Sampling with Arbitrary Sampling and Reconstruction Spaces and Oblique Dual Frame Vectors*

Yonina C. Eldar†

April 23, 2002

Abstract

This paper introduces a general framework for sampling and reconstruction procedures based on a consistency requirement, introduced by Unser and Aldroubi in [29]. The procedures we develop allow for almost arbitrary sampling and reconstruction spaces, as well as arbitrary input signals. We first derive a nonredundant sampling procedure. We then introduce the new concept of oblique dual frame vectors, that lead to frame expansions in which the analysis and synthesis frame vectors are not constrained to lie in the same space. Based on this notion, we develop a redundant sampling procedure that can be used to reduce the quantization error when quantizing the measurements prior to reconstruction.

Index Terms— Sampling, consistent sampling, frames, dual frame vectors, oblique projections, oblique pseudoinverse.

MRC numbers— 94A20, 65F20, 42C15

*To appear in *Journal of Fourier Analysis and its Applications*.

†Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge MA, 02139. E-mail: yonina@alum.mit.edu. Tel: (617) 577-5857, Fax: (617) 253-8495.

This work is supported in part by BAE Systems Cooperative Aggreement RP6891 under Army Research Laboratory Grant DAAD19-01-2-0008, by the Army Research Laboratory Collaborative Technology Alliance through BAE Systems Subcontract RK78554, and by Texas Instruments through the TI Leadership University Consortium.

1 Introduction

Many methods exist for representing a signal f by a sequence of numbers, which can be interpreted as measurements of f. The classical approach is to choose the measurements as samples of f. A more recent approach [29, 22, 3, 27, 28, 12] is to consider measurements that can be expressed as inner products of f with a set of vectors that span a subspace $\mathcal{S}$, which is referred to as the sampling space. The problem then is to reconstruct f from these measurements, using a set of vectors that span a subspace $\mathcal{W}$, which we refer to as the reconstruction space. If f does not lie in $\mathcal{W}$, then it cannot be perfectly reconstructed using only reconstruction vectors that span $\mathcal{W}$. Therefore, if we allow for signals out of $\mathcal{W}$, then we must relax the requirement for perfect reconstruction.

Given a reconstruction method, we can always choose a sampling method so that the reconstructed signal is closest to f in an l_2-sense. However, this requires the sampling space $\mathcal{S}$ to be equal to the reconstruction space $\mathcal{W}$. If the sampling scheme is such that $\mathcal{S} \neq \mathcal{W}$, then the minimal-error approximation cannot be obtained. Therefore, our problem is to construct a good approximation of f given both a sampling method and a reconstruction method.

In [29] the authors introduce the concept of consistent reconstruction, in which the reconstructed signal is in general not equal to f, but nonetheless yields the same measurements. Based on this requirement, they derive a sampling procedure for the special case in which f lies in L_2, and $\mathcal{S}$ and $\mathcal{W}$ are generated by integer translates of appropriately chosen functions.

In this paper we extend the results of [29] in several ways. First, we expand their results to a broader framework that does not require $\mathcal{S}$ and $\mathcal{W}$ to be generated by integer translates, and does not require f to lie in L_2, but rather can be applied to arbitrary subspaces of an arbitrary Hilbert space. This framework leads to some new sampling theorems, as well as further insight into the results of [29]. We also develop a geometric interpretation of the sampling and reconstruction scheme that provides further insight into the problem. Second, we develop *redundant* sampling procedures in which the measurements constitute an overcomplete representation of f. These measurements correspond to inner products of f with a frame for $\mathcal{S}$, and reconstruction is obtained using a frame for $\mathcal{W}$. To obtain a consistent reconstruction of f in this case, we develop a generalization of the well known dual frame operator [6], which we refer to as an *oblique dual frame operator*. The corresponding frame vectors are referred to as the *oblique dual frame vectors*. As we show, these frame vectors have properties that are very similar to those of the conventional dual frame vectors. However, in contrast with the dual frame vectors, they are not constrained to lie in the same space as the original frame vectors. Thus, using oblique dual frame vectors we can extend the notion of a frame expansion to include redundant expansions in which the analysis and synthesis frame vectors lie in different spaces.

By allowing for arbitrary sampling and reconstruction spaces, the sampling algorithms can be greatly simplified in many cases with only a minor increase in approximation error [29, 27, 28, 30, 4, 5]. Using oblique dual frame vectors we can further simplify the sampling and reconstruction processes while still retaining the flexibility of choosing the spaces almost arbitrarily, due to the extra degrees of freedom offered by the use of frames that allow us to construct frames with prescribed properties [15, 1]. Furthermore, if the measurements are quantized prior to reconstruction, then as we show the average power of the reconstruction error using this redundant procedure can be reduced by the redundancy of the frame in comparison with the nonredundant procedure.

For simplicity of exposition the results in this paper are derived for the finite-dimensional case; however, the results can be extended to include the infinite-dimensional case as well.

This paper is organized as follows. In Section 2 we consider the consistency requirement in detail, and develop a geometric interpretation of the sampling and reconstruction scheme. Section 3 considers explicit reconstruction methods. The aliasing and reconstruction error resulting from our general scheme are analyzed in Section 4. Section 5 considers nonredundant sampling. An example illustrating the reconstruction is given in Section 6. Section 7 considers redundant sampling procedures. In Section 8 we introduce the notion of oblique dual frame vectors and discuss their key properties, and in Section 10 we develop a redundant sampling procedure that can be used to reduce the quantization error.

2 Consistent Reconstruction

We denote vectors in an arbitrary Hilbert space $\mathcal{H}$ by lowercase letters, and the elements of a vector $c \in \mathbb{C}^N$ by $c[k]$. The inner product between vectors $x, y \in \mathcal{H}$ is denoted by $\langle x, y\rangle$. $P_{\mathcal{S}}$ denotes the orthogonal projection operator onto the space $\mathcal{S}$, I_N denotes the $N \times N$ identity matrix, and $\mathcal{N}(\cdot)$ and $\mathcal{R}(\cdot)$ denote the null space and range space of the corresponding operator, respectively.

2.1 Consistency Condition

Suppose we are given measurements $c[k]$ of a signal f that lies in an arbitrary Hilbert space $\mathcal{H}$. The measurements $c[k] = \langle s_k, f\rangle$ are obtained by taking the inner products of f with a set of N sampling vectors $\{s_k, 1 \leq k \leq N\}$ that span an M-dimensional subspace $\mathcal{S} \subseteq \mathcal{H}$, which is referred to as the sampling space. We construct an approximation $\hat{f}$ of f using a given set of N reconstruction vectors $\{w_k, 1 \leq k \leq N\}$ that span an M-dimensional subspace $\mathcal{W} \subseteq \mathcal{H}$, which we refer to as the reconstruction space. In the case of nonredundant sampling $N = M$ so that the sampling and reconstruction vectors form a basis for $\mathcal{S}$ and $\mathcal{W}$, respectively; in the case of redundant sampling $N > M$ and the sampling and reconstruction vectors form a frame for $\mathcal{S}$ and $\mathcal{W}$, respectively. We do not require the sampling space $\mathcal{S}$ and the reconstruction space $\mathcal{W}$ to be equal.

The reconstruction $\hat{f}$ has the form $\hat{f} = \sum_{k=1}^{N} d[k] w_k$ for some coefficients $d[k]$ that are a linear transformation of the measurements $c[k]$, so that $d = Hc$ for some H. With W and S denoting the set transformations[1] corresponding to the vectors w_k and s_k respectively,

$$\hat{f} = \sum_{k=1}^{N} d[k] w_k = Wd = WHc = WHS^* f. \tag{1}$$

The sampling and reconstruction scheme is illustrated in Fig. 1.

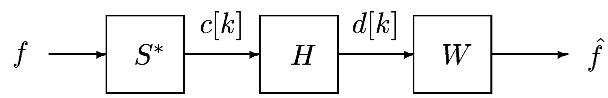


Figure 1: General sampling and reconstruction scheme.

Since $\hat{f}$ given by (1) always lies in $\mathcal{W}$, if $f \notin \mathcal{W}$, then $\hat{f} \neq f$. Because we are allowing the space of signals $\mathcal{H}$ to be larger than $\mathcal{W}$, we must replace the requirement for perfect reconstruction of $f \notin \mathcal{W}$ with a less stringent requirement. Therefore, our problem is to choose H in Fig. 1 so that $\hat{f}$ is a good approximation of f. In particular, we require that if $f \in \mathcal{W}$, then $\hat{f} = f$. To this end we must have that $\mathcal{W} \cap \mathcal{S}^{\perp} = \{0\}$. For suppose that x is a nonzero signal in $\mathcal{W} \cap \mathcal{S}^{\perp}$. Then $c[k] = \langle s_k, x \rangle = 0$ for all k, and clearly x cannot be reconstructed from the measurements $c[k]$. Consequently, throughout the paper we explicitly assume that $\mathcal{W} \cap \mathcal{S}^{\perp} = \{0\}$. Since $\mathcal{W}$ and $\mathcal{S}$ have the same dimension, this implies that $\mathcal{H} = \mathcal{W} \oplus \mathcal{S}^{\perp}$.

The sampling procedures we develop are based on a consistency requirement, introduced by Unser and Aldroubi in [29]. The idea is to construct a *consistent reconstruction* $\hat{f}$ of f that has the property that if we measure it using the measurement vectors s_k, then the measurements will be equal to the measurements $c[k]$ of f. Thus, our problem reduces to finding H in Fig. 1 such that

$$S^* \hat{f} = S^* WHS^* f = S^* f. \tag{2}$$

[1]A *set transformation* $X\colon \mathbb{C}^N \to \mathcal{H}$ corresponding to $\{x_k \in \mathcal{H}, 1 \le k \le N\}$ is defined by $Xa = \sum_{k=1}^{N} a[k] x_k$ for any $a \in \mathbb{C}^N$. From the definition of the adjoint $X^*\colon \mathcal{H} \to \mathbb{C}^N$ it follows that if $a = X^* y$, then $a[k] = \langle x_k, y \rangle$.

Theorem 1 below asserts that (2) is satisfied for all $f \in \mathcal{H}$ with $\mathcal{W} \cap \mathcal{S}^{\perp} = \{0\}$ if and only if $G = WHS^*$ is an oblique[2] projection [16, 2, 20] with $\mathcal{R}(G) = \mathcal{W}$ and $\mathcal{N}(G) = \mathcal{S}^{\perp}$, denoted by $E_{\mathcal{W}\mathcal{S}^{\perp}}$. The oblique projection $E_{\mathcal{W}\mathcal{S}^{\perp}}$ is defined as the the unique operator satisfying

$$\begin{aligned} E_{\mathcal{W}\mathcal{S}^{\perp}} w &= w \text{ for any } w \in \mathcal{W}; \\ E_{\mathcal{W}\mathcal{S}^{\perp}} v &= 0 \text{ for any } v \in \mathcal{S}^{\perp}. \end{aligned} \tag{3}$$

Theorem 1. *Let $\{c[k] = \langle s_k, f \rangle\}$ denote measurements of $f \in \mathcal{H}$ with sampling vectors $\{s_k\}$ that span an M-dimensional subspace $\mathcal{S} \subseteq \mathcal{H}$, and let the reconstruction vectors $\{w_k\}$ span an M-dimensional subspace $\mathcal{W} \subseteq \mathcal{H}$ such that $\mathcal{W} \cap \mathcal{S}^{\perp} = \{0\}$. Then $\hat{f}$ is a consistent reconstruction of f if and only if $\hat{f} = E_{\mathcal{W}\mathcal{S}^{\perp}} f$.*

Proof. Suppose that $\hat{f} = WHS^* f$ is a consistent reconstruction of f so that (2) is satisfied, and let $G = WHS^*$. Then for all $f \in \mathcal{H}$,

$$G^2 f = WHS^*WHS^* f = WHS^* \hat{f} = WHS^* f = Gf, \tag{4}$$

and G is a projection operator. Since $G = WHS^*$, $\mathcal{N}(G) \supseteq \mathcal{N}(S^*) = \mathcal{S}^{\perp}$ and $\mathcal{R}(G) \subseteq \mathcal{R}(W) = \mathcal{W}$. Suppose that $f \in \mathcal{N}(G)$ but $S^* f \neq 0$. Then $S^* \hat{f} = S^* G f = 0$ contradicting (2), so that $\mathcal{N}(G) = \mathcal{S}^{\perp}$. Now, let $f \in \mathcal{W}$. Then from (2) we have that $S^*(Gf - f) = 0$ so that $Gf - f \in \mathcal{N}(S^*) = \mathcal{S}^{\perp}$. But we also have that $Gf - f \in \mathcal{W}$. Since $\mathcal{W} \cap \mathcal{S}^{\perp} = \{0\}$, $Gf - f = 0$ for all $f \in \mathcal{W}$, and $\mathcal{R}(G) = \mathcal{W}$.

Next, suppose that $\hat{f} = Gf$ where $G = WHS^* = E_{\mathcal{W}\mathcal{S}^{\perp}}$. Then $Gf = f$ for any $f \in \mathcal{W}$, and $S^* \hat{f} = S^* G f = S^* f$. For $f \in \mathcal{S}^{\perp}$, $S^* f = Gf = 0$ so that $S^* \hat{f} = S^* f$. Since $\mathcal{H} = \mathcal{W} \oplus \mathcal{S}^{\perp}$, we conclude that $S^* \hat{f} = S^* f$ for all $f \in \mathcal{H}$. $\square$

[2]An oblique projection is a projection operator E satisfying $E^2 = E$ that is not necessarily Hermitian. The notation $E_{\mathcal{W}\mathcal{S}^{\perp}}$ denotes an oblique projection with range space $\mathcal{W}$ and null space $\mathcal{S}^{\perp}$. If $\mathcal{W} = \mathcal{S}$, then $E_{\mathcal{W}\mathcal{S}^{\perp}}$ is an orthogonal projection onto $\mathcal{W}$ which we denote by $P_{\mathcal{W}}$.

As a corollary of Theorem 1 we have that if $\mathcal{W} \cap \mathcal{S}^{\perp} = \{0\}$, then a consistent reconstruction $\hat{f}$ of a signal $f \in \mathcal{W}$ is always equal to f.

Theorem 1 describes the form of the unique consistent reconstruction if it exists, however it does not establish its existence. In Section 3 we show that a consistent reconstruction can always be obtained, and we derive explicit reconstruction procedures. Before we consider the detailed methods, in the next section we present a geometric interpretation of the sampling and reconstruction that provide further insight into the problem.

2.2 Geometric Interpretation

Let us first consider the case of perfect reconstruction for signals in $\mathcal{W}$. Thus, we would like to determine conditions under which any $f \in \mathcal{W}$ can be reconstructed from the measurements $c[k] = \langle f, s_k \rangle$. We first note that sampling f with measurement vectors in $\mathcal{S}$, is equivalent to sampling the orthogonal projection of f onto $\mathcal{S}$, denoted by $f_{\mathcal{S}} = P_{\mathcal{S}} f$. This follows from the relation

$$\langle s_k, f \rangle = \langle P_{\mathcal{S}} s_k, f \rangle = \langle s_k, P_{\mathcal{S}} f \rangle. \tag{5}$$

We may therefore decompose the sampling process into two stages, as illustrated in Fig. 2. In the first stage the signal f is (orthogonally) projected onto the sampling space $\mathcal{S}$, and in the second stage the projected signal $f_{\mathcal{S}}$ is measured. Since $f_{\mathcal{S}} \in \mathcal{S}$ and the vectors s_k span $\mathcal{S}$, $f_{\mathcal{S}}$ is uniquely determined by the measurements $c[k]$. Therefore, knowing $c[k]$ is equivalent to knowing $f_{\mathcal{S}}$.

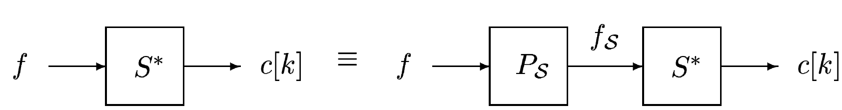
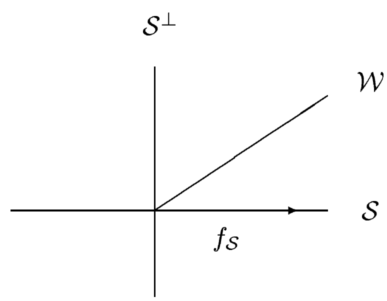
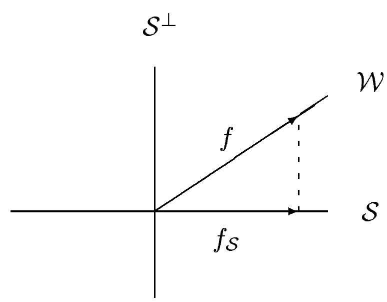


Figure 2: Decomposition of the sampling process into two stages.

In view of the interpretation of Fig. 2, our problem can be rephrased as follows. Can we reconstruct a signal in $\mathcal{W}$, given the orthogonal projection of the signal onto $\mathcal{S}$, with $\mathcal{W} \cap \mathcal{S}^{\perp} = \{0\}$? Fig. 3(a) depicts the orthogonal projection of an unknown signal $f \in \mathcal{W}$ onto $\mathcal{S}$, denoted $f_{\mathcal{S}}$. The problem then is to determine f from this projection. Since the direction of $\mathcal{W}$ is known, there is only one vector in $\mathcal{W}$ whose orthogonal projection onto $\mathcal{S}$ is $f_{\mathcal{S}}$; this vector is illustrated in Fig. 3(b). Thus, from this geometrical interpretation we conclude that for $\mathcal{W} \cap \mathcal{S}^{\perp} = \{0\}$, perfect reconstruction of any $f \in \mathcal{W}$ from the measurements $c[k]$ is always possible.



(a)



(b)

Figure 3: Illustration of perfect reconstruction of $f \in \mathcal{W}$ from $f_{\mathcal{S}} = P_{\mathcal{S}} f$, with $\mathcal{W} \cap \mathcal{S}^{\perp} = \{0\}$ (a) orthogonal projection of unknown signal in $\mathcal{W}$ onto $\mathcal{S}$ (b) unique signal in $\mathcal{W}$ with the given projection.

We now discuss consistent reconstruction for signals $f \in \mathcal{H}$. If $\hat{f}$ is a consistent reconstruction of f, then f and $\hat{f}$ have the same measurements: $c[k] = \langle s_k, f \rangle = \langle s_k, \hat{f} \rangle$. From our previous discussion it then follows that $f_{\mathcal{S}} = \hat{f}_{\mathcal{S}}$ where $\hat{f}_{\mathcal{S}} = P_{\mathcal{S}} \hat{f}$. Thus, geometrically a consistent reconstruction $\hat{f}$ of f is a signal in $\mathcal{W}$ whose orthogonal projection onto $\mathcal{S}$ is equal to the orthogonal projection of f

onto $\mathcal{S}$, as illustrated in Fig. 4. Evidently, the consistent reconstruction is unique and always exists. We have seen in Theorem 1 that this reconstruction has a nice geometrical interpretation: It is the oblique projection of f onto $\mathcal{W}$ along $\mathcal{S}^\perp$. This interpretation is illustrated in Fig. 5, from which it is apparent that $E_{\mathcal{W}\mathcal{S}^\perp}f$ and f have the same orthogonal projection onto $\mathcal{S}$ and consequently yield the same measurements.

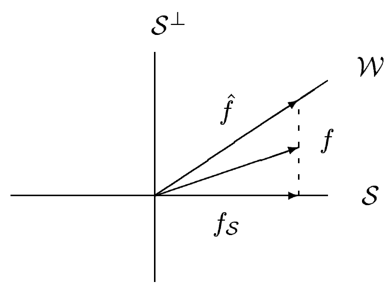


Figure 4: Illustration of consistent reconstruction of an arbitrary f from $f_\mathcal{S}$, with $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$.

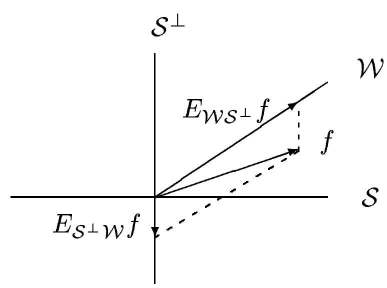


Figure 5: Decomposition of f into its components in $\mathcal{W}$ and $\mathcal{S}^\perp$ given by $E_{\mathcal{W}\mathcal{S}^\perp}f$ and $E_{\mathcal{S}^\perp\mathcal{W}}f$, respectively.

In summary, by considering a geometric interpretation of the sampling process and the consistency requirement we have demonstrated that perfect reconstruction of signals in $\mathcal{W}$ is always possible as long as $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$, and we illustrated the reconstruction geometrically. We also showed that under the same condition consistent reconstruction is always possible, and illustrated the reconstruction. It is important to note that the geometric interpretation (and Theorem 1) hold irrespective of whether the sampling process is nonredundant or redundant. In the next section we provide mathematical proof of these results and derive an explicit reconstruction scheme. Nonredundant procedures are considered in Section 5, and redundant procedures are considered in Sections 7, 8 and 10.

3 Reconstruction Scheme

3.1 Reconstruction Algorithm

From Theorem 1 and the geometric interpretation of Section 2.2 it follows that to obtain a consistent reconstruction $\hat{f}$ of f we need to determine H in Fig. 1 such that $G = WHS^* = E_{\mathcal{W}\mathcal{S}^\perp}$, *i.e.*, such that G satisfies (3). We now show that with $H = (S^*W)^\dagger$, where $(\cdot)^\dagger$ denotes the *Moore-Penrose pseudoinverse* [13],

$$\hat{f} = \sum_{k=1}^{N} d[k] w_k = Wd = W(S^*W)^\dagger c = W(S^*W)^\dagger S^* f, \tag{6}$$

is a consistent reconstruction of f for all $f \in \mathcal{H}$. To this end we prove the following proposition.

Proposition 1. *Let the vectors $\{s_k, 1 \le k \le N\}$ corresponding to S span an M-dimensional subspace $\mathcal{S} \subseteq \mathcal{H}$, and let the vectors $\{w_k, 1 \le k \le N\}$ corresponding to W span an M-dimensional subspace $\mathcal{W} \subseteq \mathcal{H}$, with $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$. Then the oblique projection onto $\mathcal{W}$ along $\mathcal{S}^\perp$ can be expressed as $E_{\mathcal{W}\mathcal{S}^\perp} = W(S^*W)^\dagger S^*$.*

Proof. We denote $G = W(S^*W)^\dagger S^*$ and show that G satisfies (3).

First, since the vectors w_k span $\mathcal{W}$, any $w \in \mathcal{W}$ can be expressed as $w = Wa$ for some $a \in \mathbb{C}^N$. Then $Gw = W(S^*W)^\dagger S^*Wa = WPa$ where from the properties of the pseudoinverse, P is an orthogonal projection onto $\mathcal{N}(S^*W)^\perp$. Since $Wx \in \mathcal{W}$ for any $x \in \mathbb{C}^N$ and $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$, $S^*Wx = 0$ if and only if $Wx = 0$, so that $\mathcal{N}(S^*W) = \mathcal{N}(W)$. Then for any $w \in \mathcal{W}$, $Gw =$

$WP_{\mathcal{N}(W)^\perp}a = Wa = w$. Next, since $s_k \in \mathcal{S}$, $S^*v = 0$ for any $v \in \mathcal{S}^\perp$, and $Gv = 0$ so that G satisfies (3) and consequently $G = E_{\mathcal{W}\mathcal{S}^\perp}$. $\square$

If $f \in \mathcal{W}$ then $\hat{f} = E_{\mathcal{W}\mathcal{S}^\perp}f = f$, and f can be perfectly reconstructed from the measurements $c[k]$ using (6). By choosing different spaces $\mathcal{H}$, $\mathcal{W}$ and $\mathcal{S}$ and using (6), we can arrive at a variety of new and interesting perfect reconstruction sampling theorems.

From (6), $\hat{f}$ is obtained by first transforming the measurements $c[k]$ into "corrected" measurements $d[k]$ corresponding to $d = (S^*W)^\dagger c = Tf$, where $T = (S^*W)^\dagger S^*$. As we now show, T has an interesting interpretation: It is the *oblique pseudoinverse* of W on $\mathcal{V} = \mathcal{N}(W)^\perp$ along $\mathcal{S}^\perp$.

3.2 Oblique Pseudoinverse

Let $T\colon \mathcal{K} \to \mathcal{U}$ be a linear transformation, and let $\mathcal{K} = \mathcal{G} \oplus \mathcal{N}(T)$ and $\mathcal{U} = \mathcal{R}(T) \oplus \mathcal{Z}$. The *oblique pseudoinverse* of T on $\mathcal{G}$ along $\mathcal{Z}$, denoted $T^\#_{\mathcal{G}\mathcal{Z}}$, is the unique transformation satisfying [23, 9]

$$TT^\#_{\mathcal{G}\mathcal{Z}} = E_{\mathcal{R}(T)\mathcal{Z}}; \tag{7}$$

$$T^\#_{\mathcal{G}\mathcal{Z}}T = E_{\mathcal{G}\mathcal{N}(T)}; \tag{8}$$

$$\mathcal{R}(T^\#_{\mathcal{G}\mathcal{Z}}) = \mathcal{G}. \tag{9}$$

As can be verified [9], (7)–(9) imply that $T^\#_{\mathcal{G}\mathcal{Z}}$ inverts T between $\mathcal{G}$ and $\mathcal{R}(T)$, while nulling out any vector in $\mathcal{Z}$. This interpretation is illustrated in Fig. 6, from which it follows that the pseudoinverse $T^\dagger$ is a special case of the oblique pseudoinverse $T^\#_{\mathcal{G}\mathcal{Z}}$ for which $\mathcal{G} = \mathcal{N}(T)^\perp$ and $\mathcal{Z} = \mathcal{R}(T)^\perp$.

Proposition 2. *Let the vectors $\{s_k, 1 \le k \le N\}$ corresponding to S span an M-dimensional subspace $\mathcal{S} \subseteq \mathcal{H}$, and let the vectors $\{w_k, 1 \le k \le N\}$ corresponding to W span an M-dimensional subspace $\mathcal{W} \subseteq \mathcal{H}$, with $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$. Then the oblique pseudoinverse of W on $\mathcal{V} = \mathcal{N}(W)^\perp$ along $\mathcal{S}^\perp$ can be expressed as*

$$W^\#_{\mathcal{V}\mathcal{S}^\perp} = (S^*W)^\dagger S^*. \tag{10}$$

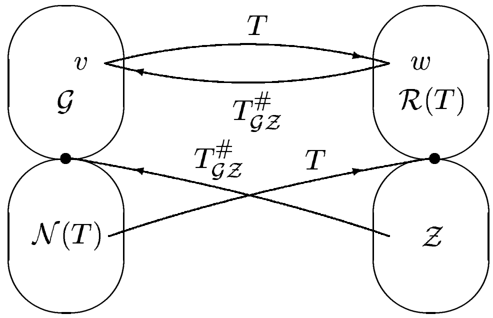


Figure 6: The action of T and $T^{\#}_{\mathcal{G}\mathcal{Z}}$ on the subspaces $\mathcal{G}, \mathcal{N}(T), \mathcal{R}(T)$ and $\mathcal{Z}$.

Proof. We need to show that $W^{\#}_{\mathcal{V}\mathcal{S}^\perp}$ given by (10) satisfies (7)–(9), *i.e.*,

$$WW^{\#}_{\mathcal{V}\mathcal{S}^\perp} = E_{\mathcal{W}\mathcal{S}^\perp}; \tag{11}$$

$$W^{\#}_{\mathcal{V}\mathcal{S}^\perp} W = P_{\mathcal{V}}; \tag{12}$$

$$\mathcal{R}(W^{\#}_{\mathcal{V}\mathcal{S}^\perp}) = \mathcal{V}. \tag{13}$$

The fact that $W^{\#}_{\mathcal{V}\mathcal{S}^\perp}$ satisfies (11) follows immediately from Proposition 1. To prove that $W^{\#}_{\mathcal{V}\mathcal{S}^\perp}$ satisfies (12) we note that from the properties of the pseudoinverse, $W^{\#}_{\mathcal{V}\mathcal{S}^\perp}W = (S^*W)^\dagger S^*W = P_{\mathcal{N}(S^*W)^\perp}$. But since $\mathcal{N}(S^*W) = \mathcal{N}(W) = \mathcal{V}^\perp$, $(S^*W)^\dagger S^*W = P_{\mathcal{V}}$. Finally, $\mathcal{R}(W^{\#}_{\mathcal{V}\mathcal{S}^\perp}) = \mathcal{R}((S^*W)^\dagger S^*) = \mathcal{R}((S^*W)^\dagger)$ since $\mathcal{N}((S^*W)^\dagger)^\perp = \mathcal{R}(S^*W) = \mathcal{R}(S^*)$ because for any $x \in \mathcal{H}$, $S^*x = S^*E_{\mathcal{W}\mathcal{S}^\perp}x$ and $E_{\mathcal{W}\mathcal{S}^\perp}x \in \mathcal{W} = \mathcal{R}(W)$. Thus $\mathcal{R}(W^{\#}_{\mathcal{V}\mathcal{S}^\perp}) = \mathcal{R}((S^*W)^\dagger) = \mathcal{N}(S^*W)^\perp = \mathcal{N}(W)^\perp = \mathcal{V}$, and $W^{\#}_{\mathcal{V}\mathcal{S}^\perp}$ satisfies (13). $\square$

Comparing (10) with (6) we see that $\hat{f} = Wd$ where $d = W^{\#}_{\mathcal{V}\mathcal{S}^\perp}f$. Thus $d[k] = \langle v_k, f\rangle$ where v_k are the vectors corresponding to $(W^{\#}_{\mathcal{V}\mathcal{S}^\perp})^*$. Since $\mathcal{R}((W^{\#}_{\mathcal{V}\mathcal{S}^\perp})^*) = \mathcal{N}(W^{\#}_{\mathcal{V}\mathcal{S}^\perp})^\perp = \mathcal{S}$, the vectors v_k

lie in $\mathcal{S}$. Furthermore, from (11) $WW^{\#}_{\mathcal{V}\mathcal{S}^{\perp}} = E_{\mathcal{W}\mathcal{S}^{\perp}}$, so that any $f \in \mathcal{S}$ can be expressed as

$$f = E_{\mathcal{S}\mathcal{W}^{\perp}} f = (E_{\mathcal{W}\mathcal{S}^{\perp}})^* f = (W^{\#}_{\mathcal{V}\mathcal{S}^{\perp}})^* W^* f = \sum_{k=1}^{N} b[k] v_k, \tag{14}$$

where $b = W^* f$, and the vectors v_k span $\mathcal{S}$.

Therefore, in the case of nonredundant sampling *i.e.,* $N = M$, the vectors v_k form a basis for $\mathcal{S}$, and in the case of redundant sampling, *i.e.,* $N > M$, the vectors v_k form a frame for $\mathcal{S}$. These basis and frame vectors have special properties which we discuss in Sections 5 and 8, respectively. Specifically, in Section 5 we show that in the case of nonredundant sampling, the vectors v_k form a basis for $\mathcal{S}$ that is biorthogonal to the basis vectors w_k. In Section 8 we show that in the case of redundant sampling, the vectors v_k form a frame for $\mathcal{S}$ which we define as *the oblique dual frame*, which has properties analogous to the dual frame vectors.

4 Aliasing and Error Bounds

Since in general $f \notin \mathcal{W}$, the reconstruction (6) may result in aliasing in $\hat{f}$, which occurs when components of f that lie out of $\mathcal{W}$ are aliased into $\hat{f}$. A very nice and intuitive way to think about aliasing was proposed in [18] in the context of multiresolution spaces in terms of the norm of the "out-of-space" component. Let Γ denote the sampling operator defined by $\hat{f} = \Gamma f$, which in our case is equal to $\Gamma = E_{\mathcal{W}\mathcal{S}^{\perp}}$. Then the aliasing norm is defined as [18, 17]

$$A_{\Gamma} = \sup_{f \in \mathcal{W}^{\perp}} \frac{||\Gamma f||}{||f||} = \sup_{f \in \mathcal{W}^{\perp}} \frac{||E_{\mathcal{W}\mathcal{S}^{\perp}} f||}{||f||}. \tag{15}$$

From (15), $A_{\Gamma} = 0$ only if $E_{\mathcal{W}\mathcal{S}^{\perp}} = 0$ for all $f \in \mathcal{W}^{\perp}$ which implies that $\mathcal{S} = \mathcal{W}$. To avoid aliasing when $\mathcal{S} \neq \mathcal{W}$, we can first orthogonally project f onto $\mathcal{W}$, and then measure the projection. The measurements are then $c = S^* P_{\mathcal{W}} f$, so that $c[k] = \langle t_k, f \rangle$ where $t_k = P_{\mathcal{W}} s_k$ and consequently

$t_k \in \mathcal{W}$; as we expect the effective sampling space is equal to the reconstruction space.

When $\mathcal{S} \neq \mathcal{W}$ we can obtain a bound on A_Γ using the fact that for any $f \in \mathcal{H}$ [30]

$$||E_{\mathcal{W}\mathcal{S}^\perp} f|| \leq \frac{1}{\cos(\theta_{\mathcal{W}\mathcal{S}})} ||f||, \tag{16}$$

where the angle $\theta_{\mathcal{W}\mathcal{S}}$ between $\mathcal{S}$ and $\mathcal{W}$ is defined as [29]

$$\cos(\theta_{\mathcal{W}\mathcal{S}}) = \inf_{f \in \mathcal{W}, ||f||=1} ||P_{\mathcal{S}} f||. \tag{17}$$

Thus,

$$A_\Gamma \leq \frac{1}{\cos(\theta_{\mathcal{W}\mathcal{S}})}. \tag{18}$$

As we expect intuitively, the bound decreases as the angle between the spaces $\mathcal{S}$ and $\mathcal{W}$ decreases, in which case $\mathcal{S}$ is "close" to $\mathcal{W}$.

The norm of the reconstruction error $f - E_{\mathcal{W}\mathcal{S}^\perp} f$ can be bounded based on results derived in [29],

$$||f - P_{\mathcal{W}} f|| \leq ||f - E_{\mathcal{W}\mathcal{S}^\perp} f|| \leq \frac{1}{\cos(\theta_{\mathcal{W}\mathcal{S}})} ||f - P_{\mathcal{W}} f||, \tag{19}$$

where $||f - P_{\mathcal{W}} f||$ is the minimal norm of the reconstruction error corresponding to the case in which $\mathcal{W} = \mathcal{S}$. From (19) we see that there is a penalty for the flexibility offered by choosing $\mathcal{S}$ (almost) arbitrarily: The norm of the reconstruction error for $f \notin \mathcal{W}$ is increased. However, in many practical applications this increase in error is very small [28, 30, 4, 5].

5 Reconstruction From Nonredundant Measurements

Suppose that the sampling vectors $\{s_k, 1 \leq k \leq M\}$ form a basis for $\mathcal{S}$ and the reconstruction vectors $\{w_k, 1 \leq k \leq M\}$ form a basis for $\mathcal{W}$. Then, as we now show, S^*W is invertible so that the general reconstruction formula (6) reduces to

$$\hat{f} = \sum_{k=1}^{M} d[k]w_k = Wd = W(S^*W)^{-1}S^*f. \tag{20}$$

Proposition 3. *Let the vectors $\{s_k, 1 \leq k \leq M\}$ corresponding to S denote a basis for an M-dimensional subspace $\mathcal{S}$ of $\mathcal{H}$, and let the vectors $\{w_k, 1 \leq k \leq M\}$ corresponding to W denote a basis for an M-dimensional subspace $\mathcal{W}$ of $\mathcal{H}$. Then S^*W is invertible if and only if $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$.*

Proof. Suppose that S^*W is invertible, and let $x \in \mathcal{W} \cap \mathcal{S}^\perp$. Since $x \in \mathcal{S}^\perp$, $S^*x = 0$. But since $x \in \mathcal{W}$, $x = Wa$ for some $a \in \mathbb{C}^M$. Thus, $S^*x = S^*Wa = 0$. Because S^*W is invertible, we must have $a = 0$ so that $x = Wa = 0$.

Conversely, suppose that $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$. Let $x \neq 0$ be a vector in $\mathcal{N}(S^*W)$ so that $S^*Wx = 0$. Since the vectors w_k are linearly independent, $y = Wx \neq 0$ and therefore $y \in \mathcal{N}(S^*) = \mathcal{R}(S)^\perp = \mathcal{S}^\perp$. In addition, $y \in \mathcal{R}(W) = \mathcal{W}$. Therefore $y = 0$, which in turn implies that $x = 0$ contradicting our assumption. Thus, S^*W is invertible. $\square$

The resulting measurement and reconstruction scheme is depicted in Fig. 7. Note, that since $\hat{f}$ is unique and the vectors w_k are linearly independent, the coefficients $d[k]$ are also unique.

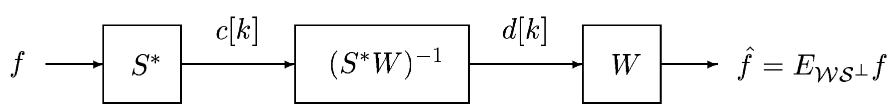


Figure 7: Consistent reconstruction of f using nonredundant sampling vectors s_k and nonredundant reconstruction vectors w_k, with $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$.

We may interpret the reconstruction scheme of Fig. 7 in terms of a basis expansion for signals in $\mathcal{W}$. Since for $f \in \mathcal{W}$, $\hat{f} = f$, any $f \in \mathcal{W}$ can be represented as $f = \sum_{k=1}^{M} d[k]w_k$ where $d[k] = \langle v_k, f\rangle$ and the vectors $v_k \in \mathcal{S}$ correspond to $V = (W^{\#}_{\mathcal{V}\mathcal{S}^\perp})^* = S(W^*S)^{-1}$. We have already seen in Section 3.2 that these vectors form a basis for $\mathcal{S}$. Since $V^*W = (S^*W)^{-1}S^*W = I_M$, these basis vectors have the property that they are biorthogonal to w_k: $\langle v_k, w_m \rangle = \delta_{km}$. Therefore Fig. 7 provides an explicit method for constructing basis vectors for an arbitrary space $\mathcal{S}$ with $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$, that are biorthogonal to the basis vectors w_k.

6 Bandlimited Sampling of Time-Limited Sequences

To illustrate the details of the sampling and reconstruction scheme of Fig. 7, we now consider an example in which $\mathcal{H}$ is the space of sequences $x[n]$ such that $x[n] = 0$ for $n < 0, n \geq N$, $\mathcal{W}$ is the space of sequences $x[n]$ such that $x[n] = 0$ for $n < 0, n \geq M$ where $M = 2M' + 1 < N$, and $\mathcal{S}$ is the space of "bandlimited" sequences $x[n]$ such that $X[k] = 0$ for $M' < k < N - M'$, where $X[k], 0 \leq k \leq N-1$ denotes the N point DFT of $x[n]$. The bases for $\mathcal{S}$ and $\mathcal{W}$ are chosen as the sequences $s_k[n], 0 \leq k \leq M-1$ and $w_k[n], 0 \leq k \leq M-1$ respectively, given by $s_k[n] = e^{j2\pi(k-M')n/N}$ for $0 \leq n \leq N-1$ and 0 otherwise, and $w_k[n] = \delta[k-n]$.

Consider an arbitrary sequence $f[n]$ in $\mathcal{H}$. The measurements $c[k], 0 \leq k \leq M-1$ of $f[n]$ are

$$c[k] = \langle s_k, f \rangle = \sum_{n=0}^{N-1} s_k^*[n] f[n] = \sum_{n=0}^{N-1} f[n] e^{-j2\pi(k-M')n/N} = F[((k-M'))_N], \tag{21}$$

where $F[k], 0 \leq k \leq N-1$ is the N point DFT of $f[n]$, and $((p))_N = p \bmod N$. Thus, the measurements $c[k]$ are the M lowpass DFT coefficients of the N point DFT of $f[n]$. To obtain a consistent reconstruction of $f[n]$ we need to determine $(S^*W)^{-1}$. The kmth element of S^*W is

$$\langle s_k, w_m \rangle = \sum_{n=0}^{N-1} s_k^*[n] w_m[n] = s_k^*[m] = Z^{km} B^m, \tag{22}$$

where $Z = e^{-j2\pi/N}$ and $B = e^{j2\pi M'/N}$. We can therefore express S^*W in the form

$$S^*W = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & Z & Z^2 & \cdots & Z^{M-1} \\ & & \vdots & & \\ 1 & Z^{M-1} & Z^{2(M-1)} & \cdots & Z^{(M-1)^2} \end{bmatrix} D. \tag{23}$$

Eq. (23) is the product of a Vandermonde matrix and a diagonal matrix D with nonzero diagonal elements $B^m, 0 \leq m \leq M-1$. Therefore, S^*W is always invertible which implies by Proposition 3 that $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$. We can compute the inverse of S^*W using any of the formulas for the inverse of a Vandermonde matrix (see *e.g.,* [21, 25]). The corrected measurements $d[k]$ are then given by the elements of $d = (S^*W)^{-1}c$ where c is the vector with elements $c[k]$ given by (21), and $\hat{f}[n] = \sum_{k=0}^{N-1} w_k[n]d[k] = d_n$ for $0 \leq n \leq M-1$ and 0 otherwise. The consistency requirement implies that $\hat{F}[((k-M'))_N] = F[((k-M'))_N]$ for $0 \leq k \leq M-1$, where $\hat{F}[k]$ is the N point DFT of $\hat{f}[n]$. Thus $\hat{f}[n]$ is a "time-limited" sequence that has the same lowpass DFT coefficients as $f[n]$.

In [9, 7] we develop a systematic method for constructing signals with prescribed properties. In particular, we consider constructing a signal in $\mathcal{H}$ with specified properties in both $\mathcal{W}$ and $\mathcal{S}$. Using these methods we can generalize our construction here to produce a signal with specified lowpass coefficients *and* specified values on a given time interval.

Now, suppose that $f[n]$ is a length M sequence in $\mathcal{W}$, and we are given M lowpass DFT coefficients $F[((k-M'))_N]$, $0 \leq k \leq M-1$. We can then perfectly reconstruct $f[n]$ from these coefficients using the method described above. This implies the intuitive result that a time-limited discrete-time sequence can be reconstructed from a lowpass segment of its DFT transform. This result is the analogue for the finite length discrete-time case of Papoulis' theorem [24], which implies that a time-limited function can be recovered from a lowpass segment of its Fourier transform. The reconstruction based on Papoulis' theorem is typically obtained using iterative algorithms such as those discussed in [24, 26]. By choosing appropriate sampling and reconstruction vectors in the general scheme of Fig. 7, we obtained a finite length discrete-time version of this theorem together with a simple non-iterative reconstruction method. This example illustrates the type of procedure that might be followed in using our framework to generate new sampling theorems.

7 Reconstruction From Redundant Measurements

Suppose now that we are given a set of redundant measurements $\tilde{c}[k] = \langle x_k, f\rangle$ of a signal $f \in \mathcal{H}$, where the vectors $\{x_k, 1 \leq k \leq N\}$ form a frame for $\mathcal{S}$ and reconstruction is obtained using the reconstruction vectors $\{y_k, 1 \leq k \leq N\}$ which form a frame for $\mathcal{W}$: A set of vectors $\{y_k \in \mathcal{W}, 1 \leq k \leq N\}$ forms a frame for $\mathcal{W}$ if there exists constants $A > 0$ and $B < \infty$ such that

$$A||x||^2 \leq \sum_{k=1}^{N} |\langle x, y_k\rangle|^2 \leq B||x||^2, \tag{24}$$

for all $x \in \mathcal{W}$ [6]. Although in principle N maybe infinite, we assume throughout that N is finite. The lower bound in (24) ensures that the vectors y_k span $\mathcal{W}$; thus we must have $N \geq M$. If $N < \infty$, then the right hand inequality of (24) is always satisfied with $B = \sum_{k=1}^{N} \langle y_k, y_k\rangle$. Thus, any finite set of vectors that spans $\mathcal{W}$ is a frame for $\mathcal{W}$. If the bounds $A = B$ in (24), then the frame is called a tight frame. If in addition $A = B = 1$, then we call the frame a normalized tight frame. The redundancy of the frame is defined as $r = N/M$.

From the general reconstruction formula (6), $\hat{f}$ is obtained using the frame vectors y_k by transforming the measurements $\tilde{c}[k]$ into corrected measurements $\tilde{d} = (X^*Y)^\dagger\tilde{c}$, as depicted in Fig. 8. An alternative form of Fig. 8, that will be used in Section 10, can be obtained by noting that any frame Y for $\mathcal{W}$ can be expressed as $Y = WZ$ where W corresponds to an arbitrary basis for $\mathcal{W}$, and Z: $\mathbb{C}^M \rightarrow \mathbb{C}^N$ has rank M, *i.e.*, $ZZ^\dagger = I_M$. Similarly, any frame X for $\mathcal{S}$ can be expressed as $X = ST$ where S corresponds to an arbitrary basis for $\mathcal{S}$, and T: $\mathbb{C}^M \rightarrow \mathbb{C}^N$ satisfies $TT^\dagger = I_M$.

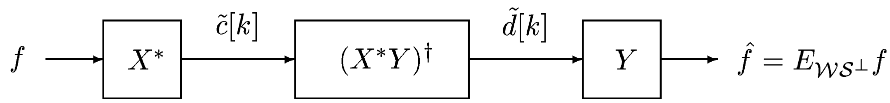


Figure 8: Consistent reconstruction of f using redundant sampling vectors x_k and redundant reconstruction vectors y_k, with $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$.

Then

$$(X^*Y)^\dagger X^* = (T^*S^*WZ)^\dagger T^*S^*. \tag{25}$$

To simplify (25) we rely on the following lemma.

Lemma 1. *Let A be an $m \times n$ matrix and let B be an $n \times k$ matrix. If $\mathcal{R}(B) = \mathcal{N}(A)^\perp$, $\mathcal{N}(AB) = \mathcal{N}(B)$ and $\mathcal{R}(AB) = \mathcal{R}(A)$, then*

$$(AB)^\dagger = B^\dagger A^\dagger. \tag{26}$$

In particular if A and B both have rank n, then (26) is satisfied.

Proof. The lemma is proven in a straightforward manner by showing that under the conditions of the lemma, $B^\dagger A^\dagger$ satisfies the Moore-Penrose conditions [13]. □

Since T^* and Z both have rank M and from Proposition 3, S^*W is invertible, it follows from Lemma 1 that

$$(T^*S^*WZ)^\dagger = Z^\dagger (S^*W)^{-1} (T^\dagger)^*. \tag{27}$$

Substituting into (25),

$$(X^*Y)^\dagger X^* = Z^\dagger (S^*W)^{-1} (T^\dagger)^* T^* S^* = Z^\dagger (S^*W)^{-1} S^*, \tag{28}$$

where we used the fact that $(T^\dagger)^* T^* = I_M$. From (28) it follows that we can obtain the redundant corrected measurements $\tilde{d}[k]$ directly from the nonredundant corrected measurements

$d = (S^*W)^{-1}S^*f = (S^*W)^{-1}c$, via $\tilde{d} = Z^\dagger d$, where $c[k] = \langle s_k, f\rangle$ are the nonredundant measurements obtained using the vectors s_k. This interpretation is illustrated in Fig. 9.

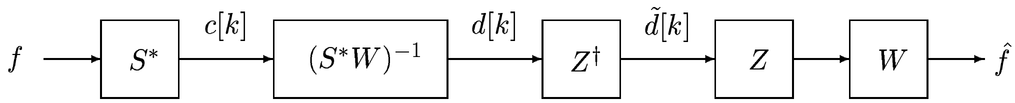


Figure 9: Equivalent representation of Fig. 8.

We have seen that the nonredundant sampling scheme of Fig. 7 can be interpreted as a basis expansion of $f \in \mathcal{W}$ in terms of a biorthogonal basis for $\mathcal{S}$. We now show that the redundant sampling scheme of Fig. 8 can be interpreted as a frame expansion of $f \in \mathcal{W}$ in terms of the *oblique dual frame vectors* on $\mathcal{S}$. Furthermore, although the redundant coefficients $\tilde{d}[k]$ are not unique, based on the properties of the oblique dual frame vectors we will show that the sampling scheme of Fig. 8 results in coefficients $\tilde{d}[k]$ with minimal l_2-norm.

In the next section we introduce the oblique dual frame vectors and discuss their key properties.

8 Oblique Dual Frame Vectors

8.1 Definition of the oblique dual frame vectors

Definition 1. *Let the vectors $\{y_k \in \mathcal{W}, 1 \leq k \leq N\}$ corresponding to Y denote a frame for an M-dimensional subspace $\mathcal{W}$ of $\mathcal{H}$, and let $\mathcal{S}$ be an M-dimensional subspace of $\mathcal{H}$ with $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$. Then the oblique dual frame vectors of y_k on $\mathcal{S}$ are the frame vectors $\{\tilde{y}_k^{\mathcal{S}} \in \mathcal{S}, 1 \leq k \leq N\}$ corresponding to the oblique dual frame operator $(Y^{\#}_{\mathcal{V}\mathcal{S}^\perp})^*$ where $\mathcal{V} = \mathcal{N}(Y)^\perp$.*

Note that from the discussion following Proposition 2, the vectors $\tilde{y}_k^{\mathcal{S}}$ form a frame for $\mathcal{S}$. As we show in the next section, these frame vectors have properties which are analogous to the properties of the conventional dual frame vectors [19, 6], and therefore justify our choice of terminology.

From (6) and Proposition 2 we have that the corrected measurement $\tilde{d}[k]$ in Fig. 8 are the inner products of f with the oblique dual frame vectors of y_k on $\mathcal{S}$: $\tilde{d}[k] = \langle \tilde{y}_k^{\mathcal{S}}, f \rangle$. Since $YY^{\#}_{\mathcal{VS}^\perp} = E_{\mathcal{WS}^\perp}$, any $f \in \mathcal{W}$ can be expressed as

$$f = E_{\mathcal{WS}^\perp} f = \sum_{k=1}^{N} \langle \tilde{y}_k^{\mathcal{S}}, f \rangle y_k. \tag{29}$$

Eq. (29) is just a frame expansion of a signal $f \in \mathcal{W}$. However, in contrast with conventional frame expansions, here the synthesis frame vectors lie in $\mathcal{W}$, while the analysis frame vectors $\tilde{y}_k^{\mathcal{S}}$ lie in an arbitrary space $\mathcal{S}$, such that $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$.

In the special case in which $\mathcal{S} = \mathcal{W}$, $Y^{\#}_{\mathcal{VS}^\perp} = Y^\dagger$ and the oblique dual frame operator reduces to the conventional dual frame operator [6]. Then any $f \in \mathcal{W}$ can be expressed as $f = \sum_{k=1}^{N} \langle \tilde{y}_k, f \rangle y_k$, where $\tilde{y}_k \in \mathcal{W}$ are the dual frame vectors [6] of y_k in $\mathcal{W}$, corresponding to $(Y^\dagger)^*$.

8.2 Properties of the Oblique Dual Frame Vectors

Given a frame y_k for $\mathcal{W}$, there are many choices of coefficients $\tilde{d}[k]$ that correspond to measurements of f using a frame for $\mathcal{S}$, and such that $E_{\mathcal{WS}^\perp} f = \sum_k \tilde{d}[k] y_k$. The particular choice $\tilde{d}[k] = \langle \tilde{y}_k^{\mathcal{S}}, f \rangle$ has some desirable properties which we now discuss.

Proposition 4. *Let $\{y_k, 1 \leq k \leq N\}$ denote a frame for an M-dimensional subspace $\mathcal{W} \subseteq \mathcal{H}$, and let $\mathcal{S} \subseteq \mathcal{H}$ denote an M-dimensional subspace such that $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$. Then from all possible coefficients $d[k]$ that satisfy*

$$E_{\mathcal{WS}^\perp} f = \sum_{k=1}^{N} d[k] y_k \tag{30}$$

for all $f \in \mathcal{H}$, the coefficients $\tilde{d}[k]$ corresponding to $\tilde{d} = Y^{\#}_{\mathcal{VS}^\perp} f$ with $\mathcal{V} = \mathcal{N}(Y)^\perp$ have minimal l_2-norm.

Proof. From (6) and Proposition 2 it follows that the coefficients $\tilde{d}[k]$ satisfy (30). Now, let d denote an arbitrary sequence satisfying (30). Then $\sum_k (d[k] - \tilde{d}[k]) y_k = 0$, which implies that

$d - \tilde{d} \in \mathcal{N}(Y)$. Since $\tilde{d} = Y^{\#}_{\mathcal{V}\mathcal{S}^\perp} f$, $\tilde{d} \in \mathcal{R}(Y^{\#}_{\mathcal{V}\mathcal{S}^\perp}) = \mathcal{N}(Y)^\perp$. Thus $d = \tilde{d} + y$ where $y \in \mathcal{N}(Y)$ so that $\langle \tilde{d}, y \rangle = 0$. Then, $||d||^2 = ||\tilde{d}||^2 + ||y||^2 \geq ||\tilde{d}||^2$, with equality if and only if $d = \tilde{d}$. $\square$

Since in Fig. 8, $\tilde{d}[k] = \langle \tilde{y}_k^{\mathcal{S}}, f \rangle$, it follows from Proposition 4 that these coefficients have minimal l_2-norm from all possible coefficients leading to consistent reconstruction.

We can consider the property stated in Proposition 4 from a slightly different point of view. Since the vectors y_k form a frame for $\mathcal{W}$, any $f \in \mathcal{W}$ can be expressed as $f = Yd$ for some d. However, since the vectors y_k are linearly dependent, d is not unique. The minimal norm coefficients are the unique coefficients that lie in $\mathcal{N}(Y)^\perp = \mathcal{V}$. We may express these coefficients as $d = Y^\dagger f$; indeed $Yd = YY^\dagger f = P_{\mathcal{W}} f = f$. Alternatively, $d = Y^{\#}_{\mathcal{V}\mathcal{S}^\perp} f$ where $\mathcal{S}^\perp$ is an arbitrary subspace of $\mathcal{H}$ such that $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$; indeed $Yd = YY^{\#}_{\mathcal{V}\mathcal{S}^\perp} f = E_{\mathcal{W}\mathcal{S}^\perp} f = f$. Thus, although the minimal norm coefficients $d[k]$ are unique, the resulting sampling vectors t_k such that $d[k] = \langle t_k, f \rangle$ are not unique. If in addition we impose the constraint that $t_k \in \mathcal{S}$, then the unique sampling vectors that result in coefficients with minimal norm correspond to $(Y^{\#}_{\mathcal{V}\mathcal{S}^\perp})^*$. This interpretation is useful in applications in which a signal $f \in \mathcal{W}$ is corrupted by noise that is known to lie in some subspace $\mathcal{S}^\perp$. By using appropriate sampling vectors in $\mathcal{S}$, we can totally eliminate this noise and at the same time recover the minimal norm coefficients.

Proposition 5. *Let $\hat{f} = \sum_{k=1}^{N} b[k] w_k$ for some vectors $\{w_k, 1 \leq k \leq N\}$ that form a frame for $\mathcal{W}$, and are to be determined. Let $\{t_k, 1 \leq k \leq N\}$ denote a given set of sampling vectors corresponding to T. Then the vectors w_k corresponding to the set transformation $(T^{\#}_{\mathcal{N}(T)^\perp \mathcal{W}^\perp})^*$ result in $\hat{f}$ with measurements $\langle t_k, \hat{f} \rangle$ that are as close as possible to $b[k]$ in a least-squares sense.*

Proof. Let z denote the measurements of $\hat{f}$ with sampling vectors t_k, so that $z = T^* W b$. Then $z \in \mathcal{R}(T^*) = \mathcal{N}(T)^\perp$. To minimize $||z - b||$ we need to choose a W such that $z = P_{\mathcal{N}(T)^\perp} b$, *i.e.*, such that $T^* W = P_{\mathcal{N}(T)^\perp}$. In addition we must have that $\mathcal{R}(W) = \mathcal{W}$. Let $W = (T^{\#}_{\mathcal{N}(T)^\perp \mathcal{W}^\perp})^*$. Then from (8), $T^* W = (T^{\#}_{\mathcal{N}(T)^\perp \mathcal{W}^\perp} T)^* = P_{\mathcal{N}(T)^\perp}$, and $\mathcal{R}(W) = \mathcal{N}(T^{\#}_{\mathcal{N}(T)^\perp \mathcal{W}^\perp})^\perp = \mathcal{W}$. $\square$

We conclude that the oblique dual frame vectors are very similar to the conventional dual frame vectors: Given a frame $\{y_k\}$ for $\mathcal{W}$, the dual frame vectors $\{\tilde{y}_k\}$ are the unique vectors in $\mathcal{W}$ such that any $f \in \mathcal{W}$ can be expressed as $f = \sum_k \langle \tilde{y}_k, f \rangle y_k$, and the coefficients $\langle \tilde{y}_k, f \rangle$ have minimal norm. Similarly, the oblique dual frame vectors of w_k on $\mathcal{S}$, with $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$, are the unique vectors in $\mathcal{S}$ such that any $f \in \mathcal{W}$ can be expressed as $f = \sum_k \langle \tilde{y}_k^{\mathcal{S}}, f \rangle y_k$, and the coefficients $\langle \tilde{y}_k^{\mathcal{S}}, f \rangle$ have minimal norm. Thus, using the concept of oblique dual frame vectors we can extend the notion of a frame expansion to the case in which the analysis frame vectors are not constrained to lie in $\mathcal{W}$, but rather may lie in an arbitrary subspace $\mathcal{S} \subseteq \mathcal{H}$, with $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$.

It is interesting to note that the oblique dual frame vectors of $\tilde{y}_k^{\mathcal{S}}$ on $\mathcal{W}$ are the vectors y_k. Thus not only do we have $f = \sum_{k=1}^N \langle \tilde{y}_k^{\mathcal{S}}, f \rangle y_k$ for any $f \in \mathcal{W}$ but also $f = \sum_{k=1}^N \langle y_k, f \rangle \tilde{y}_k^{\mathcal{S}}$ for any $f \in \mathcal{S}$.

Proposition 6. *Let $T = (Y^{\#}_{\mathcal{V}\mathcal{S}^\perp})^*$ denote the set transformation corresponding to frame vectors $\{\tilde{y}_k^{\mathcal{S}} \in \mathcal{S}, 1 \leq k \leq N\}$, where Y is a set transformation corresponding to frame vectors $\{y_k \in \mathcal{W}, 1 \leq k \leq N\}$. Then the oblique dual frame vectors of $\tilde{y}_k^{\mathcal{S}}$ on $\mathcal{W}$ are the vectors y_k.*

Proof. By definition the oblique dual frame vectors of $\tilde{y}_k^{\mathcal{S}}$ on $\mathcal{W}$ are the vectors corresponding to $(T^{\#}_{\mathcal{U}\mathcal{W}^\perp})^*$ where $\mathcal{U} = \mathcal{N}(T)^\perp$. Thus we need to show that $Y = (T^{\#}_{\mathcal{U}\mathcal{W}^\perp})^*$, or equivalently $Y^* = T^{\#}_{\mathcal{U}\mathcal{W}^\perp}$, which based on (7)–(9) reduces to proving that

$$TY^* = E_{\mathcal{R}(T)\mathcal{W}^\perp}; \tag{31}$$

$$Y^*T = P_{\mathcal{U}}; \tag{32}$$

$$\mathcal{R}(Y^*) = \mathcal{U}. \tag{33}$$

First, we note that $\mathcal{U} = \mathcal{N}(T)^\perp = \mathcal{N}((Y^{\#}_{\mathcal{V}\mathcal{S}^\perp})^*)^\perp = \mathcal{R}(Y^{\#}_{\mathcal{V}\mathcal{S}^\perp}) = \mathcal{V}$. Now, $\mathcal{R}(Y^*) = \mathcal{N}(Y)^\perp = \mathcal{V}$ so that (33) is satisfied. Next, $TY^* = (Y^{\#}_{\mathcal{V}\mathcal{S}^\perp})^* Y^* = (YY^{\#}_{\mathcal{V}\mathcal{S}^\perp})^* = (E_{\mathcal{W}\mathcal{S}^\perp})^* = E_{\mathcal{S}\mathcal{W}^\perp}$. Furthermore, $\mathcal{R}(T) = \mathcal{R}((Y^{\#}_{\mathcal{V}\mathcal{S}^\perp})^*) = \mathcal{N}(Y^{\#}_{\mathcal{V}\mathcal{S}^\perp})^\perp = \mathcal{S}$ so (31) is also satisfied. Finally, $Y^*T = Y^*(Y^{\#}_{\mathcal{V}\mathcal{S}^\perp})^* = (Y^{\#}_{\mathcal{V}\mathcal{S}^\perp}Y)^* = P_{\mathcal{V}}$ completing the proof. $\square$

9 Summary of Consistent Sampling and Reconstruction

We summarize our results regarding consistent reconstruction in the following theorem:

Theorem 2 (Consistent sampling and reconstruction). *Let $\{c[k] = \langle x_k, f\rangle, 1 \leq k \leq N\}$ denote measurements of a signal $f \in \mathcal{H}$ with sampling vectors $\{x_k, 1 \leq k \leq N\}$ that form a frame for an M-dimensional subspace $\mathcal{S} \subseteq \mathcal{H}$. Let $\{y_k, 1 \leq k \leq N\}$ denote a set of reconstruction vectors that form a frame for an M-dimensional subspace $\mathcal{W} \subseteq \mathcal{H}$, with $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$. Then any $f \in \mathcal{H}$ can be consistently reconstructed from the measurements $c[k]$ using the reconstruction vectors y_k as $\hat{f} = \sum_{k=1}^N d[k] y_k$ with $d = (X^*Y)^\dagger c$, and the consistent reconstruction is unique. In addition,*

1. *$d = Y^{\#}_{\mathcal{V}\mathcal{S}^\perp} f$ where $Y^{\#}_{\mathcal{V}\mathcal{S}^\perp}$ is the oblique pseudoinverse on $\mathcal{V} = \mathcal{N}(Y)^\perp$ along $\mathcal{S}^\perp$ so that $d[k] = \langle \tilde{y}_k^{\mathcal{S}}, f\rangle$ where the vectors $\{\tilde{y}_k^{\mathcal{S}}, 1 \leq k \leq N\}$ are the oblique dual frame vectors of y_k on $\mathcal{S}$ and correspond to $X(Y^*X)^\dagger$.*

2. *If $M = N$, then*

   (a) the coefficients $d[k]$ are unique;

   (b) $\{\tilde{y}_k^{\mathcal{S}}, 1 \leq k \leq M\}$ are the unique vectors in $\mathcal{S}$ biorthogonal to $\{w_k, 1 \leq k \leq M\}$.

3. *If $M < N$, then*

   (a) the coefficients $d[k]$ are not unique;

   (b) the coefficients $d[k]$ have minimal l_2-norm among all possible coefficients $e[k]$ such that $\hat{f} = \sum_{k=1}^N e[k] y_k$.

10 Reducing Quantization Error

One of the reasons for using redundant measurements is to reduce the average power of the quantization error, when quantizing the corrected measurements prior to reconstruction. If $\mathcal{S} = \mathcal{W}$, then it is well known that using a redundant procedure the quantization error can be reduced by the redundancy of the frame [6, 14]. We now extend this result to the case in which $\mathcal{S} \neq \mathcal{W}$. In particular, we show that we can choose a normalized tight frame y_k for $\mathcal{W}$ such that when using the redundant sampling procedure of Figs. 8 and 9 we can reduce the average power of the reconstruction error by the redundancy, in comparison with the nonredundant scheme of Fig. 7.

Let $\{w_k, 1 \leq k \leq M\}$ denote reconstruction vectors that form an orthonormal basis for $\mathcal{W}$, and let $\{s_k, 1 \leq k \leq M\}$ denote sampling vectors that form a basis for $\mathcal{S}$. From Theorem 2, $\hat{f} = \sum_k d[k]w_k$ where $d[k] = \langle v_k, f\rangle$, and $\{v_k, 1 \leq k \leq M\}$ are the vectors corresponding to $V = S(W^*S)^{-1}$. Thus,

$$\hat{f} = \sum_{k=1}^{M} \langle v_k, f\rangle w_k = \sum_{k=1}^{M} q[k]\langle \overline{v}_k, f\rangle \overline{w}_k, \tag{34}$$

where $q[k] = \sqrt{a[k]b[k]}$, $a[k] = \langle w_k, w_k\rangle = 1$, $b[k] = \langle v_k, v_k\rangle$, $\overline{w}_k = w_k/\sqrt{a[k]}$, and $\overline{v}_k = v_k/\sqrt{b[k]}$.

Assume we quantize the normalized measurements $\overline{d}[k] = \langle \overline{v}_k, f\rangle$ prior to reconstruction, and model the quantization error as an additive zero-mean white noise source, so that the quantized measurements are given by $\overline{d}[k]' = \overline{d}[k] + e[k]$ where $E(e[k]e[j]) = \sigma^2\delta_{kj}$. The reconstruction error is then $\epsilon = \sum_{k=1}^{M} q[k]e[k]\overline{w}_k$ and the average power of the reconstruction error, denoted by D, is

$$D = E(\langle \epsilon, \epsilon\rangle) = \sigma^2 \sum_{k=1}^{M} q^2[k] = \sigma^2 \sum_{k=1}^{M} b[k]. \tag{35}$$

Suppose now we use a redundant procedure so that we reconstruct the signal using a normalized tight frame $\{y_k, 1 \leq k \leq N\}$ for $\mathcal{W}$, with redundancy N/M. Then $Y = WZ$ for some Z: $\mathbb{C}^M \to \mathbb{C}^N$ such that $ZZ^* = I_M$. From Theorem 2, the sampling vectors leading to consistent reconstruction correspond to $X = (Y^{\#}_{\mathcal{V}\mathcal{S}^{\perp}})^* = S(W^*S)^{-1}Z$, so that in this case

$$\hat{f} = \sum_{k=1}^{N} \langle x_k, f\rangle y_k = \sum_{k=1}^{N} \tilde{q}[k]\langle \overline{x}_k, f\rangle \overline{y}_k, \tag{36}$$

where $\tilde{q}[k] = \sqrt{\tilde{a}[k]\tilde{b}[k]}$, $\tilde{a}[k] = \langle y_k, y_k \rangle$, $\tilde{b}[k] = \langle x_k, x_k \rangle$, $\overline{y}_k = y_k/\sqrt{\tilde{a}[k]}$, and $\overline{x}_k = x_k/\sqrt{\tilde{b}[k]}$. If we quantize the normalized redundant measurements $\langle \overline{x}_k, f \rangle$ and model the error as before, then the average power of the reconstruction error using the redundant procedure, denoted by $\tilde{D}$, is

$$\tilde{D} = \sigma^2 \sum_{k=1}^{N} \tilde{q}^2[k] = \sigma^2 \sum_{k=1}^{N} \tilde{a}[k]\tilde{b}[k]. \tag{37}$$

We now show that we can choose a normalized tight frame y_k such that $\tilde{D} = (M/N)D$.

Let $Y = W\tilde{\mathcal{F}}$, where $\tilde{\mathcal{F}}$ is an $M \times N$ matrix whose rows are equal to the first M rows of the $N \times N$ Fourier matrix $\mathcal{F}$ with elements $(1/\sqrt{N})e^{-j2\pi km/N}$. Since $YY^* = P_{\mathcal{W}}$, the corresponding vectors y_k form a normalized tight frame for $\mathcal{W}$. The oblique dual frame vectors x_k of y_k on $\mathcal{S}$ are

$$X = (Y^{\#}_{\mathcal{V}\mathcal{S}^{\perp}})^* = V(W^*V)^{-1}\tilde{\mathcal{F}} = V\tilde{\mathcal{F}}. \tag{38}$$

Let f_k denote the kth column of $\tilde{\mathcal{F}}$. From the definition of $\tilde{\mathcal{F}}$, $\langle f_k, f_k \rangle = M/N$ for all k so that,

$$a_k = \langle y_k, y_k \rangle = \langle Wf_k, Wf_k \rangle = \langle f_k, f_k \rangle = \frac{M}{N}, \tag{39}$$

since $W^*W = I_M$, and (37) reduces to

$$\tilde{D} = \sigma^2 \frac{M}{N} \sum_{k=1}^{N} \tilde{b}[k]. \tag{40}$$

Now,

$$\sum_{k=1}^{N} \tilde{b}[k] = \mathrm{Tr}(X^*X) = \mathrm{Tr}(V^*V) = \sum_{k=1}^{M} b[k]. \tag{41}$$

Substituting (41) into (40), and comparing with (35) we conclude that $\tilde{D} = (M/N)D$.

Therefore, to reduce the quantization error in the sampling and reconstruction scheme of Fig. 7, we propose the following. Instead of directly quantizing the measurements $d[k]$ in Fig. 7, we first take the N point DFT of the length M sequence of measurements $d[k]$, and then quantize the DFT coefficients. The reconstructed signal is then a linear combination of the reconstruction vectors w_k, where the coefficients are the first M values of the inverse DFT of the quantized DFT coefficients, as depicted in Fig. 10. If we take out the quantizer in Fig. 10, then $\hat{f} = E_{\mathcal{W}\mathcal{S}^\perp} f$ as in Fig. 7. However, in the presence of the quantizer, using the redundant sampling scheme of Fig. 10 the average power of the quantization error is reduced by a factor of N/M in comparison with a nonredundant scheme.

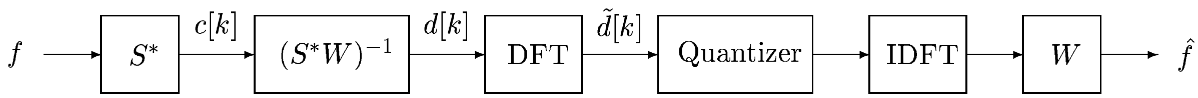


Figure 10: Reconstruction of f from quantized measurements using a redundant sampling scheme.

There are many other choices of frame vectors y_k for $\mathcal{W}$ and oblique dual frame vectors x_k on $\mathcal{S}$, that lead to a noise reduction of N/M in comparison with a basis expansion. We refer to any such frame expansion as a 'good' oblique frame expansion. We then have the following theorem.

Theorem 3. *Let $\{w_k, 1 \leq k \leq M\}$ denote an orthonormal basis for $\mathcal{W}$, and let $\{v_k, 1 \leq k \leq M\}$ denote the biorthogonal basis for $\mathcal{S}$, with $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$. Let $\{y_k, 1 \leq k \leq N\}$ denote a frame for $\mathcal{W}$, and let $\{x_k, 1 \leq k \leq N\}$ denote the oblique dual frame vectors of y_k on $\mathcal{S}$. Let $\mathcal{F}$ denote the $N \times N$ Fourier matrix, and let $\tilde{\mathcal{F}}$ denote the first M rows of $\mathcal{F}$. Then*

1. *The frame vectors corresponding to $Y = W\tilde{\mathcal{F}}$, $X = V\tilde{\mathcal{F}}$ form a good oblique frame expansion;*

2. *The frame vectors corresponding to $Y = W\tilde{\mathcal{F}}T$, $X = V\tilde{\mathcal{F}}T$ where T is a unitary circulant matrix[3] form a good oblique frame expansion.*

Proof. We already proved the first part of the theorem; it remains to prove the second part. We can immediately verify that $X = V\tilde{\mathcal{F}}T$ is in fact the oblique dual frame operator on $\mathcal{S}$ of $Y = W\tilde{\mathcal{F}}T$.

[3] *A circulant matrix is a matrix where every row (or column) is obtained by a right circular shift (by one position) of the previous row (or column).*

Since T is circulant and unitary it is diagonalized by $\mathcal{F}^*$ [8], so we can express T as $T = \mathcal{F}^*\Lambda\mathcal{F}$ where Λ is a diagonal matrix with diagonal elements λ_k with $|\lambda_k| = 1$ for all k. Then,

$$Y = W\tilde{\mathcal{F}}T = W\tilde{\mathcal{F}}\mathcal{F}^*\Lambda\mathcal{F} = W\tilde{I}\Lambda\mathcal{F}, \tag{42}$$

where $\tilde{I} = [I_M \ 0]$, and

$$Y^*Y = \mathcal{F}^*\Lambda^*\tilde{I}^*\tilde{I}\Lambda\mathcal{F} = \mathcal{F}^*\tilde{I}^*\tilde{I}\mathcal{F} = \tilde{\mathcal{F}}^*\tilde{\mathcal{F}}. \tag{43}$$

Combining (43) and (39), we have $\langle y_k, y_k\rangle = M/N$ for all k. From (37), the average power of the reconstruction error using $\{y_k, x_k\}$ is $\tilde{D} = \sigma^2 M/N \sum_{k=1}^{N} \langle x_k, x_k\rangle$. Now, $X = V\tilde{\mathcal{F}}T = V\tilde{I}\Lambda\mathcal{F}$ so

$$\sum_{k=1}^{N} \langle x_k, x_k\rangle = \mathrm{Tr}(X^*X) = \mathrm{Tr}(\mathcal{F}^*\Lambda^*\tilde{I}^*V^*V\tilde{I}\Lambda\mathcal{F}) = \mathrm{Tr}(V^*V), \tag{44}$$

and $\tilde{D} = (M/N)D$ where D is the average power of the reconstruction error using $\{w_k, v_k\}$. $\square$

Based on results derived in [8, 11, 10] we can show that Theorem 3 still holds when we replace $\mathcal{F}$ by a generalized Fourier matrix defined on a direct product of cyclic groups (*e.g.,* a Hadamard matrix), and replace T by a unitary permuted matrix whose rows and columns are all permutations of each other. This is because a permuted matrix is diagonalized by a generalized Fourier matrix, and the magnitude of the elements of an $N \times N$ generalized Fourier matrix are all equal $1/\sqrt{N}$.

11 Acknowledgment

The author wishes to thank the associate editor and the anonymous reviewers for valuable suggestions that greatly improved the presentation.

References

[1] A. Aldroubi. Portraits of frames. *Proc. Amer. Math. Soc.*, 123:1661–1668, 1995.

[2] A. Aldroubi. Oblique projections in atomic spaces. *Proc. Amer. Math. Soc.*, 124(7):2051–2060, 1996.

[3] A. Aldroubi and M. Unser. Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory. *Numer. Funct. Anal. Optimiz.*, 15:1–21, Feb. 1994.

[4] T. Blu and M. Unser. Quantitative Fourier analysis of approximation techniques: Part I - Interpolators and projectors. *IEEE Trans. Signal Processing*, 47(10):2783–2795, Oct. 1999.

[5] T. Blu and M. Unser. Quantitative Fourier analysis of approximation techniques: Part II - Wavelets. *IEEE Trans. Signal Processing*, 47(10):2796–2806, Oct. 1999.

[6] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, Philadelphia, 1992.

[7] Y. C. Eldar. Constructing signals with prescribed properties. submitted to *IEEE Signal Proc. Lett.*, Jan. 2002.

[8] Y. C. Eldar. Least-squares inner product shaping. *Linear Algebra Appl.*, to appear.

[9] Y. C. Eldar. *Quantum Signal Processing*. PhD thesis, Massachusetts Institute of Technology, Dec. 2001.

[10] Y. C. Eldar and H. Bölcskei. Geometrically uniform frames. submitted to *IEEE Trans. Inform. Theory*, August 2001; also available at `http://arXiv.org/abs/math.FA/0108096`.

[11] Y. C. Eldar and G. D. Forney, Jr. On quantum detection and the square-root measurement. *IEEE Trans. Inform. Theory*, 47:858–872, Mar. 2001.

[12] Y. C. Eldar and A. V. Oppenheim. Nonredundant and redundant sampling with arbitrary sampling and reconstruction spaces. *Proc. of SampTA*, pages 229–234, May 2001.

[13] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[14] V. K. Goyal, M. Vetterli, and N. T. Thao. Quantized overcomplete expansions in $\mathcal{R}^N$: Analysis, synthesis, and algorithms. *IEEE Trans. Inform. Theory*, 44:16–31, Jan. 1998.

[15] C. E. Heil and D. F. Walnut. Continuous and discrete wavelet transforms. *SIAM Rev.*, 31(4):628–666, Dec. 1989.

[16] K. Hoffman and R. Kunze. *Linear Algebra*. Prentice-Hall, Inc., second edition, 1971.

[17] J. A. Hogan and J. Lakey. Sampling and aliasing without translation-invariance. *Proc. of SampTA*, pages 61–66, May 2001.

[18] A. J. E. M. Janssen. The Zak transform and sampling theorems for wavelet subspaces. *IEEE Trans. Signal Processing*, 41(12):3360–3364, Dec. 1993.

[19] G. Kaiser. *A Friendly Guide to Wavelets.* Birkhauser, 1994.

[20] S. Kayalar and H. L. Weinert. Oblique projections: Formulas, algorithms, and error bounds. *Math. Contr. Signals Syst.*, 2(1):33–45, 1989.

[21] N. Macon and A. Spitzbart. Inverses of Vandermonde matrices. *American Mathematical Monthly*, 65(2):95–100, Feb. 1958.

[22] S. G. Mallat. A theory of multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11:674–693, 1989.

[23] R. D. Milne. An oblique matrix pseudoinverse. *SIAM J. Appl. Math.*, 16(5):931–944, Sep. 1968.

[24] A. Papoulis. A new algorithm in spectral analysis and bandlimited extrapolation. *IEEE Trans. Circuits Syst.*, CAS-22:735–742, 1975.

[25] F. D. Parker. Inverses of Vandermonde matrices. *American Mathematical Monthly*, 71(4):410–411, Apr. 1964.

[26] H. Stark and Y. Yang. *Vector Space Projections.* John Wiley and Sons, Inc., 1998.

[27] M. Unser. Splines: A perfect fit for signal and image processing. *IEEE Signal Processing Mag.*, pages 22–38, Nov. 1999.

[28] M. Unser. Sampling—50 years after Shannon. *IEEE Proc.*, 88:569–587, Apr. 2000.

[29] M. Unser and A. Aldroubi. A general sampling theory for nonideal acquisition devices. *IEEE Trans. Signal Processing*, 42(11):2915–2925, Nov. 1994.

[30] M. Unser and J. Zerubia. Generalized sampling: Stability and performance analysis. *IEEE Trans. Signal Processing*, 45(12):2941–2950, Dec. 1997.