{n}(w_0 - w_{n/2})k\left(\frac{n}{2}\right)^2\right] \\ &= \frac{1}{2n}\left[w_{n/2} + w_0\right]\,.\end{aligned}$$

If n is even and $k \geq n/2$, then $n-k \leq n/2$. Thus, treating $n-k$ as new k in the above argument gives the desired result. This concludes the case of even n.

Now, for odd n and $0 \leq j \leq n/2$, we have

$$\begin{aligned}
\tilde{w}_j &= \left(\frac{n-1}{2} - j\right)\tilde{p}_{(n-1)/2} + n\tilde{p}_n \\
&= \left(\frac{n-1}{2} - j\right)\frac{2}{n-1}(w_0 - w_{(n-1)/2}) + w_{(n-1)/2} \\
&= \left(1 - \frac{2j}{n-1}\right)(w_0 - w_{(n-1)/2}) + w_{(n-1)/2} \\
&= b_j(w_0 - w_{(n-1)/2}) + w_{(n-1)/2}\,.
\end{aligned}$$

For $l \in \{0, 1, \cdots, n-1\}$, we have

$$\begin{aligned}
\tilde{w}(e^{2\pi il/n}) &= \sum_{j=0}^{n} \tilde{w}_j e^{2\pi ilj/n} \\
&= \sum_{j=-(n-1)/2}^{(n-1)/2} \left\{b_j(w_0 - w_{(n-1)/2}) + w_{(n-1)/2}\right\} e^{2\pi ilj/n} \\
&= (w_0 - w_{(n-1)/2})\frac{2}{n-1}D^2\left(2\pi l/n, \frac{n-1}{2}\right) + w_{(n-1)/2}D(2\pi l/n, n).
\end{aligned}$$

To address the case of odd n, we embed the set $\mathbb{K} + n\mathbb{Z}$ in $\mathbb{Z}$. We partition $\mathbb{K} + n\mathbb{Z} = \mathbb{K}_1 \cup \mathbb{K}_2$ such that there are only even differences between all pairs of elements taken from either $\mathbb{K}_1$ or $\mathbb{K}_2$, and odd differences between pairs that contain elements from both sets. Then, since n is odd, $\mathbb{K}_1 + n = \mathbb{K}_2$.

For $j \in \mathbb{Z}$, we have

$$\sin^2\left(\frac{(n-1)}{4}\frac{2\pi}{n}j\right) = \sin^2\left(\frac{\pi j}{2} - \frac{\pi j}{2n}\right) = \begin{cases} \sin^2(\frac{\pi}{2n}j), & \text{if } j \text{ even} \\ \cos^2(\frac{\pi}{2n}j), & \text{if } j \text{ odd} \end{cases}.$$

Therefore, for $j \neq 0$, we have

$$D^2(2\pi j/n,(n-1)/2) = \begin{cases} \frac{\sin^2(\frac{\pi}{2n}j)}{\sin^2(\frac{\pi}{n}j)}, & \text{if } j \text{ even} \\ \\ \frac{\cos^2(\frac{\pi}{2n}j)}{\sin^2(\frac{\pi}{2n}j)}, & \text{if } j \text{ odd} \end{cases}$$
$$= \begin{cases} \frac{1}{4}\sec^2(\frac{\pi}{2n}j), & \text{if } j \text{ even} \\ \\ \frac{1}{4}\csc^2(\frac{\pi}{2n}j), & \text{if } j \text{ odd} \end{cases}.$$

Thus, for an odd j, we have

$$\begin{aligned} D^2(2\pi j/n,(n-1)/2) &= \frac{1}{4}\csc^2\left(\frac{\pi}{2n}j\right) \\ &= \frac{1}{4}\sec^2\left(\frac{\pi}{2n}(n-j)\right) \\ &= D^2(2\pi(n-j)/n,(n-1)/2) \end{aligned}$$

Therefore, we conclude that

$$\sum_{j,l\in\mathbb{K}} \tilde{w}(e^{2\pi i(j-l)/n}) = \sum_{j,l\in\tilde{\mathbb{K}}_1} \tilde{w}(e^{2\pi i(j-l)/n}),$$

where $\tilde{\mathbb{K}}_1 = \mathbb{K}_1 \cap S$ and $S = \{0, 1, \cdots, 2n-1, 2n\}$.

In this case, all differences between $j, l \in \tilde{\mathbb{K}}_1$ are even, and so $\tilde{w}(e^{2\pi i(j-l)/n})$ is symmetric increasing in $j - l \in 2\mathbb{Z}$.

We consider rearrangements among the set of numbers in $\{1, 2, \ldots, 2n\} \cap (\mathbb{K}_1 + 2\mathbb{Z})$. Using Lemma 2.3.6, we obtain that the sum $\sum_{j,l\in\tilde{\mathbb{K}}_1} \tilde{w}(e^{2\pi i(j-l)/n})$ is minimal if the rearranged $\tilde{\mathbb{K}}_1$ consists of evenly spaced numbers, with a difference of two modulo $2n$ for each consecutive pair.

Now we identify points in $\tilde{\mathbb{K}}_1$ which differ by n and thus map this set onto $\mathbb{K}$, to get the minimizers for the mean-square error.

From the above and from Theorem 2.2.5, we infer that if a frame $\hat{F}$ with analysis operator $\hat{V}$ is complementary to an (n,k)-frame F, then it is a minimizer of $\mathcal{E}(\hat{V},p)$ among all cyclic $(n,n-k)$-frames if and only if F is a minimzier among all cyclic (n,k)-frames. From the characterization of minimizers when $k \leq n/2$ we thus obtain the minimizers if $k \geq n/2$. □

Remark 2.3.8. As pointed out in the literature, frames can be viewed as codes over the real or complex numbers [49, 8]. If $k \geq n/2$ and if it is relatively prime to n, then an (n,k)-frame F which minimizes the upper bound for the mean-square reconstruction error is a complex BCH code as defined in [49]. The performance of these codes in the context of active error correction has already been investigated in previous works [56, 57, 58].

2.3.5 An improved upper bound

Earlier we used a convexity argument and replaced the probability vector for the burst lengths by another vector $\tilde{p}$ to get an upper bound for $\mathcal{E}(V,p)$. Now, we use a similar argument to get an improved upper bound in the case when n is a multiple of 2^m for some fixed m. We consider the case of a transmission that erases no coefficients, $n/2^m$ coefficients, $n/2^{m-1}$ coefficients, $\cdots$, $n/2$ coefficients, or all of the coefficients. We obtain an upper bound for $\mathcal{E}(V,p)$ by replacing the probability vector p by another vector $\tilde{p}$ with $\tilde{p}_j = 0$ if $j \notin \{0, n/2^m, n/2^{m-1}, \cdots, n/2, n\}$ and

$p_j \geq 0$ for $1 \leq j \leq n$ so that the weights $\tilde{w}_j$ can not decrease.

Theorem 2.3.9. *Let F be an (n, k) frame with analysis operator V. Let n be a multiple of 2^m and $p = (p_0, p_1, \cdots, p_n)$ be a probability vector. Then there exists a vector $\tilde{p} = (\tilde{p}_0, \tilde{p}_1, \cdots, \tilde{p}_n)$ with $p_j = 0$ for $j \notin \{0, n/2^m, n/2^{m-1}, \cdots, n/4, n/2, n\}, p_j \geq 0$ for $1 \leq j \leq n$ and satisfying $\tilde{p}_0 = 1 - \sum_{j=1}^{n} \tilde{p}_j$ such that the associated sequence of weights $\{w_j\}$ and $\{\tilde{w}_j\}$ agree for the indices $j \in \{0, n/2^m, n/2^{m-1}, \cdots, n/4, n/2\}$ when $0 \leq j \leq n/2$, and*

$$\mathcal{E}(V, p) \leq \mathcal{E}(V, \tilde{p}) .$$

Proof. We define $\tilde{p}_j = 0$ for $j \notin \{0, n/2^m, n/2^{m-1}, \cdots, n/4, n/2, n\}$. By the definition of the weights $\tilde{w}_j$ associated to the vector $\tilde{p}$, we obtain the following: for $2 \leq t \leq m$ and $n/2^t \leq j \leq n/2^{(t-1)} - 1$,

$$\tilde{w}_j = \left(\frac{n}{2^{(t-1)}} - j\right) \tilde{p}_{n/2^{(t-1)}} + \left(\frac{n}{2^{(t-2)}} - j\right) \tilde{p}_{n/2^{(t-2)}} + \cdots + \left(\frac{n}{2} - j\right) \tilde{p}_{n/2} + n\tilde{p}_n ,$$

for $0 \leq j \leq n/2^m - 1$,

$$\tilde{w}_j = \left(\frac{n}{2^m} - j\right) \tilde{p}_{n/2^m} + \left(\frac{n}{2^{m-1}} - j\right) \tilde{p}_{n/2^{m-1}} + \cdots + \left(\frac{n}{2} - j\right) \tilde{p}_{n/2} + n\tilde{p}_n .$$

and $\tilde{w}_{n/2} = n\tilde{p}_n$.

Now, defining $\tilde{w}_j = w_j$ for $j \in \{0, n/2^m, n/2^{m-1}, \cdots, n/4, n/2\}$, we obtain that for $1 \leq t \leq m$,

$$w_{n/2^t} = n\tilde{p}_n + n \sum_{l=1}^{t-1} \left(\frac{1}{2^{t-l}} - \frac{1}{2^t}\right) \tilde{p}_{n/2^{t-l}}$$

and

$$w_0 = n \sum_{l=0}^{m} \frac{1}{2^{m-l}} \tilde{p}_{n/2^{m-l}} .$$

From the above equations, we obtain $\tilde{p}_n = \dfrac{1}{n} w_{n/2}$ and $\tilde{p}_{n/2} = \frac{4}{n}(w_{n/4} - w_{n/2})$. For $2 \le t \le m-1$, we consider the following expression

$$\alpha(t) = 2^{t+1} w_{n/2^{n+1}} - 3(2)^t w_{n/2^t} + 2^t w_{n/2^{n-1}}\,.$$

Using the definition of w_j's and simplifying, we get

$$\alpha(t) = n\left[\sum_{i=1}^{t}(2^i-1)\tilde{p}_{n/2^{t+1-i}} - 3\sum_{i=1}^{t-1}(2^i-1)\tilde{p}_{n/2^{t-i}} + 2\sum_{i=1}^{t-2}(2^i-1)\tilde{p}_{n/2^{t-1-i}}\right].$$

By changing the indexing in the first and third term of the above expression, we obtain

$$\alpha(t) = n\left[\sum_{i=0}^{t-1}(2^{i+1}-1)\tilde{p}_{n/2^{t-i}} - 3\sum_{i=1}^{t-1}(2^i-1)\tilde{p}_{n/2^{t-i}} + 2\sum_{i=2}^{t-1}(2^{i-1}-1)\tilde{p}_{n/2^{t-i}}\right].$$

Combining the coefficients of $\tilde{p}_{n/2^{t-i}}$ and simplifying, we get $\alpha(t) = n\tilde{p}_{n/2^t}$ and therefore,

$$\tilde{p}_{n/2^t} = \frac{2^t}{3n}\left(\frac{2}{3} w_{n/2^{t+1}} - w_{n/2^t} + \frac{1}{3} w_{n/2^{t-1}}\right)$$

for all $2 \le t \le m-1$.

To find $\tilde{p}_{n/2^m}$, we consider the expression $\alpha_0 = 2^m w_0 - 2^{m+1} w_{n/2^m} + 2^m w_{n/2^{m-1}}$. Again, using the definition of w_j and simplifying, we get

$$\alpha_0 = n\left[\sum_{i=0}^{m-1} 2^i \tilde{p}_{n/2^{m-i}} - 2\sum_{i=1}^{m-1}(2^i-1)\tilde{p}_{n/2^{m-i}} + 2\sum_{i=1}^{m-2}(2^i-1)\tilde{p}_{n/2^{m-1-i}}\right].$$

By changing the indexing in the third term of the above expression and then combining the coefficients of $\tilde{p}_{n/2^{m-i}}$, we obtain $\alpha_0 = n\tilde{p}_{n/2^m}$. Therefore, $\tilde{p}_{n/2^m} = \dfrac{2^{m+1}}{n}\left(\dfrac{1}{2} w_0 - w_{n/2^m} + \dfrac{1}{2} w_{n/2^{m-1}}\right)$.

As w_j is non-negative, decreasing and convex in j for $0 \le j \le n/2$ (Lemma 2.3.2), we conclude that $\tilde{p}_{n/2^t} \ge 0$ for all $0 \le t \le m$. Finally, define $\tilde{p}_0 = 1 - \sum_{j=1}^{n} \tilde{p}_j$. As

$w_j \geq w_{(j+1)}$ for all $0 \leq j \leq (n/2-1)$, we conclude that $\tilde{p}_j \geq 0$ for all $1 \leq j \leq n$. Now, as

$$\mathcal{E}(V,p) = \frac{1}{kn} \sum_{j,l=1}^{n} w_{j-l} |G_{j,l}|^2,$$

it is enough to show that $w_j \leq \tilde{w}_j$ for each $0 \leq j \leq n/2$.

Now, we denote $n/2^m$ by r. We observe that $\tilde{w}_j$ is linear in j on $[0,r]$ and on $[2^t r, 2^{t+1} r]$ for each fixed $0 \leq t \leq m-2$. As $\tilde{w}_0 = w_0$, $\tilde{w}_r = w_r$ and w_j is decreasing and convex in j(from Lemma 2.3.2), we conclude that $\tilde{w}_j \leq w_j$ for all $0 \leq j \leq r$. A similar argument shows that for any fixed $0 \leq t \leq m-2$, $\tilde{w}_j \leq w_j$ for all $2^t r \leq j \leq 2^{t+1} r$. This proves the theorem. □

Again, we examine the special case of a cyclic (n,k)-frame F, with Grammian $G = q(S)$ with $q(e^{2\pi i j/n}) = 1$ for $j \in \mathbb{K} \subset \{0,1,2,\ldots n-1\}$, $|\mathbb{K}| = k$ and $q(e^{2\pi i j/n}) = 1$ for $j \in \{0,1,2,\ldots n-1\} \setminus \mathbb{K}$.

Lemma 2.3.10. *Let F be a cyclic (n,k) frame with Grammian $G = q(S)$, where S is the shift matrix. Let $\tilde{p}$ be as in Theorem 2.3.9. Then for any fixed $0 \leq l \leq n-1$,*

$$\tilde{w}(e^{2\pi i l/n}) = \tilde{p}_{n/2^m} D^2(2\pi l/n, n/2^m) + \tilde{p}_{n/2^{m-1}} D^2(2\pi l/n, n/2^{m-1}) + \cdots$$
$$+ \tilde{p}_{n/2} D^2(2\pi l/n, n/2) + n\tilde{p}_n D(2\pi l/n, n).$$

Proof. For notational convenience, we abbreviate $\Delta_{t,j} = \frac{n}{2^t} - |j|$. For fixed $0 \leq l \leq n-1$, using the definition of $\tilde{w}_j$, we have

$$\begin{aligned}
\tilde{w}(e^{2\pi il/n}) &= \sum_{j=0}^{n} \tilde{w}_j e^{2\pi ilj/n} \\
&= \sum_{j=-n/2+1}^{j=n/2-1} \tilde{w}_j e^{2\pi ilj/n} + \tilde{w}_{n/2} e^{\pi il} \\
&= \sum_{j=-n/2^m+1}^{j=n/2^m-1} \Delta_{m,j} \tilde{p}_{n/2^m} e^{2\pi ilj} + \sum_{j=-n/2^{m-1}+1}^{j=n/2^{m-1}-1} \Delta_{m-1,j} \tilde{p}_{n/2^{m-1}} e^{2\pi ilj} + \\
&\qquad \cdots + \sum_{j=-n/2+1}^{j=n/2-1} \Delta_{1,j} \tilde{p}_{n/2} e^{2\pi ilj} + n\tilde{p}_n \left[1 + \sum_{j=-n/2+1}^{j=n/2-1} e^{2\pi ilj} \right] \\
&= \tilde{p}_{n/2^m} D^2(2\pi l/n, n/2^m) + \tilde{p}_{n/2^{m-1}} D^2(2\pi l/n, n/2^{m-1}) + \cdots \\
&\qquad + \tilde{p}_{n/2} D^2(2\pi l/n, n/2) + n\tilde{p}_n D(2\pi l/n, n)\,.
\end{aligned}$$

□

Theorem 2.3.11. *If $k \le n/2^m$ and $\tilde{p}$ is as in Theorem 2.3.9, then a cyclic (n,k)-frame F with Grammian $G = q(S)$ and frequency support $\mathbb{K}$, $q(e^{2\pi ij/n}) = 1$ for $j \in \mathbb{K} \subset \{0,1,2,\cdots,n-1\}$, minimizes the upper bound $\mathcal{E}(V,\tilde{p})$ if and only if each pair $j,l \in \mathbb{K}, j \ne l$, satisfies $(j-l) \in 2^m\mathbb{Z} \setminus \{0\}$. If $k \ge n - n/2^m$, then F minimizes the upper bound if and only if each pair $j,l \in \{1,2,\ldots,n\} \setminus \mathbb{K}, j \ne l$, satisfies $(j-l) \in 2^m\mathbb{Z} \setminus \{0\}$. Moreover, in the case of $k \le n/2^m$, there exists an optimal frame with analysis operator $\tilde{V}_{min}$ which achieves the minimum value of the mean-square error given by*

$$\mathcal{E}(\tilde{V}_{min},\tilde{p}) = \frac{1}{n2^m}\left[w_0 + 2w_{n/2^m} + 2^{m-2}w_{n/2} + 3\sum_{l=1}^{m-2} 2^{l-1} w_{n/2^{m-l}} \right].$$

If $k \ge n - n/2^m$, then we have an optimal frame with analysis operator $\tilde{V}_{min}$ and

$$\mathcal{E}(\tilde{V}_{min},\tilde{p}) = \frac{1}{n2^m}\left[\big(2^m\frac{n-2k}{k} + 1\big) w_0 + 2w_{n/2^m} + 2^{m-2}w_{n/2} + 3\sum_{l=1}^{m-2} 2^{l-1} w_{n/2^{m-l}} \right].$$

Proof. We have

$$\mathcal{E}(V,\tilde{p}) = \frac{1}{kn^2}\sum_{j,l\in\mathbb{K}} \tilde{w}(e^{2\pi i(j-l)/n})$$

In order to minimize $\mathcal{E}(V,\tilde{p})$, we must minimize the contributions from $j \neq l$ as the contribution from $j = l$ does not depend on what $\mathbb{K}$ we pick with $|\mathbb{K}| = k$. From above Lemma, we observe that $\tilde{w}(e^{2\pi i(j-l)/n})$ is minimized if $D^2(2\pi(l - j)/n, n/2^m) = 0, D^2(2\pi(l-j)/n, n/2^{m-1}) = 0, \cdots, D^2(2\pi(l-j)/n, n/4) = 0$ and $D^2(2\pi(l-j)/n, n/2) = 0$, and this happens if $(l-j) \in 2^m\mathbb{Z} \setminus \{0\}$. Therefore, if we pick $\mathbb{K}$ in such a way that each pair $j, l \in \mathbb{K}, j \neq l$ is such that $(j-l) \in 2^m\mathbb{Z} \setminus \{0\}$, then $\mathcal{E}(V,\tilde{p})$ is minimized.

If $k \geq n - n/2^m$, then we have $n - k \leq n/2^m$. Choosing $n-k$ as new k in the above arguments, we obtain the result in this case.

Moreover, in the case of $k \leq n/2^m$ minimum value of $\mathcal{E}(V,\tilde{p})$ is given by

$$\begin{aligned}\mathcal{E}(\tilde{V}_{min},\tilde{p}) &= \frac{1}{kn^2}\left[\tilde{p}_{n/2^m}k\left(\frac{n}{2^m}\right)^2 + \tilde{p}_{n/2^{m-1}}k\left(\frac{n}{2^{m-1}}\right)^2 + \cdots + \tilde{p}_{n/2}k\left(\frac{n}{2}\right)^2 + nkn\tilde{p}_n\right] \\ &= \frac{1}{2^{2m}}\left[\sum_{l=0}^{m} 2^{2l}\tilde{p}_{n/2^{m-l}}\right].\end{aligned}$$

Substituting the values of $\tilde{p}_{n/2^{m-l}}$ for $1 \leq l \leq m-2$ in the above expression and simplifying, we obtain

$$\begin{aligned}\mathcal{E}(\tilde{V}_{min},\tilde{p}) = \frac{1}{n2^m}\Bigg[&\frac{n}{2^m}\tilde{p}_{n/2^m} + \sum_{l=1}^{m-2} 2^{l+1} w_{n/2^{m-l+1}} - 3\sum_{l=1}^{m-2} 2^l w_{n/2^{m-l}} \\ &+ \sum_{l=1}^{m-2} 2^l w_{n/2^{m-l-1}} + n2^{(m-2)}\tilde{p}_{n/2} + n2^m\tilde{p}_n\Bigg].\end{aligned}$$

Plugging the values of $\tilde{p}_{n/2^m}, \tilde{p}_{n/2}, \tilde{p}_n$ and changing the indexing in the first and third sums in the above expression, we get

$$\mathcal{E}(\tilde{V}_{min}, \tilde{p}) = \frac{1}{n2^m}\left[w_0 - 2w_{n/2^m} + w_{n/2^{m-1}} + \sum_{l=0}^{m-3} 2^{l+2} w_{n/2^{m-l}} - 3\sum_{l=1}^{m-2} 2^l w_{n/2^{m-l}} \right.$$
$$\left. + \sum_{l=2}^{m-1} 2^{l-1} w_{n/2^{m-l}} + 2^m (w_{n/4} - w_{n/2}) + 2^m w_{n/2}\right].$$

Combining the coefficients of $w_{n/2^{m-t}}$ for $0 \le t \le m$ and simplifying them, we obtain

$$\mathcal{E}(\tilde{V}_{min}, \tilde{p}) = \frac{1}{n2^m}\left[w_0 + 2w_{n/2^m} + 2^{m-2} w_{n/2} + 3\sum_{l=1}^{m-2} 2^{l-1} w_{n/2^{m-l}}\right].$$

Using the identity for complementary frames in Theorem 2.2.5, we deduce from the formula for $k \le n/2^m$ the corresponding expression for the minimal mean-square error when $k \ge n - n/2^m$. □

2.3.6 Accuracy of the upper bound for cyclic (n, k)-frames

We observe that Theorem 2.3.11 is concerned with a minimizer for $\mathcal{E}(V, \tilde{p})$, which resulted from modifying the burst-length statistics to obtain an upper bound for the mean-square error. Next, we want to compare the performance of this minimizer with the unmodified statistics given by p.

Let $\mathcal{E}(V_{min}, p)$ be the minimum of $\mathcal{E}(V, p)$, where the minimum is taken over all V belonging to cyclic (n, k)-frames. In the case of a cyclic (n, k)-frame, from Theorem 2.3.11, the minimum of $\mathcal{E}(V, \tilde{p})$ is given by

$$\mathcal{E}(\tilde{V}_{min}, \tilde{p}) = \frac{1}{n2^m}\left[w_0 + 2w_{n/2^m} + 2^{m-2} w_{n/2} + 3\sum_{l=1}^{m-2} 2^{l-1} w_{n/2^{m-l}}\right].$$

As $\mathcal{E}(V_{min}, p) \le \mathcal{E}(V, p)$ and $\mathcal{E}(V, p) \le \mathcal{E}(V, \tilde{p})$ (from Theorem 2.3.9) for any fixed

V, we conclude that $\mathcal{E}(V_{min}, p) \leq \mathcal{E}(\tilde{V}_{min}, p)$. We now obtain an upper bound for $\mathcal{E}(\tilde{V}_{min}, p) - \mathcal{E}(V_{min}, p)$.

Theorem 2.3.12. *Let p and $\tilde{p}$ be as in Theorem 2.3.9. If V_{min} and $\tilde{V}_{min}$ are as defined above, then*

$$\mathcal{E}(\tilde{V}_{min}, p) - \mathcal{E}(V_{min}, p) \leq \frac{n-k}{n^2}\left(w_1 - w_{n/2}\right) .$$

Proof. Let $\tilde{G}$ and G be the Grammian for the frames with analysis operators $\tilde{V}_{min}$ and V_{min} respectively. Using $\tilde{G}_{j,j} = G_{j,j}$ for $1 \leq j \leq n$, we obtain

$$\begin{aligned}
\mathcal{E}(\tilde{V}_{min}, p) - \mathcal{E}(V_{min}, p) &= \frac{1}{kn}\sum_{l,j=1}^{n} w_{l-j}\left[|\tilde{G}_{j,l}|^2 - |G_{j,l}|^2\right] \\
&= \frac{1}{kn}\sum_{\substack{l,j=1 \\ l\neq j}}^{n} w_{l-j}\left[|\tilde{G}_{j,l}|^2 - |G_{j,l}|^2\right] \\
&\leq \frac{1}{kn}\left[\sum_{\substack{l,j \\ |\tilde{G}_{j,l}|\geq|G_{j,l}|}} w_1\left(|\tilde{G}_{j,l}|^2 - |G_{j,l}|^2\right)\right. \\
&\qquad \left. + \sum_{\substack{l,j \\ |\tilde{G}_{j,l}|<|G_{j,l}|}} w_{n/2}\left(|\tilde{G}_{j,l}|^2 - |G_{j,l}|^2\right)\right] \\
&= \frac{1}{kn}\left[w_{n/2}\sum_{l,j=1}^{n}\left(|\tilde{G}_{j,l}|^2 - |G_{j,l}|^2\right)\right. \\
&\qquad \left. + \sum_{\substack{l,j \\ |\tilde{G}_{j,l}|\geq|G_{j,l}|}} (w_1 - w_{n/2})\left(|\tilde{G}_{j,l}|^2 - |G_{j,l}|^2\right)\right] .
\end{aligned}$$

Using relations (2.5) and (2.6) in the above expression, we get

$$\mathcal{E}(\tilde{V}_{min}, p) - \mathcal{E}(V_{min}, p) \leq \frac{1}{kn}(w_1 - w_{n/2}) \sum_{\substack{l,j=1 \\ l \neq j}}^{n} |\tilde{G}_{j,l}|^2 = \frac{n-k}{n^2} \left(w_1 - w_{n/2}\right) .$$

□

2.3.7 Example: packetization and burst lengths

We examine a particular case in which the burst-length statistics depend on just one parameter $\beta \in [0, 1]$. This parameter is related to transition probabilities in a two-state Markov model for the transmission channel, see [64]. The two states correspond to a frame coefficient being erased or being perfectly transmitted. Given that a coefficient is erased, then the conditional probability of the following coefficient being erased is β.

The problem we wish to consider is the transmission of a sequence of vectors in terms of packets of size n, given by the frame coefficients corresponding to each vector. If the packet/frame size were infinite, then a geometric distribution for burst lengths would ensue. To obtain the mean-square reconstruction error in the finite-length case, we compute the effective distribution of the burst lengths inside each packet. For simplicity, we eliminate a trivial dependence and assume that the probability of a burst occurring within each packet is equal to one. If packetization is ignored, then it is natural to assume that each position in a sequence of transmitted frame coefficients should have the same probability for being the starting point of a burst. Partitioning the sequence into consecutive packets of a given length leads to an effective distribution in which the first position in a packet has a higher probability of being the starting point compared to the others. On the other hand, the fixed packet size effectively truncates long bursts.

The length statistics of the truncated bursts depends on their starting point within a packet. Let $\{X_t : t \in \{1, 2, \cdots, n\}\}$ be a sequence of integer-valued random variables such that the value $r \in \{1, 2, \ldots, n\}$ is assumed with probability

$$P(X_t = r) = \begin{cases} (1-\beta)\beta^{r-1}, & r < n-t+1 \\ \beta^{n-t}, & r = n-t+1 \\ 0, & \text{else} \end{cases} .$$

Here, t indicates the starting point of the burst in a packet.

To compute the probability that the first position in a packet belongs to a burst, we consider the N coefficients preceding a packet together with the those in the packet, for some fixed $N \in \mathbb{N}$. Assuming that a burst could start at any $N + n$ position of the sequence with equal probability, if M is a random variable specifying at which position the burst starts *inside the packet*, then the probabilities P_N for M are

$$P_N(M = 2) = P_N(M = 3) = \cdots = P_N(M = n) = \frac{A_N}{N + n}$$

and

$$P_N(M = 1) = \frac{A_N}{N + n}(1 + \beta + \beta^2 + \cdots + \beta^N) = \frac{1 - \beta^{N+1}}{(N + n)(1 - \beta)} ,$$

where A_N is a normalization constant. Now, we have

$$\sum_{i=1}^{n} P_N(M = i) = \frac{A_N}{N + n}\left(n - 1 + \frac{1 - \beta^{N+1}}{1 - \beta}\right) = 1$$

which determines the value of A_N. After normalizing and letting $N \to \infty$, we get the probabilities $P(M = i)$ that a burst starts at position i in the packet and they are given by

$$P(M = i) = \begin{cases} \frac{1}{1+(n-1)(1-\beta)}, & i = 1 \\ \\ \frac{1-\beta}{1+(n-1)(1-\beta)}, & 2 \le i \le n \end{cases} .$$

Therefore, p, the length statistics for the burst erasure is given by $p_0 = 0$ and for $1 \le j \le n$

$$p_j = P(X_M = j) = \sum_{t=1}^{n} P(M = t) P(X_t = j) .$$

We plug in the values of $P(M = t)$ and $P(X_t = j)$ in the above expression and simplify to obtain

$$p_j = \begin{cases} \frac{(1-\beta)\beta^{j-1}}{1+(n-1)(1-\beta)}[2 + (n-1-j)(1-\beta)], & 1 \le j \le n-1 \\ \\ \frac{\beta^{n-1}}{1+(n-1)(1-\beta)}, & j = n \end{cases} .$$

From p, we get to $\tilde{p}$ as in Theorem 2.3.9, calculate $\mathcal{E}(\tilde{V}_{min}, \tilde{p})$ as an upper bound and compare this to $\mathcal{E}(\tilde{V}_{min}, p)$ for different values of m, fixed k and $r = n/2^m$. We include log-log plots of $\frac{1}{n}\left[\frac{k}{n} w_0 + \left(1 - \frac{k}{n}\right) w_{n/2}\right]$, of $\mathcal{E}(\tilde{V}_{min}, \tilde{p})$, of $\mathcal{E}(\tilde{V}_{min}, p)$ and of the expected burst length for each $m \in \{1, 2, 3, 4\}, k = 9$ and $r = 10$ in Figure 2.1. The probability vector p is calculated for each $\beta \in [0, 1]$, from which the expected burst length is deduced. We observe that $\beta = 0$ implies that $p = (0, 1, 0, \cdots, 0)$, thus the expected burst length is 1 and $\mathcal{E}(\tilde{V}_{min}, p) = \frac{k}{n^2}$. This gives the starting point of the each of the curves for $\mathcal{E}(\tilde{V}_{min}, p)$ for different values of n as k remains fixed.

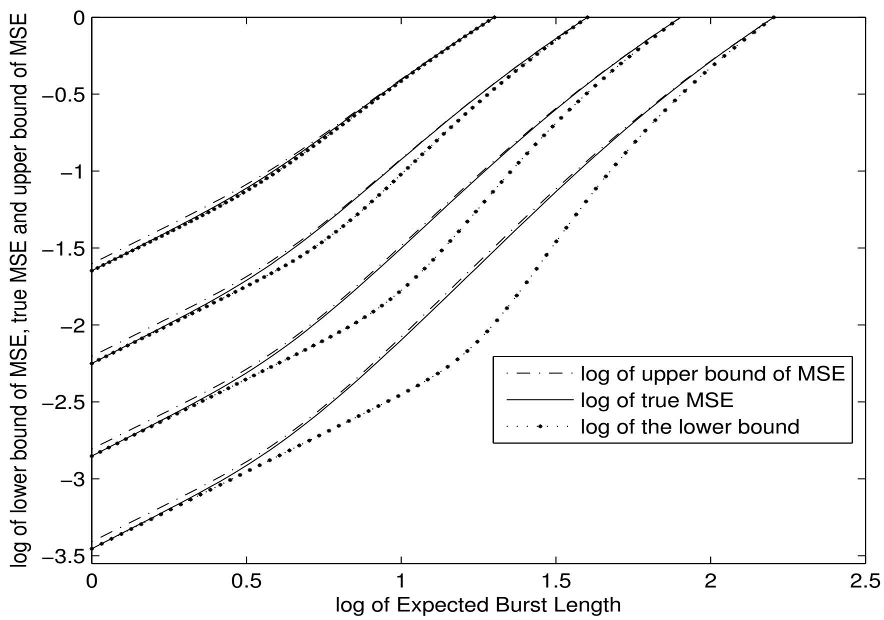

Figure 2.1: Log-log plots of upper and lower bounds and of the mean-square error depending on the expected burst length, for each frame of size $n \in \{20, 40, 80, 160\}$ in a Hilbert space of dimension $k = 9$.

In order to assess the magnitude of deviation from the mean-square reconstruction error, we plot simulated square errors for $m = 4, r = 10$ and $k = 9$ in Figure 2.2, with $10,000$ randomly generated input vectors and bursts.

As the packet length n increases to infinity, the distribution of the burst lengths in a packet converges to a geometric distribution. If the ratio k/n converges to a constant c as n diverges, then the asymptotics of the upper and lower bounds coincide (using Proposition 2 and Theorem 2.3.11). Therefore, the mean-square error satisfies $n\mathcal{E}(\tilde{V}_{min}, p) \to cw_0$. In a log-log plot, this leads to an asymptotically linear graph

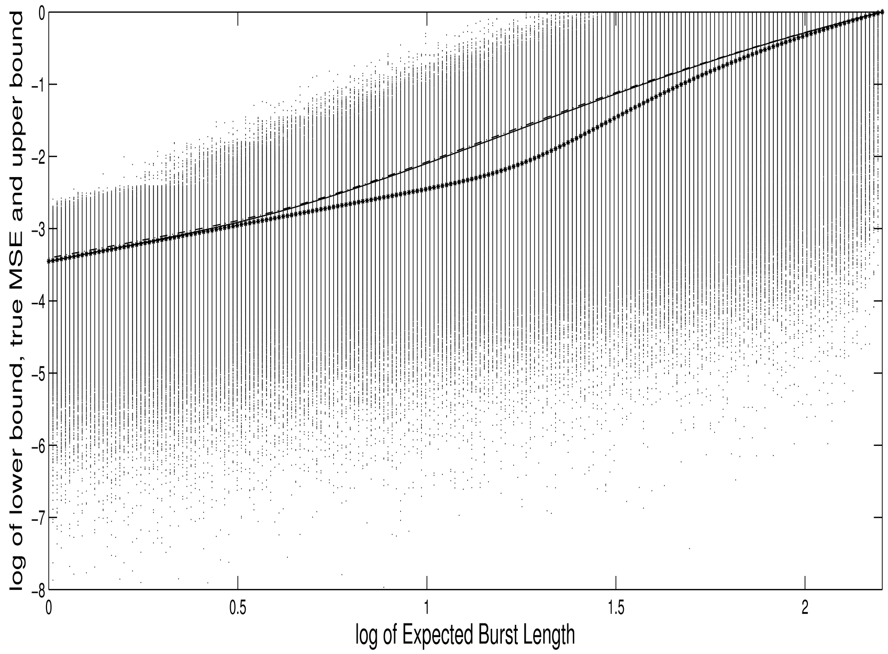

Figure 2.2: Log-log plots of upper and lower bounds and of the mean-square error depending on the expected burst length for a frame of size $n = 160$ in a Hilbert space of dimension $k = 9$, together with square errors resulting from randomly selected input vectors and bursts.

of the mean-square error. To illustrate the effect of the truncation due to the finite packet length, we have plotted the true mean-square reconstruction error and our estimates for fixed values of β and packet length $n = r2^m$ with $m = 4$, $k = r - 1$ and thus $k/n \to 1/2^m = 1/16$ in Figure 2.3.

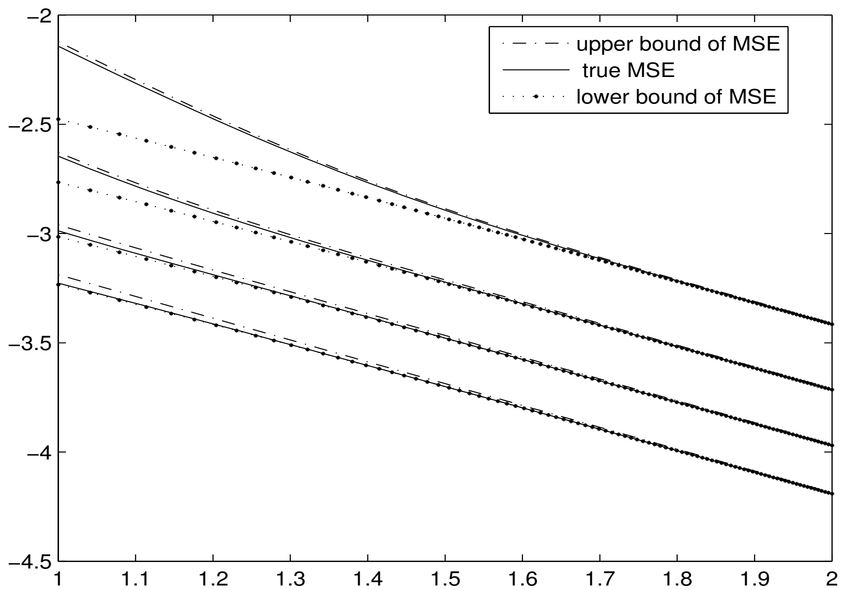

Figure 2.3: Log-log plot of upper and lower bounds and of the mean-square error depending on $r \in \{10, 11, \cdots, 100\}$, with corresponding packet lengths $n = 2^4 r$ and dimensions $k = r - 1$, so $k/n \to 1/16$, and different values of the parameter $\beta \in \{0.9, 0.8, .064, .04\}$, which determines the approximately geometric distribution of burst lengths.

CHAPTER 3

Linear Packet Encoding and Erasures

In this chapter, we study another form of structured erasures. In contrast to chapter 2, the structure of the lost data does not consist of a burst, but in an erased subset of coefficients, a packet. We consider a signal as a vector in an D-dimensional real Hilbert space. An encoding map sends this vector to N packets each having m coefficients and thus to a vector in $\mathbb{R}^{Nm}$. The range of this ecoding map is a D-dimensional suspace of $\mathbb{R}^{Nm}$ and the kernel of its adjoint is a subspace of $\mathbb{R}^{Nm}$ of dimension $M = Nm - D$. This encoded vector is transmitted through a noisy channel in the form of N packets of m linear coefficients. During the transmission, a few of the packets are lost or corrupted and thus, the receiver has less intact packets. We assume $D < Nm$ to allow for the possibility that a significant part of the signal may be recovered without the need for re-sending when a few packets are lost, corrupted or impractically delayed in the transmission. We also assume that only k packets of the transmitted vector are lost or corrupted with k being very small compared to N. For the purposes of designing an algorithm to recover the transmitted vector, we use matrix $T$of size $M \times Nm$ with entries $t_{i,j}$, which are independent realizations of a Gaussian random variable with mean 0 and variance $\frac{1}{M}$. Denoting the range of map $T^* : \mathbb{R}^M \to \mathbb{R}^{Nm}$ by R and $Ker(T)$ by K, which is an D-dimensional subspace of $\mathbb{R}^{Nm}$, we observe that $R^\perp = K$. Let $\{a_j : 1 \leq j \leq D\}$ be an orthonormal basis of $K = Ker(T)$. We denote $A = [a_1|a_2|\cdots|a_D]$ and write the rows of A as v_i for $1 \leq i \leq Nm$. Then $F = \{v_i : 1 \leq i \leq Nm\}$ is a Parseval frame for $\mathbb{R}^D$. We use the analysis operator Q for this Parseval frame F as encoder for $\mathbb{R}^D$. Thus, for any $v \in \mathbb{R}^D$, the encoded vector is $Qv \in K \subset \mathbb{R}^{Nm}$. Then we send this encoded vector through a noisy channel and at the receiving end, we receive $Qv + x$, where x is the error during the transmission. If we know x, then we will have the encoded vector Qv. We assume that only k packets of x are non-zero, where k is very small compared to N, the total number of packets. As $K = Ker(T)$, applying T to the received vector gives us $T(Qv + x) = 0 + Tx = Tx \in \mathbb{R}^M$. In order to get the measurements, we then apply T^* and a multiple of a random projection P on $\mathbb{R}^{Nm}$ of rank p.

In order to achieve the goal, we follow the following steps:

- For a fixed $x \in \mathbb{R}^{Nm}$, we show that the value of $\|Tx\|^2$ is concentrated around its expected value $\|x\|^2$ (Lemma 3.2.1 and 3.2.2).

- Let $x \in \mathbb{R}^{Nm}$ be written as $x = (x_1^t, x_2^t, \cdots, x_N^t)$ with $x_i^t \in \mathbb{R}^m$, that is, x is treated as N packets with each having m linear coefficients. Let $S \subset \{1, 2, \cdots, N\}$ be a set of indices with $\#S \leq k$ and X_S be the subspace of $\mathbb{R}^{Nm}$ consisting of all $x \in \mathbb{R}^{Nm}$ with $x_i = 0$ for all $i \notin S$. For a fixed $0 < \delta < 0.4$, $\|Tx\|^2$ is close to $\|x\|^2$ for all $x \in X_S$ with high probability(Lemma 3.2.3).

- For a fixed $0 < \delta < 0.4$ and fixed S, we then show that the length of $\|T^*Tx\|$ is close to $\|x\|$ for all $x \in X_S$ with high probability (Lemma 3.2.4).

- Using the result(Lemma 3.2.5) that for a fixed vector $z \in \mathbb{R}^{Nm}$ and a random projection P of rank p on $\mathbb{R}^{Nm}$, $\sqrt{\frac{Nm}{p}}\|Pz\|$ is close to $\|z\|$ with high probability (Theorem 3.2.6).

- Combining the above steps, we show that $\sqrt{\frac{Nm}{p}}\|PT^*Tx\|$ is close to $\|x\|$ with high probability for a fixed $0 < \delta < 0.4$ and fixed X_S (Theorem 3.2.7).

- We show that for fixed $0 < \delta < 0.2$ and any X_S, $\sqrt{\frac{Nm}{p}}PT^*T$ satisfies weak form of restricted isometry property (Theorem 3.2.8).

- For a fixed $0 < \delta < 0.2$, we provide conditions on N, k, m, M, and p that guarantee that $\sqrt{\frac{Nm}{p}}PT^*T$ satisfies the weak form of the restricted isometry property with probability $\tilde{p}$, which goes to 1 as we allow m, M, and p to go to infinity under certain conditions (Proposition 3).

In the section 3.3, we reconstruct x with only k non-zero packets as a solution to a convex optimization problem. We also show the stability of the solution in the presence of noise of small magnitude.

3.1 Notations and definitions

We generalize the concept of frames for a Hilbert space. We define a fusion frame or frame of subspaces and related concepts for a Hilbert space $\mathcal{H}$.

Definition 3.1.1. *Let J be some index set, and let $\{v_j : v_j > 0 j \in J\}$ be a family of weights. A family of closed subspaces $\{W_j : j \in J\}$ of a Hilbert space $\mathcal{H}$ is a fusion frame or a frame of subspaces with respect to $\{v_j : v_j > 0$ for each $j \in J\}$ for $\mathcal{H}$, if there exist constants $0 < C \leq D < \infty$ such that*

$$C\|x\|^2 \leq \sum_{j \in J} v_j^2 \|P_{W_j} x\|^2 \leq D\|x\|^2 \text{ for all } x \in \mathcal{H}\,, \tag{3.1}$$

where P_{W_j} is the orthogonal projection of $\mathcal{H}$ onto W_j. We call C and D the frame bounds for the fusion frame. The family $\{W_j : j \in J\}$ is called a C-tight fusion frame with respect to $\{v_j : v_j > 0 \quad \forall j \in J\}$, if in (3.1) the constants C and D can be chosen so that $C = D$, a Parseval fusion frame with respect to $\{v_j : v_j > 0 \quad \forall j \in J\}$ provided that $C = D = 1$ and an orthonormal basis of subspaces if $H = \bigoplus_{j \in J} W_j$. Moreover, we call a fusion frame $\{W_j : j \in J\}$ with respect to $\{v_j : v_j > 0 \quad \forall j \in J\}$ v-uniform, if $v := v_i = v_j$ for all $i, j \in J$.

For natural numbers m and N, we define (N, m, M)-reconstruction system for the Hilbert space $H = \mathbb{R}^M$.

Definition 3.1.2. *A family $\mathcal{V} = \{V_j : 1 \leq j \leq N\}$ of linear maps from $\mathbb{R}^m$ to $\mathbb{R}^M$ is called an (N, m, M)-reconstruction system for $\mathbb{R}^M$, if*

$$\mathcal{S}_{\mathcal{V}} := \sum_{j=1}^{N} V_j V_j^*$$

is a positive and invertible operator on $\mathbb{R}^M$*. This is called an* (N, m, M)*-protocol if*

$$\sum_{j=1}^{N} V_j V_j^* = I_M.$$

If W_j is the range of V_j, then the analysis operator of this system $\mathcal{V}$ is defined by

$$T_{\mathcal{V}} : \mathbb{R}^M \to \oplus_{j=1}^{N} W_j = W$$

via

$$T_{\mathcal{V}} y = (V_1^* y, V_2^* y, \cdots, V_N^* y)\,.$$

Its adjoint $T_{\mathcal{V}}^*$ is called the synthesis operator of the system $\mathcal{V}$, and it satisfies

$$T_{\mathcal{V}}^* : W = \oplus_{j=1}^{N} W_j \to \mathbb{R}^M$$

with

$$T_{\mathcal{V}}^*(x_1, x_2, \cdots, x_N) = \sum_{j=1}^{N} V_j x_j\,.$$

3.2 Restricted isometry property of a random matrix

In this section we construct a matrix which allows the recovery of vectors that are non-zero in only a few packets. If the signal is encoded in such a way that it is annihilated by this matrix, then recovery of sparse errors and recovery of the signal are equivalent. As in other works on the recovery of sparse vectors, we also use measure concentration as an essential tool. We consider an $M \times Nm$ random matrix

$T = (t_{i,j})$ whose entries $t_{i,j}$ are independent realizations of Gaussian random variable with mean 0 and variance $\frac{1}{M}$, i. e. ,

$$t_{i,j} \sim \mathcal{N}\left(0, \frac{1}{M}\right).$$

We write $T = [T_1, T_2, \cdots, T_{Nm}]$, where T_j is the jth column of T. By concatenating the first m columns of T, we form a matrix $V_1 \in \mathbb{R}^{M\times m}$. Similarly, we denote the matrix formed by the next m columns of T by V_2 and continue in this fashion to get $V_3, \cdots, V_N$. Thus we obtain a family $\mathcal{V} = \{V_j : 1 \leq j \leq N\}$ of linear maps from $\mathbb{R}^m$ to $\mathbb{R}^M$. This family $\mathcal{V} = \{V_j : 1 \leq j \leq N\}$ is an (N, m, M)- reconstruction system for $\mathbb{R}^M$ with T^* as the analysis operator. We first show that the expected value of the random variable $\|Tx\|^2$ for a fixed x is $\|x\|^2$ and it is highly concentrated about $\|x\|^2$ as stated in [2].

Lemma 3.2.1. *[2] Let ν be the probability measure on*

$$\mathcal{E} = \left\{T = (t_{i,j})_{\substack{1\leq i\leq M,\\ 1\leq j\leq Nm}} : t_{i,j}\text{'s are independent realizations of } \mathcal{N}\left(0, \frac{1}{M}\right)\right\}$$

induced by the probability measures on $t_{i,j}$'s. Then for any fixed $x \in \mathbb{R}^{Nm}$ with $x = (\xi_1, \xi_2, \cdots, \xi_{Nm})^t$, we have

$$\mathbb{E}[\|Tx\|^2] = \|x\|^2,$$

where the expectation is computed with respect to the measure ν.

Proof. We have

$$\begin{aligned}
\|Tx\|^2 &= \sum_{i=1}^{M}\left(\sum_{j=1}^{Nm} t_{i,j}\xi_j\right)^2 \\
&= \sum_{i=1}^{M}\left[\sum_{j=1}^{Nm}(t_{i,j})^2\xi_j^2 + \sum_{\substack{j,l=1 \\ j\neq l}}^{Nm} t_{i,j}\xi_j t_{i,l}\xi_l\right]
\end{aligned}$$

As $\mathbb{E}[t_{i,j}] = 0, \mathbb{E}[(t_{i,j})^2] = \dfrac{1}{M}$ and $\mathbb{E}[t_{i,l}t_{i,j}] = \mathbb{E}[t_{i,l}]\mathbb{E}[t_{i,j}]$ whenever $j \neq l$ (by independence), we have

$$\begin{aligned}
\mathbb{E}\left[\|Tx\|^2\right] &= \sum_{i=1}^{M}\left[\sum_{j=1}^{Nm}\mathbb{E}\left[(t_{i,j})^2\xi_j^2\right] + \sum_{\substack{j,l=1 \\ j\neq l}}^{Nm}\mathbb{E}\left[t_{i,j}\xi_j t_{i,l}\xi_l\right]\right] \\
&= \sum_{i=1}^{M}\left[\sum_{j=1}^{Nm}\xi_j^2\mathbb{E}\left[(t_{i,j})^2\right] + \sum_{\substack{j,l=1 \\ j\neq l}}^{Nm}\xi_j\xi_l\mathbb{E}\left[t_{i,j}\right]\mathbb{E}\left[t_{i,l}\right]\right] \\
&= \sum_{i=1}^{M}\left[\frac{1}{M}\sum_{j=1}^{Nm}\xi_j^2 + 0\right] \\
&= \|x\|^2 .
\end{aligned}$$

□

Next, we need that for any fixed vector $x \in \mathbb{R}^{Nm}$, the random variable $T \mapsto \|Tx\|^2$ is strongly concentrated about its expected value $\|x\|^2$ [1, 2].

Lemma 3.2.2. *[1, 2] Let ν be the probability measure on*

$$\mathcal{E} = \left\{T = (t_{i,j})_{\substack{1\leq i\leq M, \\ 1\leq j\leq Nm}} : t_{i,j}\text{'s are independent realizations of } \mathcal{N}\left(0, \frac{1}{M}\right)\right\}$$

induced by the probability measures on $t_{i,j}$'s. For any $x \in \mathbb{R}^{Nm}$ and fixed $0 < \epsilon < 1$, we have

$$\nu\left\{T \in \mathcal{E} : |\|Tx\|^2 - \|x\|^2| \geq \epsilon\|x\|^2\right\} \leq 2e^{-Mc_0(\epsilon)} \tag{3.2}$$

where $c_0(\epsilon) = \dfrac{\epsilon^2}{4} - \dfrac{\epsilon^3}{6}$.

Proof. We observe that $\|Tx\|^2 = \sum_{i=1}^{M}(Tx)_i^2$. For each $1 \leq i \leq M$, we consider the random variable

$$Y_i = (Tx)_i = \sum_{j=1}^{Nm} t_{i,j}\xi_j.$$

The expected value of Y_i is given by

$$\mathbb{E}[Y_i] = \sum_{j=1}^{Nm} \mathbb{E}[t_{i,j}]\xi_j = 0.$$

Using the independence of $t_{i,j}$ and $t_{i,l}$ whenever $j \neq l$, we compute the second moment of Y_i as follows:

$$\mathbb{E}[Y_i^2] = \sum_{j,l=1}^{Nm} E[t_{i,j}t_{i,l}]\xi_j\xi_l = \sum_{\substack{j,l=1 \\ j\neq l}}^{Nm} \mathbb{E}[t_{i,j}]\mathbb{E}[t_{i,l}]\xi_j\xi_l + \sum_{j=1}^{Nm} \mathbb{E}[t_{i,j}^2]\xi_j^2 = \frac{1}{M}\|x\|^2.$$

This shows that $Z_i = \dfrac{\sqrt{M}}{\|x\|}Y_i$ is distributed as $\mathcal{N}(0,1)$ and Z_i's are independent. Thus, the random variable $\sum_{i=1}^{M} Z_i^2$ is χ_M^2 chi-squared distributed with M degrees of freedom. Now, we bound the failure probability of one side. From the relation $\dfrac{\|x\|^2}{M}\sum_{i=1}^{M} Z_i^2 = \|Tx\|^2$, we obtain

$$\begin{aligned}\nu\left\{T \in \mathcal{E} : \|Tx\|^2 \geq (1+\epsilon)\|x\|^2\right\} &= \nu\left\{T \in \mathcal{E} : \sum_{i=1}^{M} Z_i^2 \geq (1+\epsilon)M\right\} \\ &= \nu\left\{T \in \mathcal{E} : \chi_M^2 \geq (1+\epsilon)M\right\}.\end{aligned}$$

We appeal to a concentration result below:

$$\nu\left\{T \in \mathcal{E} : \chi_M^2 \geq (1+\epsilon)M\right\} \leq e^{-\frac{M}{4}(\epsilon^2-\epsilon^3)}. \tag{3.3}$$

To prove the above relation 3.3, we observe that Z_i's are i.i.d. $\mathcal{N}(0,1)$ random variables. By Markov's inequality, we have

$$\begin{aligned}
\nu\left\{T \in \mathcal{E} : \chi_M^2 \geq (1+\epsilon)M\right\} &= \nu\left\{T \in \mathcal{E} : \sum_{i=1}^{M} Z_i^2 \geq (1+\epsilon)M\right\} \\
&= \nu\left\{T \in \mathcal{E} : e^{\lambda \sum_{i=1}^{M} Z_i^2} \geq e^{\lambda(1+\epsilon)M}\right\} \\
&\leq \frac{\mathbb{E}\left[e^{\lambda \sum_{i=1}^{M} Z_i^2}\right]}{e^{\lambda(1+\epsilon)M}} \\
&= \frac{\left(\mathbb{E}\left[e^{\lambda Z_1^2}\right]\right)^M}{e^{\lambda(1+\epsilon)M}} \\
&= e^{-\lambda(1+\epsilon)M}\left(\frac{1}{1-2\lambda}\right)^{M/2},
\end{aligned}$$

where the last step follows from evaluating the expectation of $e^{\lambda Z_1^2}$, which holds for $0 < \lambda < 1/2$. Choosing $\lambda = \dfrac{\epsilon}{2(1+\epsilon)}$ minimizes the above expression and thus, we have

$$\nu\left\{T \in \mathcal{E} : \chi_M^2 \geq (1+\epsilon)M\right\} \leq \left[(1+\epsilon)e^{-\epsilon}\right]^{M/2} \leq e^{-\frac{M}{4}(\epsilon^2-\epsilon^3)},$$

using the upper bound $1+\epsilon \leq e^{\epsilon-(\epsilon^2-\epsilon^3)/2}$. Similarly, we have

$$\nu\left\{T \in \mathcal{E} : \|Tx\|^2 \leq (1-\epsilon)\|x\|^2\right\} = \nu\left\{T \in \mathcal{E} : \chi_M^2 \leq (1-\epsilon)M\right\} \leq e^{-\frac{M}{4}(\epsilon^2-\epsilon^3)}. \tag{3.4}$$

From the definition of Z_i's and combining (3.3) and (3.4), we conclude that

$$\nu\left\{T \in \mathcal{E} : \|Tx\|^2 \leq (1-\epsilon)\|x\|^2 \text{ or } \|Tx\|^2 \geq (1+\epsilon)\|x\|^2\right\} \leq 2e^{-\frac{M}{4}(\epsilon^2-\epsilon^3)},$$

which completes the proof.

□

For $x \in \mathbb{R}^{Nm}$, we write $x = (x_1, x_2, \cdots, x_N)^t$ with $x_j^t \in \mathbb{R}^m$. Let us restrict our attention to the acton of random T on a fixed km-dimensional subspace. For any set of indices $S \subset \{1, 2, \cdots, N\}$ with $\#S \leq k$, we denote by X_S the set of all vectors $x \in \mathbb{R}^{Nm}$ satisfying $x_j = 0$ for all $j \notin S$. The strategy of the proof of the following lemma is similar to that used by Baraniuk et al. in [2].

Lemma 3.2.3. *Let ν be the probability measure on $\mathcal{E}$. Then, for any fixed $S \subset \{1, 2, \cdots, N\}$ with $\#S = k$, and any $0 < \delta < 0.4$, we have*

$$\left(1 - \frac{4\delta}{7}\right) \|x\|^2 \leq \|Tx\|^2 \leq \left(1 + \frac{4\delta}{7}\right) \|x\|^2 \text{ for all } x \in X_S \tag{3.5}$$

with

$$\nu\{T \in \mathcal{E} : (3.5) \text{ holds } \} \geq 1 - 2\left(1 + \frac{16}{\delta}\right)^{km} e^{-c_0(\delta/4)M}. \tag{3.6}$$

Proof. By scaling, we only need to show that (3.5) holds for all $\|x\| = 1, x \in X_S$. By a volume inequality for sphere packings [2], we know that for given $\delta > 0$, there exits a set $B \subset \Sigma_S = \{x \in X_S : \|x\| = 1\}$ such that

1. for every $x \in \Sigma_S$, there exists $y \in B$ such that $\|x - y\| \leq \delta/8$, so B is a $\delta/8$-net in Σ_S

2. the cardinality of B is $\leq \left(1 + \dfrac{16}{\delta}\right)^{km}$.

Using the union bound along with the relation (3.2), with probability exceeding the right hand side of (3.6), we have

$$\left(1-\frac{\delta}{4}\right)\|y\|^2 \le \|Ty\|^2 \le \left(1+\frac{\delta}{4}\right)\|y\|^2 \text{ for all } y \in B. \tag{3.7}$$

Now, we define a to be the smallest number such that

$$\|Tx\|^2 \le (1+a)\,\|x\|^2 \text{ for all } x \in X_S \text{ with } \|x\| \le 1. \tag{3.8}$$

For any $x \in \Sigma_S$, there exists $y \in B$ such that $\|x-y\| \le \delta/8$. In this case, we have

$$\|Ty\|^2\|T(x-y)\|^2 \le (1+a)\|y\|^2(1+a)\|x-y\|^2 \le (1+a)^2\left(\frac{\delta}{8}\right)^2 \tag{3.9}$$

and this implies that

$$\begin{aligned}\|Tx\|^2 &\le \|Ty\|^2 + \|T(x-y)\|^2 + 2\|Ty\|\|T(x-y)\| \\ &\le (1+\delta/4) + (1+a)\left(\frac{\delta}{8}\right)^2 + 2(1+a)\left(\frac{\delta}{8}\right).\end{aligned}$$

Since, by definition, a is the smallest number for which (3.8) holds, we obtain

$$\begin{aligned} & (1+a) \le (1+\delta/4) + (1+a)\left(\frac{\delta}{8}\right)^2 + 2(1+a)\left(\frac{\delta}{8}\right) \\ \Rightarrow\quad & a\left(1-\frac{\delta^2}{64}-\frac{\delta}{4}\right) \le \frac{\delta}{4}+\frac{\delta^2}{64} \\ \Rightarrow\quad & a \le \frac{32\delta+\delta^2}{64-16\delta-\delta^2}.\end{aligned}$$

In order to show the upper inequality in (3.5), we need to prove that for $0<\delta<0.4$,

$$\frac{32\delta+\delta^2}{64-16\delta-\delta^2} \le \frac{4\delta}{7}.$$

We consider

$$\frac{4\delta}{7} - \frac{32\delta+\delta^2}{64-16\delta-\delta^2} = \frac{\delta(32-71\delta-4\delta^2)}{7(64-16\delta-\delta^2)}. \tag{3.10}$$

As $32 - 71\delta - 4\delta^2$ and $64 - 16\delta - \delta^2$ are decreasing and positive functions of δ on the interval $(0, 0.4)$, using the relation (3.10), we conclude that the upper inequality in (3.5) holds.

Also, from triangle inequality, we have

$$\begin{aligned}\|Tx\|^2 &\geq (\|Ty\| - \|T(x-y)\|)^2 \\ &\geq \|Ty\|^2 - 2\|Ty\|\|T(x-y)\|\end{aligned}$$

Using relations (3.7), (3.9) and the upper inequality in (3.5), we obtain

$$\|Tx\|^2 \geq (1 - \delta/4) - 2\left(1 + \frac{4\delta}{7}\right)\frac{\delta}{8} = 1 - \frac{\delta}{2} - \frac{\delta^2}{7} \tag{3.11}$$

For $0 < \delta < 0.4$, we have

$$\frac{4\delta}{7} - \frac{\delta}{2} - \frac{\delta^2}{7} = \frac{\delta(1-2\delta)}{14} > 0 \tag{3.12}$$

Using (3.11) and (3.12), we conclude that the lower inequality in (3.5) holds.

□

We define map $\theta : \mathcal{E} \to \mathbb{M}_N$ given by $T \mapsto T^*T$. This map induces a probability measure ν_θ on $\mathcal{F} = \{\theta(T) = T^*T : T \in \mathcal{E}\}$. The following lemma is a consequence of the previous lemma.

Lemma 3.2.4. *For any S with $\#S = k$ and any $0 < \delta < 0.4$, we have*

$$\left(1 - \frac{4\delta}{7}\right)\|x\| \leq \|T^*Tx\| \leq \left(1 + \frac{4\delta}{7}\right)\|x\| \text{ for all } x \in X_S \tag{3.13}$$

with

$$\nu_\theta\left\{T^*T \in \mathcal{F} : (3.13) \text{ holds }\right\} \geq 1 - 2\left(1 + \frac{16}{\delta}\right)^{km} e^{-c_0(\delta/4)M}. \tag{3.14}$$

Proof. We observe that the relation (3.5) is equivalent to

$$\begin{aligned}&\langle(1-4\delta/7)Ix, x\rangle \leq \langle T^*Tx, x\rangle \leq \langle(1+4\delta/7)Ix, x\rangle \text{ for all } x \in X_S\\ \Rightarrow&(1-4\delta/7)I_{X_S} \leq T^*T_{X_S} \leq (1+4\delta/7)I_{X_S}\,.\end{aligned}$$

This shows that $(1-4\delta/7)$ and $(1+4\delta/7)$ are the smallest and largest eigenvalues of $T^*T_{X_S}$. Therefore, $(1-4\delta/7)^2$ and $(1+4\delta/7)^2$ are the smallest and largest eigenvalues of $(T^*T_{X_S})^2$. This implies that

$$\begin{aligned}&(1-4\delta/7)^2 I_{X_S} \leq (T^*T_{X_S})^2 \leq (1+4\delta/7)^2 I_{X_S}\\ \Rightarrow&\langle(1-4\delta/7)^2 Ix, x\rangle \leq \langle T^*TT^*Tx, x\rangle \leq \langle(1+4\delta/7)^2 Ix, x\rangle \text{ for all } x \in X_S\\ \Rightarrow&(1-4\delta/7)^2\|x\|^2 \leq \|T^*Tx\|^2 \leq (1+4\delta/7)^2\|x\|^2 \text{ for all } x \in X_S\\ \Rightarrow&(1-4\delta/7)\|x\| \leq \|T^*Tx\| \leq (1+4\delta/7)\|x\| \text{ for all } x \in X_S\,.\end{aligned}$$

As ν_θ is induced by the map θ and relation (3.5) holds with the same probability as given in (3.6), lemma is proved. □

Now, for the random projection, we have the following lemma:

Lemma 3.2.5. *[3, 46] Let $z \in \mathbb{R}^{Nm}$ be a non-zero vector and let $\mu_{Nm,p}$ be the invariant probability measure on the Grassmannian $G_p(\mathbb{R}^{Nm})$ of p-dimensional subspaces in $\mathbb{R}^{Nm}$. For $V \in G_p(\mathbb{R}^{Nm})$, let z_V be the orthogonal projection of z onto V. Then, for any $0 < \epsilon < 1$,*

$$\begin{aligned}&\mu_{Nm,p}\left\{V \in G_p(\mathbb{R}^{Nm}) : \sqrt{\frac{Nm}{p}}\|z_V\| \geq (1-\epsilon)^{-1}\|z\|\right\} \leq e^{-\epsilon^2 p/4} + e^{-\epsilon^2 Nm/4} \text{ and}\\ &\mu_{Nm,p}\left\{V \in G_p(\mathbb{R}^{Nm}) : \sqrt{\frac{Nm}{p}}\|z_V\| \leq (1-\epsilon)\|z\|\right\} \leq e^{-\epsilon^2 p/4} + e^{-\epsilon^2 Nm/4}\,.\end{aligned}$$

Let $\mathcal{O}(Nm,p)$ denote the set of orthogonal projections of rank p from $\mathbb{R}^{Nm}$. As $\mu_{Nm,p}$ is the probability measure on the Grassmannian $G_p(\mathbb{R}^{Nm})$, the map $\psi : G_p(\mathbb{R}^{Nm}) \to \mathcal{O}(Nm,p)$ given by $V \mapsto P_V$, where P_V is the orthogonal projection of $\mathbb{R}^{Nm}$ onto the subspace V, induces a probability measure μ_ψ on $\mathcal{O}(Nm,p)$. Thus, using lemma 3.2.5, we have that for any fixed $z \in \mathbb{R}^{Nm}$,

$$\mu_\psi \left\{ P \in \mathcal{O}(Nm,p) : (1-\epsilon)\|z\| \le \sqrt{\frac{Nm}{p}}\|Pz\| \le (1-\epsilon)^{-1}\|z\| \right\} \ge 1 - 2\left(e^{-\epsilon^2 p/4} + e^{-\epsilon^2 Nm/4}\right). \tag{3.15}$$

Now, we consider the product measure $\lambda = \mu_\psi \times \nu_\theta$ on $\mathcal{O}(Nm,p) \times \mathcal{F}$. The map $(P, T^*T) \mapsto PT^*T$ induces a probability measure σ on

$$\mathcal{G} = \{PT^*T : P \in \mathcal{O}(Nm,p), T^*T \in \mathcal{F}\}.$$

Theorem 3.2.6. *Let $T \in \mathcal{E}$ and $P \in \mathcal{O}(Nm,p)$. For any set $S \subset \{1,2,\cdots,N\}$ of indices, let X_S denote the set of all vectors $x = (x_1, x_2, \cdots, x_N)^t$ with $x_j^t \in \mathbb{R}^m$ in $\mathbb{R}^{Nm}$ satisfying $x_j = 0$ for all j outside of S. Then, for any S with $\#S = k$, a fixed vector $x \in X_S$ and for any $0 < \delta < 0.4$, we have*

$$\sigma\left(\left\{PT^*T : (1-\delta/8)(1-4\delta/7)\|x\| \le \sqrt{\frac{Nm}{p}}\|PT^*Tx\| \le \frac{(1+4\delta/7)}{(1-\delta/8)}\|x\|\right\}\right) \ge \left(1 - 2\left(1+\frac{16}{\delta}\right)^{km} e^{-c_0(\delta/4)M}\right)\left(1 - 2(e^{-\delta^2 p/256} + e^{-\delta^2 Nm/256})\right) \tag{3.16}$$

Proof. For a fixed vector $x \in X_S$,

$$(1-4\delta/7)(1-\delta/8)\|x\| \le \sqrt{\frac{Nm}{p}}\|PT^*Tx\| \le \frac{(1+4\delta/7)}{(1-\delta/8)}\|x\|$$

is true whenever both of the following relations hold:

$$(1-\delta/8)\|T^*Tx\| \leq \sqrt{\frac{Nm}{p}}\|PT^*Tx\| \leq \frac{1}{(1-\delta/8)}\|T^*Tx\|$$

and

$$(1-4\delta/7)\|x\| \leq \|T^*Tx\| \leq (1+4\delta/7)\|x\|\,.$$

Thus, for a fixed $x \in X_S$, we have

$$\begin{aligned}
&\sigma\left(\left\{PT^*T : (1-\delta/8)(1-4\delta/7)\|x\| \leq \sqrt{\frac{Nm}{p}}\|PT^*Tx\| \leq \frac{(1+4\delta/7)}{(1-\delta/8)}\|x\|\right\}\right)\\
&\geq \sigma\left(\left\{PT^*T : (1-\delta/8)\|T^*Tx\| \leq \sqrt{\frac{Nm}{p}}\|PT^*Tx\| \leq (1-\delta/8)^{-1}\|T^*Tx\| \text{ and}\right.\right.\\
&\qquad (1-4\delta/7)\|x\| \leq \|T^*Tx\| \leq (1+4\delta/7)\|x\|\})\\
&= \mu_\psi\left(\left\{P \in \mathcal{O}(Nm,p) : (1-\delta/8)\|z\| \leq \sqrt{\frac{Nm}{p}}\|Pz\| \leq (1-\delta/8)^{-1}\|z\|\right\}\right)\\
&\qquad \times \nu_\phi\left(\{T^*T \in \mathcal{F} : (1-4\delta/7)\|x\| \leq \|T^*Tx\| \leq (1+4\delta/7)\|x\|\}\right).
\end{aligned}$$

Using lemma 3.2.4 and relation (3.15), we obtain that the RHS of the above inequality is

$$\geq \left(1-2\left(1+\frac{16}{\delta}\right)^{km} e^{-c_0(\delta/4)M}\right)\left(1-2(e^{-\delta^2 p/256}+e^{-\delta^2 Nm/256})\right).$$

□

In order to avoid writing complicated expressions, let us denote

$$\alpha(\delta) = 2\left(1+\frac{8}{\delta}\right)^{km} e^{-c_0(\delta/2)M} \tag{3.17}$$

$$\beta(\delta) = 2\left(e^{-\delta^2 p/4}+e^{-\delta^2 Nm/4}\right) \tag{3.18}$$

Using the above notations, the relation (3.16) becomes

$$\sigma\left(\left\{PTT^* : (1-\delta/8)(1-4\delta/7)\|x\| \le \sqrt{\frac{N}{p}}\|PTT^*x\| \le \frac{(1+4\delta/7)}{(1-\delta/8)}\|x\| \text{ does not hold}\right\}\right)$$
$$\le \alpha(\delta/2)+\beta(\delta/8)\,. \tag{3.19}$$

Theorem 3.2.7. *Let T, P and X_S be as above. For any $0 < \delta < 0.4$, we have*

$$(1-\delta)\|x\| \le \sqrt{\frac{Nm}{p}}\|PT^*Tx\| \le \frac{(1+4\delta/7)}{(1-\delta/8)^2}\|x\|, \text{ for all } x \in X_S \tag{3.20}$$

with probability

$$\sigma\{PT^*T : P \in \mathcal{O}(Nm,p), T^*T \in \mathcal{F}\} \ge 1-\left(1+\frac{16}{\delta}\right)^{km}(\alpha(\delta/2)+\beta(\delta/8))\ , \tag{3.21}$$

where $\alpha(\delta)$ and $\beta(\delta)$ are defined by 3.17 and 3.18.

Proof. The idea of the proof is based on the arguments made by Baraniuk et al. in [2]. As PT^*T is linear, it is enough to prove (3.20) in the case when $\|x\| = 1$. By a volume inequality for sphere packings [2], we know that for given $\delta > 0$, there exits a set $B \subset \Sigma_S = \{x \in X_S : \|x\| = 1\}$ such that

1. for every $x \in \Sigma_S$, there exists $y \in B$ such that $\|x - y\| \le \delta/8$, so B is a $\delta/8$-net in Σ_S
2. the cardinality of B is $\le \left(1+\frac{16}{\delta}\right)^{km}$.

Using the union bound along with the relation (3.19), with probability exceeding the right hand side of (3.21), we have

$$(1-\delta/8)(1-4\delta/7)\|y\| \le \sqrt{\frac{Nm}{p}}\|PT^*Ty\| \le \frac{(1+4\delta/7)}{(1-\delta/8)}\|y\| \text{ for all } y \in B\,. \tag{3.22}$$

Define A as the smallest number such that

$$\sqrt{\frac{Nm}{p}}\|PT^*Tx\| \leq A\|x\|, \text{ for all } x \in X_S, \|x\| \leq 1\,. \tag{3.23}$$

We need to show that $A \leq \dfrac{(1+4\delta/7)}{(1-\delta/8)^2}$. For any $x \in X_S$ with $\|x\| = 1$, we can pick a $y \in B$ such that $\|x-y\| \leq \delta/8$. In this case, using (3.22) and (3.23), we have

$$\begin{aligned}\sqrt{\frac{Nm}{p}}\|PT^*Tx\| &\leq \sqrt{\frac{Nm}{p}}\|PT^*Ty\| + \sqrt{\frac{Nm}{p}}\|PT^*T(x-y)\| \\ &\leq \frac{(1+4\delta/7)}{(1-\delta/8)}\|y\| + A\|(x-y)\| \\ &\leq \frac{(1+4\delta/7)}{(1-\delta/8)} + A\frac{\delta}{8}\end{aligned}$$

By definition of A, we obtain

$$\begin{aligned}A &\leq \frac{(1+4\delta/7)}{(1-\delta/8)} + A\frac{\delta}{8} \\ \Rightarrow A &\leq \frac{(1+4\delta/7)}{(1-\delta/8)^2}\end{aligned}$$

Thus, the upper inequality in (3.20) holds. For the lower inequality, we have

$$\begin{aligned}\sqrt{\frac{Nm}{p}}\|PT^*Tx\| &\geq \sqrt{\frac{Nm}{p}}\|PT^*Ty\| - \sqrt{\frac{Nm}{p}}\|PT^*T(x-y)\| \\ &\geq (1-\delta/8)(1-4\delta/7)\|y\| - \frac{(1+4\delta/7)}{(1-\delta/8)^2}\|(x-y)\| \\ &\geq (1-\delta/8)(1-4\delta/7) - \frac{(1+4\delta/7)}{(1-\delta/8)^2}\frac{\delta}{8} \\ &= \frac{3584 - 3840\delta + 680\delta^2 - 103\delta^3 + 4\delta^4}{56(8-\delta)^2}\,.\end{aligned} \tag{3.24}$$

For $0 < \delta < 0.4$, we have

$$\begin{aligned}&\frac{3584 - 3840\delta + 680\delta^2 - 103\delta^3 + 4\delta^4}{56(8-\delta)^2} - (1-\delta) \\ &= \frac{\delta\,(640 - 272\delta - 47\delta^2 + 4\delta^3)}{56(8-\delta)^2} > 0\end{aligned} \tag{3.25}$$

Thus, from (3.24) and (3.25), we conclude that the upper inequality in (3.20) holds.

□

Theorem 3.2.8. *Let T, P and X_S be as above. For any $0 < \delta < 0.2$, we have*

$$c\|x\|^2 \leq \frac{Nm}{p}\|PT^*Tx\|^2 \leq d\|x\|^2 \text{ for all } x \in X_S, \tag{3.26}$$

where $c \geq 1 - (\sqrt{2}-1)$ and $d \leq 1 + (\sqrt{2}-1)$, with probability

$$\geq 1 - \left(1 + \frac{16}{\delta}\right)^{km} \left(\alpha(\delta/2) + \beta(\delta/8)\right), \tag{3.27}$$

where $\alpha(\delta)$ and $\beta(\delta)$ are defined by 3.17 and 3.18.

Proof. From theorem 3.2.7, it is enough to prove that

$$\frac{(1+4\delta/7)}{(1-\delta/8)^2} \leq 2^{1/4}$$

and $(1-\delta)^2 \leq 2 - \sqrt{2}$ for all $0 < \delta < 0.2$.

We have

$$\begin{aligned} &\frac{(1+4\delta/7)}{(1-\delta/8)^2} \leq 2^{1/4} \\ \Rightarrow &1 - 2^{1/4} + \frac{4\delta}{7} + \frac{\delta}{(2)2^{3/4}} - \frac{\delta^2}{(32)2^{3/4}} \leq 0 \end{aligned}$$

The roots of this quadratic equation are $\frac{4}{7}2^{3/4}\left(16+(7)2^{1/4}\right) - \frac{4}{7}2^{3/4}\sqrt{256+(273)2^{1/4}}$ and $\frac{4}{7}2^{3/4}\left(16+(7)2^{1/4}\right) + \frac{4}{7}2^{3/4}\sqrt{256+(273)2^{1/4}}$. Thus the above inequality holds for all $0 < \delta < \frac{4}{7}2^{3/4}\left(16+(7)2^{1/4}\right) - \frac{4}{7}2^{3/4}\sqrt{256+(273)2^{1/4}}$ and hence for all $0 < \delta < 0.2$.

Also, $(1-\delta)^2$ is decreasing on the interval $\delta \in (0, 0.2)$ and for $\delta = 0.2, (1-\delta)^2 \geq 2 - \sqrt{2}$. Hence the result. □

We know that for each subspace X_S with $S \subset \{1, 2, \cdots, N\}, |S| = k$, the matrix $\sqrt{\frac{Nm}{p}} PT^*T$ fails to satisfy (3.26) with probability

$$\leq \left(1 + \frac{16}{\delta}\right)^{km} (\alpha(\delta/2) + \beta(\delta/8)) .$$

There are $\binom{N}{k} \leq (eN/k)^k$ such subspaces. Hence, for any k-sparse x, (3.26) will fail to hold with probability

$$\leq \left(\frac{eN}{k}\right)^k \left(1 + \frac{16}{\delta}\right)^{km} \left(2\left(1 + \frac{16}{\delta}\right)^{km} e^{-c_0(\delta/4)M} + 2\left(e^{-\delta^2 p/256} + e^{-\delta^2 Nm/256}\right)\right) \tag{3.28}$$

$$= 2e^{k\ln(\frac{eN}{k})} \left[e^{-Mc_0(\delta/4) + 2km\ln(1+\frac{16}{\delta})} + e^{-p\delta^2/256 + km\ln(1+\frac{16}{\delta})} + e^{-Nm\delta^2/256 + km\ln(1+\frac{16}{\delta})}\right] \tag{3.29}$$

Proposition 3. *For any fixed $0 < \delta < 0.2$, let N and k be fixed with $\frac{N}{k} \geq \frac{256\ln(1+16/\delta)}{\delta^2}$. For any k-sparse x, (3.26) will fail to hold with probability $\leq \tilde{p}$, which goes to zero as M, m, and p go to infinity provided $\frac{m}{M} =$ constant, $c_0(\delta/4) > 2k\frac{m}{M}\ln(1 + \frac{16}{\delta})$ and $\frac{p}{m} =$ constant $\geq \frac{256k\ln(1+16/\delta)}{\delta^2}$.*

Proof. We note that $\phi = \sqrt{\frac{Nm}{p}} PT^*T$ remains a fixed multiple of PT^*T even if m and p tend to infinity. Let us consider each term one by one on the RHS of the above relation. We observe that the expression $k\ln(\frac{eN}{k})$ is a fixed constant and appears in each of the exponents in the above relation. Using the condition on ratio of N and k and from exponent of the third term in above relation, we have

$$\frac{Nm\delta^2}{256} > km\ln(1 + \frac{16}{\delta}).$$

This implies that the third term goes to zero as $m \to \infty$. Similarly, using the conditions on the ratio of p and m, we see the second term also goes to 0 as $m \to \infty$.

Now, we consider the exponent of the first term: Using the condition that $c_0(\delta/4) > 2k\frac{m}{M}\ln(1+\frac{16}{\delta})$, we conclude that

$$-Mc_0(\delta/4) + k\ln(\frac{eN}{k}) + 2km\ln(1+\frac{16}{\delta}) = k\ln(\frac{eN}{k}) + M\left[-c_0(\delta/4) + 2k\frac{m}{M}\ln(1+\frac{16}{\delta})\right]$$

goes to $-\infty$ as M and m go to ∞ with $\frac{M}{m}$. This shows that the first term also goes to zero. $\square$

3.3 Recovery in the case of sparse error

Now, suppose we have $\phi = \sqrt{\frac{Nm}{p}}PT^*T$ such that ϕ satisfies (3.26) for all $x \in X_S$ and for all X_S with $|S| \leq k$.

Suppose $W_1, W_2, \cdots, W_N$ are subspaces of a Hilbert space H with $\dim(W_i) = m$ for all $1 \leq i \leq N$. Define W as the direct sum of these subspaces, $W = \oplus_{i=1}^N W_i$. For any $x = (x_1, x_2, \cdots, x_N) \in W, x_i \in W_i$, we define the mixed $l_{q,r}$ norms as

$$\|x\|_{q,r} = \left(\sum_{i=1}^N \|x_i\|_q^r\right)^{1/r} .$$

When the parameter q is omitted, we mean that $q = 2$:

$$\|x\|_r = \left(\sum_{i=1}^N \|x_i\|_2^r\right)^{1/r} .$$

Suppose that we observe

$$y = \phi x\,, \tag{3.30}$$

where $x = (x_1, x_2, \cdots, x_N) \in W$ is a vector we wish to reconstruct, $y \in W$ are available measurements, and ϕ is a known linear operator on W. We also assume that the dimension of the range of W is $p < Nm$. Now, we ask whether it is possible to reconstruct x with good accuracy.

Definition 3.3.1. *For each integer $k = 1, 2, \cdots$, define the weak restricted isometry constant δ_k of an operator ϕ as the smallest number such that*

$$(1 - \delta_k)\|x\|_2^2 \leq \|\phi x\|_2^2 \leq (1 + \delta_k)\|x\|_2^2 \tag{3.31}$$

holds for all k-sparse vectors. A vector $x = (x_1, x_2, \cdots, x_N) \in W$ is said to be k-sparse if the cardinality of $\{i : x_i \neq 0\}$ is $\leq k$.

Let x^* be the solution to

$$\min_{\bar{x} \in W} \|\bar{x}\|_1 \text{ subject to } \phi\bar{x} = y\,. \tag{3.32}$$

For $x = (x_1, x_2, \cdots, x_N) \in W$, we have $(\|x_1\|_2, \|x_2\|_2, \cdots, \|x_N\|_2) \in \mathbb{R}^N$ and we denote this by $\tilde{x}$. We will compare the reconstruction x^* with the *best sparse approximation* one could obtain if one knew exactly the locations and amplitudes of the k-largest entries of $\tilde{x}$. We denote this approximation by $x^{(k)}$, i. e., the vector x with all but the k-largest components set to zero.

We observe

$$y = \phi x + z\,, \tag{3.33}$$

where z is an unknown noise term. In this context, we try to reconstruct x as the solution to the convex optimization problem

$$\min_{\bar{x} \in W} \|\bar{x}\|_1 \text{ subject to } \|y - \phi\bar{x}\|_2 \leq \epsilon\,. \tag{3.34}$$

The following theorem is an extension of a result by [21] and its proof uses the ideas from the proof of the original result presented in this paper by Candes.

Theorem 3.3.2. *Assume that* $\delta_{2k} < \sqrt{2} - 1$ *and* $\|z\|_2 \leq \epsilon$. *Then the solution to (3.34) obeys*

$$\|x^* - x\|_2 \leq C_0 k^{-1/2} \|x - x^{(k)}\|_1 + C_1 \epsilon \tag{3.35}$$

with the some constants C_0 *and* C_1 *given explicitly below.*

First, we prove the following lemma:

Lemma 3.3.3. *For all* $x, x' \in W$ *supported on disjoint subsets* $S, S' \subset \{1, 2, \cdots, N\}$ *with* $|S| \leq k$ *and* $|S'| \leq k'$, *we have*

$$|\langle \phi x, \phi x' \rangle| \leq \delta_{k+k'} \|x\|_2 \|x'\|_2 \,.$$

Proof. The proof of this lemma follows by essentially the same arguments given by Candes[21]. Suppose x and x' are unit vectors with disjoint supports. We have

$$(1 - \delta_{k+k'}) \|x \pm x'\|_2^2 \leq \|\phi(x \pm x')\|_2^2 \leq (1 + \delta_{k+k'}) \|x \pm x'\|_2^2 \,.$$

As $S \cap S' = \emptyset$, $\|x \pm x'\|_2^2 = \|x\|_2^2 + \|x\|_2^2 = 2$. Therefore, we have

$$2(1 - \delta_{k+k'}) \leq \|\phi(x \pm x')\|_2^2 \leq 2(1 + \delta_{k+k'}) \,.$$

Using the parallelogram identity, we have

$$|\langle \phi x, \phi x' \rangle| = \frac{1}{4} \left[\|\phi(x + x')\|_2^2 - \|\phi(x - x')\|_2^2 \right] \leq \delta_{k+k'} \,,$$

which concludes the proof. □

Now, we prove the theorem 3.3.2:

Proof. We observe that

$$\|\phi(x^* - x)\|_2 \leq \|\phi x^* - y\|_2 + \|y - \phi x\|_2 \leq 2\epsilon\,, \tag{3.36}$$

which follows from the triangle inequality and the fact that x is feasible for the problem (3.34). We set $x^* = x+h$ and decompose h into a sum of vectors $h_{S_0}, h_{S_1}, h_{S_2}, \cdots$, each of sparsity at most k. Here, S_0 corresponds to the locations of k largest coefficients of $\tilde{x}$; S_1 corresponds to the locations of the k largest coefficients of $\widetilde{h_{S_0^c}}$; S_2 to the locations of the next k largest coefficients of $\widetilde{h_{S_0^c}}$, and so on.

First, we show that the size of h outside of $S_0 \cup S_1$ is essentially bounded by that of h on $S_0 \cup S_1$.

We note that for each $j \geq 2$,

$$\|h_{S_j}\|_2 \leq k^{1/2}\|h_{S_j}\|_\infty \leq k^{-1/2}\|h_{S_{j-1}}\|_1$$

and thus

$$\sum_{j\geq 2}\|h_{S_j}\|_2 \leq k^{-1/2}(\|h_{S_1}\|_1 + \|h_{S_2}\|_1 + \cdots) \leq k^{-1/2}\|h_{S_0^c}\|_1\,. \tag{3.37}$$

In particular, this gives us the useful estimate

$$\|h_{(S_0\cup S_1)^c}\|_2 = \|\sum_{j\geq 2} h_{S_j}\|_2 \leq \sum_{j\geq 2}\|h_{S_j}\|_2 \leq k^{-1/2}\|h_{S_0^c}\|_1\,. \tag{3.38}$$

Also, we have

$$\begin{aligned}\|x\|_1 &= \sum_{i=1}^{N} \|x_i + h_i\|_2 \\ &= \sum_{i\in S_0} \|x_i + h_i\|_2 + \sum_{i\notin S_0} \|x_i + h_i\|_2 \\ &\geq \sum_{i\in S_0} (\|x_i\|_2 - \|h_i\|_2) + \sum_{i\notin S_0} (\|x_i\|_2 - \|h_i\|_2) \\ &= \|x_{S_0}\|_1 - \|h_{S_0}\|_1 + \|h_{S_0^c}\|_1 - \|x_{S_0^c}\|_1 .\end{aligned}$$

This implies that

$$\begin{aligned}\|h_{S_0^c}\|_1 &\leq \|x\|_1 - \|x_{S_0}\|_1 + \|h_{S_0}\|_1 + \|x_{S_0^c}\|_1 \\ &\leq \|x - x_{S_0}\|_1 + \|h_{S_0}\|_1 + \|x_{S_0^c}\|_1 \\ &= 2\|x_{S_0^c}\|_1 + \|h_{S_0}\|_1 \qquad (3.39)\end{aligned}$$

Applying (3.38), then (3.39) and the Cauchy- Schwarz inequality to bound $\|h_{S_0}\|_1$ by $k^{1/2}\|h_{S_0}\|_2$, we obtain

$$\begin{aligned}\|h_{(S_0\cup S_1)^c}\|_2 &\leq k^{-1/2}\|h_{S_0^c}\|_1 && \text{[using (3.38)]} \\ &\leq k^{-1/2}[2\|x_{S_0^c}\|_1 + \|h_{S_0}\|_1] && \text{[using (3.39)]} \\ &\leq 2k^{-1/2}\|x_{S_0^c}\|_1 + k^{-1/2}k^{1/2}\|h_{S_0}\|_2 && [\text{using } \|h_{S_0}\|_1 \leq k^{1/2}\|h_{S_0}\|_2] \\ &= \|h_{S_0}\|_2 + 2e_0 \qquad \text{with } e_0 = k^{-1/2}\|x - x^{(k)}\|_1 . && (3.40)\end{aligned}$$

In the next step, we bound $\|h_{(S_0\cup S_1)^c}\|_2$. To do this, we note that $\phi h_{(S_0\cup S_1)} = \phi h - \sum_{j\geq 2} \phi h_{S_j}$ and, therefore, we have

$$\|h_{(S_0\cup S_1)}\|_2^2 = \langle \phi h_{(S_0\cup S_1)}, \phi h\rangle - \langle \phi h_{(S_0\cup S_1)}, \sum_{j\geq 2} \phi h_{S_j}\rangle .$$

From (3.36) and the weak restricted isometry property(3.31), we obtain

$$|\langle \phi h_{(S_0\cup S_1)}, \phi h\rangle| \leq \|\phi h_{(S_0\cup S_1)}\|_2 \|\phi h\|_2 \leq 2\epsilon\sqrt{1+\delta_{2k}}\|h_{(S_0\cup S_1)}\|_2 \,.$$

Also, from lemma 3.3.3, we have

$$|\langle \phi h_{S_0}, \phi h_{S_j}\rangle| \leq \delta_{2k}\|h_{S_0}\|_2\|h_{S_j}\|_2$$

and likewise for S_1 in place of S_0. As S_0 and S_1 are disjoint, we have

$$\|h_{S_0}\|_2 + \|h_{S_1}\|_2 \leq \sqrt{2}\|h_{(S_0\cup S_1)}\|_2 \,.$$

Thus, we have

$$(1-\delta_{2k})\|h_{(S_0\cup S_1)}\|_2^2 \leq \|\phi h_{(S_0\cup S_1)}\|_2^2 \leq \|h_{(S_0\cup S_1)}\|_2 (2\epsilon\sqrt{1+\delta_{2k}} + \sqrt{2}\delta_{2k}\sum_{j\geq 2}\|h_{S_j}\|_2) \,.$$

Therefore, using (3.37), we have that

$$\|h_{(S_0\cup S_1)}\|_2 \leq \alpha\epsilon + \rho k^{-1/2}\|h_{S_0^c}\|_1, \alpha = \frac{2\sqrt{1+\delta_{2k}}}{1-\delta_{2k}}, \rho = \frac{\sqrt{2}\delta_{2k}}{1-\delta_{2k}} \,. \tag{3.41}$$

From this inequality and (3.39) , we conclude that

$$\|h_{(S_0\cup S_1)}\|_2 \leq \alpha\epsilon + \rho\|h_{(S_0\cup S_1)}\|_2 + 2\rho e_0$$

$$\Rightarrow \|h_{(S_0\cup S_1)}\|_2 \leq (1-\rho)^{-1}(\alpha\epsilon + 2\rho e_0) \,.$$

Finally, we have

$$\begin{aligned}
\|h\|_2 &\leq \|h_{(S_0\cup S_1)}\|_2 + \|h_{(S_0\cup S_1)^c}\|_2 \\
&\leq 2\|h_{(S_0\cup S_1)}\|_2 + 2e_0 \\
&\leq 2(1-\rho)^{-1}(\alpha\epsilon + (1+\rho)e_0) \\
&= C_0 k^{-1/2}\|x - x^{(k)}\|_1 + C_1\epsilon
\end{aligned}$$

with C_0 and C_1 given by

$$C_0 = 2(1-\rho)^{-1}(1+\rho) \tag{3.42}$$

and

$$C_1 = 2(1-\rho)^{-1}\alpha\,, \tag{3.43}$$

where ρ and α have explicit form in 3.41 □

Now, we combine the results we have shown above to conclude the following theorem:

Theorem 3.3.4. *For any fixed $0 < \delta < 0.2$, let N and k be fixed with $\frac{N}{k} \geq \frac{256\ln(1+16/\delta)}{\delta^2}$. For any $k/2$-sparse x, let $y = \phi(x) + z$, where z satisfies $\|z\| \leq \epsilon$, then the solution to*

$$\min_{\bar{x}\in W}\|\bar{x}\|_1 \text{ subject to } \|y-\phi\bar{x}\|_2 \leq \epsilon\,. \tag{3.44}$$

will fail to obey

$$\|x^* - x\|_2 \leq C_0(k/2)^{-1/2}\|x - x^{(k/2)}\|_1 + C_1\epsilon \tag{3.45}$$

with probability $\leq \tilde{p}$, which goes to zero as $M, m,$ and p go to infinity provided $\frac{m}{M} = constant, c_0(\delta/4) > 2k\frac{m}{M}\ln(1+\frac{16}{\delta})$ and $\frac{p}{m} = constant \geq \frac{256k\ln(1+16/\delta)}{\delta^2}$, where the constants C_0 and C_1 are given explicitly above in 3.42 and 3.43.

Proof. Using Theorem 3.2.8 and relation 3.29, we obtain that ϕ fails to satisfy the weak restricted isometry property 3.31 with probability

$$\leq \tilde{p} = 2e^{k\ln(\frac{eN}{k})}\left[e^{-Mc_0(\delta/4)+2km\ln(1+\frac{16}{\delta})} + e^{-p\delta^2/256+km\ln(1+\frac{16}{\delta})} + e^{-Nm\delta^2/256+km\ln(1+\frac{16}{\delta})}\right].$$

From Theorem 3.3.2, we conclude that relation 3.34 will fail to obey 3.45 with probability $\leq \tilde{p}$. Finally, from proposition 3, we conclude that under the given condition in the above theorem, $\tilde{p}$ goes to zero as m, M, and p go to infinity.

□

If there is no noise and x is $k/2$-sparse, the following corollary shows that the recovery via l_1-minimization precisely gives us x.

Corollary 3.3.5. *For any fixed* $0 < \delta < 0.2$*, let* N *and* k *be fixed with* $\frac{N}{k} \geq \frac{256\ln(1+16/\delta)}{\delta^2}$*. For any* $k/2$*-sparse* x*, let* $y = \phi(x) + z$*, where* z *satisfies* $\|z\| \leq \epsilon$*, then the solution to*

$$\min_{\bar{x}\in W} \|\bar{x}\|_1 \text{ subject to } \phi\bar{x} = y$$

will fail to obey

$$\|x^* - x\|_1 \leq C_0 \|x - x^{(k/2)}\|_1 \tag{3.46}$$

and

$$\|x^* - x\|_2 \leq C_0 (k/2)^{-1/2} \|x - x^{(k/2)}\|_1 \tag{3.47}$$

with probability $\leq \tilde{p}$*, which goes to zero as* M, m *and* p *go to infinity provided* $\frac{m}{M} =$ *constant*$, c_0(\delta/4) > 2k\frac{m}{M}\ln(1 + \frac{16}{\delta})$ *and* $\frac{p}{m} =$ *constant* $\geq \frac{256k\ln(1+16/\delta)}{\delta^2}$*, where the constant* C_0 *is given explicitly above in 3.42 and 3.43.*

Bibliography

[1] D. Achlioptas. Database-friendly random projections. *Proc. ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, pages 274–281, 2001.

[2] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3):253–263, 2008.

[3] A. Barvinok. Math 710: Measure concentration. Lecture notes, 2005.

[4] J.J. Benedetto and M. Fickus. Finite normalized tight frames. *Advances in Computational Mathematics*, 18:357–385, 2003.

[5] R. Bernardini and R. Rinaldo. Efficient reconstruction from frame-based multiple descriptions. *IEEE Transactions on Signal Processing*, 8:3282–3296, 2005.

[6] B. C. Berndt and B. P. Yeap. Explicit evaluations and reciprocity theorems for finite trigonometric sums. *Advances in Applied Mathematics Journal*, 29(3):358–385, 2002.

[7] A. Betten, M. Braun, H. Fripertinger, A. Kerber, A. Kohnert, and A. Wassermann. *Error-Correcting Linear Codes*. Springer, Berlin, 2006.

[8] R. E. Blahut. *Algebraic Methods for Signal Processing and Communications Coding*. Springer-Verlag, New York, NY, USA, 1992.

[9] B. G. Bodmann. Optimal linear transmission by loss-insensitive packet encoding. *Applied and Computational Harmonic Analysis*, 22:274–285, 2007.

[10] B. G. Bodmann. Frames as codes. In *Finite Frames: Theory and Applications*(*Applied and Numerical Harmonic Analysis*), pages 241–266. Birkhäuser, Boston, USA, 2012.

[11] B. G. Bodmann, J. Cahill, and P. G. Casazza. Fusion frames and the restricted isometry property. *Numerical Functional Analysis and Optimization*, 33(7–9):770–790, 2012.

[12] B. G. Bodmann and V. I. Paulsen. Frames, graphs and erasures. *Linear Algebra and its Applications*, 404:118–146, 2005.

[13] B. G. Bodmann and P. K. Singh. Burst erasures and the mean-square error for cyclic parseval frames. *IEEE Transactions on Information Theory*, 57(7):4622–4635, 2011.

[14] B.G. Bodmann, D.W. Kribs, and V.I. Paulsen. Decoherence-insensitive quantum communication by optimal C^*-encoding. *IEEE Transactions on Information Theory*, 53:4738–4749, 2007.

[15] R. C. Bose and D. K. Ray-Chaudhuri. On a class of error correcting binary group codes. *Information and Control*, 3(1):68–79, 1960.

[16] P. Boufounos, G. Kutynoik, and H. Rauhut. Sparse recovery from combined fusion frame measurements. *IEEE Transactions on Information Theory*, 57(6):3864 – 3876, 2011.

[17] E. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489– 509, 2006.

[18] E. Candes, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59:1207–1223, 2006.

[19] E. Candes, M. Rudelson, T. Tao, and R. Vershynin. Error correction via linear programming. *In 46th Annual Symposium on Foundations of Computer Science (FOCS), Pittsburgh, PA*, pages 295–308, 2005.

[20] E. Candes and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, 52(12):5406–5425, 2006.

[21] E.J. Candes. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathematique*, 346:589–592, 2008.

[22] E.J. Candes and T. Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 51(12):4203–4215, 2005.

[23] P. G. Casazza, M. Fickus, J. Kovačević, M. T. Leon, and J. C. Tremain. A physical interpretation for finite tight frames. In *Harmonic Analysis and Applications: In Honor of John J. Benedetto*, pages 51–78. Birkhäuser, Boston, USA, 2006.

[24] P. G. Casazza and J. Kovačević. Equal-norm tight frames with erasures. *Advances in Computational Mathematics*, 18:387–430, 2003.

[25] P. G. Casazza and G. Kutyniok. Robustness of fusion frames under erasures of subspaces and of local frame vectors. In *Radon transforms, geometry, and wavelets*, volume 464 of *Contemporary Mathematics*, pages 149–160. American Mathematical Society, Providence, RI, 2007.

[26] P.G. Casazza. The art of frame theory. *Taiwanese Journal of Mathematics*, 4(2):129–201, 2000.

[27] P.G. Casazza and G. Kutyniok. Frames of subspaces. In *Wavelets, Frames and Operator Theory*, Contemp. Math. 345, pages 87–113. American Mathematical Society, Providence, RI, 2004.

[28] P.G. Casazza, G. Kutyniok, and S. Li. Fusion frames and distributed processing. *Applied and Computational Harmonic Analysis*, 25:114–132, 2008.

[29] S. Dasgupta and A. Gupta. An elementary proof of a theorem of Johnson and Lindenstrauss. *Random Structures Algorithms*, 22(1):60–65, 2003.

[30] I. Daubechies, A. Grossmann, and Y. Meyer. Painless nonorthogonal expansions. *Journal of Mathematical Physics*, 27:1271–1283, 1986.

[31] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

[32] R. J. Duffin and A. C. Schaeffer. A class of nonharmonic fourier series. *Transactions of the American Mathematical Society*, 72:341–366, 1952.

[33] E. O. Elliott. Estimates of error rates for codes on burst-noise channels. *Bell System Technical Journal*, 42:1977–1997, 1963.

[34] M. Fickus, B.D. Johnson, K. Kornelson, and K.A. Okoudjou. Convolutional frames and the frame potential. *Applied and Computational Harmonic Analysis*, 19:77–91, 2005.

[35] A. Gabay, M. Kieffer, and P. Duhamel. Joint source-channel coding using real BCH codes for robust image transmission. *IEEE Transctions on Image Processing*, 16(6):1568–1583, 2007.

[36] E. N. Gilbert. Capacity of a burst-noise channel. *Bell System Technical Journal*, 39:1253–1265, 1960.

[37] V. K. Goyal, J. Kovačević, and J. A. Kelner. Quantized frame expansions with erasures. *Applied and Computational Harmonic Analysis*, 10(3):203–233, 2001.

[38] V. K. Goyal, M. Vetterli, and N. T. Thao. Quantized overcomplete expansions in $\mathbb{R}^n$: analysis, synthesis, and algorithms. *IEEE Transactions on Information Theory*, 44(1):16–31, 2001.

[39] A. Hocquenghem. Codes correcteurs d'erreurs. *Chiffres, Revue Assoc. Franc. Calcul*, 2:147– 156, 1959.

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

[41] R. W. Heath Jr., H. Bölcskei, and A. Paulraj. Space-time signaling and frame theory. *Proceedings of ICASSP, IV, Salt Lake City, UT*, pages 2445–2448, 2001.

[42] R. W. Heath Jr. and A. J. Paulraj. Linear dispersion codes for MIMO systems based on frame theory. *IEEE Transactions on Signal Processing .*, 50:2429–2441, 2002.

[43] D. Kalra. Complex equiangular cyclic frames and erasures. *Linear Algebra and its Applications*, 419:373–399, 2006.

[44] J. Kovačević, L. Dragotti, and V. Goyal. Filter bank frame expansions with erasures. *IEEE Transactions on Information Theory*, 48(6):1439–1450, 2002.

[45] 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.

[46] M. Ledoux. *The Concentration of Measure Phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2001.

[47] J. Lopez and D. Han. Optimal dual frames for erasures. *Linear Algebra and its Applications*, 432:471–482, 2010.

[48] A. C. Lozano, J. Kovačević, and M. Andrews. Quantized frame expansions in a wireless environment. *Proceedings of Data Compression Conference, Snowbird, UT*, pages 480–489, 2002.

[49] T. G. Marshall. Coding of real-number sequences for error correction: A digital signal processing problem. *IEEE Journal on Selected Areas in Communications*, 2(2):381–392, 1984.

[50] P. G. Massey. Optimal reconstruction systems for erasures and for the q-potential. *Linear Algebra and its Applications*, 431(8):1302–1316, 2009.

[51] P. G. Massey, M. A. Ruiz, and D. Stojanoff. The structure of minimizers of the frame potential on fusion frames. *Journal of Fourier Analysis and Applications*, 16(4):514–543, 2010.

[52] C. E. Mueller and F. B. Weissler. Hypercontractivity for the heat semi-group for ultraspherical polynomials and on the n-sphere. *Journal of Functional Analysis*, 48:252–283, 1982.

[53] P. Oswald. Frames and space splittings in Hilbert spaces. Lecture notes, Part 1, Bell Labs, 1997.

[54] D. Petz. Spectral scale of self-adjoint operators and trace inequalities. *Journal of Mathematical Analysis and Applications*, 109:74–82, 1985.

[55] M. Püschel and J. Kovačević. Real, tight frames with maximal robustness to erasures. *Proceedings of Data Compression Conference, Snowbird, UT*, pages 63–72, 2005.

[56] G. Rath and C. Guillemott. Performance analysis and recursive syndrome decoding of DFT codes for bursty erasure recovery. *IEEE Transacations on Signal Processing*, 51(5):1335–1350, 2003.

[57] G. Rath and C. Guillemott. Frame-theoretic analysis of DFT codes with erasures. *IEEE Transacations on Signal Processing*, 52(2):447–460, 2004.

[58] G. Rath and C. Guillemott. Recent advances in DFT codes based quantized frame expansions for erasure channels. *Digital Signal Processing*, 14:332–354, 2004.

[59] G. R. Redinbo. Decoding real block codes: Activity detection, wiener estimation. *IEEE Transactions on Information Theory*, 46(2):609–623, 2000.

[60] I. S. Reed and G. Solomon. Polynomial codes over certain finite fields. *Journal of the Society for Industrial and Applied Mathematics*, 8(2):300–304, 1960.

[61] B. Sklar. *Digital Communications: Fundamentals and Applications*. Prentice Hall, Englewood Cliffs, NJ, 2nd edition, 2001.

[62] T. Strohmer and R.W. Heath Jr. Grassmannian frames with applications to coding and communication. *Applied and Computational Harmonic Analysis*, 14:257–275, 2003.

[63] W. Sun. G-frames and g-Riesz bases. *Journal of Mathematical Analysis and Applications*, 322(1):437– 452, 2006.

[64] G. Tan and T. Herfet. On the architecture of erasure error recovery under strict delay constraints. *14th European Wireless Conference, Prague*, pages 1–7, 2008.

[65] J. A. Tropp, I. S. Dhillon, R. W. Heath Jr., and T. Strohmer. Designing structured tight frames via alternating projection. *IEEE Transactions on Information Theory*, 51(1):188–209, 2005.

[66] S. B. Wicker and V. K. Bhargava, editors. *Reed-Solomon Codes and Their Applications*. IEEE Press, Piscataway, NJ, 1994.