Optimal Noise Suppression: A Geometric Nature of Pseudoframes for Subspaces

Shidong Li* and Hidemitsu Ogawa†

*San Francisco State University
Department of Mathematics
San Francisco, CA 94132
shidong@sfsu.edu

†Tokyo Institute of Technology
Department of Computer Science
2-12-1 Ookayama, Meguro-ku
Tokyo 152-8552, Japan
ogawa@og.cs.titech.ac.jp

Abstract

Pseudoframes for subspaces (PFFS) is a notion of frame-like expansions for a subspace $\mathcal{X}$ in a separable Hilbert space [3]. The spanning nature of the sequences $\{x_n\}$ and $\{x_n^*\}$ in a PFFS (relative to the subspace $\mathcal{X}$) is generally very different from that of a frame. Incidentally, a PFFS constitutes generally a nonorthogonal projections onto $\mathcal{X}$. The directions of the projection determine the geometric meanings and its applications of an PFFS. PFFS also provides a mean for the construction of nonorthogonal projections that arises in various linear reconstruction problems. This article is aimed at elaborations on such geometrical properties and demonstration of natural needs of nonorthogonal projections in applications and how PFFS can be applied, particularly for optimal noise suppressions. In this specific application, we show that PFFS is not only natural and sufficient but also necessary for generating an optimal solution among the class of all linear and series-based methods.

Key Words: Pseudoframes for subspaces, frames, nonorthogonal projections, optimal noise suppression.

Math Subject Classification: 42-02, 42C15.
1 Introduction

Let $\mathcal{X}$ be a closed subspace of a separable Hilbert space $\mathcal{H}$. Let $\{x_n\} \subseteq \mathcal{H}$ be a Bessel sequence with respect to $\mathcal{X}$ and $\{x_n^*\} \subseteq \mathcal{H}$ be a Bessel sequence (in $\mathcal{H}$). We say $\{x_n\}$ is a pseudo-frame for the subspace $\mathcal{X}$ (PFFS) w.r.t. $\{x_n^*\}$ if

$$\forall f \in \mathcal{X}, \quad f = \sum_n \langle f, x_n \rangle x_n^*.$$  \hspace{1cm} (1)

The most distinct property of PFFSs is that none of the sequences $\{x_n\}$ and $\{x_n^*\}$ are required to be in $\mathcal{X}$. Consequently, the decomposition sequence $\{x_n\}$ and the reconstruction sequence $\{x_n^*\}$ are not generally in the same subspace either. PFFS is thus non-symmetric in general. Positions of the two sequences in (1) are not inter-changeable, a property known as non-commutativity. The construction of PFFSs are therefore more involved than that of traditional frames. There are two different directions of the construction - one from given $\{x_n\}$ to find $\{x_n^*\}$; the other from given $\{x_n^*\}$ to determine $\{x_n\}$. Details about the construction can be found in [3].

As a simple example, let us re-visit the following:

**Example 1.1 Biorthogonal Basis:** Let $\{x_n^*\}$ be a Riesz basis for a closed subspace $\mathcal{X} \subseteq \mathcal{H}$. We have then a unique (biorthogonal) dual sequence $\{x_n^0\} \subseteq \mathcal{X}$ such that $\langle x_m^0, x_n^* \rangle = \delta_{mn}$ and

$$\forall f \in \mathcal{X}, \quad f = \sum_n \langle f, x_n^* \rangle x_n^0 = \sum_n \langle f, x_n \rangle x_n^*.$$  \hspace{1cm} (2)

In the context of biorthogonal bases within $\mathcal{X}$, this is usually the end of it. However, consider now a function $\Delta x_n \in \mathcal{X}^\perp \subseteq \mathcal{H}$, and let

$$x_n = x_n^0 + \Delta x_n.$$  \hspace{1cm} (3)

We see obviously that

1. $\langle x_m, x_n^*\rangle = \langle x_m^0 + \Delta x_m, x_n^*\rangle = \langle x_m^0, x_n^*\rangle + 0 = \delta_{mn}$, and

2. for all $f \in \mathcal{X}$,

$$\sum_n \langle f, x_n \rangle x_n^* = \sum_n \langle f, x_n^0 + \Delta x_n \rangle x_n^* = f.$$  \hspace{1cm} (3)
The conclusion is that there are infinitely many biorthogonal duals \( \{ x_n^* \} \) to a Riesz basis \( \{ x_n \} \) of \( \mathcal{X} \), so long if \( \{ x_n \} \) is allowed to go beyond \( \mathcal{X} \). This is the very nature and flexibility of PFFSs.

We also noted that biorthogonal duals \( \{ x_n^* \} \) to the given biorthogonal basis \( \{ x_n \} \) of \( \mathcal{X} \) can be constructed by a different way. Assume that \( \mathcal{P} \) is a projection operator onto \( \mathcal{X}^\perp \) in the direction of \( \mathcal{N}(\mathcal{P}) \), i.e., \( \mathcal{P} \) is a nonorthogonal projection in general, then the sequence \( \{ x_n \} \) given by

\[
x_n = x_n^0 + \mathcal{P} y_n, \quad \forall n
\]

would be always biorthogonal to \( \{ x_n \} \), for any \( y_n \in \mathcal{H} \).

This seems to have no difference than that of (2), but the projection direction of \( \mathcal{P} \) would yield different updating component to \( x_n^0 \) (for the same element \( y_n \)). Indeed, the projection direction makes some differences in the geometric behavior of the corresponding PFFS.

To see the projection nature of PFFSs, let us recall the most fundamental characterization of pseudo-frame for subspaces. Let us define \( U : \mathcal{X} \to l^2(\mathbb{Z}) \) by

\[
\forall f \in \mathcal{X}, \quad U f = \{ \langle f, x_n \rangle \},^1
\]

and define \( V : l^2(\mathbb{Z}) \to \mathcal{H} \) such that

\[
\forall c \equiv \{ c(n) \} \in l^2(\mathbb{Z}), \quad V c = \sum_n c(n)x_n^*.
\]

Then the following characterization of PFFS holds.

**Theorem 1.2** [3] Let \( \{ x_n \} \) and \( \{ x_n^* \} \) be two sequences in \( \mathcal{H} \) (not necessarily in \( \mathcal{X} \)). Assume that \( \{ x_n \} \) is a Bessel sequence w.r.t. the subspace \( \mathcal{X} \), and \( \{ x_n^* \} \) is a Bessel sequence in \( \mathcal{H} \). Let \( U \) be defined by (4), and \( V \) be defined by (5). Suppose that \( \mathcal{P} \) is a projection from \( \mathcal{H} \) onto \( \mathcal{X} \). Then \( \{ x_n \} \) is a pseudo-frame for \( \mathcal{X} \) w.r.t. \( \{ x_n^* \} \) if and only if

\[
V U \mathcal{P} = \mathcal{P}.
\]

Therefore, the constructions of PFFS all start from (6), and

\[
\forall f \in \mathcal{H}, \quad \mathcal{P} f = \sum_n \langle \mathcal{P} f, x_n \rangle x_n^*
\]

\[
= \sum_n \langle f, \mathcal{P} x_n^\perp \rangle x_n x_n^*.
\]

^1Note that if \( \{ x_n \} \) is a Bessel sequence in \( \mathcal{H} \), then \( U \) is well-defined on \( \mathcal{H} \).
Evidently, the projection operator $\mathcal{P}$ can be an arbitrary projection onto $\mathcal{X}$. $\mathcal{P}$ inevitably plays a key role of the geometric properties of an PFFS.

In a special case, if we assume that $\{x_n\}$ is also a Bessel sequence in $\mathcal{H}$, then PFFS can be applied to $\mathcal{H}$, i.e.,

$$\forall f \in \mathcal{H}, \sum_n \langle f, x_n \rangle x_n^*$$

is well defined, and for a properly chosen $\mathcal{P}$ in a PFFS, (7) corresponds to a special nonorthogonal projection onto $\mathcal{X}$ as we will discuss in this article.

The organization of the article is as follows. Section 2 discusses a $\mathcal{P}$-consistent property of a Bessel sequence $\{x_n\}$ and its relation to PFFSs. A necessary and sufficient conditions for a Bessel sequence $\{x_n\}$ to be $\mathcal{P}$-consistent is provided. The notion of $\mathcal{P}$-consistent PFFS is also defined. Necessary and sufficient conditions are also presented. In Section 3, constructions of nonorthogonal projections via PFFS are discussed in detail. The principle of consistency is more appropriate for the construction of projections through PFFS. Focusing in Section 4, we present a major application of PFFSs in optimal noise suppression. Using the geometric nature of PFFSs, we show that PFFS is not only natural but also necessary for general noise suppression, among linear and series-based methods. We prove that regular frame/basis-based method will not be able to suppress the noise to the maximum extent.

2 $\mathcal{P}$-consistent sequence $\{x_n\}$ and PFFS

It turns out that for a given subspace $\mathcal{X}$, the selection of the projection direction (onto $\mathcal{X}$) and the sequence $\{x_n\}$ in a PFFS can make some differences. Such differences will not merely reflect in geometric interpretations. More importantly, they make the implementation of the projection simpler.

Let $\mathcal{P}$ be a bounded linear projection onto $\mathcal{X}$ along $\mathcal{N}(\mathcal{P})$ denoted by $\mathcal{P} \equiv \mathcal{P}_{x, \mathcal{N}(\mathcal{P})}$. Following Unser and Aldroubi [9], we define the “consistent principle” as follows:

Definition 2.1 Let $\{x_n\}$ be a Bessel sequence. We say that $\{x_n\}$ is $\mathcal{P}$-consistent for $\mathcal{X}$ if

$$U \mathcal{P} = U.$$
Note that in the context of sampling, the “consistent principle” \( U\mathcal{P} = U \) implies that the projection approximation \( \mathcal{P} f \) of \( f \) (in \( \mathcal{X} \)) and the original \( f \) have the consistent/same “sampling” values, or the same inner product values with the sequence \( \{ x_n \} \). It is also clear that the consistent principle is also the interpolation principle.

Indeed, a Bessel sequence \( \{ x_n \} \) can be \( \mathcal{P} \)-consistent for the subspace \( \mathcal{X} \). But, having a \( \mathcal{P} \)-consistent sequence \( \{ x_n \} \) is certainly not the purpose. Eventually, this is to be integrated with a PFFS for \( \mathcal{X} \) for reconstructions. We shall see that the notion of \( \mathcal{P} \)-consistency is more useful in the construction of nonorthogonal projections via PFFS, and thereby useful in optimal noise suppression.

But first, let us examine what make a sequence \( \{ x_n \} \) \( \mathcal{P} \)-consistent for \( \mathcal{X} \).

**Theorem 2.2** Let \( \mathcal{X} \) be a closed subspace of \( \mathcal{H} \), and let \( \mathcal{P} \) be a projection onto \( \mathcal{X} \). A Bessel sequence \( \{ x_n \} \subseteq \mathcal{H} \) is \( \mathcal{P} \)-consistent if and only if

\[
\mathcal{N}(\mathcal{P}) \subseteq \mathcal{P}\mathcal{F}\{ x_n \}^\perp.
\]

**Proof:** Assume that \( U\mathcal{P} = U \). Then \( \mathcal{N}(\mathcal{P}) \subseteq \mathcal{N}(U\mathcal{P}) = \mathcal{N}(U) = \mathcal{P}\mathcal{F}\{ x_n \}^\perp \).

Conversely, if \( \mathcal{N}(\mathcal{P}) \subseteq \mathcal{P}\mathcal{F}\{ x_n \}^\perp \), then \( \mathcal{N}(\mathcal{P}) \subseteq \mathcal{N}(U) \), and \( U\mathcal{P}_{\mathcal{N}(\mathcal{P}),\mathcal{X}} = 0 \). Hence,

\[
U\mathcal{P} = U (I - (I - \mathcal{P})) = U - U\mathcal{P}_{\mathcal{N}(\mathcal{P}),\mathcal{X}} = U.
\]

Next, let us put a \( \mathcal{P} \)-consistent sequence in a PFFS. Under the assumptions that \( \mathcal{P}^*(\mathcal{P}\mathcal{F}\{ x_n \}) = \mathcal{N}(\mathcal{P})^\perp \) and that \( \mathcal{R}(U\mathcal{P}) \) is closed, there are PFFS (for reconstructions) corresponding to a given \( \mathcal{P} \)-consistent sequence \( \{ x_n \} \). Such a PFFS also provides a nonorthogonal projection onto \( \mathcal{X} \) for all functions in \( \mathcal{H} \):

\[
\mathcal{P} = VU\mathcal{P} = VU \text{ on } \mathcal{H}.
\]

That is, if \( U\mathcal{P} = U \), (7) mentioned earlier is a special PFFS that constitutes a nonorthogonal projection onto \( \mathcal{X} \) along \( \mathcal{N}(\mathcal{P}) \) (where \( \mathcal{N}(\mathcal{P}) = \mathcal{P}\mathcal{F}\{ x_n \}^\perp \), see also Theorem 2.3).

With a slight abuse of notation, we shall also term a PFFS satisfying \( U\mathcal{P} = U \) a \( \mathcal{P} \)-consistent PFFS.
Evidently, the choice of the projection $\mathcal{P}$ and the span of $\{x_n\}$ both affect geometric behaviors of a PFFS. We shall see a couple of such properties of PFFS in the following.

**Theorem 2.3** Let $\{x_n\}$ and $\{x^*_n\}$ be Bessel sequences in $\mathcal{H}$. Let $\mathcal{P}$ be a projection onto $\mathcal{X}$. Then $\{x_n\}$ and $\{x^*_n\}$ form a $\mathcal{P}$-consistent PFFS for $\mathcal{X}$ if and only if

\[ VU = \mathcal{P} \quad \text{and} \quad \mathcal{N}(\mathcal{P}) = \mathcal{P} \{x_n\}^\perp. \]

**Proof:** Assume $U\mathcal{P} = U$. Then for all $f \in \mathcal{H}$, $\langle f, x_n \rangle = \langle Uf, e_n \rangle = \langle U\mathcal{P}f, x_n \rangle$, where $\{e_n\}$ is the standard ONB of $l^2$. Assume also that $\{x_n\}$ and $\{x^*_n\}$ form a PFFS for $\mathcal{X}$. We then have $VV = \mathcal{P}$ by (6). For any $h \in \mathcal{N}(\mathcal{P})$, and for all $n \in \mathbb{Z}$,

\[ \langle h, x_n \rangle = \langle \mathcal{P} h, x_n \rangle = 0. \]

Hence, $h \in \mathcal{P} \{x_n\}^\perp$, and $\mathcal{N}(\mathcal{P}) \subseteq \mathcal{P} \{x_n\}^\perp$. Meantime, for all $g \in \mathcal{P} \{x_n\}^\perp$,

\[ \mathcal{P}g = \sum_{n} \langle \mathcal{P}g, x_n \rangle x^*_n = \sum_{n} \langle g, x_n \rangle x^*_n = 0, \]

implying $g \in \mathcal{N}(\mathcal{P})$. Consequently, $\mathcal{N}(\mathcal{P}) = \mathcal{P} \{x_n\}^\perp$.

Conversely, if $UU = \mathcal{P}$, and $\mathcal{N}(\mathcal{P}) = \mathcal{P} \{x_n\}^\perp$. Then, $UU \mathcal{P} = \mathcal{P}$, and $\mathcal{N}(\mathcal{P}) \subseteq \mathcal{P} \{x_n\}^\perp$. By Theorem 2.2, $U\mathcal{P} = U$. \qed

**Remark:**

a. To have $U\mathcal{P} = U$, all one needs is that $\mathcal{N}(\mathcal{P}) \subseteq \mathcal{P} \{x_n\}^\perp$, cf., Theorem 2.2. However, for PFFSs, $U\mathcal{P} = U$ is equivalent to $\mathcal{N}(\mathcal{P}) = \mathcal{P} \{x_n\}^\perp$.

b. In general, $UU = \mathcal{P}$ does not imply $U\mathcal{P} = U$. $UU = \mathcal{P}$ would imply the $\mathcal{P}$-consistency provided that $V$ is left invertible (or $\mathcal{N}(V) = \{0\}$, or $V$ is one-to-one).

There is also another by-product for $\mathcal{P}$-consistent PFFS. Let us start from a more general scenario.

**Proposition 2.4** Let $\{x_n\}$ be a PFFS for $\mathcal{X}$ w.r.t. $\{x^*_n\}$. Then

\[ \forall g \in \mathcal{R}(\mathcal{P}^*), \quad g = \sum_{n} \langle g, x^*_n \rangle \mathcal{P}^* x_n. \]
The proof uses the fact that $\mathcal{P}^* = \mathcal{P}^* U^* V^*$ from the PFFS assumption. Notice that under the assumption of $\mathcal{P}$-consistency, $\mathcal{R}(\mathcal{P}^*) = \mathcal{N}(\mathcal{P})^\perp = \overline{\mathcal{P}} \{ x_n \}$. We have therefore the following corollary.

**Corollary 2.5** Let $\{ x_n \}$ and $\{ x_n^* \}$ be Bessel sequences in $\mathcal{H}$. If they form a $\mathcal{P}$-consistent PFFS for $\mathcal{X}$, then

$$\forall g \in \overline{\mathcal{P}} \{ x_n \}, \quad g = \sum_n \langle g, x_n^* \rangle x_n. \quad (8)$$

**Remark:** In case of a $\mathcal{P}$-consistent PFFS, Corollary 2.5 also asserts that $\{ x_n \}$ is necessarily a frame sequence (i.e., a frame for its closed linear span $\overline{\mathcal{P}} \{ x_n \}$). The upper frame bound of $\{ x_n \}$ is given by the Bessel assumption, and the lower frame bound is implied by the upper Bessel bound of the sequence $\{ x_n^* \}$ and the equation (8). However, different from regular frames, that $\{ x_n \}$ is a frame sequence does not imply that $\{ x_n^* \}$ in (8) is a regular dual frame of $\{ x_n \}$ in $\overline{\mathcal{P}} \{ x_n \}$.

### 3 Construction of nonorthogonal projections via PFFS

It doesn’t take a lot to realize that in the context of sampling the consistent principle is exactly the interpolation principle in numerical analysis. As is well-known, interpolation problems often result in “consistent” approximations that pass through the interpolation points, but offset substantially at non-sample points. This, when measured by a common norm, implies that the approximation can be substantially different from the original function. Such a difference can go well beyond tolerance. One can think of such a phenomenon in polynomial interpolations for a visual comprehension.

While the non-orthogonal projection aspect of PFFS does yield a consistency principle in a special case, we comment that the real value of consistent principle lies in its convenience of generating a given non-orthogonal projection via PFFS.

In the following, we shall use geometric properties of PFFS to demonstrate how non-orthogonal projections $\mathcal{P}$ can be constructed via PFFS.

If a PFFS is $\mathcal{P}$-consistent for $\mathcal{X}$, then $VU = \mathcal{P}_{x, \overline{\mathcal{P}} \{ x_n \}^\perp}$, implying that the PFFS gives rise to an nonorthogonal projection as shown in Theorem 2.3. The
following theorem deals with a “dual” problem of what is addressed above. Namely, let \( P \) be any projection onto \( X \). One can actually construct such a \( P \) through an (and any) PFFS for \( X \).

**Theorem 3.1** Let \( P \) be a projection onto \( X \). Let \( \{x_n\} \) and \( \{x_n^*\} \) be Bessel sequences in \( \mathcal{H} \), and assume that \( \{x_n\} \) and \( \{x_n^*\} \) form a PFFS for \( X \). Then there is a sequence \( \{x_n^{(1)}\} \subseteq \mathcal{N}(P) \) such that \( \{x_n^{(1)}\} \) and \( \{x_n^*\} \) form a \( P \)-consistent PFFS for \( X \), and generate the projection by

\[
\forall f \in \mathcal{H}, \quad Pf = \sum_n \langle f, x_n^{(1)} \rangle x_n^*.
\]

**Proof:** Since \( P \) is a projection onto \( X \), we have \( X + \mathcal{N}(P) = \mathcal{H} \) and \( X \cap \mathcal{N}(P) = \{0\} \). Consequently, \( X \perp \perp \mathcal{N}(P) \perp = \mathcal{H} \) (we have used “\( \perp \perp \)” as the (nonorthogonal) direct sum of two subspaces). Therefore, since \( x_n \in \mathcal{H} \), it can be decomposed into

\[
\forall n \in \mathbb{Z}, \quad x_n = x_n^{(1)} + x_n^{(2)},
\]

where \( x_n^{(1)} \in \mathcal{N}(P) \) and \( x_n^{(2)} \in X^\perp \), and

\[
\forall f \in \mathcal{H}, \quad Pf = \sum_n \langle Pf, x_n \rangle x_n^* = \sum_n \langle f, P^*(x_n^{(1)} + x_n^{(2)}) \rangle x_n^* = \sum_n \langle f, x_n^{(1)} \rangle x_n^*.
\]

This implies, by the definition of PFFS, that \( \{x_n^{(1)}\} \) and \( \{x_n^*\} \) form a PFFS for \( X \), and obviously, they are also a \( P \)-consistent PFFS for \( X \). The \( P \)-consistency can be seen either from direct verification that \( U^{(1)}P = U^{(1)} \), or it is a simple fact that \( \mathcal{P} \{x_n^{(1)}\} = \mathcal{N}(P) \). Here \( U^{(1)} \) corresponds to \( \{x_n^{(1)}\} \) and is similarly defined as in (4).

Theorem 3.1 asserts that one may construct a projection onto \( X \) starting from any PFFS for \( X \). We mention that it is also possible to construct a (nonorthogonal) projection \( P \) onto \( X \) directly via PFFS theory.

**Theorem 3.2** Let \( P \) be a given projection onto \( X \). Let \( \{x_n^*\} \) be a Bessel sequence in \( \mathcal{H} \) such that \( \mathcal{R}(V) \) is closed and \( \mathcal{P} \{x_n^*\} \supseteq X \). Then the class of all
sequences \( \{x_n\} \) such that \( \sum_n \langle f, x_n \rangle x_n^* = \mathcal{P} f \) for all \( f \in \mathcal{H} \) is given by

\[
x_n = \mathcal{P}^* \left( x_n^\dagger + y_n - \sum_m \langle x_m^\dagger, x_m^* \rangle y_m \right), \quad \forall n \in \mathbb{Z}, \tag{9}
\]

where \( x_n^\dagger \equiv (V^*)^\dagger e_n \), the pseudo-inverse of \( V^* \) acting on the standard ONB \( \{e_n\} \) in \( l^2 \), and \( \{y_n\} \) is an arbitrary Bessel sequence in \( \mathcal{H} \).

This result is a special case of one direction of the construction of PFFS in [3]. For the proof of Theorem 3.2, we recall two construction results from [3].

**Proposition 3.3** [3] Let \( \{x_n^\dagger\} \) be a Bessel sequence in \( \mathcal{H} \) such that \( \mathcal{R}(V) \) is closed and \( \mathcal{R}(x_n^\dagger) \supseteq \mathcal{X} \). Then the set of all linear operator \( U \) satisfying (6) is given by

\[
U = V^\dagger \mathcal{P} + W - V^\dagger WV^\dagger, \tag{10}
\]

where \( V^\dagger \) is the pseudo-inverse of \( V \), and \( W : \mathcal{H} \to l^2 \) is a varying bounded linear operator. Moreover, let \( U \) be given by (10), and define

\[
x_n = U^* e_n \quad \forall n \in \mathbb{Z}, \tag{11}
\]

where \( \{e_n\} \) is the standard basis of \( l^2 \). Then \( \{x_n\} \) is a dual PFFS sequence for \( \mathcal{X} \) w.r.t \( \{x_n^\dagger\} \), and (11) gives all dual PFFS sequences by changing \( W \) in (10).

The next formula follows directly from **Proposition 3.3**:

**Corollary 3.4** [3] Let \( \{x_n^\dagger\} \) be defined by the pseudo-inverse of \( V^* \) through \( x_n^\dagger = (V^*)^\dagger e_n \) and let \( \{y_n\} \) be a Bessel sequence in \( \mathcal{H} \) with which a bounded linear operator \( W \) is defined by

\[
\forall c \in l^2, \quad Wc = \sum_n c(n) y_n. \tag{12}
\]

Then, all PFFS sequences \( \{x_n\} \) for \( \mathcal{X} \) w.r.t. \( \{x_n^\dagger\} \) are given by

\[
x_n = \mathcal{P}^* x_n^\dagger + y_n - \sum_m \langle x_m^\dagger, x_m^* \rangle \mathcal{P}^* y_m, \quad \forall n \in \mathbb{Z}. \tag{13}
\]

**Proof of Theorem 3.2:** Notice that our essential goal is to construct a \( \mathcal{P} \)-consistent PFFS for \( \mathcal{X} \). For this purpose, all we need to make sure is that \( \mathcal{P}_{\mathcal{X}(\mathcal{P})^\perp, \mathcal{X}^\perp} x_n = x_n \) for all \( n \) from Corollary 3.4. Since the first and the last terms
of \( x_n \) in (13) are already in \( \mathcal{N}(\mathcal{P})^\perp \), we may replace \( \{y_n\} \) by \( \{\mathcal{P}_{\mathcal{N}(\mathcal{P})^\perp} y_n\} \) and (9) follows.

We hereby comment on a work on “oblique dual frames”, cf. [1] where the idea is to construct a dual frame \( \{x_n^*\} \) to a given frame \( \{x_n^*\} \) for \( \mathcal{X} = \mathcal{P}\{x_n^*\} \). Essentially, “oblique dual frame” is a (frame) sequence \( \{x_n\} \subseteq \mathcal{N}(\mathcal{P})^\perp \). We mention though our construction is more general since \( \mathcal{P}\{x^*\} \supseteq \mathcal{X} \). While PFFS certainly include the case of \( \mathcal{X} = \mathcal{P}\{x_n^*\} \) and that \( \{x_n^*\} \) is a frame for \( \mathcal{X} \) as considered in [1], our PFFS considers a broader class of frame-like expansions for the subspace \( \mathcal{X} \).

4 Noise suppression: why PFFS is natural and necessary

We show in this section that PFFS is not only a natural but also a necessary tool for general additive noise suppression (among linear operators) [8]. To see the necessity of PFFS in noise suppression, we shall show that regular frames can not achieve noise suppression purposes to the maximum extend, unless noises are merely contained in the orthogonal complement of the signal subspace. The spirit of this section is partly presented by Ogawa, et al. in [5]. We demonstrate here in substantially different and much simpler ways, making use of the nonorthogonal geometric nature of PFFS. We show that the analysis and derivation of optimal (noise suppression) solutions naturally leads to PFFSs.

4.1 The noise operator and the noise subspace

For noise suppression purpose only, we shall consider a signal model

\[ g = f + n, \]

where a signal \( f \in \mathcal{X} \) is contaminated by an additive noise \( n \). We comment that for other linear models \( g = Af + n \) (where \( A : \mathcal{H} \rightarrow \mathcal{H} \)), noise suppression can be done either in exactly the same way by considering \( \mathcal{X} = \mathcal{R}(A) \), or by using various projection filter approaches [4], [6], [11], [10], where similar results with the need of nonorthogonal projections hold.
Our purpose is to find bounded linear operators $B$ such that

$$J[B] \equiv E_n\|Bg - f\|^2$$

is minimized. We shall also assume naturally that if $g \in \mathcal{X}$, then $B$ does not create distortion to $g$, i.e., $Bg = g$. In other words, we are to minimize $J[B]$ subject to $B\mathcal{P} = \mathcal{P}$.

Since

$$J[B] = E_n\|Bg - f\|^2 = E_n\|Bn\|^2 = E_n \left( tr(Bn \otimes \overline{Bn}) \right) = tr \left( B E_n[n \otimes \overline{n}] B^* \right), \tag{15}$$

where $(f \otimes \overline{g})$ is an operator known as the Neumann-Shatten product [7], with which for given $f$ and $g$ in $\mathcal{H},$

$$(f \otimes \overline{g})h \equiv \langle h, g \rangle f, \quad \forall h \in \mathcal{H}.$$  

We see that the operator $E_n[n \otimes \overline{n}]$ naturally arises. Following previous treatment [11], we term this operator $Q$, namely,

$$Q \equiv E_n[n \otimes \overline{n}].$$

Hence (15) becomes

$$J[B] = tr(BQB^*). \tag{16}$$

$Q$ is obviously the correlation operator of the noise ensemble $n$. $Q$ is therefore nonnegative, and the range of $Q$, $\overline{\mathcal{R}(Q)}$ (with the closure) depicts the general noise subspace that we want to analyze and to eliminate to the maximum extent.

### 4.2 Solution

Consider an observation operator for the signal model (14):

$$T \equiv \mathcal{P} + Q. \tag{17}$$

Obviously, $\overline{\mathcal{R}(T)}$ is the observation subspace where the signal in $\mathcal{X}$ is seen and the noise in $\overline{\mathcal{R}(Q)}$ is added, i.e.,

$$\overline{\mathcal{R}(T)} = \mathcal{X} + \overline{\mathcal{R}(Q)}.$$
Here $+$ stands for simply the sum of two linear subspaces. As a general theorem due to Ogawa, we can find a direct sum decomposition of $\overline{\mathcal{R}(T)}$ which plays an ultimate role in the need for a PFFS as a nonorthogonal projection.

**Lemma 4.1** Let $S \subseteq \mathcal{H}$ be a closed subspace and $A$ be a positive semi-definite operator defined on $\mathcal{H}$. Then

$$S + \overline{\mathcal{R}(A)} = S + \overline{AS^\perp}.$$ 

**Proof:** If $S = \mathcal{H}$, the lemma clearly holds. We shall prove the case when $S \subset \mathcal{H}$. This takes several steps:

a. We shall show that

$$S \cap AS^\perp = \{0\}. \quad (18)$$

For any $u \in S \cap AS^\perp$, there exists $0 \neq v \in S^\perp$ such that $u = Av$ and $\langle u, v \rangle = 0$.

Hence, $\langle Av, v \rangle = 0$, implying that $u = Av = 0$ since $A$ is nonnegative.

Note that in the proof of (18), if it happens that $S$ is not closed, the result still holds. We shall use this comment in the next step.

b. Let us establish $S + \overline{AS^\perp} = S + \overline{\mathcal{R}(A)}$. Take an element $u \in S^\perp \cap (AS^\perp)^\perp$.

By (18) and treating $AS^\perp$ as $S$ in (18),

$$Au \in AS^\perp \cap A(AS^\perp)^\perp = \{0\}.$$ 

Therefore, $u \in \mathcal{N}(A)$ and $S^\perp \cap (AS^\perp)^\perp \subseteq \mathcal{N}(A)$. Take the orthogonal complement of this last relation, we have

$$S + \overline{AS^\perp} \supseteq \overline{\mathcal{R}(A)},$$

which in turn implies $S + \overline{AS^\perp} \supseteq S + \overline{\mathcal{R}(A)}$.

Meantime, since $AS^\perp \subseteq \mathcal{R}(A)$, taking the closure and adding $S$ to this relation, we obtain the other enclosing relationship $S + \overline{AS^\perp} \subseteq S + \overline{\mathcal{R}(A)}$. We have thus shown $S + \overline{AS^\perp} = S + \overline{\mathcal{R}(A)}$.

c. Finally, let us prove

$$S \cap \overline{AS^\perp} = \{0\}.$$ 

Let $u \in S \cap \overline{AS^\perp}$. Because of (18), there exists $v \in S^\perp$ such that for any $\epsilon > 0$,

$$\|u - Av\| \leq \epsilon. \quad (19)$$
Moreover,  \( \langle u, v \rangle = 0 \). So now,

\[
|\langle Av, v \rangle| = |\langle Av - u, v \rangle + \langle u, v \rangle| \\
= |\langle Av - u, v \rangle| \\
\leq \|Av - u\|\|v\| \leq \varepsilon\|v\|.
\]

We see that \( Av = 0 \) since \( A \geq 0 \). This together with (19) implies that \( \|u\| \leq \varepsilon \) or \( u = 0 \).

Steps b and c prove the assertion. \( \blacksquare \)

The following corollary is thereby immediate since \( Q \) is a nonnegative operator:

**Corollary 4.2** Let \( T \) and \( Q \) be defined in (17), and \( \mathcal{X} \) be the signal subspace of interests. Then

\[
\overline{\mathcal{R}(T)} = \mathcal{X} + \overline{Q\mathcal{X}^\perp}.
\]

The subspace \( \overline{Q\mathcal{X}^\perp} \) (or \( Q\mathcal{X}^\perp \)) in Corollary 4.2 plays an important role for the optimal noise suppression purposes. We have the following result similar to that in [5]:

**Theorem 4.3** Suppose \( B\mathcal{P} = \mathcal{P} \). Then \( B \) minimizes \( J[B] \) if and only if \( \mathcal{N}(B) \supseteq \overline{Q\mathcal{X}^\perp} \).

**Proof:** a. We shall first show that \( B \) is optimal if and only if \( B\mathcal{P} = \mathcal{P} \), and \( BQ = CP \) for some operator \( C \) and an orthogonal projection \( P = P^* \).

For the sufficiency, assume that \( B_0 \) satisfies \( B_0\mathcal{P} = \mathcal{P} \) (and hence \( B_0P = P \) since \( \mathcal{P}P = P \)), and \( B_0Q = CP \) as stated earlier. Then

\[
B_0QB_0^* = CPB_0^* = C(B_0P)^* = CP^* = CP.
\] (20)

Since the left-hand side of (20) is self-adjoint, for any \( B \) satisfying \( B\mathcal{P} = \mathcal{P} \), we have

\[
B_0QB_0^* = PC^* = BPC^* = B(CP)^* = B(B_0Q)^* = BQB_0^*.
\] (21)

(21) also implies

\[
B_0QB_0^* = B_0QB^*.
\] (22)
With (21), (22) and in reference of (16), it is straightforward to verify that

\[ J[B] - J[B_0] = \text{tr} \left( (B - B_0)Q(B - B_0)^* \right) \geq 0. \] (23)

Therefore, \( B_0 \) yields the least \( J \) and is an optimal solution.

Conversely, assume that \( B \) is an optimal solution satisfying \( BP = \mathcal{P} \). Then \( J[B] = J[B_0] \), where \( B_0 \) is the optimal solution seen in the last paragraph. Calculations in (20) to (23) all hold, and (23) implies

\[ \text{tr} \left( (B_0 - B)Q(B_0 - B)^* \right) = 0 \quad \iff \quad (B_0 - B)Q = 0, \]

and hence \( BQ = B_0Q = CP \) as asserted.

b. We show next that \( \mathcal{N}(B) \supseteq Q\mathcal{X}^\perp \) if and only if there exists an operator \( C \) such that \( BQ = CP \).

(\( \Rightarrow \)) If \( \mathcal{N}(B) \supseteq Q\mathcal{X}^\perp \), then \( \mathcal{N}(B) \supseteq Q\mathcal{X}^\perp \), and \( BQ\mathcal{X}^\perp = \{0\} \). This is to say \( \mathcal{N}(P) \subseteq \mathcal{N}(BQ) \), which further implies that there is always an operator \( C \) for which the equation \( BQ = CP \) holds.

(\( \Leftarrow \)) If \( BQ = CP \), then \( BQ\mathcal{X}^\perp = CP\mathcal{X}^\perp = \{0\} \). Therefore, \( \mathcal{N}(B) \supseteq Q\mathcal{X}^\perp \), implying \( \mathcal{N}(B) \supseteq Q\mathcal{X}^\perp \) since \( \mathcal{N}(B) \) is closed.

Remark: a. Theorem 4.3 shows that \( Q\mathcal{X}^\perp \) is the subspace that contains the most of noises outside of \( \mathcal{X} \), and by eliminating \( Q\mathcal{X}^\perp \), \( B \) provides the optimal noise suppression.

b. We also comment that since \( \mathcal{N}(B) \) is closed, \( \mathcal{N}(B) \supseteq Q\mathcal{X}^\perp \) is equivalent to \( \mathcal{N}(B) \supseteq Q\mathcal{X}^\perp \).

4.3 PFFS is a natural solution

Let \( \mathcal{H} = \overline{\mathcal{R}(T)^\perp} + \mathcal{R}(T)^\perp \), where \( \mathcal{R}(T)^\perp \) is any fixed complementary subspace of \( \overline{\mathcal{R}(T)} \) in \( \mathcal{H} \). Then Corollary 4.2 gives rise to

\[ \mathcal{H} = (\mathcal{X} + Q\mathcal{X}^\perp) + \mathcal{R}(T)^\perp = \mathcal{X} + (Q\mathcal{X}^\perp + \mathcal{R}(T)^\perp). \]

In light of Theorem 4.3, the following result is immediate.

Theorem 4.4 Let

\[ \mathcal{Y} = Q\mathcal{X}^\perp + \mathcal{R}(T)^\perp. \] (24)
Assume that $P$ is an projection onto $\mathcal{X}$ with $\mathcal{N}(P) = \mathcal{Y}$, i.e., $P = P_{x,y}$. Then any $P$-consistent PFFS $\{x_n, x^*_n\}$ for $\mathcal{X}$ is naturally an optimal solution. Namely, if $B$ is defined by

$$\forall g \in \mathcal{H}, \quad Bg = \sum_n \langle g, x_n \rangle x^*_n,$$

then $B$ minimizes $J[B]$ subject to $BP = P$.

**Proof:** Let $\{x_n\}$ and $\{x^*_n\}$ be a pair of $P$-consistent PFFS for $\mathcal{X}$, and let $U$ and $V$ be defined by (4) and (5), respectively. Then $B = VU = P$. Therefore, $BP = P$, and $\mathcal{N}(B) = \mathcal{N}(P) = \mathcal{Y} \subseteq \overline{Q\mathcal{X}^\perp}$.

**Remark:** This is what we mean by PFFS being a natural solution. In fact, PFFS is so convenient that as long as one starts from a PFFS for $\mathcal{X}$, by setting $\mathcal{N}(P) = \mathcal{Y}$ as defined in (24), an optimal noise suppression is obtained immediately. We discuss in the following three methods of calculating the solution $Bg$ using PFFS.

**The calculation of $Bg$ via PFFS**

We shall explain in a moment that PFFS is also fundamentally necessary for general noise suppression purposes among linear solutions, particularly when compared with regular frame approaches. But first, let us examine how PFFS can be convenient as an optimal (noise suppression) solution. Three different approaches using PFFS are examined:

(i) **Starting from any PFFS**

Let $\{x_n\}$ and $\{x^*_n\}$ be a PFFS for $\mathcal{X}$, then the optimal noise suppression solution goes as follows:

$$\forall g \in \overline{R(T)}, \quad Bg = \sum_n \langle P_{x,y} g, x_n \rangle x^*_n = \sum_n \langle g, P_{y^\perp, x^\perp} x_n \rangle x^*_n; \quad (25)$$

(ii) **Using Theorem 3.1**

Let again $\{x_n\}$ and $\{x^*_n\}$ be a PFFS for $\mathcal{X}$. Essentially as we seen in Theorem 3.1, we may also decompose the sequence $\{x_n\}$ into $\{x_n^{(1)}\} \subseteq \mathcal{Y}^\perp$ and $\{x_n^{(2)}\} \subseteq \mathcal{Y}^\perp$. 

15
$\mathcal{X}$. Then

$$\forall g \in \overline{\mathcal{R}(T)}, \quad Bg = \sum_n \langle g, x_n^{(1)} \rangle x_n^*,$$

(26)

(iii) **Using Theorem 3.2**

The third approach is by the construction of Theorem 3.2. Namely, we construct a solution directly by letting $\{x_n^*\}$ be a Bessel sequence such that $\overline{\mathcal{P}\{x_n^*\}} \supseteq \mathcal{X}$, and (see details in Theorem 3.2) construct PFFS-dual sequences $\{x_n\}$ by, as in (9):

$$x_n = \mathcal{P}_{y^\perp,x^\perp} \left( x_n^* + y_n - \sum_m \langle x_n^*, x_m^* \rangle y_m \right), \quad \forall n \in \mathbb{Z}. \quad (27)$$

For such an $\{x_n\}$, an optimal noise suppression solution becomes

$$\forall g \in \overline{\mathcal{R}(T)}, \quad Bg = \sum_n \langle g, x_n \rangle x_n^*. \quad (28)$$

We mention that while the first term of (27), $\{\mathcal{P}_{y^\perp,x^\perp}x_n^\perp\}$, is sufficient to serve as a PFFS-dual sequence $\{x_n\}$, other terms in the equation are additional flexibilities for the construction of more regular functions $x_n$. This point of view is demonstrated explicitly in the construction of biorthogonal wavelets using PFFS [2].

Let us not forget to mention that the reason why PFFS is natural to achieve such noise suppression goals is because that $\{x_n\}$ is allowed to go beyond the subspace $\mathcal{X}$, a property that regular frame lacks.

This can be seen in all three optimal noise suppression solutions (25), (26) and (28). The suppression of noise is due to the fact that $\overline{\mathcal{P}\{x_n\}}$ is essentially steered to be perpendicular to $\mathcal{Y}$ that contains the noise subspace. This is especially clear in (28) without loss of the generality seen in (25) and (26) as well. And the steering possibility of $\overline{\mathcal{P}\{x_n\}}$ lies in the fact that $\{x_n\}$ does not have to be in $\mathcal{X}$ and that $\overline{\mathcal{P}\{x_n\}}$ does not have to even contain $\mathcal{X}$, which is impossible in the context of frames, but natural in PFFSs.

### 4.4 PFFS is necessary for optimal noise suppression purposes

Among linear and series-based solutions, particularly when compared with conventional frame or basis based approaches, PFFS is also necessary to suppress
additive noises to the maximum extent. This is because that regular frame approaches will not do.

**Proposition 4.5** [5] There is a regular frame solution to the noise suppression problem with optimal \( B = VU \) if and only if

\[
\overline{Q\mathcal{X}^\perp} \subseteq \mathcal{X}^\perp.
\]  

(29)

Here, \( U \) and \( V \) are defined by (4) and (5) by the frame and dual frame sequences \( \{x_n\} \) and \( \{x_n^*\} \) in \( \mathcal{X} \).

**Proof:** Here is a simple proof. If \( \{x_n\} \) and \( \{x_n^*\} \) are a frame pair in \( \mathcal{X} \), then \( B = VU \) is an orthogonal projection onto \( \mathcal{X} \) and \( BP = P \). Therefore, if condition (29) holds, \( BQ\mathcal{X}^\perp = \{0\} \) and \( B \) is optimal. Conversely, if \( B \) is optimal, then \( \overline{Q\mathcal{X}^\perp} \subseteq \mathcal{N}(B) = \mathcal{X}^\perp \).

**Proposition 4.5** asserts that if the removable noise subspace \( \overline{Q\mathcal{X}^\perp} \) is contained in \( \mathcal{X}^\perp \), then a regular frame can be used for an optimal solution. In general, however, \( \overline{Q\mathcal{X}^\perp} \not\subseteq \mathcal{X}^\perp \). An optimal linear solution that suppresses the noise to the maximum extent has to go beyond a frame approach.

Even if \( \overline{Q\mathcal{X}^\perp} \subseteq \mathcal{X}^\perp \), there is still an added benefit using PFFS. In such cases, an optimal solution using PFFS satisfies that \( s\mathcal{P}\{x_n^*\} \supseteq \mathcal{X} \) and \( s\mathcal{P}\{x_n\} = \mathcal{X} \). There are not only additional freedom in the choices of \( \{x_n^*\} \), one can also see that any regular frame solution would simply be a part of the PFFS class.

**Most general solutions**

While a PFFS for \( \mathcal{X} \) in \( \mathcal{H} \) provides a natural solution to the optimal noise suppression problem, and while regular frame approaches will not be able to suppress the noise if \( \overline{Q\mathcal{X}^\perp} \not\subseteq \mathcal{X}^\perp \), PFFSs simply provide a class (the most useful class) of solutions. This can be seen in the general solution. By **Theorem 4.3**, the most general solution is clearly given by

\[
\forall g \in \mathcal{H}, \quad Bg = \begin{cases} 
g, & \text{if } g \in \mathcal{X}, \\
0, & \text{if } g \in \overline{Q\mathcal{X}^\perp}, \\
Wg, & \text{if } g \in \mathcal{R}(T)^-, \end{cases}
\]
where $W : \mathcal{H} \to \mathcal{H}$ is an arbitrary bounded linear operator, and $\mathcal{R}(T)^-\parallel$ is any fixed complementary subspace of $\mathcal{R}(T)$ in $\mathcal{H}$.

It is clear that if $W = 0$, then a solution can be constructed by a PFFS for $\mathcal{X}$ in $\mathcal{H}$ since $\mathcal{N}(B) = \mathcal{Y}$ as defined in (24). We note that for the purpose of noise suppressions, the observation $g \in \overline{\mathcal{R}(T)}$, $W$ may not make any contribution to the noise suppression problem. It is very natural to have $W = 0$, and hence the PFFS solutions given by Theorem 4.4.

**Deriving all series-based optimal solutions via PFFS**

Nevertheless, we also comment that all series-based optimal solutions can be constructed by PFFS of $\mathcal{X}$ in $\mathcal{H}_1 \equiv \mathcal{X} + Q\mathcal{X}^{\perp} = \overline{\mathcal{R}(T)}$. In such cases, let $\mathcal{Y} \equiv \overline{Q\mathcal{X}^{\perp}}$. Then any $\mathcal{P}_{x,y}$-consistent PFFS for $\mathcal{X}$ in $\mathcal{H}_1$ is an optimal solution. Conversely, any series-based optimal solution can be constructed by a $\mathcal{P}_{x,y}$-consistent PFFS for $\mathcal{X}$ in $\mathcal{H}_1$.

**References**


