Optimal Tight Frames and Quantum Measurement

Yonina C. Eldar, *Student Member, IEEE*, and G. David Forney, Jr., *Fellow, IEEE*

***Abstract*—Tight frames and rank-one quantum measurements are shown to be intimately related. In fact, the family of normalized tight frames for the space in which a quantum-mechanical system lies is precisely the family of rank-one generalized quantum measurements on that space. Using this relationship, frame-theoretical analogs of various quantum-mechanical concepts and results are developed. The analog of a least-squares quantum measurement is a tight frame that is closest in a least-squares sense to a given set of vectors. The least-squares tight frame is found for both the case in which the scaling of the frame is specified (constrained least-squares frame (CLSF)) and the case in which the scaling is chosen to minimize the least-squares error (unconstrained least-squares frame (ULSF)). The well-known canonical frame is shown to be proportional to the ULSF and to coincide with the CLSF with a certain scaling.**

***Index Terms*—Canonical frames, least-squares frame, least-squares quantum measurement, Neumark's theorem, tight frames.**

I. Introduction

FRAMES are generalizations of bases which lead to redundant signal expansions [1], [2]. A frame for a Hilbert space $\mathcal{U}$ is a set of not necessarily linearly independent vectors that spans $\mathcal{U}$ and has some additional properties. Frames were first introduced by Duffin and Schaeffer [1] in the context of nonharmonic Fourier series, and play an important role in the theory of nonuniform sampling [1]–[3]. Recent interest in frames has been motivated in part by their utility in analyzing wavelet expansions [4], [5].

Many efforts have been made to construct bases with specified properties. Since the conditions on bases are quite stringent, in many applications it is hard to find "good" bases. The conditions on frame vectors are usually not as stringent, allowing for increased flexibility in their design [4], [6]. For example, frame expansions admit signal representations that are localized in both time and frequency [5], as well as sparse representations [7].

Frame expansions have many other desirable properties. The coefficients may be computed with less precision than the coefficients in a basis expansion for a given desired reconstruction precision [5]; the effect of additive noise on the coefficients on the reconstructed signal is reduced in comparison with a basis expansion [5], [8]–[10]; and the coefficients are more robust to quantization degradations [11], [12]. Recently, frames have been applied to the development of modern uniform and nonuniform sampling techniques [13], to various detection problems [14], [15], and to the analysis and design of packet-based communication systems [16].

A *tight frame* is a special case of a frame for which the reconstruction formula is particularly simple. As we show in Section IV, a tight frame expansion of a signal is reminiscent of an orthogonal basis expansion, even though the frame vectors in the expansion are linearly dependent. Tight frames are particularly popular, and will be the focus of this paper.

Frame-like expansions have been developed and used in a wide range of disciplines. Many connections between frame theory and various signal processing techniques have been recently discovered and developed. For example, the theory of frames has been used to analyze and design oversampled filter banks [17]–[20] and error-correction codes [21]. Wavelet families have been used in quantum mechanics and many other areas of theoretical physics, particularly in the study of semiclassical approximations to quantum mechanics [5].

In this paper, we explore yet another connection between quantum mechanics and tight frames. Specifically, we show that the family of (normalized) tight frames for a subspace $\mathcal{U}$ in which a quantum-mechanical system is known to lie is precisely the family of possible positive operator-valued measures (POVMs) on $\mathcal{U}$. Exploiting this equivalence, we can apply ideas and results derived in the context of quantum measurement to the theory of frames and *vice versa*.

We begin in Section III by characterizing quantum measurements. With each rank-one quantum measurement we associate a measurement matrix. Using the measurement matrix representation, we give a simple and constructive proof of Neumark's theorem [22], [23], which relates general quantum measurements to orthogonal measurements. We then discuss the problem of constructing measurements optimized to distinguish between a set of nonorthogonal pure quantum states.

We then follow a similar path in Section IV for tight frames. We associate a frame matrix with every tight frame, which as we show has essentially the same properties as a quantum measurement matrix. Next, we derive an analog of Neumark's theorem for tight frames, which expresses tight frame vectors as orthogonal projections of a set of orthogonal vectors in a larger space. Finally, motivated by the construction of optimal quantum measurements, we consider the problem of constructing optimal tight frames for a subspace $\mathcal{U}$ from a given set of vectors that span $\mathcal{U}$.

The problem of frame design has received relatively little attention in the frame literature. Typically, in applications, the frame vectors are chosen rather than optimized. A popular frame

Manuscript received June 12, 2001; revised October 16, 2001.

Y. C. Eldar is with the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: yonina@mit.edu).

G. D. Forney, Jr. is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: forneyd@mediaone.net).

Communicated by P. W. Shor, Associate Editor for Quantum Information Theory.

Publisher Item Identifier S 0018-9448(02)01060-X.

construction from a given set of vectors is the canonical frame [8], [19], [24]–[27], first proposed in the context of wavelets in [28]. The canonical frame is relatively simple to construct, can be determined directly from the given vectors, in many cases of interest inherits stability and symmetry properties of the original vector set [24], [27], and plays an important role in wavelet theory [29]–[31]. Some optimality properties of canonical frames have been discussed in [26].

In Section V, we systematically construct optimal frames from a given set of vectors. Motivated by the least-squares measurement [32] derived for quantum detection, we seek a tight frame consisting of frame vectors that minimize the sum of the squared norms of the error vectors, where the ith error vector is defined as the difference between the ith given vector and the ith frame vector. We consider both the case in which the scaling of the frame is specified and the case in which the scaling is such that the error is minimized. When the scaling is specified, the optimizing frame is referred to as the constrained least-squares frame (CLSF), and when the scaling is chosen to minimize the error, the optimizing frame is referred to as the unconstrained least-squares frame (ULSF).

In Section VII, we show that the canonical frame vectors are proportional to the ULSF vectors, and that they coincide with the CLSF vectors with a specific choice of scaling.

Before proceeding to the detailed development, in Section II, we first provide an overview of the notation and some mathematical preliminaries.

II. Preliminaries

In this section, we briefly review elements of linear algebra that are common to both signal processing and quantum mechanics. Our main goal is to characterize "transjectors" (partial isometries) using the singular value decomposition (SVD).

A. Hilbert Spaces and Operators

In both signal processing and quantum mechanics, the setting we consider is a finite-dimensional subspace $\mathcal{U}$ of a complex Hilbert space $\mathcal{H}$. The elements of $\mathcal{H}$ are called vectors. We will assume, for notational convenience, that $\mathcal{H}$ is finite-dimensional, with $\dim \mathcal{H} = k$; then, by appropriate choice of coordinates, we can identify $\mathcal{H}$ with $\mathbb{C}^k$.

In signal processing, the elements of $\mathcal{H}$ are regarded as column vectors and denoted, e.g., by $x \in \mathcal{H}$. Then x^* denotes the row vector which is the conjugate transpose of x. The inner product of two vectors is a complex number, denoted, e.g., by $\langle x, y\rangle = x^* y$. An outer product of two vectors such as xy^* is a rank-one matrix, which as an operator takes $z \in \mathcal{H}$ to $xy^* z = \langle y, z\rangle x \in \mathcal{H}$.

The Dirac bra-ket notation of quantum mechanics expresses such concepts very nicely; however, recognizing that it is unfamiliar, we do not rely on it in this paper. Nonetheless, to assist the reader unfamiliar with this notation in reading the quantum literature, we will give the bra-ket equivalents for various expressions in this section.

In the bra-ket notation, the elements of $\mathcal{H}$ are "ket" vectors, denoted, e.g., by $|x\rangle \in \mathcal{H}$. The corresponding "bra" vector $\langle x|$ is an element of the dual space $\mathcal{H}^*$ and may be regarded as the conjugate transpose of $|x\rangle$. The inner product of two vectors is a complex number denoted by $\langle x|y\rangle$. An outer product of two vectors such as $|x\rangle\langle y|$ is a rank-one matrix, which as an operator takes $|z\rangle \in \mathcal{H}$ to $|x\rangle\langle y||z\rangle = \langle y|z\rangle|x\rangle \in \mathcal{H}$.

An operator on $\mathcal{H}$ is a continuous linear transformation A: $\mathcal{H} \to \mathcal{H}$. The adjoint of a linear operator A is the unique operator A^* such that $\langle x, Ay\rangle = \langle A^* x, y\rangle$ for all $x, y \in \mathcal{H}$. If the elements of $\mathcal{H}$ are column vectors, then a linear operator A is represented by a square matrix, and its adjoint is represented by the conjugate transpose A^*, since

$$\langle x, Ay\rangle = x^* A y = (A^* x)^* y = \langle A^* x, y\rangle.$$

An operator A is called *Hermitian* if it is self-adjoint; i.e., if $A^* = A$.

An orthogonal (Hermitian) projector P is a Hermitian operator on $\mathcal{H}$ such that $P^2 = P$; all projections used in this paper will be orthogonal projections. The eigenvalues of P are all equal 0 or 1, and P has an orthonormal set of eigenvectors. If $\{u_i\}$ is a set of orthonormal eigenvectors corresponding to the nonzero eigenvalues of P, then the subspace $\mathcal{U} \subseteq \mathcal{H}$ spanned by the set $\{u_i\}$ is the range of P, and we write the orthogonal projector as $P_{\mathcal{U}}$. A one-dimensional orthogonal projector has a single normalized eigenvector u and may be written as the outer product $P_u = uu^*$ (or $P_u = |u\rangle\langle u|$ in bra-ket notation); then, P_u projects any $x \in \mathcal{H}$ into the projection $P_u x = \langle u, x\rangle u$ (or $|u\rangle\langle u|x\rangle$). An r-dimensional orthogonal projector $P_{\mathcal{U}}$ may be written as the sum of r one-dimensional orthogonal projectors, $P_{\mathcal{U}} = \sum_i P_{u_i}$, where $\{u_i\}$ is any orthonormal basis for $\mathcal{U}$.

B. Transjectors (Partial Isometries)

Let F be a rank-r matrix whose columns are a set of n vectors $\varphi_i \in \mathcal{H}$. It is well known in signal processing (but not as well known in quantum mechanics[1]) that any such matrix F has an SVD $F = U\Sigma V^*$, where U is a unitary matrix whose columns $\{u_i \in \mathcal{H}\}$ are eigenvectors of the Hermitian operator $T = FF^*$, V is a unitary matrix whose columns $\{v_i \in \mathbb{C}^n\}$ are eigenvectors of the Hermitian matrix $S = F^* F$ (the Gram matrix of inner products), and Σ is a real diagonal matrix whose r nonzero values σ_i, called the *singular values* of F, are the positive square roots of the nonzero eigenvalues of either S or T. Thus, we may write $F = \sum_{i=1}^r \sigma_i u_i v_i^*$ (or $F = \sum_{i=1}^r \sigma_i |u_i\rangle\langle v_i|$), a sum of r rank-1 outer products.

An outer product such as $u_i v_i^*$ (or $|u_i\rangle\langle v_i|$) is called a one-dimensional *transjector*. The transjector $u_i v_i^*$ takes a basis vector $v_i \in \mathbb{C}^n$ to the corresponding basis vector $u_i \in \mathcal{H}$. By linear superposition, it therefore takes a general element $x = \sum_j \langle v_j, x\rangle v_j \in \mathbb{C}^n$ to $u_i v_i^* x = \langle v_i, x\rangle u_i \in \mathcal{H}$. Similarly, the adjoint transjector $v_i u_i^*$ takes $y = \sum_j \langle u_j, y\rangle u_j \in \mathcal{H}$ to $v_i u_i^* y = \langle u_i, y\rangle v_i \in \mathbb{C}^n$.

The subspace spanned by the r orthonormal eigenvectors $u_i \in \mathcal{H}$ corresponding to the r nonzero eigenvalues of $S = F^* F$ will be denoted as $\mathcal{U} \subseteq \mathcal{H}$, and the subspace spanned by the r orthonormal eigenvectors $v_i \in \mathbb{C}^n$ corresponding to the r nonzero eigenvalues of $T = FF^*$ will be denoted as $\mathcal{V} \subseteq \mathbb{C}^n$. The image

[1]The SVD has sometimes been presented in quantum mechanics as a corollary of the polar decomposition (e.g., in [33, Appendix A]).

of F is $\mathcal{U}$, and the image of F^* is $\mathcal{V}$; the kernel of F is the orthogonal complement $\mathcal{V}^\perp$ of $\mathcal{V}$, and the kernel of F^* is $\mathcal{U}^\perp$. F operates by first performing an orthonormal expansion of $\mathbb{C}^n$ using the orthonormal basis $\{v_i\}$, scaling each component by σ_i, and then "transjecting" to $\mathcal{U} \subseteq \mathcal{H}$ by replacing each v_i by the corresponding u_i. F^* similarly "transjects" from $\mathcal{H}$ to $\mathcal{V} \subseteq \mathbb{C}^n$.

A rank-r matrix F is called an *r-dimensional transjector* if its r nonzero singular values are all equal to 1. In other words, $F = UZ_rV^*$, where U and V are unitary and

$$Z_r = \overbrace{\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]}^{n}. \tag{1}$$

Equivalently, $FF^* = U(Z_rZ_r^*)U^* = P_{\mathcal{U}}$ is an r-dimensional orthogonal projector onto an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$ with an orthonormal basis $\{u_i \in \mathcal{H}, 1 \le i \le r\}$ (the $\mathcal{U}$-basis) consisting of the first r columns of U, and $F^*F = V(Z_r^*Z_r)V^* = P_{\mathcal{V}}$ is an r-dimensional orthogonal projector onto an r-dimensional subspace $\mathcal{V} \subseteq \mathbb{C}^n$ with an orthonormal basis $\{v_i \in \mathbb{C}^n, 1 \le i \le r\}$ (the $\mathcal{V}$-basis) consisting of the first r columns of V.

The SVD $F = UZ_rV^*$ thus reduces to a sum of r one-dimensional transjectors (outer products)

$$F = \sum_{i=1}^{r} u_i v_i^*. \tag{2}$$

An r-dimensional transjector F is also called a *partial isometry*, because it is an isometry (distance-preserving transformation) between the subspaces $\mathcal{U} \subseteq \mathcal{H}$ and $\mathcal{V} \subseteq \mathbb{C}^n$. Indeed, if v, $v' \in \mathcal{V}$ and $u = Fv$, $u' = Fv'$, then

$$\langle u, u'\rangle = u^*u' = v^*F^*Fv' = v^*P_{\mathcal{V}}v' = v^*v' = \langle v, v'\rangle \tag{3}$$

so inner products and *a fortiori* squared norms and distances are preserved. Similarly, if $u, u' \in \mathcal{U}$, then $\langle F^*u, F^*u'\rangle = \langle u, u'\rangle$. However, inner products are not preserved if u, $u' \notin \mathcal{U}$ or v, $v' \notin \mathcal{V}$.

This discussion is summarized in the following theorem.

Theorem 1 (Transjectors (Partial Isometries)): The following statements are equivalent for a matrix F whose columns are n vectors in a complex Hilbert space $\mathcal{H}$:

1) F is a transjector (partial isometry) between r-dimensional subspaces $\mathcal{U} \subseteq \mathcal{H}$ and $\mathcal{V} \subseteq \mathbb{C}^n$;
2) $FF^* = P_{\mathcal{U}}$ for an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$;
3) $F^*F = P_{\mathcal{V}}$ for an r-dimensional subspace $\mathcal{V} \subseteq \mathbb{C}^n$.

A transjector F between r-dimensional subspaces $\mathcal{U} \subseteq \mathcal{H}$ and $\mathcal{V} \subseteq \mathbb{C}^n$ may be expressed as $F = UZ_rV^*$, where U is a unitary matrix whose first r columns $\{u_i, 1 \le i \le r\}$ are an orthonormal basis for $\mathcal{U}$, V is an $n \times n$ unitary matrix whose first r columns $\{v_i, 1 \le i \le r\}$ are an orthonormal basis for $\mathcal{V}$, and Z_r is given by (1). Equivalently, $F = \sum_{i=1}^{r} u_i v_i^*$.

A transjector $F: \mathbb{C}^n \to \mathcal{U}$ (resp., $F^*: \mathcal{H} \to \mathcal{V}$) is an isometry if restricted to $\mathcal{V}$ (resp., $\mathcal{U}$).

III. Quantum Measurement

In this section, we present some elements of the theory of quantum measurement following [32] and unpublished work in [34]. In the remainder of the paper, we will develop analogous results for tight frames.

A quantum system in a pure state is characterized by a normalized vector ψ in a Hilbert space $\mathcal{H}$. Information about a quantum system is extracted by subjecting the system to a measurement. In quantum theory, the outcome of a measurement is inherently probabilistic, with the probabilities of the outcomes of any conceivable measurement determined by the state vector $\psi \in \mathcal{H}$.

A quantum measurement is described by a collection of Hermitian operators $\{Q_i\}$ on $\mathcal{H}$, where the index i corresponds to a possible measurement outcome. The laws of quantum mechanics impose certain mathematical constraints on the measurement operators.

In the simplest case, the measurement operators are rank-one operators and have the outer-product form $Q_i = \mu_i\mu_i^*$ for some nonzero vectors $\mu_i \in \mathcal{H}$. Such measurements will be called rank-one measurements, and the vectors μ_i will be called the measurement vectors.

If the state vector is ψ, then the probability of observing the ith outcome is

$$p(i) = \langle \psi, Q_i\psi\rangle = |\langle \mu_i, \psi\rangle|^2. \tag{4}$$

To ensure that the probabilities $p(i)$ sum to 1 for any normalized $\psi \in \mathcal{H}$, we impose the constraint

$$\sum_i Q_i = I_{\mathcal{H}} \tag{5}$$

where $I_{\mathcal{H}}$ is the identity operator on $\mathcal{H}$; then

$$\sum_i p(i) = \left\langle \psi, \sum_i Q_i\psi \right\rangle = \langle \psi, \psi\rangle = 1. \tag{6}$$

We distinguish between standard (von Neumann) measurements and generalized measurements, or POVMs. In a standard measurement, the measurement operators $\{Q_i\}$ form a complete set of orthogonal projectors. Thus,

$$Q_iQ_i = Q_i \tag{7}$$

$$Q_iQ_j = 0, \qquad \text{if } i \neq j \tag{8}$$

$$\sum_i Q_i = I_{\mathcal{H}}. \tag{9}$$

If the measurement is rank-one, so that $Q_i = \mu_i\mu_i^*$, then (7) and (8) imply that $\langle \mu_i, \mu_j\rangle = \delta_{ij}$, while (9) implies that

$$x = I_{\mathcal{H}}x = \sum_i \langle \mu_i, x\rangle \mu_i, \qquad \forall x \in \mathcal{H} \tag{10}$$

so the measurement vectors $\{\mu_i\}$ form an orthonormal basis for $\mathcal{H}$.

Sometimes a generalized measurement is a more efficient way of obtaining information about the state of a quantum system than a standard measurement [23]. A generalized measurement consists of a set $\{Q_i\}$ of nonnegative Hermitian operators, not necessarily projectors, that satisfy $\sum_i Q_i = I_{\mathcal{H}}$.

Such a set of operators is termed a POVM. If the measurement is rank-one so that $Q_i = \mu_i\mu_i^*$, then the measurement vectors μ_i must satisfy

$$\sum_i \mu_i\mu_i^* = I_{\mathcal{H}}. \tag{11}$$

A POVM is more general than a standard measurement in that the measurement vectors μ_i are not required to be either normalized or orthogonal.

It can be shown that a generalized measurement on a quantum system can be implemented by introducing an auxiliary system and performing standard measurements on the combined system. We will discuss this property in Section III-B in the context of Neumark's theorem; in Section IV-B we show that this property has an analog for tight frames.

A. Measurement Matrices

A rank-one POVM acting on an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$ in which the system to be measured is known *a priori* to lie is defined by a set of n measurement vectors $\{\mu_i,\, 1 \le i \le n\}$ that satisfy

$$\sum_{i=1}^{n} \mu_i\mu_i^* = P_{\mathcal{U}} \tag{12}$$

i.e., the n operators $Q_i = \mu_i\mu_i^*$ must be a resolution of the identity[2] on $\mathcal{U}$.

The *measurement matrix* M corresponding to a set of measurement vectors $\mu_i \in \mathcal{U}$ is defined as the matrix of columns μ_i [32]. We have immediately from (12) that

$$MM^* = P_{\mathcal{U}}. \tag{13}$$

It then follows from Theorem 1 that a measurement matrix M with n columns in $\mathcal{H}$ corresponds to a rank-one POVM acting on an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$ if and only if M is a transjector (partial isometry) between $\mathcal{U}$ and an r-dimensional subspace $\mathcal{V} \subseteq \mathbb{C}^n$. Thus, M has all the properties enumerated in Theorem 1.

A measurement matrix M represents a standard measurement if and only if its n columns are orthonormal; i.e., if and only if its Gram matrix satisfies $M^*M = I_n$. Then M has rank n, $\mathcal{U}$ has dimension n, $\mathcal{V} = \mathbb{C}^n$, and $M = UZ_nV^*$ for unitary U and V, where Z_n is given by

$$Z_n = \left[\begin{array}{c} I_n \\ \hline 0 \end{array}\right]. \tag{14}$$

We summarize the properties of measurement matrices in the following theorem.

Theorem 2 (Measurement Matrices): The following statements are equivalent for a matrix M whose columns are n vectors in a complex Hilbert space $\mathcal{H}$:

1) M is a measurement matrix corresponding to a rank-one POVM acting on an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$;
2) M is a transjector (partial isometry) between r-dimensional subspaces $\mathcal{U} \subseteq \mathcal{H}$ and $\mathcal{V} \subseteq \mathbb{C}^n$;
3) $MM^* = P_{\mathcal{U}}$ for an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$;
4) $M^*M = P_{\mathcal{V}}$ for an r-dimensional subspace $\mathcal{V} \subseteq \mathbb{C}^n$.

A measurement matrix M corresponding to a rank-one POVM acting on an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$ may be expressed as $M = UZ_rV^*$, where U is a unitary matrix whose first r columns $\{u_i,\, 1 \le i \le r\}$ are an orthonormal basis for $\mathcal{U}$, V is an $n \times n$ unitary matrix whose first r columns $\{v_i,\, 1 \le i \le r\}$ are an orthonormal basis for $\mathcal{V}$, and Z_r is given by (1). Equivalently, $M = \sum_{i=1}^{r} u_i v_i^*$.

A measurement matrix M is an isometry if restricted to $\mathcal{V}$.

A measurement matrix M whose columns are n vectors in $\mathcal{H}$ represents a standard measurement if and only if its rank is n. Then $M = UZ_nV^*$, where Z_n is given by (14), and $M^*M = I_n$.

B. Neumark's Theorem

Neumark's theorem [22], [23] guarantees that any POVM with measurement vectors $\mu_i \in \mathcal{U}$ can be realized by a set of orthonormal vectors $\tilde{\mu}_i$ in an extended space $\tilde{\mathcal{U}}$ such that $\mathcal{U} \subseteq \tilde{\mathcal{U}}$, so that $\mu_i = P_{\mathcal{U}}\tilde{\mu}_i$.

Using the measurement matrix characterization of a POVM and the SVD, we now obtain a simple statement and proof of Neumark's theorem. Moreover, our proof is constructive; we explicitly construct a set of orthogonal measurement vectors such that their orthogonal projections onto $\mathcal{U}$ are the original measurement vectors. In Section IV-B, we use this construction to extend a tight frame into an orthogonal basis for a larger space.

Theorem 3 (Neumark's Theorem): Let M be a rank-r measurement matrix of an arbitrary POVM, with n columns in a complex Hilbert space $\mathcal{H}$. In other words, M is a transjector between an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$ and an r-dimensional subspace $\mathcal{V} \subseteq \mathbb{C}^n$. Then there exists a standard (von Neumann) measurement with measurement matrix $\tilde{M}$ which is a transjector between an expanded n-dimensional subspace $\tilde{\mathcal{U}} \supseteq \mathcal{U}$ in a possibly expanded complex Hilbert space $\tilde{\mathcal{H}} \supseteq \mathcal{H}$ and $\mathbb{C}^n$, and whose orthogonal projection onto $\mathcal{U}$ is $M = P_{\mathcal{U}}\tilde{M}$.

Proof: Using Theorem 2, we may express M as $M = UZ_rV^*$. Let u_i and v_i denote the columns of U and V respectively. Assume that $\mathcal{H}$ is finite-dimensional, and let $k = \dim \mathcal{H}$.

We distinguish between the case $k \ge n$ (i.e., M has at least as many rows as columns), and the case $k < n$ (i.e., M has more columns than rows).

In the case $k \ge n$, define $\tilde{M} = \sum_{i=1}^{n} u_i v_i^*$; then $\tilde{\mathcal{U}} \subseteq \mathcal{H}$ is the n-dimensional subspace spanned by $\{u_i,\, 1 \le i \le n\}$. The orthogonal projection of $\tilde{M}$ onto $\mathcal{U}$ is

$$P_{\mathcal{U}}\tilde{M} = \sum_{j=1}^{m} u_j u_j^* \sum_{i=1}^{n} u_i v_i^* = \sum_{i=1}^{m} u_i v_i^* = M. \tag{15}$$

Moreover, the columns of $\tilde{M}$ are orthonormal, since its Gram matrix is

$$\tilde{M}^*\tilde{M} = \sum_{j=1}^{n} v_j u_j^* \sum_{i=1}^{n} u_i v_i^* = \sum_{i=1}^{n} v_i v_i^* = I_n. \tag{16}$$

[2]Often these operators are supplemented by an orthogonal projection $Q_0 = P_{\mathcal{U}^\perp} = I_{\mathcal{H}} - P_{\mathcal{U}}$ onto the orthogonal subspace $\mathcal{U}^\perp \subseteq \mathcal{H}$, so that $\sum_{i=0}^{m} Q_i = I_{\mathcal{H}}$—i.e., the augmented POVM is a resolution of the identity on $\mathcal{H}$.

In the case $k < n$, first embed $\mathcal{U}$ in an n-dimensional space $\tilde{\mathcal{U}}$ in an expanded complex Hilbert space $\tilde{\mathcal{H}} \supseteq \mathcal{H}$, and let $\{\tilde{u}_i, 1 \leq i \leq n\}$ be an orthonormal basis for $\tilde{\mathcal{U}}$ of which the first m vectors are the $\mathcal{U}$-basis. Then proceed as before, using $\tilde{u}_i$ in place of u_i. □

It is instructive to consider the matrix representation of $\tilde{M}$ in both cases. Recall that $M = UZ_rV^*$, where Z_r is given by (1).

In the case $k \geq n$, we construct $\tilde{M}$ simply by extending the identity matrix along the diagonal; $\tilde{M} = UZ_nV^*$ where Z_n is given by (14). Thus, when $k \geq n$, the left and right unitary matrices in the SVD of M and $\tilde{M}$ are the same, and are equal to U and V, respectively.

If $k = n$, then $Z_n = I_n$ and $\tilde{M} = UV^*$.

In the case $k < n$, we first replace the left unitary matrix U by $\tilde{U}$, and thus replace k by $\tilde{k} = n$; then $\tilde{U}$ is an $n \times n$ unitary matrix whose first r columns are the $\mathcal{U}$-basis (where we append $n-k$ zeros to each basis vector u_i). We then define $\tilde{M} = \tilde{U}V^*$.

Examples of the construction of the orthogonal measurement vectors associated with a given POVM along the lines of this proof will be given in Section IV-B, in the context of frames.

C. Optimal Quantum Measurements

We now recapitulate some results on optimal quantum measurements according to various criteria, which will be relevant to the construction of optimal tight frames.

Let $\{\psi_i, 1 \leq i \leq n\}$ be a collection of $n \leq k$ normalized vectors ψ_i in a k-dimensional complex Hilbert space $\mathcal{H}$, representing different states of a quantum system. In general, these vectors are nonorthogonal and span an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$. The vectors are linearly independent if $r = n$.

To distinguish between the different states, we subject the system to a measurement. For our measurement, we restrict our attention to POVMs consisting of n rank-one operators of the form $Q_i = \mu_i\mu_i^*$ with measurement vectors $\mu_i \in \mathcal{U}$. We do not require the vectors μ_i to be orthogonal or normalized. However, to constitute a POVM on $\mathcal{U}$ the measurement vectors must satisfy (12).

If the states are prepared with equal prior probabilities, then the probability of detection error using the measurement vectors μ_i is given from (4) by

$$P_e = 1 - \frac{1}{n}\sum_{i=1}^{n} |\langle \mu_i, \psi_i \rangle|^2. \tag{17}$$

If the vectors μ_i are orthonormal, then choosing $\mu_i = \psi_i$ results in $P_e = 0$. However, if the given vectors are not orthonormal, then no measurement can distinguish perfectly between them. Therefore, a fundamental problem in quantum mechanics is to construct measurements optimized to distinguish between a set of nonorthogonal pure quantum states.

This problem may be formulated as a quantum detection problem, so that the measurement vectors are chosen to minimize the probability of detection error, or more generally, minimize the Bayes cost. Necessary and sufficient conditions for an optimum measurement minimizing the Bayes cost have been derived [35]–[37]. However, except in some particular cases [37]–[39], obtaining a closed-form analytical expression for the optimal measurement directly from these conditions is a difficult and unsolved problem.

An alternative approach proposed in [32] is to choose a different optimality criterion, namely, a squared-error criterion, and to seek measurement vectors that minimize this criterion. Specifically, the measurement vectors are chosen to minimize the sum of the squared norms of the error vectors, where the ith error vector is defined as the difference between the ith state vector and the ith measurement vector. The optimizing measurement is referred to as the *least-squares measurement* (LSM).

The problem of finding a set of orthonormal measurement vectors that minimize the squared-error criterion when the states are linearly independent was first solved in [40]. A more general, independent development that covers the cases of linearly dependent states and nonorthogonal measurement vectors appears in [32].

It turns out that the LSM problem has a simple closed-form solution which has many desirable properties. Its construction is relatively simple; it can be determined directly from the given collection of states; it minimizes the probability of detection error when the states exhibit certain symmetries [32]; it is "pretty good" when the states to be distinguished are equally likely and almost orthogonal [41]; it achieves a probability of error within a factor of two of the optimal probability of error [42]; and it is asymptotically optimal [43].

In the next section, we will develop a relationship between POVMs and tight frames. We then apply ideas and results derived in the context of quantum detection to the construction and characterization of tight frames. In particular, we will apply the squared-error criterion developed in [32] to the construction of optimal tight frames.

IV. Tight Frames

Frames, which are generalization of bases, were introduced in the context of nonharmonic Fourier series by Duffin and Schaeffer [1] (see also [2]). Recently, the theory of frames has been expanded [4], [5], [8], [6], in part due to the utility of frames in analyzing wavelet decompositions. Here we will focus on tight frames, which have particularly nice properties.

Let $\{\varphi_i, 1 \leq i \leq n\}$ denote a set of n vectors in an r-dimensional subspace $\mathcal{U}$ of a Hilbert space $\mathcal{H}$. The vectors φ_i form a *tight frame* for $\mathcal{U}$ if there exists a constant $\beta > 0$ such that

$$\sum_{i=1}^{n} |\langle x, \varphi_i \rangle|^2 = \beta^2 \|x\|^2 \tag{18}$$

for all $x \in \mathcal{U}$ [8]. If $\beta = 1$, the tight frame is said to be *normalized*; otherwise, it is said to be *β-scaled*.[3]

[3]More generally, the vectors φ_i form a *frame* for $\mathcal{U}$ if there exist constants $0 < \alpha \leq \beta < \infty$ such that

$$\alpha^2\|x\|^2 \leq \sum_{i=1}^{n} |\langle x, \varphi_i \rangle|^2 \leq \beta^2\|x\|^2$$

for all $x \in \mathcal{U}$ [8]. The lower bound ensures that the vectors φ_i span $\mathcal{U}$; thus, we must have $n \geq r$. If $n < \infty$, then the right-hand inequality is always satisfied with $\beta^2 = \sum_{i=1}^{n}\langle \varphi_i, \varphi_i \rangle$. Thus, any finite set of vectors that spans $\mathcal{U}$ is a frame for $\mathcal{U}$. A tight frame is a special case of a frame for which $\alpha = \beta$.

Of course, any orthonormal basis for $\mathcal{U}$ is a normalized tight frame for $\mathcal{U}$. However, there also exist tight frames for $\mathcal{U}$ with $n > r$, which are necessarily linearly dependent. The *redundancy* of the tight frame is defined as $\rho = n/r$.

Since

$$\sum_{i=1}^{n} |\langle x, \varphi_i \rangle|^2 = \sum_{i=1}^{n} x^* \varphi_i \varphi_i^* x = \left\langle x, \left(\sum_i \varphi_i \varphi_i^* \right) x \right\rangle \tag{19}$$

the fact that (18) holds for all $x \in \mathcal{U}$ implies that

$$\sum_{i=1}^{n} \varphi_i \varphi_i^* = \beta^2 P_{\mathcal{U}}. \tag{20}$$

Conversely, if the vectors $\varphi_i \in \mathcal{U}$ satisfy (20), then (19) implies that (18) is satisfied for all $x \in \mathcal{U}$. We conclude that a set of n vectors $\varphi_i \in \mathcal{U}$ forms a tight frame for $\mathcal{U}$ if and only if the vectors satisfy (20) for some $\beta > 0$.

Comparing (20) with (12), we conclude the following.

Theorem 4 (Tight Frames): A set of vectors $\varphi_i \in \mathcal{U}$ forms a β-scaled tight frame for $\mathcal{U}$ if and only if the scaled vectors $\beta^{-1}\varphi_i$ are the measurement vectors of a rank-one POVM on $\mathcal{U}$. In particular, the vectors φ_i form a normalized tight frame for $\mathcal{U}$ if and only if they are the measurement vectors of a rank-one POVM on $\mathcal{U}$.

This fundamental relationship between rank-one quantum measurements and tight frames will be the basis for the developments in subsequent sections. In the next section, we define frame matrices in analogy to the measurement matrices of quantum mechanics. We then use Neumark's theorem to extend tight frames to orthogonal bases. Motivated by the LSM of quantum mechanics, in Section V we address the problem of constructing optimal tight frames.

A. Frame Matrices

In analogy to the measurement matrix, we define the *frame matrix* F as the matrix of columns φ_i, where the vectors φ_i form a tight frame for $\mathcal{U}$. From (20) it then follows that

$$FF^* = \beta^2 P_{\mathcal{U}}. \tag{21}$$

The properties of a frame matrix F follow immediately from Theorems 4 and 2.

Theorem 5 (Frame Matrices): For a matrix F whose columns are n vectors in a complex Hilbert space $\mathcal{H}$ and for a constant $\beta > 0$, the following statements are equivalent:

1) F is the frame matrix of a β-scaled tight frame for an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$;
2) $\beta^{-1}F$ is a transjector (partial isometry) between r-dimensional subspaces $\mathcal{U} \subseteq \mathcal{H}$ and $\mathcal{V} \subseteq \mathbb{C}^n$;
3) $FF^* = \beta^2 P_{\mathcal{U}}$ for an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$;
4) $F^*F = \beta^2 P_{\mathcal{V}}$ for an r-dimensional subspace $\mathcal{V} \subseteq \mathbb{C}^n$.

A frame matrix F of a β-scaled tight frame for an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$ may be expressed as $F = \beta U Z_r V^*$, where U is a unitary matrix whose first r columns $\{u_i, 1 \le i \le r\}$ are an orthonormal basis for $\mathcal{U}$, V is an $n \times n$ unitary matrix whose first r columns $\{v_i, 1 \le i \le r\}$ are an orthonormal basis for $\mathcal{V}$, and Z_r is given by (1). Equivalently, $F = \beta \sum_{i=1}^{r} u_i v_i^*$.

A frame matrix F of a β-scaled tight frame is an isometry if restricted to $\mathcal{V}$ and scaled by β^{-1}.

A frame matrix F of a β-scaled tight frame whose columns are n vectors in $\mathcal{H}$ represents an orthogonal basis for $\mathcal{U}$ (i.e., is an *orthogonal frame matrix*) if and only if its rank is n. Then $F = \beta U Z_n V^*$, where Z_n is given by (14), and $F^*F = \beta^2 I_n$; i.e., all frame vectors have squared norm β^2.

If the vectors $\{\varphi_i, 1 \le i \le n\}$ form a tight frame for $\mathcal{U}$, then any $x \in \mathcal{U}$ may be expressed as a linear combination of these vectors: $x = \sum_{i=1}^{n} a_i \varphi_i$. When $n > r$, the coefficients in this expansion are not unique. A possible choice is $a_i = \beta^{-2}\langle \varphi_i, x \rangle$, because

$$\beta^{-2} \sum_{i=1}^{n} \langle \varphi_i, x \rangle \varphi_i = \beta^{-2} FF^* x = P_{\mathcal{U}} x = x. \tag{22}$$

The vectors $\beta^{-2}\varphi$ are defined as the *dual frame vectors*. This choice of coefficients has the property that among all possible coefficients it has the minimal norm [8], [44].

The expansion of (22) is reminiscent of an expansion of x in terms of an orthonormal basis for $\mathcal{U}$. However, whereas the vectors in an orthonormal expansion are linearly independent, the vectors φ_i in (22) are linearly dependent when $n > r$.

B. Neumark's Theorem and Construction of Tight Frames

Neumark's theorem (Theorem 3) was derived based on the properties of measurement matrices. Since, by Theorem 4, frame matrices of tight frames have essentially the same properties as measurement matrices of rank-one POVMs, we can now obtain an equivalent of Neumark's theorem for tight frames. The proof is essentially the same as the proof of Theorem 3, so we omit it.

Theorem 6 (Neumark's Theorem for Tight Frames): Let F be a rank-r frame matrix, with n columns in a complex Hilbert space $\mathcal{H}$ that span an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$. Then there exists an orthogonal frame matrix $\tilde{F}$ with equal-norm orthogonal columns that span an expanded n-dimensional subspace $\tilde{\mathcal{U}} \supseteq \mathcal{U}$ in a possibly expanded complex Hilbert space $\tilde{\mathcal{H}} \supseteq \mathcal{H}$ such that the orthogonal projection $P_{\mathcal{U}}\tilde{F}$ of $\tilde{F}$ onto $\mathcal{U}$ is F.

We remark that given a set of equal-norm orthogonal vectors in $\tilde{\mathcal{U}} \supseteq \mathcal{U}$, their orthogonal projections onto $\mathcal{U}$ will always form a tight frame for $\mathcal{U}$ [6]. Combining this result with Theorem 6, we can conclude that a set of vectors forms a tight frame for $\mathcal{U}$ if and only if the vectors can be expressed as an orthogonal projection onto $\mathcal{U}$ of a set of orthogonal vectors with equal norm in a larger space $\tilde{\mathcal{U}}$ containing $\mathcal{U}$.

Starting with a given frame matrix F in $\mathcal{U}$, the proof of Theorem 3 gives a concrete construction of an orthogonal frame matrix $\tilde{F}$ in $\tilde{\mathcal{U}} \supseteq \mathcal{U}$ such that $P_{\mathcal{U}}\tilde{F} = F$. We now give two examples of this construction. We consider first an example in which $\dim \mathcal{H} < n$, and then one in which $\dim \mathcal{H} > n$.

Example 1: Consider the four frame vectors $\varphi_1 = [0.35 \quad -0.61]^*$, $\varphi_2 = [0.61 \quad 0.35]^*$, $\varphi_3 = [0.5 \quad -0.5]^*$, and $\varphi_4 = [0.5 \quad 0.5]^*$. The frame matrix associated with this frame is

$$F = \begin{bmatrix} 0.35 & 0.61 & 0.5 & 0.5 \\ -0.61 & 0.35 & -0.5 & 0.5 \end{bmatrix}; \tag{23}$$

we may check that F is indeed the frame matrix of a tight frame since $FF^* = I_2$.

We wish to construct an orthogonal frame matrix $\tilde{F}$ such that $F = P_{\mathcal{U}}\tilde{F}$. In the proof of Theorem 3 for the case $\dim \mathcal{H} < n$, we constructed an $n \times n$ unitary matrix $\tilde{F}$ using the SVD $F = U\Sigma V^*$. Using this construction here, we obtain

$$\begin{aligned} U &= \begin{bmatrix} 0.5 & -0.87 \\ -0.87 & -0.5 \end{bmatrix} \\ \Sigma &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \\ V &= \begin{bmatrix} 0.70 & 0 & 0.70 & 0 \\ 0 & -0.70 & 0 & -0.70 \\ 0.68 & -0.18 & -0.68 & 0.18 \\ -0.18 & -0.68 & 0.18 & 0.68 \end{bmatrix}. \end{aligned} \tag{24}$$

We now define the extended frame matrix $\tilde{U}$ in accordance with the proof of Theorem 3. The first two columns of $\tilde{U}$ are uniquely defined as the first two columns of U with zeros appended. The remaining two columns are arbitrary, as long as the resulting $\tilde{U}$ is unitary. A possible choice is

$$\tilde{U} = \begin{bmatrix} 0.5 & -0.87 & 0 & 0 \\ -0.87 & -0.5 & 0 & 0 \\ 0 & 0 & 0.5 & -0.87 \\ 0 & 0 & -0.87 & -0.5 \end{bmatrix}. \tag{25}$$

Then

$$\tilde{F} = \tilde{U}V^* = \begin{bmatrix} 0.35 & 0.61 & 0.5 & 0.5 \\ -0.61 & 0.35 & -0.5 & 0.5 \\ 0.35 & 0.61 & -0.5 & -0.5 \\ -0.61 & 0.35 & 0.5 & -0.5 \end{bmatrix}. \tag{26}$$

We may immediately verify that $\tilde{F}^*\tilde{F} = I_4$; i.e., $\tilde{F}$ represents an orthonormal set of vectors.

Since the columns of F span a two-dimensional Hilbert space $\mathcal{U} = \mathcal{H}$, the orthogonal projection onto this space is given by

$$P_{\mathcal{U}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{27}$$

and, indeed, $F = P_{\mathcal{U}}\tilde{F}$.

Example 2: We now consider an example in which $\dim \mathcal{H} > n$. The construction of $\tilde{F}$ is simpler than in the previous case because we do not have to extend $\mathcal{H}$. Consider the three frame vectors $\varphi_1 = \frac{1}{2}[1\ 1\ 1]^*$, $\varphi_2 = \frac{1}{2}[-1\ 1\ 1]^*$, and $\varphi_3 = \frac{1}{2}[\sqrt{2}\ 0\ 0]^*$. The frame matrix associated with this frame is

$$F = \frac{1}{2}\begin{bmatrix} 1 & -1 & \sqrt{2} \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}. \tag{28}$$

In order to verify that F is indeed the frame matrix of a tight frame, we again determine the SVD $F = U\Sigma V^*$, which yields

$$\begin{aligned} U &= \begin{bmatrix} 0.58 & 0.82 & 0 \\ 0.58 & -0.4 & 0.7 \\ 0.58 & -0.4 & -0.7 \end{bmatrix} \\ \Sigma &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ V &= \begin{bmatrix} 0.87 & 0 & 0.5 \\ 0.29 & -0.82 & -0.5 \\ 0.4 & 0.58 & -0.7 \end{bmatrix}. \end{aligned} \tag{29}$$

From Theorem 5 we conclude that F is indeed the frame matrix of a tight frame since its nonzero singular values are all equal to 1; i.e., F is a transjector. A basis for the subspace $\mathcal{U}$ spanned by the columns of F is the two vectors

$$\begin{aligned} u_1 &= [0.58 \quad 0.58 \quad 0.58]^* \\ u_2 &= [0.82 \quad -0.4 \quad -0.4]^*. \end{aligned} \tag{30}$$

Thus, $P_{\mathcal{U}}$ is given by

$$P_{\mathcal{U}} = \sum_{i=1}^{2} u_i u_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \end{bmatrix} \tag{31}$$

and, indeed, $FF^* = P_{\mathcal{U}}$.

We now define an extended frame matrix $\tilde{F}$ such that $F = P_{\mathcal{U}}\tilde{F}$ and $\tilde{F}^*\tilde{F} = I_3$. From the proof of Theorem 3, we have

$$\begin{aligned} \tilde{F} &= UZ_3V^* = UV^* = F + u_3 v_3^* \\ &= \begin{bmatrix} 0.5 & -0.5 & 0.7 \\ 0.85 & 0.15 & -0.5 \\ 0.15 & 0.85 & 0.5 \end{bmatrix} \end{aligned} \tag{32}$$

where

$$u_3 = [0 \quad 0.7 \quad -0.7]^*, \quad v_3 = [0.5 \quad -0.5 \quad 0.7]^*. \tag{33}$$

Since $P_{\mathcal{U}} u_3 v_3^* = 0$, we have immediately that $F = P_{\mathcal{U}}\tilde{F}$.

V. Optimal Tight Frames

It is often of interest to construct a tight frame from a given set of vectors $\{\psi_i,\ 1 \le i \le n\}$. Using the LSM developed in the context of quantum detection [32], we now propose a systematic method of constructing optimal tight frames from a given set of vectors.

Thus, we wish to construct a tight frame of vectors $\{\varphi_i,\ 1 \le i \le n\}$ from a given set of vectors $\{\psi_i,\ 1 \le i \le n\}$ that span an r-dimensional space $\mathcal{U} \subseteq \mathcal{H}$. A reasonable approach is to find a set of vectors $\varphi_i \in \mathcal{U}$ that are "closest" to the vectors ψ_i in the least-squares sense. Thus, we seek vectors φ_i that minimize the squared error E, defined by

$$E = \sum_{i=1}^{n} \langle e_i,\ e_i \rangle \tag{34}$$

where e_i denotes the ith error vector

$$e_i = \psi_i - \varphi_i \tag{35}$$

subject to the constraint (20).

We may wish to constrain the scaling β in (20), e.g., we may seek a normalized tight frame with $\beta = 1$. The optimal frame in this case is referred to as the CLSF. Alternatively, we may choose the vectors $\{\varphi_i\}$ and β to satisfy (20) and to minimize the squared error E of (34). The optimal frame is then referred to as the ULSF.

A. Constrained and Unconstrained Least-Squares Frame

With F and Ψ denoting the $k \times n$ matrices of columns φ_i and ψ_i, respectively, the squared error E of (34) and (35) may be expressed as

$$E = \mathrm{Tr}((\Psi - F)^*(\Psi - F)) = \mathrm{Tr}((\Psi - F)(\Psi - F)^*) \quad (36)$$

and the constraint (20) may be restated as

$$FF^* = \beta^2 P_{\mathcal{U}}. \quad (37)$$

Employing the SVD $\Psi = U\Sigma V^*$, we rewrite the squared error E of (36) as

$$\begin{aligned} E &= \mathrm{Tr}((\Psi - F)(\Psi - F)^*) \\ &= \mathrm{Tr}(U^*(\Psi - F)(\Psi - F)^* U) \\ &= \sum_{i=1}^{k} \langle d_i, d_i \rangle, \end{aligned} \quad (38)$$

where

$$d_i = (\Psi - F)^* u_i. \quad (39)$$

The vectors $\{u_i, 1 \le i \le r\}$ form an orthonormal basis for $\mathcal{U}$. Therefore, the orthogonal projection operator onto $\mathcal{U}$ is given by

$$P_{\mathcal{U}} = \sum_{i=1}^{r} u_i u_i^*. \quad (40)$$

Essentially, we want to construct a map F^* such that the images of the maps defined by Ψ^* and F^* are as close as possible in the squared norm sense, subject to the constraint

$$FF^* = \beta^2 \sum_{i=1}^{r} u_i u_i^*. \quad (41)$$

The SVD of Ψ^* is given by $\Psi^* = V\Sigma^* U^*$. Consequently

$$\Psi^* u_i = \begin{cases} \sigma_i v_i, & 1 \le i \le r \\ 0, & r+1 \le i \le k \end{cases} \quad (42)$$

where 0 denotes the zero vector. Denoting the image of u_i under F^* by $a_i = F^* u_i$, for any choice of F satisfying the constraint (41), we have

$$\langle a_i, a_i \rangle = u_i^* FF^* u_i = \begin{cases} \beta^2, & 1 \le i \le r \\ 0, & r+1 \le i \le k \end{cases} \quad (43)$$

and

$$\langle a_i, a_j \rangle = u_i^* FF^* u_j = 0, \qquad i \ne j. \quad (44)$$

Thus, the vectors a_i, $1 \le i \le r$ are mutually orthogonal with $\langle a_i, a_i \rangle = \beta^2$ and $a_i = 0$, $r+1 \le i \le k$. Combining (42) and (43), we may express d_i as

$$d_i = \begin{cases} \sigma_i v_i - a_i, & 1 \le i \le r \\ 0, & r+1 \le i \le k. \end{cases} \quad (45)$$

Our problem, therefore, reduces to finding a set of r orthogonal vectors a_i with norm β that minimize

$$\begin{aligned} E &= \sum_{i=1}^{r} \langle d_i, d_i \rangle \\ &= \sum_{i=1}^{r} \langle \sigma_i v_i - a_i, \sigma_i v_i - a_i \rangle \\ &= r\beta^2 + \sum_{i=1}^{r} \sigma_i^2 - 2\sum_{i=1}^{r} \Re\{\langle a_i, v_i \rangle\} \end{aligned} \quad (46)$$

where the vectors v_i are orthonormal. For any choice of β

$$\Re\{\langle a_i, v_i \rangle\} \le |\langle a_i, v_i \rangle| \le \langle a_i, a_i \rangle^{1/2} \langle v_i, v_i \rangle^{1/2} = \beta \quad (47)$$

with equality if and only if $a_i = \beta v_i$. Thus, the vectors a_i minimizing E are $a_i = \beta v_i$, $1 \le i \le r$.

If β is fixed, then the optimal frame matrix F, denoted by $\hat{F}_c$, satisfies

$$\hat{F}_c^* u_i = \begin{cases} \beta v_i, & 1 \le i \le r \\ 0, & r+1 \le i \le k. \end{cases} \quad (48)$$

Consequently the CLSF vectors are the columns of

$$\hat{F}_c = \beta \sum_{i=1}^{r} u_i v_i^* = \beta U Z_r V^* \quad (49)$$

where Z_r is defined by (1). We may express $\hat{F}_c$ directly in terms of Ψ as

$$\hat{F}_c = \beta\Psi\left((\Psi^*\Psi)^{1/2}\right)^\dagger = \beta\left((\Psi\Psi^*)^{1/2}\right)^\dagger \Psi \quad (50)$$

where $(\cdot)^\dagger$ denotes the Moore–Penrose pseudo-inverse [45]. The residual squared error is then

$$E_{\min}^c = \sum_{i=1}^{r} (\beta - \sigma_i)^2 \langle v_i, v_i \rangle = \sum_{i=1}^{r} (\beta - \sigma_i)^2. \quad (51)$$

We note that the CLSF vectors $\hat{\varphi}_i^c$ which are the columns of $\hat{F}_c$ satisfy

$$\langle \hat{\varphi}_i^c, \psi_i \rangle = [\hat{F}_c^* \Psi]_{ii} = \beta [\Psi^*\Psi]_{ii}^{1/2} \quad (52)$$

where $[\cdot]_{ii}$ denotes the iith element of the matrix. This relation may be used to derive bounds on the inner products $\langle \hat{\varphi}_i^c, \psi_i \rangle$ in terms of the inner products $\langle \psi_i, \psi_j \rangle$; see [43].

To derive the ULSF, we further minimize E of (46) with respect to β. Substituting the optimal vectors $a_i = \beta v_i$, $1 \le i \le r$ back into (46), we choose β to minimize

$$E_{\min}^c = \sum_{i=1}^{r} (\beta - \sigma_i)^2.$$

The optimal value of β, denoted by $\hat{\beta}$, is given by

$$\hat{\beta} = \frac{1}{r} \sum_{i=1}^{r} \sigma_i = \frac{1}{r} \mathrm{Tr}\left((\Psi^*\Psi)^{1/2}\right) \quad (53)$$

and the ULSF vectors $\hat{\varphi}_i^u$ are the columns of

$$\hat{F}_u = \hat{\beta} \sum_{i=1}^{r} u_i v_i^* = \hat{\beta}\Psi\left((\Psi^*\Psi)^{1/2}\right)^\dagger = \hat{\beta}\left((\Psi\Psi^*)^{1/2}\right)^\dagger \Psi. \quad (54)$$

The residual squared error is

$$E^u_{\min} = \sum_{i=1}^{r} (\hat{\beta} - \sigma_i)^2. \tag{55}$$

The CLSF and the ULSF vectors can be expressed in a unified manner as the columns of the least-squares frame (LSF) matrix

$$\hat{F} = \alpha \sum_{i=1}^{r} u_i v_i^* = \alpha \Psi \left((\Psi^*\Psi)^{1/2} \right)^{\dagger} = \alpha \left((\Psi\Psi^*)^{1/2} \right)^{\dagger} \Psi \tag{56}$$

where in the CLSF $\alpha = \beta$, and in the ULSF $\alpha = \hat{\beta}$ given by (53). In the sequel, when the value of α is immaterial, we will refer to the LSF which encompasses both the CLSF and the ULSF.

Note that if the singular values σ_i of F are distinct, then the vectors u_i, $1 \leq i \leq r$ are unique (up to a phase factor $e^{j\theta_i}$). Given the vectors u_i, the vectors v_i are uniquely determined, so the optimal frame vectors corresponding to $\hat{F}$ are unique. If, on the other hand, there are repeated singular values, then the corresponding eigenvectors are not unique. Nonetheless, the choice of singular vectors does not affect $\hat{F}$. Indeed, if the vectors corresponding to a repeated singular value σ are $\{u_j\}$, then $\sum_j u_j u_j^*$ is an orthogonal projection onto the corresponding eigenspace, and therefore is the same regardless of the choice of the vectors $\{u_j\}$. Thus,

$$\sum_j u_j v_j^* = \frac{1}{\sigma} \sum_j u_j u_j^* \Psi \tag{57}$$

independent of the choice of $\{u_j\}$, and the optimal frame is unique.

B. Optimal Orthogonal Basis, the CLSF, and the ULSF

In the previous section, we sought the β-scaled tight frame that minimizes the least-squares error. We may similarly seek the optimal orthogonal vectors with norm β of the same form. We now explore the connection between the resulting optimal vectors both in the case of linearly independent vectors ψ_i ($r = n$), and in the case of linearly dependent vectors ($r < n$).

Linearly Independent Vectors: If the vectors ψ_i are linearly independent and, consequently, Ψ has full column rank (i.e., $r = n$), then the LSF (56) reduces to

$$\hat{F} = \alpha \Psi (\Psi^*\Psi)^{-1/2}. \tag{58}$$

The optimal frame vectors $\hat{\varphi}_i$ that are the columns of $\hat{F}$ are mutually orthogonal with equal norm α, since their Gram matrix is

$$\hat{F}^*\hat{F} = \alpha^2 (\Psi^*\Psi)^{-1/2} \Psi^*\Psi (\Psi^*\Psi)^{-1/2} = \alpha^2 I. \tag{59}$$

Thus, the optimal LSF is, in fact, an optimal orthogonal basis for $\mathcal{U}$.

Linearly Dependent Vectors: If the vectors ψ_i are linearly dependent, so that the matrix Ψ does not have full column rank (i.e., $r < n$), then the n frame vectors φ_i cannot be mutually orthogonal since they span an r-dimensional subspace. We now try to gain some insight into the optimal frame vectors in this case. Our problem is to find a set of vectors that are as close as possible to the n vectors ψ_i, which lie in an r-dimensional subspace $\mathcal{U}$. We now show that these vectors are related to the optimal frame vectors through an orthogonal projection onto the subspace $\mathcal{U}$, spanned by the vectors ψ_i.

To see this, suppose we seek a set of orthogonal vectors $\tilde{\varphi}_i \in \mathcal{H}$ with $\langle \tilde{\varphi}_i, \tilde{\varphi}_i \rangle = \beta^2$ that are as close as possible to the vectors ψ_i. From Theorem 5, we have that

$$\sum_{i=1}^{n} \tilde{\varphi}_i \tilde{\varphi}_i^* = \beta^2 P_{\tilde{\mathcal{U}}} \tag{60}$$

where $\tilde{\mathcal{U}} \supseteq \mathcal{U}$ is the space spanned by the vectors $\tilde{\varphi}_i$.

Since there are at most r orthogonal vectors in $\mathcal{U}$, imposing an orthogonality constraint forces the optimal orthogonal vectors $\tilde{\varphi}_i$ to lie partly in the orthogonal complement $\mathcal{U}^{\perp}$. Each vector then has a component in $\mathcal{U}$, $\tilde{\varphi}_i^{\mathcal{U}}$, and a component in $\mathcal{U}^{\perp}$, $\tilde{\varphi}_i^{\mathcal{U}^{\perp}}$. Using (60), the component in $\mathcal{U}$ satisfies

$$\sum_{i=1}^{n} \tilde{\varphi}_i^{\mathcal{U}} (\tilde{\varphi}_i^{\mathcal{U}})^* = \sum_{i=1}^{n} P_{\mathcal{U}} \tilde{\varphi}_i \tilde{\varphi}_i^* P_{\mathcal{U}} = \beta^2 P_{\mathcal{U}} P_{\tilde{\mathcal{U}}} P_{\mathcal{U}} = \beta^2 P_{\mathcal{U}} \tag{61}$$

where the last equality follows from the fact that $\mathcal{U} \subseteq \tilde{\mathcal{U}}$. Now we rewrite the error E of (34) as

$$\begin{aligned} E &= \sum_{i=1}^{n} \left\langle \psi_i - \tilde{\varphi}_i^{\mathcal{U}} - \tilde{\varphi}_i^{\mathcal{U}^{\perp}}, \psi_i - \tilde{\varphi}_i^{\mathcal{U}} - \tilde{\varphi}_i^{\mathcal{U}^{\perp}} \right\rangle \\ &= \sum_{i=1}^{n} \left(\left\langle \psi_i - \tilde{\varphi}_i^{\mathcal{U}}, \psi_i - \tilde{\varphi}_i^{\mathcal{U}} \right\rangle + \left\langle \tilde{\varphi}_i^{\mathcal{U}^{\perp}}, \tilde{\varphi}_i^{\mathcal{U}^{\perp}} \right\rangle \right) \end{aligned} \tag{62}$$

since $\langle \psi_i - \tilde{\varphi}_i^{\mathcal{U}}, \tilde{\varphi}_i^{\mathcal{U}^{\perp}} \rangle = 0$. From (61)

$$\begin{aligned} \sum_{i=1}^{n} \left\langle \tilde{\varphi}_i^{\mathcal{U}^{\perp}}, \tilde{\varphi}_i^{\mathcal{U}^{\perp}} \right\rangle &= \sum_{i=1}^{n} \langle \tilde{\varphi}_i, \tilde{\varphi}_i \rangle - \sum_{i=1}^{n} \langle \tilde{\varphi}_i^{\mathcal{U}}, \tilde{\varphi}_i^{\mathcal{U}} \rangle \\ &= n\beta^2 - \mathrm{Tr}\left(\sum_{i=1}^{n} \tilde{\varphi}_i^{\mathcal{U}} \left(\tilde{\varphi}_i^{\mathcal{U}} \right)^* \right) \\ &= n\beta^2 - \mathrm{Tr}(\beta^2 P_{\mathcal{U}}) = (n-r)\beta^2 \end{aligned} \tag{63}$$

independent of the choice of vectors $\tilde{\varphi}_i$. Thus, minimization of E is equivalent to minimization of

$$E' = \sum_{i=1}^{n} \left\langle \psi_i - \tilde{\varphi}_i^{\mathcal{U}}, \psi_i - \tilde{\varphi}_i^{\mathcal{U}} \right\rangle + (n-r)\beta^2. \tag{64}$$

Furthermore, from (61), the vectors $\tilde{\varphi}_i^{\mathcal{U}}$ form a β-scaled tight frame for $\mathcal{U}$.

If β is fixed, then choosing the orthogonal vectors with equal norm β that minimize E is equivalent to choosing an optimal β-scaled tight frame for $\mathcal{U}$. The optimal orthogonal vectors are not unique; however, their orthogonal projections onto $\mathcal{U}$ are unique and are just the β-scaled CLSF vectors. We may choose the projections of the optimal orthogonal vectors onto $\mathcal{U}^{\perp}$ arbitrarily, as long as the resulting n vectors are orthogonal with norm β. A convenient choice is

$$\hat{\tilde{F}}_c = \beta \sum_{i=1}^{n} u_i v_i. \tag{65}$$

Indeed, Theorem 6 shows that the optimal orthogonal vectors are just a realization of the CLSF vectors. This theorem guarantees that any β-scaled tight frame may be realized by a set of orthogonal vectors with norm β in an extended space such that their orthogonal projections onto the smaller space are the given frame vectors. Denoting by $\hat{\varphi}_i^c$ and $\hat{\hat{\varphi}}_i^c$ the optimal β-scaled frame vectors and the optimal orthogonal vectors with norm β, respectively, (64) asserts that

$$\hat{\varphi}_i^c = P_{\mathcal{U}}\hat{\hat{\varphi}}_i^c. \tag{66}$$

If β is chosen to minimize the least-squares error, then we need to further minimize E' with respect to β. Substituting $\tilde{\varphi}_i^{\mathcal{U}} = \hat{\varphi}_i^c$ back into (64), and using the fact that $\langle \tilde{\varphi}_i^{\mathcal{U}}, \psi_i\rangle = \beta\sigma_i$, the optimal value of β is chosen to minimize

$$\begin{aligned} E'' &= \sum_{i=1}^{r} \langle \tilde{\varphi}_i^{\mathcal{U}}, \tilde{\varphi}_i^{\mathcal{U}}\rangle - 2\sum_{i=1}^{r} \Re\{\langle \tilde{\varphi}_i^{\mathcal{U}}, \psi_i\rangle\} + (n-r)\beta^2 \\ &= n\beta^2 - 2\beta\sum_{i=1}^{r}\sigma_i. \end{aligned} \tag{67}$$

Minimizing E'' with respect to β, the optimal value of β, denoted by $\hat{\hat{\beta}}$, is

$$\hat{\hat{\beta}} = \frac{1}{n}\sum_{i=1}^{r}\sigma_i = \frac{r}{n}\hat{\beta} = \frac{\hat{\beta}}{\rho} \tag{68}$$

where $\hat{\beta}$ is defined by (53) and ρ is the redundancy of the frame. Thus, the optimal projections are the columns of $(1/\rho)\hat{F}_u$, where $\hat{F}_u$ is the frame matrix of the ULSF vectors.

We conclude that choosing a set of orthogonal vectors with unconstrained norm that minimize E is equivalent to choosing an optimal unconstrained tight frame for $\mathcal{U}$ and scaling these optimal frame vectors by $1/\rho$. The optimal unconstrained orthogonal vectors are not unique; however, their orthogonal projections onto $\mathcal{U}$ are unique and are proportional to the optimal unconstrained tight frame vectors. We may choose the projections of the optimal orthogonal vectors onto $\mathcal{U}^\perp$ arbitrarily, as long as the resulting n vectors are orthogonal with norm $\hat{\beta}/\rho$. A convenient choice is

$$\hat{\hat{F}}_u = \frac{\hat{\beta}}{\rho}\sum_{i=1}^{n} u_i v_i^*. \tag{69}$$

We summarize our results regarding the CLSF and the ULSF in the following theorem.

Theorem 7 (Least Squares Frame (LSF)): Let $\{\psi_i\}$ be a set of n vectors in a k-dimensional complex Hilbert space $\mathcal{H}$ that span an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$. Let $\{\hat{\varphi}_i\}$ denote the optimal n frame vectors that minimize the least-squares error defined by (34) and (35), subject to the constraint (20). Let $\Psi = U\Sigma V^*$ be the rank-r $k\times n$ matrix whose columns are the vectors ψ_i, and let $\hat{F}$ be the $k \times n$ frame matrix whose columns are the vectors $\hat{\varphi}_i$. Then the unique optimal $\hat{F}$ is given by

$$\begin{aligned} \hat{F} &= \alpha\sum_{i=1}^{r} u_i v_i^* = \alpha U Z_r V^* = \alpha\Psi\left((\Psi^*\Psi)^{1/2}\right)^\dagger \\ &= \alpha\left((\Psi\Psi^*)^{1/2}\right)^\dagger \Psi, \end{aligned}$$

where u_i and v_i denote the columns of U and V, respectively, Z_r is defined by (1), and

1) if β in (20) is specified then $\alpha = \beta$ and the resulting frame vectors are defined as the constrained LSF vectors;
2) if β is chosen to minimize the least-squares error then $\alpha = \hat{\beta}$ where $\hat{\beta} = (1/r)\,\mathrm{Tr}((\Psi^*\Psi)^{1/2})$, and the resulting frame vectors are defined as the unconstrained LSF vectors.

The residual squared error is given by

$$E_{\min} = \sum_{i=1}^{r}(\alpha - \sigma_i)^2$$

where $\{\sigma_i,\, 1 \le i \le r\}$ are the nonzero singular values of Ψ.

In addition,

1) if $r = n$
   a) $\hat{F} = \alpha\Psi(\Psi^*\Psi)^{-1/2}$;
   b) $\hat{F}^*\hat{F} = \alpha^2 I_n$, and the corresponding frame vectors are orthogonal with norm α;
2) if $r < n$
   a) if β is fixed then
      i) the constrained LSF vectors may be realized by the β-scaled optimal orthogonal frame matrix
      $$\hat{\hat{F}} = \beta\sum_{i=1}^{n} u_i v_i^* = \beta U Z_n V^*;$$
      ii) the action of the two optimal vector sets in the subspace $\mathcal{U}$ is the same;
   b) if β is chosen to minimize the least-squares error then the unconstrained LSF vectors may be realized by the optimal orthogonal frame matrix
   $$\hat{\hat{F}} = (\hat{\beta}/\rho)\sum_{i=1}^{n} u_i v_i^* = (\hat{\beta}/\rho) U Z_r V^*$$
   where $\rho = n/r$.

VI. Connection With the Polar Decomposition

We now show that LSF is related to the polar decomposition (PD) of the matrix Ψ.

Let Ψ denote a $k \times n$ matrix, where $k \ge n$. Then Ψ has a *polar decomposition (PD)* [46], [47]

$$\Psi = HY \tag{70}$$

where H is a $k \times n$ partial isometry that satisfies $H^*H = I_n$, and $Y = (\Psi^*\Psi)^{1/2}$. The Hermitian factor Y is always unique; the partial isometry H is unique if and only if Ψ has full column rank.

If $\Psi = U\Sigma V^*$ is the SVD of Ψ, then a natural choice for H is

$$H = U Z_n V^* \tag{71}$$

where Z_n is given by (14). If $r = n$, then this choice of H is unique. Otherwise, H is not unique; however, its orthogonal projection onto the column space $\mathcal{U}$ of Ψ is unique and is given by [48]

$$H_{\mathcal{U}} = P_{\mathcal{U}} H = U Z_r V^* = \Psi\left((\Psi^*\Psi)^{1/2}\right)^\dagger \tag{72}$$

where Z_r is given by (1).

Comparing (72) with (56), we conclude that the LSF is proportional to the (unique) orthogonal projection onto $\mathcal{U}$ of the partial isometry in a PD of Ψ, and can, therefore, be computed very efficiently by use of the many known efficient algorithms for computing the PD (see, e.g., [45], [49], [46], [50]).

Recently, the truncated PD (TPD), a variation on the PD, has been introduced [51] and has proved to be useful for various estimation and detection problems. As we now show, the columns of the TPD of a matrix Ψ are just the closest normalized frame vectors to the columns ψ_i of Ψ.

Let $\Psi = U\Sigma V^*$ denote an arbitrary $k \times n$ matrix with rank r. Then the order-p TPD of Ψ is the factorization

$$P_{\mathcal{U}_p}\Psi = [UZ_pV^*][V\Sigma^* Z_pV^*] = \tilde{H}\tilde{Y} \tag{73}$$

where $P_{\mathcal{U}_p}$ is the orthogonal projection onto the space spanned by the first p singular vectors u_i of Ψ. From (73), it follows that the left-hand matrix in the order-r TPD of Ψ is just the optimal normalized frame matrix $\hat{F}_c$ corresponding to $\beta = 1$. Similarly, the left-hand matrix in the order-p TPD of Ψ, with $p < r$, is the optimal normalized frame matrix corresponding to the vectors $P_{\mathcal{U}_p}\psi_i$.

Since the LSF is related to the PD of Ψ, properties of the optimal frame vectors can be deduced from properties of the PD (see, e.g., [46], [47], [49], [52]). For example, the CLSF corresponding to two vector sets $\{\varphi_i\}$ and $\{\mu_i\}$ are the same if and only if the corresponding frame matrices satisfy $FM^* = (FF^*)^{1/2}(MM^*)^{1/2}$ [52].

VII. Canonical Frames

A popular frame construction from a given set of vectors is the canonical frame. Given a set of vectors $\{\psi_i, 1 \le i \le n\}$, the *canonical frame* associated with these vectors is the frame corresponding to the frame matrix [8], [19], [24], [25]

$$F = \Psi\left((\Psi^*\Psi)^{1/2}\right)^{\dagger}. \tag{74}$$

Comparing (74) with (56), we see immediately that the canonical frame vectors are just the normalized tight-frame vectors that are closest in a least-squares sense to the vectors $\{\psi_i\}$. Furthermore, the β-scaled tight frame vectors for fixed β that are closest to the vectors $\{\psi_i\}$ are the canonical frame vectors scaled by β.

From Theorem 7, it follows that the canonical frame vectors are the tight-frame vectors that minimize the least-squares error only if $\hat{\beta} = 1$, i.e., only if $\sum_{i=1}^r \sigma_i = r$. Otherwise, the canonical frame is no longer the optimal tight frame in a least-squares sense. However, if we simply scale each of the canonical frame vectors by $\hat{\beta}$, then the resulting frame minimizes the least-squares error among all possible tight frames.

We summarize our results regarding canonical frames in the following theorem.

Theorem 8 (Canonical Frames): Let $\{\psi_i\}$ be a set of n vectors in a k-dimensional complex Hilbert space $\mathcal{H}$ that span an r-dimensional subspace $\mathcal{U} \subseteq \mathcal{H}$. Let $\Psi = U\Sigma V^*$ be the rank-r $k \times n$ matrix whose columns are the vectors ψ_i. Let u_i and v_i denote the columns of the unitary matrices U and V, respectively, let $\{\sigma_i, 1 \le i \le r\}$ denote the nonzero singular values of Ψ, and let Z_r be defined as in (1). Let $\{\varphi_i\}$ be the n canonical frame vectors associated with the vectors ψ_i, and let F denote the matrix of columns φ_i. Then

$$F = UZ_rV^* = \Psi\left((\Psi^*\Psi)^{\dagger}\right)^{1/2} = \left((\Psi\Psi^*)^{1/2}\right)^{\dagger}\Psi.$$

In addition, we have the following.

1) If $r = n$
   a) the canonical frame vectors form an orthonormal basis for $\mathcal{U}$;
   b) the canonical frame vectors are the closest orthonormal vectors to the vectors $\{\psi_i\}$, in a least-squares sense;
   c) if $\sum_{i=1}^r \sigma_i = r$, then the canonical frame vectors are the closest orthogonal vectors with equal norm to the vectors $\{\psi_i\}$, in a least-squares sense;
   d) define the scaled canonical frame vectors $\varphi_i' = \beta\varphi_i$. Then
      i) the scaled canonical frame vectors are the closest orthogonal vectors with norm β to the vectors $\{\psi_i\}$, in a least-squares sense;
      ii) if $\beta = (1/r)\sum_{i=1}^r \sigma_i$, then the scaled canonical frame vectors are the closest orthogonal vectors with equal norm to the vectors $\{\psi_i\}$, in a least-squares sense.
2) If $r < n$,
   a) the canonical frame vectors form a tight frame for $\mathcal{U}$;
   b) the canonical frame vectors are the closest normalized tight frame vectors to the vectors $\{\psi_i\}$, in a least-squares sense;
   c) if $\sum_{i=1}^r \sigma_i = r$, then the canonical frame vectors are the closest tight frame vectors to the vectors $\{\psi_i\}$, in a least-squares sense;
   d) define the scaled canonical frame vectors $\varphi_i' = \beta\mu_i$; then
      i) the scaled canonical frame vectors are the closest β-scaled tight frame vectors to the vectors $\{\psi_i\}$, in a least-squares sense;
      ii) if $\beta = (1/r)\sum_{i=1}^r \sigma_i$, then the scaled canonical frame vectors are the closest tight frame vectors to the vectors $\{\psi_i\}$, in a least-squares sense.

Acknowledgment

The authors are grateful to H. Bölcskei for comments and encouragement. Y. C. Eldar wishes to thank A. V. Oppenheim for his support.

References

[1] R. J. Duffin and A. C. Schaeffer, “A class of nonharmonic Fourier series,” *Trans. Amer. Math. Soc.*, vol. 72, pp. 314–366, 1952.
[2] R. M. Young, *An Introduction to Nonharmonic Fourier Series*. New York: Academic, 1980.
[3] J. J. Benedetto, “Irregular sampling and frames,” in *Wavelets—A Tutorial in Theory and Applications*, C. K. Chui, Ed. Boca Raton, FL: CRC, 1992, pp. 445–507.
[4] C. E. Heil and D. F. Walnut, “Continuous and discrete wavelet transforms,” *SIAM Rev.*, vol. 31, no. 4, pp. 628–666, Dec. 1989.

[5] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inform. Theory*, vol. 36, pp. 961–1005, Sept. 1990.

[6] A. Aldroubi, "Portraits of frames," *Proc. Amer. Math. Soc.*, vol. 123, pp. 1661–1668, 1995.

[7] S. O. Aase, J. H. Husøy, K. Skretting, and K. Engan, "Optimized signal expansions for sparse representations," *IEEE Trans. Signal Processing*, vol. 49, pp. 1087–1096, May 2001.

[8] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[9] N. J. Munch, "Noise reduction in tight Weyl-Heisenberg frames," *IEEE Trans. Inform. Theory*, vol. 38, pp. 608–616, Mar. 1992.

[10] Y. C. Eldar, "Sampling and reconstruction in arbitrary spaces and oblique dual frame vectors," *J. Fourier Analys. Appl.*, submitted for publication.

[11] Z. Cvetković and M. Vetterli, "Overcomplete expansions and robustness," in *Proc. IEEE TFTS-96*, Paris, France, June 1996, pp. 325–328.

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

[13] A. Aldroubi and K. Gröchenig, "Non-uniform sampling and reconstruction in shift-invariant spaces," SIAM Rev., to be published.

[14] Y. C. Eldar and A. V. Oppenheim, "Orthogonal and projected orthogonal matched-filter detection," *IEEE Trans. Signal Processing*, submitted for publication.

[15] Y. C. Eldar, "Quantum signal processing," Ph.D dissertation, Mass. Inst. Technol. (MIT), Cambridge, Dec. 2001.

[16] V. K. Goyal, J. Kovačević, and J. A. Kelner, "Quantized frame expansions with erasures," *Appl. Comp. Harmonic Anal.*, vol. 10, pp. 203–233, May 2001.

[17] Z. Cvetković, "Overcomplete expansions for digital signal processing," Ph.D. dissertation, Univ. California, Berkeley, Dec. 1995.

[18] Z. Cvetković and M. Vetterli, "Oversampled filter banks," *IEEE Trans. Signal Processing*, vol. 46, pp. 1245–1255, May 1998.

[19] H. Bölcskei, "Oversampled filter banks and predictive subband coders," Ph.D. dissertation, Vienna Univ. Technol., Vienna, Austria, Nov. 1997.

[20] H. Bölcskei, F. Hlawatsch, and H. G. Feichtinger, "Frame-theoretic analysis of oversampled filter banks," *IEEE Trans. Signal Processing*, vol. 46, pp. 3256–3268, Dec. 1998.

[21] P. J. S. G. Ferreira, "Mathematics for multimedia signal processing II—Discrete finite frames and signal reconstruction," in *Signal Processing for Multimedia*, J. S. Byrnes, Ed. Amsterdam, The Netherlands: IOS Press, 1999, pp. 35–54.

[22] M. A. Neumark, "On a representation of additive operator set functions," *C. R. (Dokl.) Acad. Sci. URSS (N.S.)*, vol. 41, pp. 359–361, 1943.

[23] A. Peres, "Neumark's theorem and quantum inseparability," *Found. Phys.*, vol. 20, no. 12, pp. 1441–1453, 1990.

[24] H. Bölcskei and A. J. E. M. Janssen, "Gabor frames, unimodularity, and window decay," *J. Fourier Anal. Appl.*, vol. 6, no. 3, pp. 255–276, 2000.

[25] A. J. E. M. Janssen and H. Bölcskei, "Equivalence of two methods for constructing tight Gabor frames," *IEEE Signal Processing Lett.*, vol. 7, pp. 79–82, Apr. 2000.

[26] A. J. E. M. Janssen and T. Strohmer. Characterization and computation of canonical tight windows for Gabor frames. [Online]. Available: http://xxx.lanl.gov/abs/math.FA/0010245

[27] Y. C. Eldar and H. Bolcskei. (2001, Aug.) Geometrically uniform frames [Online]. Available: http://arXiv.org/abs/math.FA/0108096

[28] Y. Meyer, "Ondelettes et functions splines," Seminaire EDP, Ecole Polytechnique, Paris, France, Dec. 1986.

[29] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Appl. Math.*, pp. 909–996, 1988.

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

[31] M. Unser and A. Aldroubi, "Families of multiresolution and wavelet spaces with optimal properties," *Numer. Funct. Anal. Optimiz.*, vol. 14, pp. 417–446, 1993.

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

[33] C. King and M. B. Ruskai, "Minimal entropy of states emerging from noisy quantum channels," *IEEE Trans. Inform. Theory*, vol. 47, pp. 192–209, Jan. 2001.

[34] Y. C. Eldar and G. D. Forney, Jr., "On measurements and density operators," in preparation.

[35] A. S. Holevo, "Statistical decisions in quantum theory," *J. Multivar. Anal.*, vol. 3, pp. 337–394, Dec. 1973.

[36] H. P. Yuen, R. S. Kennedy, and M. Lax, "Optimum testing of multiple hypotheses in quantum detection theory," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 125–134, Mar. 1975.

[37] C. W. Helstrom, *Quantum Detection and Estimation Theory*. New York: Academic, 1976.

[38] M. Charbit, C. Bendjaballah, and C. W. Helstrom, "Cutoff rate for the m-ary PSK modulation channel with optimal quantum detection," *IEEE Trans. Inform. Theory*, vol. 35, pp. 1131–1133, Sept. 1989.

[39] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, "Optimum measurements for discrimination among symmetric quantum states and parameter estimation," *Int. J. Theor. Phys.*, vol. 36, pp. 1269–1288, 1997.

[40] A. S. Holevo, "On asymptotically optimal hypothesis testing in quantum statistics," *Theor. Probab. and Appl.*, vol. 23, pp. 411–415, 1978.

[41] P. Hausladen and W. K. Wootters, "A 'pretty good' measurement for distinguishing quantum states," *J. Mod. Opt.*, vol. 41, pp. 2385–2390, 1994.

[42] H. Barnum and E. Knill. Reversing quantum dynamics with near-optimal quantum and classical fidelity. [Online]. Available: http://xxx.lanl.gov/abs/quant-ph/0004088.

[43] P. Hausladen, R. Josza, B. Schumacher, M. Westmorel, and W. K. Wootters, "Classical information capacity of a quantum channel," *Phys. Rev. A*, vol. 54, pp. 1869–1876, Sept. 1996.

[44] G. Kaiser, *A Friendly Guide to Wavelets*. Boston, MA: Birkhauser, 1994.

[45] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[46] N. J. Higham, "Computing the polar decomposition—With applications," *SIAM J. Sci. Stat. Comput.*, vol. 7, pp. 1160–1174, 1986.

[47] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[48] Y. C. Eldar, "Least-squares inner product shaping," *Linear Alg. Appl.*, to be published.

[49] P. Zieliński and K. Ziętak, "The polar decomposition—Properties, applications and algorithms," *Matematyka Stosowana*, vol. 38, pp. 23–40, 1995.

[50] A. A. Dubrulle, "An optimum iteration for the matrix polar decomposition," *Electron. Trans. Numer. Anal.*, vol. 8, pp. 21–25, 1999.

[51] O. Besson and P. Stoica, "Exponential signals with time-varying amplitude: Parameter estimation via polar decomposition," *Signal Processing*, vol. 66, pp. 27–43, 1998.

[52] A. Markiewicz, "Simultaneous polar decomposition of rectangular complex matrix matrices," *Linear Alg. Appl.*, vol. 289, pp. 279–284, 1999.