Sampling theory, oblique projections and a question by Smale and Zhou

Jorge Antezana and Gustavo Corach

September 12, 2005

Abstract

Keywords: sampling, frames, oblique projection, angles and compatibility .
AMS classification: Primary 47A30, 94A20 Secondary: 47B32, 68P05, 46E22.

1 Introduction

A recent paper by S. Smale and D. X. Zhou [23] on Shannon sampling theorem deals with reducing noise in the sampling data, using probability estimates and a measure of the richness of the data. Related problems have also recently studied by S. Li and H. Ogawa [21], Y. Eldar [14], O. Christensen and Y. Eldar [7] and Y. Eldar and T. Werther [15], among others. In this note we show that the use of oblique projections in Hilbert spaces together with a convenient modification of the measure of richness of the data used by Smale and Zhou, give an error estimation which answer a question raised in their paper (see [23] Remark p. 303). Our method works as well for abstract Hilbert spaces as for reproducing kernel Hilbert spaces.
Given a complete metric space X and a reproducing kernel Hilbert space $\mathcal{H}$ of function defined on X, consider discrete subsets $\overline{t}$, $\overline{x}$ of X and define, as Smale and Zhou, the closed subspaces

$$\mathcal{H}_{k,\ \overline{t}} = \overline{\operatorname{span}\left\langle k_t :\ t \in \overline{t}\right\rangle} \qquad \text{and} \qquad \mathcal{H}_{k,\ \overline{x}} = \overline{\operatorname{span}\left\langle k_x :\ x \in \overline{x}\right\rangle}.$$

of $\mathcal{H}$, where k is the kernel of $\mathcal{H}$. Suppose that $\{k_t\}_{t\in\overline{t}}$ is a frame for $\mathcal{H}_{k,\ \overline{t}}$ and $\{k_x\}_{x\in\overline{x}}$ is a frame for $\mathcal{H}_{k,\ \overline{x}}$ and let F, G be their synthesis operators. Denote $K_{\overline{t},\overline{t}} = F^*F$, $K^{\overline{x},\overline{x}} = GG^*$ and $K_{\overline{x},\overline{t}} = K^*_{\overline{t},\overline{x}} = G^*F$. Finally, let D_ω be the diagonal operator with respect to the canonical basis of $\ell^2(\overline{x})$ defined by a sequence $\omega = \{\omega_x\}_{x\in\overline{x}}$ of positive numbers with $\inf \omega_x > 0$. Given $f \in \mathcal{H}$ and $y = \{f(x)\}_{x\in\overline{x}}$ (the sampling data) it is proved that the solution of the minimization problem

$$f_{\alpha,\omega} = \arg \min_{h\in\mathcal{H}_{k,\ \overline{t}}} \Big(\sum_{x\in\overline{x}} \omega_x |h(x) - f(x)|^2 + \alpha \|h\|^2_{\mathcal{H}} \Big),$$

is $f_{\alpha,\omega} = \sum_{t\in\bar{t}} L_{\alpha,\omega}(y)(t)\ k_t$, where $L_{\alpha,\omega} = \Big(K_{\bar{t},\bar{x}} D_\omega K_{\bar{x},\bar{t}} + \alpha K_{\bar{t},\bar{t}}\Big)^\dagger K_{\bar{t},\bar{x}} D_\omega$.
In addition, if $\overline{x}$ provides rich data with respect to $\bar{t}$ and ω (see definition 4.1), then

$$\|f_{\alpha,\omega} - f\|_{\mathcal{H}} \le \Big\|(1 - P_{D,\mathcal{H}_{k,\ \bar{t}}})(f)\Big\|_{\mathcal{H}} + \left(\alpha \frac{\|K_{\bar{t},\bar{t}}\|}{\gamma(D_\omega^{1/2} K_{\bar{x},\bar{t}}) + \alpha\ \gamma(K_{\bar{t},\bar{t}})}\right) \|f\|_{\mathcal{H}}$$

where $D = (K^{\bar{x},\bar{x}} + \alpha P_{\mathcal{H}_{k,\ \bar{t}}})$ and $P_{D,\mathcal{H}_{k,\ \bar{t}}}$ is the orthogonal projection onto $\mathcal{H}_{k,\ \bar{t}}$ with respect to the semi-inner product on $\mathcal{H}$ defined by D (see details in section 2).

2 Preliminaries

Let $\mathcal{H}$ be a separable Hilbert space, $L(\mathcal{H})$ the algebra of bounded linear operators on $\mathcal{H}$ and $L(\mathcal{H})^+$ the cone of positive (semi-definite) operators. For an operator $A \in L(\mathcal{H})$, we denote by $R(A)$ the range or image of A, $N(A)$ the nullspace of A, $\sigma(A)$ the spectrum of A, A^* the adjoint of A, $\|A\|$ the usual norm of A and, if $R(A)$ is closed, $A^\dagger$ the Moore-Penrose pseudoinverse of A. Given a closed subspace $\mathcal{S}$ of $\mathcal{H}$, we denote by $P_{\mathcal{S}}$ the orthogonal (i.e. selfadjoint) projection onto $\mathcal{S}$.

Angle between subspaces and reduced minimum modulus

We need the following two definitions of angles between subspaces in a Hilbert space; they are due, respectively, to Friedrichs and Dixmier (see [12] and [16], and the excellent survey by Deutsch [11]).

Definition 2.1. Given two closed subspaces $\mathcal{M}$ and $\mathcal{N}$, the *Friedrichs angle* between $\mathcal{M}$ and $\mathcal{N}$ is the angle in $[0, \pi/2]$ whose cosine is defined by

$$c\,[\,\mathcal{M},\,\mathcal{N}\,] = \sup\Big\{\,|\,\langle x,\,y\rangle\,|\,:\ x \in \mathcal{M} \ominus (\mathcal{M}\cap\mathcal{N}),\ y \in \mathcal{N} \ominus (\mathcal{M}\cap\mathcal{N}) \text{ and } \|x\| = \|y\| = 1\Big\}.$$

The *Dixmier angle* between $\mathcal{M}$ and $\mathcal{N}$ is the angle in $[0, \pi/2]$ whose cosine is defined by

$$c_0\,[\,\mathcal{M},\,\mathcal{N}\,] = \sup\Big\{\,|\,\langle x,\,y\rangle\,|\,:\ x \in \mathcal{M},\ y \in \mathcal{N} \text{ and } \|x\| = \|y\| = 1\Big\}.$$

The next proposition collects the results on angles which are relevant to our work.

Proposition 2.2. *Let $\mathcal{M}$ and $\mathcal{N}$ be two closed subspaces of $\mathcal{H}$. Then*

1. $c\,[\,\mathcal{M},\,\mathcal{N}\,] = c_0\,[\,\mathcal{M} \ominus (\mathcal{M}\cap\mathcal{N}),\,\mathcal{N}\,] = c_0\,[\,\mathcal{M},\,\mathcal{N} \ominus (\mathcal{M}\cap\mathcal{N})\,]$.
2. $c\,[\,\mathcal{M},\,\mathcal{N}\,] = c\,\big[\,\mathcal{M}^\perp,\,\mathcal{N}^\perp\,\big]$
3. $c\,[\,\mathcal{M},\,\mathcal{N}\,] < 1$ *if and only if* $\mathcal{M} + \mathcal{N}$ *is closed.*
4. $\mathcal{H} = \mathcal{M}^\perp + \mathcal{N}^\perp$ *if and only if* $c_0\,[\,\mathcal{M},\,\mathcal{N}\,] < 1$.

Definition 2.3. Given $T \in L(\mathcal{H})$, the *reduced minimum modulus* $\gamma(T)$ is defined by

$$\gamma(T) = \inf\{\|Tx\| : \|x\| = 1 \ , \ x \in N(T)^\perp\} \tag{1}$$

It is well known that $\gamma(T) = \gamma(T^*) = \gamma(T^*T)^{1/2}$. Also, it can be shown that an operator T has closed range if and only if $\gamma(T) > 0$. In this case, $\gamma(T) = \|T^\dagger\|^{-1}$.

The next result has been proved in [2] (See also [5]).

Proposition 2.4. *If $A, B \in L(\mathcal{H})$ have closed ranges, then*

$$\gamma(A)\gamma(B) \ s\,[\,N(A),\ R(B)\,] \le \gamma(AB) \le \|A\|\,\|B\| \ s\,[\,N(A),\ R(B)\,]\,. \tag{2}$$

In particular, AB has closed range if and only if $s\,[\,N(A),\ R(B)\,] > 0$.

D-selfadjoint projections and compatibility

Any $D \in L(\mathcal{H})^+$ defines a bounded, positive and sesquilinear form $\langle \xi, \eta \rangle_D = \langle D\xi, \eta \rangle$, $\xi,\ \eta \in \mathcal{H}$. We say that $C \in L(\mathcal{H})$ is *D-selfadjoint* if $DC = C^*D$. Consider the set of D-selfadjoint projections whose range is exactly $\mathcal{S}$:

$$\mathcal{P}(D, \mathcal{S}) = \{Q \in \mathcal{Q} : R(Q) = \mathcal{S}, DQ = Q^*D\}.$$

A pair $(D, \mathcal{S})$ is called *compatible* if $\mathcal{P}(D, \mathcal{S})$ is not empty. In this case, there exists a distinguished projection $P_{D,\mathcal{S}} \in \mathcal{P}(D, \mathcal{S})$ whose nullspace is $D^{-1}(\mathcal{S}^\perp) \ominus \big(D^{-1}(\mathcal{S}^\perp) \cap \mathcal{S}^\perp\big)$.

In the following theorem we present several results about compatibility, taken from [8] and [9].

Theorem 2.5. *Given $D \in L(\mathcal{H})^+$, let $\mathcal{S}$ be a closed subspace of $\mathcal{H}$ such that the pair $(D, \mathcal{S})$ is compatible. Then*

1. *$\mathcal{P}(D, \mathcal{S})$ has a unique element if and only if $N(D) \cap \mathcal{S} = \{0\}$.*
2. *$P_{D,\mathcal{S}}$ has minimal norm in $\mathcal{P}(D, \mathcal{S})$, i.e. $\|P_{D,\mathcal{S}}\| = \min\{\ \|Q\| : Q \in \mathcal{P}(D, \mathcal{S})\}$.*

The reader is referred to [8], [9] and [10] for several applications of $P_{D,\mathcal{S}}$ (see also Hassi and Nordström [17]).

Frames

We introduce some basic facts about frames in Hilbert spaces. For complete descriptions of frame theory and applications, the reader is referred to the review by Heil and Walnut [18] or the books by Young [25] and Christensen [6].

Definition 2.6. Let $\mathcal{H}$ be a separable Hilbert space, $\mathcal{W}$ a closed subspace of $\mathcal{H}$ and $\mathcal{F} = \{f_n\}_{n\in\mathbb{N}}$ a sequence in $\mathcal{W}$. The sequence $\mathcal{F}$ is called a *frame* for the subspace $\mathcal{W}$ if there exist numbers $A, B > 0$ such that, for every $f \in \mathcal{W}$,

$$A\|f\|^2 \leq \sum_{n\in\mathbb{N}} |\langle f, f_n\rangle|^2 \leq B\|f\|^2 \tag{3}$$

The optimal constants A, B for equation (3) are called the *frame bounds* for $\mathcal{F}$. $\mathcal{F}$ is a *tight frame* if $A = B$, and it is a *Parseval frame* if $A = B = 1$.

Associated with $\mathcal{F}$ there is an operator $T : \ell^2 \to \mathcal{H}$ such that $T(e_n) = f_n$, where $\{e_n\}_{n\in\mathbb{N}}$ denotes the canonical basis of ℓ^2. This operator is called the *synthesis operator* of $\mathcal{F}$. In the case of finite dimensional frames we assume that the domain of the synthesis operator is $\mathbb{C}^m$ where m is the number of vectors of the frame. The adjoint $T^* \in L(\mathcal{H}, \ell^2)$ of T, given by $T^*(f) = \sum_{n\in\mathbb{N}}\langle f, f_n\rangle e_n$, is called *analysis operator* of $\mathcal{F}$, and the operator $S = TT^*$ is usually called the *frame operator* of $\mathcal{F}$. Observe that

$$Sf = \sum_{n\in\mathbb{N}} \langle f, f_n\rangle f_n \qquad f \in \mathcal{W}. \tag{4}$$

It follows from (3) that $A.P_{\mathcal{W}} \leq S \leq B.P_{\mathcal{W}}$, so that $S|_{\mathcal{W}}$ is invertible in $L(\mathcal{W})$.

3 Sampling in abstract Hilbert spaces

Throughout this section $\mathcal{M}$ and $\mathcal{W}$ denote closed subspaces of $\mathcal{H}$, $\{f_n\}$ and $\{g_n\}$ are frames for the subspaces $\mathcal{W}$ and $\mathcal{M}$ respectively, with synthesis operators F and G. Finally $\{e_n\}$ denotes the canonical orthonormal basis of ℓ^2.

The next notion is an extension of one introduced by Smale and Zhou [23] in reproducing kernel Hilbert spaces.

Definition 3.1. *We say that G provides rich data with respect to F if*

$$\inf_{z\in N(F)^\perp} \|G^*F(z)\| > 0.$$

This notion is related to a decomposition of $\mathcal{H}$ as a sum of $\mathcal{W}^\perp$ and $\mathcal{M}$

Proposition 3.2. *The following statements are equivalent:*

1. *G provides rich data with respect to F.*
2. *G^*F is a closed range operator and $\mathcal{W} \cap \mathcal{M}^\perp = \{0\}$.*
3. *$c_0\left[\mathcal{W}, \mathcal{M}^\perp\right] < 1$.*
4. *$c\left[\mathcal{W}^\perp, \mathcal{M}\right] < 1$ and $\mathcal{W}^\perp + \mathcal{M} = \mathcal{H}$.*

Proof.

$1 \Rightarrow 2$ The condition of rich data implies that the restriction of G^*F to $N(F)^\perp$ is injective, so $\mathcal{W} \cap \mathcal{M}^\perp = \{0\}$. On the other hand, since $N(G^*F) = N(F)$, it also implies that $\gamma(G^*F) > 0$. Hence G^*F has closed range.

$2 \Rightarrow 3$ It is a direct consequence of Proposition 2.4.

$3 \Rightarrow 4$ By proposition 2.2, $c_0\left[\mathcal{W}, \mathcal{M}^\perp\right] = c\left[\mathcal{W}, \mathcal{M}^\perp\right]$ and $c\left[\mathcal{W}, \mathcal{M}^\perp\right] = c\left[\mathcal{W}^\perp, \mathcal{M}\right]$. So, $c\left[\mathcal{W}, \mathcal{M}^\perp\right] < 1$. This in particular shows that $\mathcal{W}^\perp + \mathcal{M}$ is closed. Therefore, since $c_0\left[\mathcal{W}, \mathcal{M}^\perp\right] < 1$ also implies that $\mathcal{W} \cap \mathcal{M}^\perp = \{0\}$,

$$\mathcal{W}^\perp + \mathcal{M} = \overline{\mathcal{W}^\perp + \mathcal{M}} = (\mathcal{W} \cap \mathcal{M}^\perp)^\perp = \mathcal{H}.$$

$4 \Rightarrow 1$ Firstly, note that $\mathcal{W} \cap \mathcal{M}^\perp = (\mathcal{W}^\perp + \mathcal{M})^\perp = \{0\}$. Hence, $N(G^*F) = N(F)$. So, by Proposition 2.4,

$$\begin{aligned} \inf_{x \in N(F)^\perp} \|G^*F(x)\| &= \gamma(G^*F) \geq \gamma(G^*)\ \gamma(F)\ s\left[R(F), N(G^*)\right] \\ &= \gamma(G^*)\ \gamma(F)\ s\left[\mathcal{W}, \mathcal{M}^\perp\right] > 0. \end{aligned}$$

■

Remark 3.3. Note that the above proposition emphasizes the fact that the hypothesis of rich data only depends on the subspaces and not on the particular frames chosen for each subspace. ▲

The next result shows that D-selfadjoint projections, for a convenient D, play a relevant role in certain minimization problems.

Proposition 3.4. *Let $f \in \mathcal{H}$, $y = \{y_n\} = G^*(f)$ the sampling data, and assume that G provides rich data with respect to F. Then, the solution of the minimization problem*

$$f_{\mathcal{W}} = \arg\min_{h \in \mathcal{W}} \sum_{n=1}^{\infty} |\langle h,\ g_n\rangle - y_n|^2$$

is given by $f_{\mathcal{W}} = P_{GG^*,\mathcal{W}}(f)$. In particular

$$\|f_{\mathcal{W}} - f\| = \|(1 - P_{GG^*,\mathcal{W}})(f)\| \leq s\left[\mathcal{W}, (GG^*)^{-1}(\mathcal{W}^\perp)\right]^{-1} \|f\|.$$

Proof. First of all, note that

$$\begin{aligned} f_{\mathcal{W}} &= \arg\min_{h \in \mathcal{W}} \sum_{n=1}^{\infty} |\langle h,\ g_n\rangle - y_n|^2 \\ &= F\Big(\arg\min_{z \in N(F)^\perp} \Big\|G^*F(z) - y\Big\|_{\ell^2}\Big) \\ &= F\Big(\arg\min_{z \in N(G^*F)^\perp} \Big\|G^*F(z) - y\Big\|_{\ell^2}\Big), \end{aligned}$$

where we have used that $N(G^*F) = N(F)$ because G provides rich data with respect to F. Since $\arg \min_{z \in N(G^*F)^\perp} \left\| G^*F(z) - y \right\|_{\ell^2} = (G^*F)^\dagger(y) = (G^*F)^\dagger G^*(f)$, we get

$$f_{\mathcal{W}} = F(G^*F)^\dagger G^*(f). \tag{5}$$

Let $Q = F(G^*F)^\dagger G^*$. It is easy to see that Q is a projection whose range is $\mathcal{W}$. On the other hand, as $(G^*F)^\dagger = (G^*FF^*G)^\dagger F^*G$,

$$(GG^*)Q = (GG^*)F(G^*FF^*G)^\dagger F^*GG^*$$

which shows that $(GG^*)Q$ is selfadjoint, and therefore Q is (GG^*)-selfadjoint. Finally, since $N(GG^*) \cap \mathcal{W} = N(G^*) \cap \mathcal{W} = \mathcal{M}^\perp \cap \mathcal{W} = \{0\}$, by Theorem 2.5, there exists only one (GG^*)-selfadjoint projection onto $\mathcal{W}$. Hence, $Q = P_{GG^*, \mathcal{W}}$.

■

Now, we are interested in estimating the reconstruction error when we use a perturbed data instead of the original one.

Proposition 3.5. *Let $y = \{y_n\} = G^*(f)$ and $f_{\mathcal{W}}$ as in Proposition 3.4. Suppose that $\widehat{f}_{\mathcal{W}}$ is the vector obtained when we use $\widehat{y} = \{\widehat{y}_n\}$ instead of the original data $y = \{y_n\}$, then:*

$$\|\widehat{f}_{\mathcal{W}} - f_{\mathcal{W}}\| \leq \|F(G^*F)^\dagger\| \|y - \widehat{y}\| \leq \frac{\|F\|}{\gamma(G^*F)} \|y - \widehat{y}\| \leq \frac{\|F\|}{\gamma(G)\gamma(F) c_0 \left[\mathcal{W}, \mathcal{M}^\perp\right]} \|y - \widehat{y}\|$$

Proof. It follows from equation (5), the definition of reduced minimum modulus, and Proposition 2.4. ■

3.1 The weighted regularized case

Let $\alpha > 0$ and $\{\omega_n\}$ a sequence of positive numbers bounded from above. In this subsection D_ω will denote the diagonal bounded operator on ℓ^2 defined by $D_\omega(e_n) = \omega_n \, e_n$. Before stating our first result, we need to modify the definition of rich data in the presence of the diagonal operator D_ω.

Definition 3.6. *We say that G provides rich data with respect to F and D_ω if*

$$\inf_{z \in N(F)^\perp} \|D_\omega^{1/2} G^* F(z)\| > 0.$$

Remark 3.7. As in the non-regularized case, G provides rich data with respect to F and D_ω if and only if $c_0 \left[\mathcal{W}, \mathcal{M}^\perp\right] < 1$. However, if the operator D_ω is not invertible, the property of having rich data depends not only on the subspaces $\mathcal{W}$ and $\mathcal{M}$, but also on the operators G, F and D_ω. ▲

Now, we are ready to state the first result of this subsection:

Proposition 3.8. *Let $f \in \mathcal{H}$ and $y = \{y_n\} = G^*(f)$ be the sampling data. Then the solution of the minimization problem:*

$$f_{\mathcal{W},\alpha,\omega} = \arg\min_{h\in\mathcal{W}} \Big(\sum_{n=1}^{\infty} \omega_n |\langle h,\ g_n\rangle - y_n|^2 + \alpha\|h\|^2 \Big), \tag{6}$$

is given by $f_{\mathcal{W},\alpha,\omega} = L_{\alpha,\omega}(f)$, where $L_{\alpha,\omega}$ is defined by

$$L_{\alpha,\omega} = F\Big(F^*GD_\omega G^*F + \alpha F^*F\Big)^\dagger F^*GD_\omega G^*.$$

In particular, if we assume that G provides rich data with respect to F and D_ω, we get the following estimation of the reconstruction error:

$$\|f_{\mathcal{W},\alpha,\omega} - f\| \le \Big\|(1 - P_{(GG^*+\alpha P_{\mathcal{W}}),\mathcal{W}})(f)\Big\| + \left(\alpha \frac{\|F\|^2}{\gamma(D^{1/2}G^*F)^2 + \alpha\ \gamma(F)^2}\right)\|f\| \tag{7}$$

Remark 3.9. The first term appears if f does not belong to $\mathcal{W}$. On the other hand, the term $\alpha\left(\frac{\|F\|^2}{\gamma(D^{1/2}G^*F)^2 + \alpha\ \gamma(F)^2}\right)\|f\|$ tends to zero as $\alpha \to 0$ and only depends on the regularization. ▲

Before proving Proposition 3.8, we need the following norm estimation.

Lemma 3.10. *Suppose that G provides rich data with respect to F and D_ω. Then,*

$$\left\|\Big(F^*(GD_\omega G^* + \alpha P_{\mathcal{W}})F\Big)^\dagger\right\| \le \frac{1}{\gamma(D^{1/2}G^*F)^2 + \alpha\ \gamma(F)^2}.$$

Proof. Note that $N(F^*GD_\omega G^*F) = N(F^*F) = N(F^*GD_\omega G^*F - \alpha F^*F)$. Then

$$\begin{aligned}\gamma(F^*(GD_\omega G^* + \alpha P_{\mathcal{W}})F) &= \gamma(F^*GD_\omega G^*F + \alpha F^*F)\\ &\ge \gamma(F^*GD_\omega G^*F) + \gamma(\alpha F^*F)\\ &= \gamma(D_\omega^{1/2}G^*F)^2 + \alpha\gamma(F)^2.\end{aligned}$$

■

Proof of proposition 3.8. Let $\mathcal{L} = \ell^2 \oplus \mathcal{H}$ and consider the operator $T : \ell^2 \to \mathcal{L}$ defined by

$$T(z) = \begin{pmatrix} D_\omega^{1/2}G^*F \\ \alpha^{1/2}F \end{pmatrix}(z) = \Big(D_\omega^{1/2}G^*F(z)\Big) \oplus \Big(\alpha^{1/2}F(z)\Big).$$

In term of the operator T, using the fact that $N(T) = N(F)$, the least square problem stated in (6) can be rewritten in the following way:

$$\begin{aligned} f_{\mathcal{W},\alpha,\omega} &= \arg\min_{h\in\mathcal{W}} \Big(\sum_{n=1}^{\infty} \omega_n |\langle h,\ g_n\rangle - y_n|^2 + \alpha\|h\|^2 \Big)\\ &= F\left(\arg\min_{z\in N(T)^\perp} \Big\|T(z) - (D_\omega^{1/2}(y) \oplus 0)\Big\|\right).\end{aligned}$$

Therefore, $f_{\mathcal{W},\alpha,\omega} = FT^\dagger(D_\omega^{1/2}(y) \oplus 0)$. Using the identity $A^\dagger = (A^*A)^\dagger A^*$, we get

$$\begin{aligned} T^\dagger &= \left(\begin{pmatrix} F^*GD_\omega^{1/2} & \alpha^{1/2}F^* \end{pmatrix} \begin{pmatrix} D_\omega^{1/2}G^*F \\ \alpha^{1/2}F \end{pmatrix} \right)^\dagger \begin{pmatrix} F^*GD_\omega^{1/2} & \alpha^{1/2}F^* \end{pmatrix} \\ &= \Big(F^*GD_\omega G^*F + \alpha F^*F \Big)^\dagger \begin{pmatrix} F^*GD_\omega^{1/2} & \alpha^{1/2}F^* \end{pmatrix}. \end{aligned}$$

Hence,

$$\begin{aligned} f_{\mathcal{W},\alpha,\omega} &= F\Big(F^*GD_\omega G^*F + \alpha F^*F \Big)^\dagger \begin{pmatrix} F^*GD_\omega^{1/2} & \alpha^{1/2}F^* \end{pmatrix} (D_\omega^{1/2}(y) \oplus 0) \\ &= F\Big(F^*GD_\omega G^*F + \alpha F^*F \Big)^\dagger F^*GD_\omega G^*(f) = L_{\alpha,\omega}(f) \end{aligned}$$

Straightforward computations show that $\Big(F\Big(F^*(GD_\omega G^* + \alpha P_{\mathcal{W}})F \Big)^\dagger (F^*(GD_\omega G^* + \alpha P_{\mathcal{W}})) = P_{(GG^*+\alpha P_{\mathcal{W}}),\mathcal{W}}$. Using this fact and Lemma 3.10 we obtain:

$$\begin{aligned} \|f_{\mathcal{W},\alpha,\omega} - f\| &= \Big\| \Big(F\big(F^*GD_\omega G^*F + \alpha F^*F\big)^\dagger F^*GD_\omega G^* - I \Big)(f) \Big\| \\ &= \Big\| \Big(F\big(F^*(GD_\omega G^* + \alpha P_{\mathcal{W}})F\big)^\dagger F^*GD_\omega G^* - I \Big)(f) \Big\| \\ &\leq \left\| \Big(F\Big(F^*(GD_\omega G^* + \alpha P_{\mathcal{W}})F\Big)^\dagger (F^*(GD_\omega G^* + \alpha P_{\mathcal{W}})) - I \Big)(f) \right\| \\ &\quad + \left\| \Big(F\Big(F^*(GD_\omega G^* + \alpha P_{\mathcal{W}})F\Big)^\dagger \alpha F^*P_{\mathcal{W}} \Big)(f) \right\| \\ &= \Big\| (I - P_{(GG^*+\alpha P_{\mathcal{W}}),\mathcal{W}})(f) \Big\| + \left\| \Big(F\Big(F^*(GD_\omega G^* + \alpha P_{\mathcal{W}})F\Big)^\dagger \alpha F^*P_{\mathcal{W}} \Big)(f) \right\| \\ &\leq \Big\| (I - P_{(GG^*+\alpha P_{\mathcal{W}}),\mathcal{W}})(f) \Big\| + \left\| \Big(F^*(GD_\omega G^* + \alpha P_{\mathcal{W}})F\Big)^\dagger \right\| \ \alpha\|F\|^2 \ \|f\| \\ &\leq \Big\| (1 - P_{(GG^*+\alpha P_{\mathcal{W}}),\mathcal{W}})(f) \Big\| + \left(\alpha \frac{\|F\|^2}{\gamma(D_\omega^{1/2}G^*F) + \alpha\ \gamma(F)^2} \right) \|f\| \end{aligned}$$

■

As before, we also want an estimation of the error produced by a perturbation of the sampling data.

Proposition 3.11. *Let $y = \{y_n\} = G^*(f)$ and $f_{\mathcal{W},\alpha,\omega}$ as in Proposition 3.8 and suppose that $\widehat{f}_{\mathcal{W},\alpha,\omega}$ is the vector obtained if we use $\widehat{y} = \{\widehat{y}_n\}$ instead of the original data $y = \{y_n\}$. Then*

$$\begin{aligned} \|\widehat{f}_{\mathcal{W},\alpha,\omega} - f_{\mathcal{W},\alpha,\omega}\| &\leq \left\| F\Big(F^*(GD_\omega G^* + \alpha P_{\mathcal{W}})F\Big)^\dagger F^*GD_\omega \right\| \|y - \widehat{y}\| \\ &\leq \left(\frac{\|F\|\ \|\omega\|_\infty \|F^*G\|}{\gamma(D_\omega^{1/2}G^*F) + \alpha\ \gamma(F)^2} \right) \|y - \widehat{y}\| \end{aligned}$$

Proof. Note that

$$\begin{aligned}\|\widehat{f}_{\mathcal{W},\alpha,\omega} - f_{\mathcal{W},\alpha,\omega}\| &\leq \left\| \left(F \left(F^*(GD_\omega G^* + \alpha P_{\mathcal{W}})F \right)^\dagger F^* G D_\omega \right)(y - \widehat{y}) \right\| \\ &\leq \|F\| \, \|F^*G\| \, \|\omega\|_\infty \left\| \left(F^*(GD_\omega G^* + \alpha P_{\mathcal{W}})F \right)^\dagger \right\| \|y - \widehat{y}\|.\end{aligned}$$

Therefore, the desired estimation follows by Lemma 3.10. ■

4 Sampling in reproducing kernel Hilbert spaces

Let X be a complete metric space and $\mathcal{K}$ a Hilbert space. Given a function $K : X \to \mathcal{K}$, for every $\eta \in \mathcal{K}$ we define $f_\eta(x) = \langle \eta, \ K(x) \rangle$. Let $\mathcal{H}$ be the space of all the functions obtained in this way. Defining $T : \mathcal{K} \to \mathcal{H}$ by

$$T(\eta) = \langle \eta, \ K(\cdot) \rangle \,,$$

the space $\mathcal{H}$ may be endowed with the norm

$$\|f\|_{\mathcal{H}} = \inf\{\|v\| : \ f = Tv\}.$$

In this way, T becomes an isometry and $\mathcal{H}$, with the inner product associated to the norm $\|\cdot\|_{\mathcal{H}}$, becomes a Hilbert space isomorphic to $N(T)^\perp$. Let $k : X \times X \to \mathbb{C}$ be the kernel defined by

$$k(x_1, x_2) = \langle K(x_2), K(x_1) \rangle_{\mathcal{K}} \,.$$

Then:

- $k_x(\cdot) = k(\cdot, x) \in \mathcal{H}$ for every $x \in X$.
- For every $x \in X$ and every $f \in \mathcal{H}$ the identity $f(x) = \langle f, k_x \rangle_{\mathcal{H}}$ holds.

A Hilbert space of functions defined on a complete metric space with such a kernel is called *reproducing kernel Hilbert space* (RKHS). It is well known that the existence of a reproducing kernel is equivalent to the fact that every point evaluation be a continuous functional ([3] [22]).

In this section we shall translate our results on sampling in abstract Hilbert spaces to RKHS. Following Smale and Zhou's notation, let $\overline{t}$ and $\overline{x}$ be discrete subsets of X and define

$$\mathcal{H}_{k, \ \overline{t}} = \overline{\operatorname{span} \left\langle k_t : \ t \in \overline{t} \right\rangle} \qquad \text{and} \qquad \mathcal{H}_{k, \ \overline{x}} = \overline{\operatorname{span} \left\langle k_x : \ x \in \overline{x} \right\rangle}.$$

We shall assume that $\{k_t\}_{t \in \overline{t}}$ and $\{k_x\}_{x \in \overline{x}}$ are frames for $\mathcal{H}_{k, \ \overline{t}}$ and $\mathcal{H}_{k, \ \overline{x}}$ respectively. If F denotes the synthesis operator of $\{k_t\}_{t \in \overline{t}}$ and G denotes the synthesis operator of $\{k_x\}_{x \in \overline{x}}$ we shall consider the following operators:

$$K_{\overline{t},\overline{t}} = F^*F \qquad K^{\overline{x},\overline{x}} = GG^* \qquad \text{and} \qquad K_{\overline{x},\overline{t}} = K^*_{\overline{t},\overline{x}} = G^*F.$$

Finally, $\{\omega_x\}_{x \in \overline{x}}$ is a sequence of positive numbers bounded from above, and D_ω the corresponding diagonal operator with respect to the canonical basis of $\ell^2(\overline{x})$.
Let us begin with the notion of rich data in this setting:

Definition 4.1. *We say that $\overline{x}$ provides rich data with respect to $\overline{t}$ and ω if*

$$\inf_{z\in N(K_{\overline{t},\overline{t}})^\perp} \|D_\omega^{1/2} K_{\overline{x},\overline{t}}(z)\| > 0.$$

Equivalently, if the operator $D_\omega^{1/2} K_{\overline{x},\overline{t}}$ has closed range and $\mathcal{H}_{k,\ \overline{t}} \cap \mathcal{H}_{k,\ \overline{x}}{}^\perp = \{0\}$.

Now, we are ready to rewrite Propositions 3.8 and 3.11 in this setting:

Proposition 4.2. *Given $f \in \mathcal{H}$ and $y = \{f(x)\}_{x\in\overline{x}}$ (the sampling data), the solution of the minimization problem*

$$f_{\alpha,\omega} = \arg\min_{h\in\mathcal{H}_{k,\ \overline{t}}} \Big(\sum_{x\in\overline{x}} \omega_x |h(x) - f(x)|^2 + \alpha\|h\|^2_{\mathcal{H}}\Big), \tag{8}$$

is given by

$$f_{\alpha,\omega} = \sum_{t\in\overline{t}} L_{\alpha,\omega}(y)(t)\ k_t,$$

where $L_{\alpha,\omega} : \ell^2(\overline{x}) \to \ell^2(\overline{t})$ is the operator defined by

$$L_{\alpha,\omega} = \Big(K_{\overline{t},\overline{x}} D_\omega K_{\overline{x},\overline{t}} + \alpha K_{\overline{t},\overline{t}}\Big)^\dagger K_{\overline{t},\overline{x}} D_\omega.$$

In particular, if we assume that $\overline{x}$ provides rich data with respect to $\overline{t}$ and ω, we get the following estimation of the reconstruction error:

$$\|f_{\alpha,\omega} - f\|_{\mathcal{H}} \le \Big\|(1 - P_{D,\mathcal{H}_{k,\ \overline{t}}})(f)\Big\|_{\mathcal{H}} + \left(\alpha \frac{\|K_{\overline{t},\overline{t}}\|}{\gamma(D_\omega^{1/2} K_{\overline{x},\overline{t}}) + \alpha\ \gamma(K_{\overline{t},\overline{t}})}\right) \|f\|_{\mathcal{H}} \tag{9}$$

where $D = (K^{\overline{x},\overline{x}} + \alpha P_{\mathcal{H}_{k,\ \overline{t}}})$.

Proposition 4.3. *Let $y = \{f(x)\}_{x\in\overline{x}}$ and $f_{\alpha,\omega}$ as in Proposition 4.2. Suppose that $\widehat{f}_{\alpha,\omega}$ is the vector obtained if we use $\widehat{y} = \{\widehat{y}_x\}_{x\in\overline{x}}$ instead of the original data $y = \{f(x)\}_{x\in\overline{x}}$. Then:*

$$\|\widehat{f}_{\alpha,\omega} - f_{\alpha,\omega}\|_{\mathcal{H}} \le \left(\frac{\|K_{\overline{t},\overline{t}}\|^{1/2}\|\omega\|_\infty \|K_{\overline{t},\overline{x}}\|}{\gamma(D_\omega^{1/2} K_{\overline{x},\overline{t}}) + \alpha\ \gamma(K_{\overline{t},\overline{t}})}\right) \Big\|\ y - \widehat{y}\ \Big\|_{\ell^2(\overline{x})}.$$

Concluding remarks. As we have already mentioned in the Introduction, one of the motivations of this work is a question posed in [23] by Smale and Zhou. In that paper, they ask for an error estimation if the sampled vector does not belong to the subspace $\mathcal{H}_{k,\ \overline{t}}$. Inequality (9) is a posible answer. Moreover, if $f \in \mathcal{H}_{k,\ \overline{t}}$, estimation (9) slightly improves their inequality:

$$\|f_{\alpha,\omega} - f\|_{\mathcal{H}} \le \left(\alpha \frac{\|K_{\overline{t},\overline{t}}\|}{\gamma(D_\omega^{1/2} K_{\overline{x},\overline{t}})}\right) \|f\|_{\mathcal{H}}$$

Note that in the general setting of the previous section, Proposition 3.8 answers an equivalent question.

Observe also that in [23] the sequences $\{k_t\}_{t\in\overline{t}}$ and $\{k_x\}_{x\in\overline{x}}$ are supposed to be Riesz bases; however, in the sequel [24] Smale and Zhou weaken the hypothesis and the sequences are supposed to be frames for the entire space; in the present approach both sequences only need to be frames for the subspaces $\mathcal{H}_{k,\ \overline{t}}$ and $\mathcal{H}_{k,\ \overline{x}}$ respectively. ▲

References

[1] J. A. Antezana, G. Corach, M. Ruiz and D. Stojanoff, Weighted projections and Riesz frames, Linear Algebra Appl. 402 (2005), 367-389.

[2] J. A. Antezana, G. Corach, M. Ruiz and D. Stojanoff, Nullspaces and frames, J. Math. Anal. Appl. 309 (2005), 709-723.

[3] N. Aronszajn, Theory of reproducing kernels, Trans. Amer. Math. Soc. 68 (1950), 337-404.

[4] A. Ben-Israel and T. N. E. Greville, Generalized inverses. Theory and applications. Second edition. CMS Books in Mathematics/Ouvrages de Mathmatiques de la SMC, 15. Springer-Verlag, New York, 2003

[5] R. Bouldin; The product of operators with closed range, Tohoku Math. J. 25 (1973), 359-363.

[6] O. Christensen, An introduction to frames and Riesz bases, Birkhäuser, Boston, 2003.

[7] O. Christensen, Y. Eldar, Oblique dual frames and shift-invariant spaces, Appl. Comput. Harmon. Anal. 17 (2004) 4868.

[8] G. Corach, A. Maestripieri and D. Stojanoff , Oblique projections and Schur complements, Acta Sci. Math. (Szeged), 67 (2001) 337-356.

[9] G. Corach, A. Maestripieri and D. Stojanoff, Generalized Schur complements and oblique projections, Linear Algebra and its Applications 341 (2002), 259-272.

[10] G. Corach, A. Maestripieri and D. Stojanoff, Oblique projections and abstract splines, J. Approx. Theory 117 (2002), 189–206.

[11] F. Deutsch, The angle between subspaces in Hilbert space, in "Approximation theory, wavelets and applications" (S. P. Singh, editor), Kluwer, Netherlands, 1995, 107-130.

[12] J. Dixmier, Etudes sur les variétés et opérateurs de Julia, avec quelques applications, Bull. Soc. Math. France 77 (1949), 11-101.

[13] R. J. Duffin and A. C. Schaeffer; A class of nonharmonic Fourier series. Trans. Amer. Math. Soc. 72, (1952). 341-366.

[14] Y. C. Eldar, Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors, J. Fourier Anal. Appl. 9 (2003), 77-96.

[15] Y. C. Eldar and T. Werther, General Framework for Consistent Sampling in Hilbert Spaces, International Journal of Wavelets, Multiresolution, and Information Processing, 3 (2005), 347-359.

[16] K. Friedrichs, On certain inequalities and characteristic value problems for analytic functions and for functions of two variables. I. Trans. Amer. Math. Soc. 41 (1937), 321-364 .

[17] S. Hassi, K. Nordström; On projections in a space with an indefinite metric, Linear Algebra Appl. 208/209 (1994), 401-417.

[18] C. E. Heil and D.F. Walnut, Continuous and discrete wavelet transforms, SIAM Rev. 31 (1989), 628-666.

[19] D. A. Herrero, Approximation of Hilbert space operators. Vol. 1. Second edition. Pitman Research Notes in Mathematics Series, 224, New York, 1989.

[20] S. Izumino, The product of operators with closed range and an extension of the reverse order law, Tohoku Math. J. 34 (1982), 43-52.

[21] S. Li and H. Ogawa, Pseudoframes for subspaces with applications, J. Fourier Anal. Appl. 10 (2004), 409-431.

[22] S. Saitoh, Theory of reproducing kernels and its applications, London, Longman (1988).

[23] S. Smale and D. X. Zhou, Shannon sampling and function reconstruction from point values, Bull. Amer. Math. Soc. 41 (2004), 279-305.

[24] S. Smale and D. X. Zhou, Shannon sampling II: Connections to learning theory, Appl. Comput. Harmon. Anal. (2005).

[25] R. M. Young, An introduction to nonharmonic Fourier series (revised first edition) Academic Press, San Diego, 2001.