A Minimum Squared-Error Framework for Generalized Sampling

Yonina C. Eldar and Tsvi G. Dvorkind*

Abstract

We treat the problem of reconstructing a signal from its non-ideal samples where the sampling and reconstruction spaces as well as the class of input signals can be arbitrary subspaces of a Hilbert space. Our formulation is general, and includes as special cases reconstruction from finitely many samples as well as uniform-sampling of continuous-time signals, which are not necessarily bandlimited. To obtain a good approximation of the signal in the reconstruction space from its samples, we suggest two design strategies that attempt to minimize the squared-norm error between the signal and its reconstruction. The approaches we propose differ in their assumptions on the input signal: If the signal is known to lie in an appropriately chosen subspace, then we propose a method that achieves the minimal squared-error. On the other hand, when the signal is not restricted, we show that the minimal-norm reconstruction cannot generally be obtained. Instead, we suggest minimizing the worst-case squared-error between the reconstructed signal, and the best possible (but usually unattainable) approximation of the signal within the reconstruction space. We demonstrate both theoretically and through simulations that the suggested methods can outperform the consistent reconstruction approach previously proposed for this problem.

1   Introduction

Digital signal processing entails representing a signal by a set of coefficients and relies on the existence of methods for reconstructing the signal from its samples. The most common setting considered in the sampling literature is that introduced by the Shannon-Whittaker sampling theorem, in which the input signal is assumed to be bandlimited and the samples of the signal are ideal, *i.e.,* they are equal to the signal values at a set of sampling points. The reconstructed signal is also a bandlimited function, generated by integer shifts of the sinc interpolation kernel. In practice, however, the input signal is never perfectly bandlimited, and the sampling process may not be ideal. Another drawback of the Shannon paradigm is the difficulty in implementing the infinite sinc interpolating kernel, which has slow decay.

To overcome these limitations of the traditional sampling framework, a more recent approach is to consider a generalized sampling scheme, in which the samples are obtained by first linearly pre-processing

the signal. The non-ideal samples can then be represented as the inner products of the input signal x with a set of sampling vectors (associated with the acquisition device), which span the sampling space $\mathcal{S}$ [1, 2, 3, 4, 5, 6, 7]. Examples include multiresolution decompositions [2, 8], and spline decompositions [3]. Reconstruction is obtained by taking linear combinations of a set of reconstruction vectors that span a space $\mathcal{W}$. Since in this framework the reconstructed signal is constrained to lie in $\mathcal{W}$, if x is not in $\mathcal{W}$ to begin with, then perfect reconstruction cannot be obtained, regardless of the sampling and reconstruction method. A natural question that arises from this formulation of the sampling problem is whether the samples can be processed prior to reconstruction such that the reconstructed signal $\hat{x}$ is close to x in some sense.

In this paper, we design reconstruction strategies for the generalized sampling scheme, where we treat the problem of reconstruction from finitely-many samples and from uniform samples of a prefiltered continuous-time signal in a unified way. The only constraints we impose on the problem are that the sampling process is linear and bounded, and the reconstruction is constrained to a subspace $\mathcal{W}$ of an arbitrary Hilbert space $\mathcal{H}$. However, we do not require any specific constraints on the spaces involved.

To ensure that the reconstruction $\hat{x}$ is close to x we may try to minimize the squared-norm of the reconstruction error $\hat{x} - x$. If the reconstruction space $\mathcal{W}$ is contained in the sampling space $\mathcal{S}$, then by proper pre-processing of the samples the minimal squared-error approximation of x in the space $\mathcal{W}$, given by the orthogonal projection $P_{\mathcal{W}}x$, can be obtained. However, as we show in Section 3, if $\mathcal{S}$ does not contain the subspace $\mathcal{W}$, then the squared-error cannot be minimized over the entire space $\mathcal{H}$ of input signals.

The sampling framework we consider here was first treated in [1] for the case in which the sampling and reconstruction spaces are shift-invariant (SI) subspaces of L_2, *i.e.,* spaces generated by translates of an appropriately chosen function. The reconstruction was obtained by first processing the samples by a digital correction filter, designed such that $\hat{x}$ is a *consistent reconstruction* of x, namely a reconstruction with the property that it yields the same samples as x. Fast iterative methods leading to consistent reconstruction were developed in [9]. The consistent approach was then generalized in [5, 7, 10] to a more general class of sampling and reconstruction spaces, as well as arbitrary input Hilbert spaces $\mathcal{H}$. Under a direct-sum condition on the spaces, the consistent reconstruction is given by $\hat{x} = E_{\mathcal{W}\mathcal{S}^\perp}x$ where $E_{\mathcal{W}\mathcal{S}^\perp}$ is the oblique projection onto $\mathcal{W}$ along the orthogonal compliment of $\mathcal{S}$. Note, however, that the fact that x and $\hat{x}$ yield the same samples does not necessarily imply that $\hat{x}$ is close to x. In fact, for an input x not in $\mathcal{W}$, the norm of the resulting reconstruction error $\hat{x} - x$ can be made arbitrarily large, if $\mathcal{S}$ is close to $\mathcal{W}^\perp$.

To obtain a good reconstruction for cases in which the consistent method leads to large errors, we suggest two alternative strategies that differ in their assumptions on the signal x. We first treat the case in which x is known to lie in a subspace of $\mathcal{H}$, and show that if the subspace is chosen appropriately, then the squared-error can be minimized over all signals in that space leading to a reconstruction that is closer to x than the consistent method. When the input signal can be any vector in $\mathcal{H}$, and the minimal error approximation cannot be achieved, we suggest minimizing a worst-case error measure over all bounded norm signals. We first consider minimizing the worst-case squared error. This approach turns out to be

overconservative resulting in the trivial solution $\hat{x} = 0$. To counterbalance the conservative behavior of this minimax strategy, we develop a competitive approach, similar in spirit to the methods of [11, 12], in which $\hat{x}$ is designed to minimize the worst-case *regret* instead of the worst-case squared-error; The regret is defined as the difference between the squared-error of $\hat{x}$ and the minimal attainable error in the ideal case when $\mathcal{W} \subseteq \mathcal{S}$. The minimax regret solution turns out to be linear, and is given by the double orthogonal projection $\hat{x} = P_{\mathcal{W}} P_{\mathcal{S}} x$ onto $\mathcal{S}$ and $\mathcal{W}$. In contrast with the consistent approach which can result in an arbitrarily large reconstruction error, the regret strategy has the desirable property that the squared-error is bounded by twice the norm of x. In the case of SI subspaces of L_2, our methods can be implemented using linear time-invariant (LTI) discrete-time filters. A particularly efficient implementation of these filters is possible in spline spaces, based on the results of [3, 13, 14, 15].

We present a detailed comparison of the regret and consistent reconstruction methods, by analyzing the error resulting from both strategies. In particular, we show that if the spaces $\mathcal{S}$ and $\mathcal{W}$ are sufficiently far apart, or if x has enough energy in $\mathcal{S}$, then the minimax regret reconstruction is preferable to the consistent approach. Our theoretical results are also demonstrated through simulations in Section 8.

The paper is organized as follows. The general sampling framework we treat in the paper is introduced in Section 2 together with some mathematical preliminaries. Section 3 shows that the minimal squared-error reconstruction cannot be obtained in general. In Section 4 we eliminate the dependency on the signal x by minimizing the error over a subspace of $\mathcal{H}$. Two minimax approaches to sampling and reconstruction are introduced in Section 5: minimax squared-error and minimax regret. We first treat the problem of linear reconstruction, and then suggest a nonlinear design criterion whose optimal solution turns out to be linear. The special case of sampling and reconstruction in SI subspaces is discussed in Section 6. In Section 7 we analyze the reconstruction error resulting from the minimax regret approach, and compare it with the error from the consistent strategy. Simulation results are presented in Section 8.

2 Problem Formulation and Preliminaries

2.1 Sampling Formulation

We denote vectors in an arbitrary Hilbert space $\mathcal{H}$ by lowercase letters, and the elements of a sequence $c \in \ell_2$ by $c[n]$. The orthogonal projection operator onto a closed subspace $\mathcal{A}$ of $\mathcal{H}$ is denoted by $P_{\mathcal{A}}$, and the orthogonal complement of $\mathcal{A}$ is denoted by $\mathcal{A}^{\perp}$. The *Moore-Penrose pseudo inverse* [16] of a bounded transformation T is denoted by $T^{\dagger}$, T^* is the adjoint of T, and $\mathcal{N}(T)$ and $\mathcal{R}(T)$ are the null space and range space, respectively. The inner product between vectors $x, y \in \mathcal{H}$ is denoted as $\langle x, y \rangle$, and is linear in the second argument; $\|x\|^2 = \langle x, x \rangle$ is the squared norm of x. The direct sum between two closed subspaces $\mathcal{W}$ and $\mathcal{S}$ is written as $\mathcal{W} \oplus \mathcal{S}$, and is the sum set $\{w + v;\ w \in \mathcal{W}, v \in \mathcal{S}\}$ with the property $\mathcal{W} \cap \mathcal{S} = \{0\}$.

The oblique projection[1] [17] onto $\mathcal{W}$ along $\mathcal{S}^\perp$ is denoted by $E_{\mathcal{W}\mathcal{S}^\perp}$, and is defined as the unique operator satisfying

$$
\begin{aligned}
E_{\mathcal{W}\mathcal{S}^\perp} w &= w \text{ for any } w \in \mathcal{W}; \\
E_{\mathcal{W}\mathcal{S}^\perp} v &= 0 \text{ for any } v \in \mathcal{S}^\perp.
\end{aligned}
\tag{1}
$$

A set transformation $V : \ell_2 \to \mathcal{H}$ corresponding to frame vectors[2] $\{v_n\} \in \mathcal{H}$ is defined by $Va = \sum_n a[n]v_n$ for all $a \in \ell_2$. From the definition of the adjoint, if $a = V^*y$, then $a[n] = \langle v_n, y \rangle$.

We consider a general sampling problem in a Hilbert space $\mathcal{H}$, in which the goal is to reconstruct a signal $x \in \mathcal{H}$ from a sequence of samples $\{c[n]\}$. Our formulation of the problem allows for a broad class of sampling strategies where the basic constraint we impose on the sampling process is that it is bounded and linear. From the Riesz representation theorem, the samples can be modelled as the inner products of the signal x with a set of sampling vectors $\{s_n\}$, so that $c[n] = \langle s_n, x \rangle$, and the problem is to reconstruct the signal x from its given samples $c[n]$.

In principle, x can be defined in a space $\mathcal{H}$ that is larger than the sampling space $\mathcal{S}$, spanned by the vectors $\{s_n\}$. Therefore, our problem is inherently ill-posed. To resolve this issue we constrain the reconstruction to a closed subspace $\mathcal{W}$ of $\mathcal{H}$. Choosing a set of vectors $\{w_n\}$ that span $\mathcal{W}$, the reconstruction $\hat{x}$ of x has the form

$$
\hat{x} = \sum_n d[n]w_n
\tag{2}
$$

for some coefficients $d[n]$ that are a transformation of the samples $c[n]$. Denoting by d and c the sequences in ℓ_2 with elements $d[n]$ and $c[n]$ respectively, we have that $d = H(c)$ for some transformation $H : \ell_2 \to \ell_2$, which can be nonlinear. Using set transformations we can express the sequence of samples as $c = S^*x$, and the reconstruction as

$$
\hat{x} = Wd = WH(S^*x),
\tag{3}
$$

where $S : \ell_2 \to \mathcal{H}$ and $W : \ell_2 \to \mathcal{H}$ are the set transformations corresponding to the vectors $\{s_n\}$ and $\{w_n\}$, respectively. The sampling and reconstruction scheme is illustrated in Fig. 1.
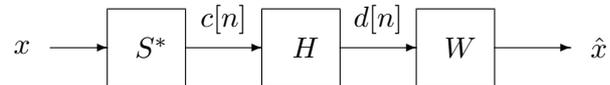
Figure 1: General sampling and reconstruction scheme.

A special case of Fig. 1, which we consider in detail in Section 6, is when $\{s_n = s(t - n)\}$ and $\{w_n =$

---

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

[2]Frames are defined in Section 2.2; for frame vectors the sum $\sum_n a[n]v_n$ is guaranteed to converge.

$w(t-n)\}$ are vectors corresponding to uniform shifts of real generators in L_2. In this setting, the sampling and reconstruction scheme of Fig. 1 can be formulated in terms of linear-time invariant (LTI) filters, as depicted in Fig. 2.
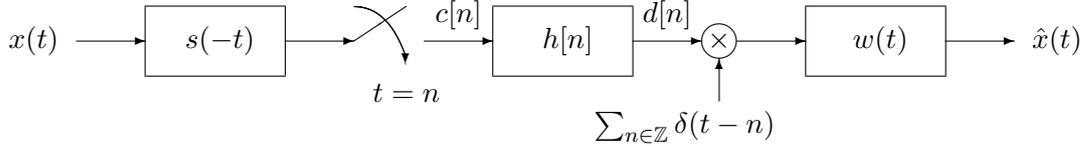
Figure 2: Sampling and reconstruction in shift-invariant spaces.

If x is in $\mathcal{W}$, and $\mathcal{W}$ and $\mathcal{S}$ satisfy the direct-sum condition

$$\mathcal{H} = \mathcal{W} \oplus \mathcal{S}^\perp, \tag{4}$$

then it was shown in [5, 7, 10] that x can be perfectly reconstructed from the samples $c[n]$ by choosing $H(c) = (S^*W)^\dagger c$. With this choice of H it follows from (7) below that the reconstruction $\hat{x}$ is equal to

$$\hat{x} = E_{\mathcal{W}\mathcal{S}^\perp} x. \tag{5}$$

If $x \notin \mathcal{W}$, then it cannot be perfectly reconstructed using only vectors in $\mathcal{W}$ since $\hat{x}$ given by (2) is always an element of $\mathcal{W}$. Therefore, in this case the reconstruction of (5) no longer equals x. Nonetheless it has the property that it is a *consistent reconstruction* [1], namely, it yields the same samples as x: $S^*x = S^*\hat{x}$. However, the fact that x and $\hat{x}$ of (5) have the same samples, does not guarantee that $\hat{x}$ is close to x. In fact, using the relation $x = E_{\mathcal{W}\mathcal{S}^\perp} x + E_{\mathcal{S}^\perp\mathcal{W}} x$ we can express the reconstruction error as $\hat{x} - x = E_{\mathcal{S}^\perp\mathcal{W}} x$, which can have arbitrarily large norm if $\mathcal{S}^\perp$ is close to $\mathcal{W}$. Therefore, our problem is to choose the transformation H in Fig. 1 such that $\hat{x}$ is a good approximation of x.

In the following sections we propose different strategies for designing H which attempt to control the squared-norm of the reconstruction error $\hat{x} - x$, and evaluate their performance. In particular we show that in many cases we can choose H such that $\hat{x}$ is a better approximation to x than the consistent reconstruction method. The solutions to all the criteria we define turn out to be linear, leading to a nice filtering interpretation in the SI case. Before proceeding to our detailed developments, we next summarize the mathematical background and hypotheses.

2.2 Mathematical Preliminaries

In order to make the sampling problem of Fig. 1 well posed, we need several mathematical hypotheses. First, we would like to ensure that the sampling is stable so that the sequence of samples obtained by $c[n] = \langle s_n, x \rangle$ has finite energy for any finite-energy $x \in \mathcal{H}$. Additionally, for the reconstruction to be well defined, the

sum $\sum_n d[n]w_n$ must converge. Both properties can be satisfied by choosing the vectors $\{s_n\}$ and $\{w_n\}$ such that they form frames for their closed span, which we denote by $\mathcal{S}$ and $\mathcal{W}$ respectively.

Definition 1 *([18]).* A family of vectors $\{h_n\}$ in a Hilbert space $\mathcal{H}$ is called a frame for a subspace $\mathcal{A} \subseteq \mathcal{H}$ if there exist constants $0 < A \leq B < \infty$ such that

$$A\|y\|^2 \leq \sum_n |\langle y, h_n \rangle|^2 \leq B\|y\|^2, \tag{6}$$

for all $y \in \mathcal{A}$.

The lower bound in (6) ensures that the vectors $\{h_n\}$ span $\mathcal{A}$; thus the number of frame elements, which we denote by N, must be at least as large as the dimension of $\mathcal{A}$. If $N < \infty$, then the right hand inequality of (6) is always satisfied with $B = \sum_n \langle h_n, h_n \rangle$. Thus, any finite set of vectors that spans $\mathcal{A}$ is a frame for $\mathcal{A}$.

If the sampling vectors $\{s_n\}$ form a frame for $\mathcal{S}$, then it follows immediately from the upper bound (6) that the sequence of samples $c[n] = \langle s_n, x \rangle$ is in ℓ_2 for any signal x that has bounded norm, and therefore the sampling process is stable. Similarly, if the vectors $\{w_n\}$ form a frame, then the sum $\sum_n d[n]w_n$ converges for any sequence $d \in \ell_2$ [19].

Set transformations corresponding to frame sequences have some nice properties, which we will exploit in our derivations. In particular, if $\{s_n\}$ is a frame sequence for $\mathcal{S}$ with set transformation S, then S is bounded and $\mathcal{R}(S) = \mathcal{S}$. This implies that S and W in Fig. 1 are both bounded. The overall sampling and reconstruction scheme is then guaranteed to be stable if we choose H as a bounded transformation.

Another useful result on set transformations is given in the following lemma.

Lemma 1. *[20, Lemma 3.3] Let $S : \ell_2 \rightarrow \mathcal{H}$ and $W : \ell_2 \rightarrow \mathcal{H}$ be bounded transformations on $\mathcal{H}$ with $\mathcal{R}(S) = \mathcal{S}$ and $\mathcal{R}(W) = \mathcal{W}$, where $\mathcal{H} = \mathcal{W} \oplus \mathcal{S}^\perp$. Then*

1. $\mathcal{N}(S^*W) = \mathcal{N}(W)$;

2. $(S^*W)^\dagger$ is a bounded operator from ℓ_2 to ℓ_2;

3. $(S^*W)^\dagger S^*W$ is equal to $P_{\mathcal{N}(W)^\perp}$.

Using part 2 of the lemma, we can obtain an explicit construction of the oblique projection $E_{\mathcal{W}\mathcal{S}^\perp}$ [10],[20]:

$$E_{\mathcal{W}\mathcal{S}^\perp} = W\,(S^*W)^\dagger\,S^* \tag{7}$$

where S and W are bounded operators with $\mathcal{R}(S) = \mathcal{S}$ and $\mathcal{R}(W) = \mathcal{W}$. As a special case, the orthogonal projection $P_{\mathcal{W}}$ can be written as

$$P_{\mathcal{W}} = W\,(W^*W)^\dagger\,W^*. \tag{8}$$

In some cases it is useful that $\mathcal{W}$ and $\mathcal{S}$ have the same dimension. The concept of dimension is well-defined in finite spaces. In the infinite-dimensional case, this condition can be made precise by requiring that there exits a bijective (injective and surjective) transformation $B\colon \mathcal{W} \to \mathcal{S}$ from $\mathcal{W}$ to $\mathcal{S}$, or equivalently, that $\mathcal{W}$ and $\mathcal{S}$ are isomorphic. One way to guarantee that such an isomorphism exists is to impose the direct-sum condition

$$\mathcal{H} = \mathcal{W} \oplus \mathcal{S}^{\perp}, \tag{9}$$

as incorporated in the following proposition.

Proposition 1. *Let $\mathcal{S}$ and $\mathcal{W}$ be closed subspaces of a Hilbert space $\mathcal{H}$ with $\mathcal{H} = \mathcal{W} \oplus \mathcal{S}^{\perp}$. Then $P_{\mathcal{S}}\colon \mathcal{W} \to \mathcal{S}$ is bijective.*

Proof. We begin by showing that $P_{\mathcal{S}}$ is injective. If $P_{\mathcal{S}}w = 0$ for some $w \in \mathcal{W}$, then $w \in \mathcal{S}^{\perp}$. But since $\mathcal{W} \cap \mathcal{S}^{\perp} = \{0\}$ we conclude that $w = 0$ and $P_{\mathcal{S}}$ is injective. To show that $P_{\mathcal{S}}$ is surjective, let $s \in \mathcal{S}$ be arbitrary. From (9) we can write $s = w + v$ where $w \in \mathcal{W}$ and $v \in \mathcal{S}^{\perp}$. Since $s \in \mathcal{S}$,

$$s = P_{\mathcal{S}}s = P_{\mathcal{S}}w + P_{\mathcal{S}}v = P_{\mathcal{S}}w, \tag{10}$$

and $P_{\mathcal{S}}$ is surjective. $\qquad\square$

In our analysis of the reconstruction error in Section 7 we will use the concept of an angle between two closed subspaces $\mathcal{A}_1$ and $\mathcal{A}_2$ of a Hilbert space $\mathcal{H}$ [1, 21]:

$$
\begin{aligned}
\cos\left(\mathcal{A}_1, \mathcal{A}_2\right) &= \inf_{f \in \mathcal{A}_1, \|f\|=1} \left\| P_{\mathcal{A}_2} f \right\| \\
\sin\left(\mathcal{A}_1, \mathcal{A}_2\right) &= \sup_{f \in \mathcal{A}_1, \|f\|=1} \left\| P_{\mathcal{A}_2^{\perp}} f \right\|,
\end{aligned}
\tag{11}
$$

and the relations [1]

$$
\begin{aligned}
\sin\left(\mathcal{A}_1, \mathcal{A}_2\right) &= \sin\left(\mathcal{A}_2^{\perp}, \mathcal{A}_1^{\perp}\right) \\
\cos\left(\mathcal{A}_1, \mathcal{A}_2\right) &= \cos\left(\mathcal{A}_2^{\perp}, \mathcal{A}_1^{\perp}\right) \\
\sin^2\left(\mathcal{A}_1, \mathcal{A}_2\right) &= 1 - \cos^2\left(\mathcal{A}_1, \mathcal{A}_2\right).
\end{aligned}
\tag{12}
$$

3 Minimal Squared-Error Reconstruction

A straightforward strategy to designing a reconstruction that is close to x is to minimize the squared-error $\|\hat{x} - x\|^2$. In this approach, the transformation H is the solution to the problem

$$\min_{H} \|\hat{x} - x\|^2 = \min_{H} \|WH(S^*x) - x\|^2. \tag{13}$$

Since $\hat{x} = WH(S^*x)$, for any choice of x

$$\|\hat{x} - x\|^2 = \|WH(S^*x) - P_{\mathcal{W}}x\|^2 + \|P_{\mathcal{W}^\perp}x\|^2 \geq \|P_{\mathcal{W}^\perp}x\|^2. \tag{14}$$

In the special case in which $\mathcal{W} \subseteq \mathcal{S}$, the bound (14) can be achieved with

$$H(c) = (W^*W)^\dagger W^* S(S^*S)^\dagger c. \tag{15}$$

Indeed, with this choice of $H(c)$,

$$\hat{x} = W(W^*W)^\dagger W^* S(S^*S)^\dagger S^* x = P_{\mathcal{W}} P_{\mathcal{S}} x = P_{\mathcal{W}} x, \tag{16}$$

where we used the representation (8) of $P_{\mathcal{W}}$ and $P_{\mathcal{S}}$, and the last equality follows from the fact that $\mathcal{W} \subseteq \mathcal{S}$. However, as we now show, when $\mathcal{W}$ is not contained in $\mathcal{S}$, the lower bound cannot be achieved for all x with a transformation $H(c)$ that depends only on the given samples $c = S^*x$.

Proposition 2. *Let $H : \ell_2 \to \ell_2$ be any solution to*

$$\min_H \|\hat{x} - x\|^2 = \min_H \|WH(S^*x) - x\|^2,$$

where W and S are bounded transformations with $\mathcal{R}(W) = \mathcal{W}$, $\mathcal{R}(S) = \mathcal{S}$, and $\mathcal{W} \nsubseteq \mathcal{S}$. Then for arbitrary choices of x, $H(S^*x)$ cannot achieve the lower bound of (14).

Proof. To prove the proposition, suppose to the contrary that there exits a solution $H(c)$ achieving the lower bound that depends only on the available samples $c = S^*x$. Consider the signal x defined by $x = x_{\mathcal{S}^\perp} + x_{\mathcal{W}}$ where $x_{\mathcal{S}^\perp}$ is in $\mathcal{S}^\perp$ but not in $\mathcal{W}^\perp$ (such a vector always exists since $\mathcal{W} \nsubseteq \mathcal{S}$) and $x_{\mathcal{W}} \in \mathcal{W}$. For this choice, $S^*x = S^*x_{\mathcal{W}} = c$ so that

$$WH(S^*x) = WH(S^*x_{\mathcal{W}}). \tag{17}$$

On the other hand, since H achieves the lower bound in (14), we must have $WH(S^*x) = P_{\mathcal{W}}x$ and $WH(S^*x_{\mathcal{W}}) = P_{\mathcal{W}}x_{\mathcal{W}} = x_{\mathcal{W}}$ which implies that $P_{\mathcal{W}}x_{\mathcal{S}^\perp} = 0$, or $x_{\mathcal{S}^\perp} \in \mathcal{W}^\perp$, contradicting our assumption. $\square$

To circumvent the problem associated with minimizing the squared-error we develop two strategies which differ in their assumptions on x. In the first approach we take advantage of prior information on x in the form of inclusion into a properly chosen subspace. This knowledge will allow us to directly minimize the squared-error, as we show in Section 4. The second strategy treats the squared-error criterion over the entire space. In this case, we eliminate the dependency of the squared-error on x by considering a worst-case error measure.

4 Minimal Squared-Error Reconstruction on a Subspace

We have seen in Proposition 2 that if $\mathcal{W} \nsubseteq \mathcal{S}$, then the lower bound in (14) cannot be achieved *for all $x \in \mathcal{H}$*. However, this does not preclude the possibility of achieving the bound for a subset of input signals. Indeed, if we consider only signals $x \in \mathcal{W}$, then from (5) it follows that under the direct-sum condition $\mathcal{H} = \mathcal{W} \oplus \mathcal{S}^\perp$, perfect reconstruction is possible with $H(c) = (S^*W)^\dagger c$. We now generalize this result to a broader class of input signals.

Suppose that x is known to lie in a subspace $\mathcal{A}$ such that

$$\mathcal{H} = \mathcal{A} \oplus \mathcal{S}^\perp. \tag{18}$$

From the mathematical discussion of Section 2.2 it follows that there is a linear bijection between $\mathcal{S}$ and $\mathcal{A}$. As we now show, this implies that the minimal error reconstruction can be achieved by proper processing of the samples.

Theorem 1. *Consider the problem*

$$\min_H \|\hat{x} - x\|^2 = \min_H \|WH(S^*x) - x\|^2, \quad x \in \mathcal{A}$$

where $\mathcal{A} \subseteq \mathcal{H}$ is a closed subspace such that $\mathcal{H} = \mathcal{A} \oplus \mathcal{S}^\perp$ and W, S are bounded operators with $\mathcal{R}(W) = \mathcal{W}, \mathcal{R}(S) = \mathcal{S}$. A possible solution is $H(c) = H_\mathcal{A} c$ where

$$H_\mathcal{A} = (W^*W)^\dagger W^* A (S^*A)^\dagger, \tag{19}$$

and A is any bounded operator with $\mathcal{R}(A) = \mathcal{A}$. With this choice, $\hat{x}$ is the minimal-norm solution $\hat{x} = P_\mathcal{W} x$.

Before proving the theorem, we note that from Lemma 1 the pseudoinverse $(S^*A)^\dagger$ is a well defined bounded operator. Furthermore, it is shown in [20, Lemma 3.7] that the operator $A(S^*A)^\dagger$ is independent of the choice of A.

Proof. We begin by noting that since $x \in \mathcal{A}$, it can be expressed as $x = Ay$ for some $y \in \mathcal{N}(A)^\perp$. In addition,

$$c = S^*x = S^*Ay. \tag{20}$$

Multiplying both sides of (20) by $(S^*A)^\dagger$ and using Lemma 1, we have that

$$(S^*A)^\dagger c = (S^*A)^\dagger (S^*A)y = P_{\mathcal{N}(A)^\perp} y = y. \tag{21}$$

We conclude that the only vector in $\mathcal{A}$ with samples given by c is the vector

$$x = Ay = A(S^*A)^\dagger c, \tag{22}$$

so that given c we can reconstruct the vector x exactly. Once we know x, the approximation in $\mathcal{W}$ minimizing the squared-error is

$$\hat{x} = P_\mathcal{W} x = P_\mathcal{W} A(S^*A)^\dagger c = W H_\mathcal{A} c, \tag{23}$$

where we used the representation (8) of $P_\mathcal{W}$. Finally, since $c = S^*x$,

$$\hat{x} = P_\mathcal{W} A \left(S^*A\right)^\dagger S^* x = P_\mathcal{W} E_{\mathcal{A}\mathcal{S}^\perp} x = P_\mathcal{W} x, \tag{24}$$

where we used the fact that from (7), $A\left(S^*A\right)^\dagger S^* = E_{\mathcal{A}\mathcal{S}^\perp}$, and since $x \in \mathcal{A}$, $E_{\mathcal{A}\mathcal{S}^\perp} x = x$. $\qquad\square$

We now consider two special choices of $\mathcal{A}$. First, assume that $\mathcal{A} = \mathcal{S}$. In this case, the condition $\mathcal{H} = \mathcal{A} \oplus \mathcal{S}^\perp$ is always satisfied, and from Theorem 1,

$$H_\mathcal{A} = (W^*W)^\dagger W^* S(S^*S)^\dagger. \tag{25}$$

In Theorem 2 we will see that this solution is equivalent to the minimax regret transformation, developed in Sections 5.1.2 and 5.2. This implies that the regret approach minimizes the squared-error over all $x \in \mathcal{S}$.

As another example, suppose that $\mathcal{H} = \mathcal{W} \oplus \mathcal{S}^\perp$, and let $\mathcal{A} = \mathcal{W}$. With this choice,

$$H_\mathcal{A} = P_{\mathcal{N}(\mathcal{W})^\perp} (S^*W)^\dagger = (S^*W)^\dagger, \tag{26}$$

where we used the fact that $\mathcal{R}\left((S^*W)^\dagger\right) = \mathcal{N}(S^*W)^\perp = \mathcal{N}(W)^\perp$; the last equality follows from Lemma 1. Thus, $H_\mathcal{A}$ is equal to the consistent reconstruction transformation, which agrees with the fact that the consistent strategy minimizes the squared-norm error over all values of $x \in \mathcal{W}$.

4.1 Geometric Interpretation

We have seen that if we know that x is in $\mathcal{A}$, where $\mathcal{A} \oplus \mathcal{S}^\perp = \mathcal{H}$, then we can always obtain the minimal norm reconstruction $\hat{x} = P_\mathcal{W} x$ given the samples $c[n]$. We now provide a geometrical interpretation of this result.

We first note that sampling x with sampling vectors in $\mathcal{S}$, is equivalent to sampling its orthogonal projection onto $\mathcal{S}$, denoted by $x_\mathcal{S} = P_\mathcal{S} x$. This follows from the relation

$$\langle s_n, x \rangle = \langle P_\mathcal{S} s_n, x \rangle = \langle s_n, P_\mathcal{S} x \rangle. \tag{27}$$

Since $x_{\mathcal{S}} \in \mathcal{S}$ and the vectors $\{s_n\}$ span $\mathcal{S}$, $x_{\mathcal{S}}$ is uniquely determined by the samples $c[n]$. Therefore, knowing $c[n]$ is equivalent to knowing $x_{\mathcal{S}}$. The reconstruction problem then becomes that of reconstructing a signal in $\mathcal{A}$ from its orthogonal projection $x_{\mathcal{S}}$ onto a subspace $\mathcal{S}$. In Fig. 3 we illustrate the fact that there is only one vector in $\mathcal{A}$ whose orthogonal projection onto $\mathcal{S}$ is $x_{\mathcal{S}}$. In our setup we are constrained to obtain a reconstruction in $\mathcal{W}$. But, since we can determine x from $x_{\mathcal{S}}$, we can also determine $P_{\mathcal{W}}x$, which is the minimal-norm reconstruction in $\mathcal{W}$.
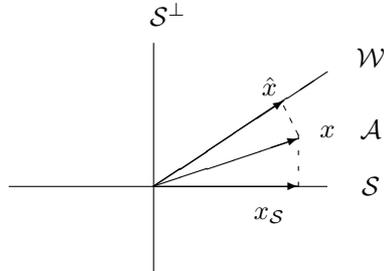
Figure 3: Illustration of minimal-norm reconstruction $\hat{x} = P_{\mathcal{W}}x$ of $x \in \mathcal{A}$ from $x_{\mathcal{S}} = P_{\mathcal{S}}x$, with $\mathcal{H} = \mathcal{A} \oplus \mathcal{S}^{\perp}$.

5   Robust Squared-Error Criteria

We now consider the general formulation of the sampling problem in which x is an arbitrary vector in $\mathcal{H}$. In this setting we propose two different strategies. In the first approach, we seek a linear transformation H that is independent of the samples c, which minimizes a worst-case error measure. The second method allows for more general nonlinear choices of H. Interestingly, we show that these two design strategies lead to the same solution.

5.1   Linear Minimax Design

5.1.1   Minimax Reconstruction

One approach to eliminating the dependence of the squared-error on x is to minimize a worst-possible error. However, if x can grow without bound, then so can the squared error. Therefore, to formulate a well-posed criterion, we minimize the worst-case error on the set of bounded-norm signals x.

To further simplify the problem, we first treat the case in which $H(c)$ is constrained to be linear so that $H(c) = Hc$ with H independent of c, leading to the following minimax design criterion:

$$\min_{H} \max_{\|x\| \leq L} \|WHS^*x - x\|^2, \tag{28}$$

where $L > 0$ is some positive (finite) constant. Note that the criterion (28) ignores the fact that the true value of x satisfies $S^*x = c$. Instead, for each possible value x_0 of x we try to obtain a small error with

11

respect to the samples that we would obtain for $x = x_0$, rather than the given samples c.

It turns out that the criterion (28) is too conservative and results in the trivial solution $H = 0$. To see this, we note that

$$\min_{H} \max_{\|x\| \leq L} \|WHS^*x - x\|^2 \geq \max_{\|x\| \leq L} \min_{H} \|WHS^*x - x\|^2 \geq \max_{\|x\| \leq L} \|P_{\mathcal{W}^\perp}x\|^2 = L^2, \tag{29}$$

where the first inequality results from exchanging the order of the minimum and maximum, and the second inequality stems from the fact that the best approximation to x in $\mathcal{W}$ is $P_{\mathcal{W}}x$. Now, with $H = 0$, we have that

$$\max_{\|x\| \leq L} \|WHS^*x - x\|^2 = \max_{\|x\| \leq L} \|x\|^2 = L^2, \tag{30}$$

and therefore $H = 0$ is a solution to (28).

5.1.2 Minimax Regret Reconstruction

To counterbalance the conservative behavior of the minimax approach, instead of minimizing the worst-case squared-norm error, we now consider minimizing the worst-case *regret*. The regret is defined as the difference between the squared-norm error and the smallest possible error $\|e_{\mathrm{OPT}}(x)\|^2$, where

$$e_{\mathrm{OPT}}(x) = P_{\mathcal{W}^\perp}x. \tag{31}$$

Beginning with a linear reconstruction, our problem is to seek an H that is a solution to

$$\min_{H} \max_{\|x\| \leq L} \left\{ \|WHS^*x - x\|^2 - \|P_{\mathcal{W}^\perp}x\|^2 \right\}. \tag{32}$$

The reconstruction is then given by $\hat{x} = WHc$. Expressing x as $x = P_{\mathcal{W}}x + P_{\mathcal{W}^\perp}x$, we have that

$$\|WHS^*x - x\|^2 = \|WHS^*x - P_{\mathcal{W}}x\|^2 + \|P_{\mathcal{W}^\perp}x\|^2, \tag{33}$$

and the problem (32) becomes

$$\min_{H} \max_{\|x\| \leq L} \|WHS^*x - P_{\mathcal{W}}x\|^2. \tag{34}$$

A solution to (34) is given in the following theorem.

Theorem 2. *Consider the problem*

$$\min_{H} \max_{\|x\| \leq L} \|WHS^*x - P_{\mathcal{W}}x\|^2,$$

12

where W and S are bounded transformations with $\mathcal{R}(W) = \mathcal{W}$ and $\mathcal{R}(S) = \mathcal{S}$. A possible solution is

$$H_{\text{REG}} = (W^*W)^\dagger W^* S (S^*S)^\dagger,$$

which is independent of L. The resulting reconstruction is $\hat{x} = W H_{\text{REG}} c = P_{\mathcal{W}} P_{\mathcal{S}} x$.

Proof. The maximum squared-error norm can be bounded as

$$\max_{\|x\|\leq L} \|WHS^*x - P_{\mathcal{W}}x\|^2 \geq \max_{\|x\|\leq L, x\in\mathcal{S}^\perp} \|WHS^*x - P_{\mathcal{W}}x\|^2 = \max_{\|x\|\leq L, x\in\mathcal{S}^\perp} \|P_{\mathcal{W}}x\|^2 = \max_{\|x\|\leq L} \|P_{\mathcal{W}}P_{\mathcal{S}^\perp}x\|^2, \tag{35}$$

where we used the fact that for $x \in \mathcal{S}^\perp$ we have $WHS^*x = 0$. Thus,

$$\min_H \max_{\|x\|\leq L} \|WHS^*x - P_{\mathcal{W}}x\|^2 \geq \max_{\|x\|\leq L} \|P_{\mathcal{W}}P_{\mathcal{S}^\perp}x\|^2. \tag{36}$$

We now show that with $H = H_{\text{REG}}$, the lower bound in (36) is achieved. Indeed, in this case,

$$WHS^*x = W (W^*W)^\dagger W^* S (S^*S)^\dagger S^*x = P_{\mathcal{W}}P_{\mathcal{S}}x, \tag{37}$$

and

$$\|WHS^*x - P_{\mathcal{W}}x\|^2 = \|P_{\mathcal{W}}(I - P_{\mathcal{S}})x\|^2 = \|P_{\mathcal{W}}P_{\mathcal{S}^\perp}x\|^2. \tag{38}$$

$\square$

Using Lemma 1 it follows that both of the operators $(W^*W)^\dagger$ and $(S^*S)^\dagger$ are bounded, so that H_{REG} of Theorem 2 is a bounded operator.

An interesting feature of the minimax regret reconstruction $\hat{x}$ of Theorem 2 is that it does not depend on the norm bound L. Therefore, $\hat{x} = P_{\mathcal{W}}P_{\mathcal{S}}x$ minimizes the worst-case regret error over all bounded inputs x, regardless of the norm of x. Furthermore, in the derivation of the minimax regret reconstruction we do not require the direct-sum condition $\mathcal{H} = \mathcal{W} \oplus \mathcal{S}^\perp$, which is necessary in the development of the consistent approach [5, 7, 10].

Another desirable property of the minimax regret reconstruction is that the resulting reconstruction error is always bounded by twice the norm of x. Specifically, expressing the error as

$$x - \hat{x} = P_{\mathcal{W}}x + P_{\mathcal{W}^\perp}x - P_{\mathcal{W}}P_{\mathcal{S}}x = P_{\mathcal{W}}P_{\mathcal{S}^\perp}x + P_{\mathcal{W}^\perp}x, \tag{39}$$

we have that

$$\|\hat{x} - x\|^2 = \|P_{\mathcal{W}}P_{\mathcal{S}^\perp}x\|^2 + \|P_{\mathcal{W}^\perp}x\|^2 \leq 2\|x\|^2; \tag{40}$$

tighter bounds on the reconstruction error are derived in Section 7.

5.2 Nonlinear Minimax Regret Design

Next, we consider minimizing the worst-case regret over all possible values of x that are consistent with the given samples $S^*x = c$, which results in the problem

$$\min_d \max_{c=S^*x, \|x\| \leq L} \|P_{\mathcal{W}}x - Wd\|^2. \tag{41}$$

In this case the reconstruction is given by $\hat{x} = Wd$ where $d = H(c)$ is the solution to (41) and in general can depend nonlinearly on c. Interestingly, the solution to (41) is linear, and is the same as the linear minimax regret solution:

Theorem 3. *Consider the problem*

$$\min_d \max_{c=S^*x, \|x\| \leq L} \|Wd - P_{\mathcal{W}}x\|^2,$$

where W and S are bounded transformations with $\mathcal{R}(W) = \mathcal{W}$ and $\mathcal{R}(S) = \mathcal{S}$. A possible solution is

$$d = (W^*W)^\dagger W^*S\,(S^*S)^\dagger c.$$

The resulting reconstruction is $\hat{x} = P_{\mathcal{W}}P_{\mathcal{S}}x$.

Proof. First we note that any x satisfying $S^*x = c$ and $\|x\| \leq L$ is of the form $x = S(S^*S)^\dagger c + v$ for some $v \in \mathcal{G}$ where

$$\mathcal{G} = \left\{ v \mid v \in \mathcal{S}^\perp, \|v\| \leq L' \right\},$$

and $L'^2 = L^2 - \left\| S(S^*S)^\dagger c \right\|^2$. Thus,

$$\max_{c=S^*x, \|x\| \leq L} \|Wd - P_{\mathcal{W}}x\|^2 =$$
$$= \max_{v \in \mathcal{G}} \left\| Wd - P_{\mathcal{W}}S(S^*S)^\dagger c - P_{\mathcal{W}}v \right\|^2$$
$$= \max_{v \in \mathcal{G}} \|a_d - P_{\mathcal{W}}v\|^2$$
$$= \max_{v \in \mathcal{G}} \left\{ \|a_d\|^2 - 2\Re\{\langle a_d, P_{\mathcal{W}}v \rangle\} + \|P_{\mathcal{W}}v\|^2 \right\}, \tag{42}$$

where we defined $a_d = W(d - (W^*W)^\dagger W^*S(S^*S)^\dagger c)$, and $\Re\{\cdot\}$ stands for the real part. Now, the maximum in (42) is achieved when

$$\Re\{\langle a_d, P_{\mathcal{W}}v \rangle\} = -|\langle a_d, P_{\mathcal{W}}v \rangle|. \tag{43}$$

Indeed, let $v \in \mathcal{G}$ be the vector for which the maximum is achieved. If $\langle a_d, P_{\mathcal{W}}v \rangle = 0$, then (43) is trivially true. Otherwise, we can define

$$v_2 = -\frac{\langle P_{\mathcal{W}}v, a_d \rangle}{|\langle a_d, P_{\mathcal{W}}v \rangle|}v. \tag{44}$$

Clearly, $\|v\| = \|v_2\|$ and $v_2 \in \mathcal{G}$. In addition, $\|P_{\mathcal{W}}v\| = \|P_{\mathcal{W}}v_2\|$ and $\langle a_d, P_{\mathcal{W}}v_2 \rangle = -|\langle a_d, P_{\mathcal{W}}v \rangle|$ so that the objective in (42) at v_2 is larger than the objective at v unless (43) is satisfied.

Combining (43) and (42) our problem becomes

$$\min_d \max_{v \in \mathcal{G}} \left\{ \|a_d\|^2 + 2|\langle a_d, P_{\mathcal{W}}v \rangle| + \|P_{\mathcal{W}}v\|^2 \right\}. \tag{45}$$

Denoting the optimal objective value by A, and replacing the order of minimization and maximization,

$$
\begin{aligned}
A & \geq \max_{v \in \mathcal{G}} \min_d \left\{ \|a_d\|^2 + 2|\langle a_d, P_{\mathcal{W}}v \rangle| + \|P_{\mathcal{W}}v\|^2 \right\} \\
& = \max_{v \in \mathcal{G}} \|P_{\mathcal{W}}v\|^2,
\end{aligned} \tag{46}
$$

where we used the fact that $\|a_d\|^2 + 2\,|\langle a_d, P_{\mathcal{W}}v \rangle| \geq 0$ with equality for $a_d = 0$, or

$$d = (W^*W)^\dagger W^* S (S^*S)^\dagger c. \tag{47}$$

Thus, for any choice of d,

$$\min_d \max_{v \in \mathcal{G}} \|a_d - P_{\mathcal{W}}v\|^2 \geq \max_{v \in \mathcal{G}} \|P_{\mathcal{W}}v\|^2. \tag{48}$$

The proof then follows from the fact that d given by (47) achieves the lower bound (48). □

In Fig. 4 we illustrate the minimax regret reconstruction. We have seen already in Section 4.1 that knowing the samples $c[n]$ is equivalent to knowing $x_{\mathcal{S}} = P_{\mathcal{S}}x$. Thus, our reconstruction problem is that of approximating an arbitrary x in $\mathcal{H}$ from its orthogonal projection $x_{\mathcal{S}}$ onto $\mathcal{S}$, where the reconstruction is constrained to lie in $\mathcal{W}$. As illustrated in the figure, the minimax regret reconstruction chooses the orthogonal projection of $x_{\mathcal{S}}$ onto $\mathcal{W}$.
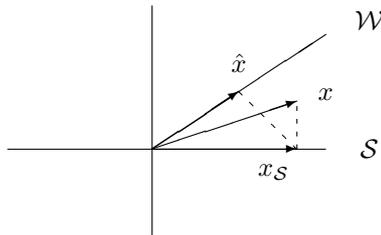
Figure 4: Illustration of minimax regret reconstruction $\hat{x} = P_{\mathcal{W}} P_{\mathcal{S}} x$ of x from $x_{\mathcal{S}} = P_{\mathcal{S}} x$.

6   Reconstruction in Shift-Invariant Spaces

The reconstruction algorithms of Theorems 1, 2 and 3 were derived for general subspaces of a Hilbert space $\mathcal{H}$. An interesting special case of this setup is when $\mathcal{H} = L_2$ and $\mathcal{S}$ and $\mathcal{W}$ are real SI subspaces, each spanned by the integer shifts of a fixed generating function [22, 23]. In this case, as we now show, H_{REG} and $H_{\mathcal{A}}$ reduce to discrete-time LTI filters, and the sampling and reconstruction can be implemented using continuous-time LTI filters.

Suppose that $x = x(t)$ is a real signal in L_2, and that

$$
\begin{aligned}
\mathcal{S} &= \left\{ f(t) \mid f(t) = \sum_n a[n]s(t-n), a \in \ell_2 \right\}; \\
\mathcal{W} &= \left\{ f(t) \mid f(t) = \sum_n a[n]w(t-n), a \in \ell_2 \right\},
\end{aligned}
\tag{49}
$$

where $s(t)$ and $w(t)$ are the real generators of $\mathcal{S}$ and $\mathcal{W}$, respectively. To ensure that the vectors $\{s_n(t) = s(t-n)\}$ and $\{w_n(t) = w(t-n)\}$ form frames for $\mathcal{S}$ and $\mathcal{W}$ respectively, we must have that [24]

$$
\begin{aligned}
\alpha \leq \sum_n |W(\omega + 2\pi k)|^2 \leq \beta, \quad \omega \in \mathcal{I}_{\mathcal{W}}; \\
\gamma \leq \sum_n |S(\omega + 2\pi k)|^2 \leq \eta, \quad \omega \in \mathcal{I}_{\mathcal{S}},
\end{aligned}
\tag{50}
$$

for some $0 < \alpha \leq \beta < \infty$ and $0 < \eta \leq \gamma < \infty$, where $W(\omega), S(\omega)$ are the continuous-time Fourier transforms of the generators $s(t), w(t)$, and $\mathcal{I}_{\mathcal{W}}, \mathcal{I}_{\mathcal{S}}$ are the set of frequencies ω for which $\sum_k |W(\omega + 2\pi k)|^2 \neq 0$ and $\sum_k |S(\omega + 2\pi k)|^2 \neq 0$, respectively. In this case, the samples $c[n]$, which are given by

$$
c[n] = \int s(t-n)x(t)dt = x(t) * s(-t)|_{t=n},
\tag{51}
$$

correspond to samples at times $t = n$ of the output of a filter with impulse response $s(-t)$, with $x(t)$ as its input. Here $g(t) * z(t)$ denotes the continuous-time convolution between the signals $g(t)$ and $z(t)$, and $y(t)|_{t=n} = y(n)$. The reconstructed signal can be viewed as the output of a filter with impulse response $w(t)$, with an impulse train whose values are the corrected measurements $d[n]$ as its input. In the minimax regret reconstruction, $d = H_{\mathrm{REG}}c$, and in the subspace approach $d = H_{\mathcal{A}}c$. We now show that in both cases the samples $d[n]$ can be obtained by filtering the samples $c[n]$ with a discrete-time LTI filter.

Denoting by S and W the set transformations corresponding to $\{s_n(t)\}$ and $\{w_n(t)\}$ respectively, it is easy to see that W^*Sa is equivalent to filtering the sequence $a[n]$ by a discrete-time LTI filter with discrete-time Fourier transform (DTFT)

$$
Q\left(e^{j\omega}\right) = \sum_{k=-\infty}^{\infty} S(\omega + 2\pi k)W^*(\omega + 2\pi k).
\tag{52}
$$

16

Similarly, the pseudoinverse $(S^*W)^\dagger$ can be represented by a filter with DTFT

$$T_{\mathcal{SW}}\left(e^{j\omega}\right) = \begin{cases} \frac{1}{Q(\omega)}, & \omega \in \mathcal{I}_{\mathcal{SW}}; \\ 0, & \omega \notin \mathcal{I}_{\mathcal{SW}}, \end{cases} \tag{53}$$

where $\omega \in \mathcal{I}_{\mathcal{SW}}$ if $Q(\omega) \neq 0$. Combining (52) and (53) it follows that the corrected samples $d = H_{\text{REG}}c$ can be obtained from c using a discrete-time filter with DTFT

$$H_{\text{REG}}\left(e^{j\omega}\right) = \begin{cases} \frac{\sum_{k=-\infty}^{\infty} S(\omega+2\pi k)W^*(\omega+2\pi k)}{\sum_{k=-\infty}^{\infty}|S(\omega+2\pi k)|^2 \sum_{k=-\infty}^{\infty}|W(\omega+2\pi k)|^2}, & \omega \in \mathcal{I}_{\mathcal{W}} \cap \mathcal{I}_{\mathcal{S}}; \\ 0, & \omega \notin \mathcal{I}_{\mathcal{W}} \cap \mathcal{I}_{\mathcal{S}}. \end{cases} \tag{54}$$

The sampling scheme of Fig. 1 then reduces to that depicted in Fig. 2.

Applying the Cauchy-Schwartz inequality to the numerator of (54), we see that the regret reconstruction filter (54) has the property that its magnitude is no larger than 1.

Next, we consider subspace reconstruction with $H_{\mathcal{A}}$ of Theorem 1, where we assume that $x(t)$ is in a SI subspace $\mathcal{A}$ generated by a function $a(t)$, such that $\mathcal{A} \oplus \mathcal{S}^\perp = \mathcal{H}$. This direct-sum condition can be verified quite easily in shift-invariant systems by exploiting the following result.

Proposition 3. *[25, Proposition 4.8] Suppose that $\{s(t-n)\}_{n\in\mathbb{Z}}$ and $\{w(t-n)\}_{n\in\mathbb{Z}}$ are frame sequences for $\mathcal{S}$ and $\mathcal{W}$, and let $\mathcal{I}_{\mathcal{W}}, \mathcal{I}_{\mathcal{S}}$ be the set of frequencies ω for which $\sum_k |W(\omega+2\pi k)|^2 \neq 0$ and $\sum_k |S(\omega+2\pi k)|^2 \neq 0$, respectively. Then $L_2 = \mathcal{W} \oplus \mathcal{S}^\perp$ if and only if $\mathcal{I}_{\mathcal{W}} = \mathcal{I}_{\mathcal{S}}$ and there exists a constant $\alpha > 0$ such that*

$$\left| \sum_{k=-\infty}^{\infty} S(\omega + 2\pi k)W^*(\omega + 2\pi k) \right| \geq \alpha, \quad \forall \omega \in \mathcal{I}_{\mathcal{W}}.$$

If we choose a generator $a(t)$ with frequency response $A(\omega)$ satisfying the condition of Proposition 3, then the subspace transformation $H_{\mathcal{A}}$ is equivalent to an LTI filter with DTFT

$$H_{\mathcal{A}}\left(e^{j\omega}\right) = \begin{cases} \frac{\sum_{k=-\infty}^{\infty} A(\omega+2\pi k)W^*(\omega+2\pi k)}{\sum_{k=-\infty}^{\infty}|W(\omega+2\pi k)|^2 \sum_{k=-\infty}^{\infty} A(\omega+2\pi k)S^*(\omega+2\pi k)}, & \omega \in \mathcal{I}_{\mathcal{W}} \cap \mathcal{I}_{\mathcal{S}}; \\ 0, & \omega \notin \mathcal{I}_{\mathcal{W}} \cap \mathcal{I}_{\mathcal{S}}, \end{cases} \tag{55}$$

where $A(\omega)$ is the continuous-time Fourier transform of $a(t)$ and we have used the fact that $\sum_{k=-\infty}^{\infty} A(\omega + 2\pi k)S^*(\omega + 2\pi k) \neq 0$ on $\mathcal{I}_{\mathcal{S}}$, since $\mathcal{A} \oplus \mathcal{S}^\perp = \mathcal{H}$.

Finally, the consistent reconstruction scheme for SI spaces, developed in [1], has the same form as Fig. 2, where the filter $H(\omega)$ is specified by

$$H_{\text{CON}}\left(e^{j\omega}\right) = \begin{cases} \frac{1}{\sum_{k=-\infty}^{\infty} S^*(\omega+2\pi k)W(\omega+2\pi k)}, & \omega \in \mathcal{I}_{\mathcal{W}} \cap \mathcal{I}_{\mathcal{S}}; \\ 0, & \omega \notin \mathcal{I}_{\mathcal{W}} \cap \mathcal{I}_{\mathcal{S}}. \end{cases} \tag{56}$$

If $\mathcal{H} = \mathcal{W} \oplus \mathcal{S}^\perp$, then from Proposition 3 the filter H_{CON} is well defined. Note, however, that although the

magnitude of the filter is bounded, it can be arbitrarily large. This is in contrast with the magnitude of the regret reconstruction filter, which, as we have seen, is no larger than 1.

6.1 Efficient Implementation with Splines

A special class of SI spaces, that is popular in applications, is the class of spline spaces which are generated by a spline function [3, 13, 14]. When $\mathcal{S}$ and $\mathcal{W}$ are spline spaces, we can use the results of [13, 14] to obtain a particularly efficient implementation of the subspace and minimax regret filters.

Suppose that $\mathcal{S}$ and $\mathcal{W}$ are SI spaces generated by splines of order n_s and n_w respectively, so that $s(t) = \beta^{n_s}(t), w(t) = \beta^{n_w}(t)$ in (49), where $\beta^m(t)$ is a symmetrical B-spline of order m, defined recursively by $\beta^{m+1}(t) = \beta^m(t) * \beta^0(t)$ with

$$\beta^0(t) = \begin{cases} 1, & t \in [-0.5, 0.5); \\ 0, & \text{otherwise.} \end{cases} \tag{57}$$

The discrete representative of a continuous spline function $\beta^m(t)$ is the discrete and symmetric finite impulse response (FIR) spline filter $b^m[n]$, which is obtained by sampling the continuous spline: $b^m[n] = \beta^m|_{t=n}$. One of the main advantages of using splines is the ability to compute many operations using the discrete representatives (see for example [15]). In particular, in our context, the infinite sums in the definition of the filters (54), (55) and (56) can be obtained explicitly.

Consider the Fourier transform $Q\left(e^{j\omega}\right) = \sum_{k=-\infty}^{\infty} S^*(\omega + 2\pi k)W(\omega + 2\pi k)$ which is the transform of the discrete-time sequence

$$q[n] = w(t) * s(-t)|_{t=n} = \int_{-\infty}^{\infty} w(t)s(t-n)dt. \tag{58}$$

If $w(t) = \beta^{n_w}(t)$ and $s(t) = \beta^{n_s}(t)$ are splines, then

$$q[n] = \int_{-\infty}^{\infty} \beta^{n_w}(t)\beta^{n_s}(n-t)dt = b^{n_w+n_s+1}[n], \tag{59}$$

where we used the symmetry of the spline function and the convolution property $\beta^{n_w}(t) * \beta^{n_s}(t) = \beta^{n_w+n_s+1}(t)$. Therefore, $Q(\omega)$ is simply the transform of $b^{n_w+n_s+1}[n]$, and $1/Q(\omega)$ is the transform of the convolutional inverse of the filter $b^m[n]$, denoted $\left\{b^{n_w+n_s+1}\right\}^{-1}[n]$. Each of the filters (54), (55) and (56) can then be implemented using discrete-time spline filters of appropriate order, as summarized in Table 1. In the table, n_w, n_s and n_a denote the orders of the spline generators of the subspaces $\mathcal{W}, \mathcal{S}$ and $\mathcal{A}$, respectively.

Note that when $n_s = n_w$ (*i.e.*, $\mathcal{S} = \mathcal{W}$) all the filters reduce to $\left\{b^{2n_w+1}\right\}^{-1}$ resulting in $\hat{x} = P_{\mathcal{W}}x$.

Table 1: Implementation of the various reconstruction filters for B-spline spaces

Reconstruction	Filter Impulse Response
Consistent	$\left\{b^{n_s+n_w+1}\right\}^{-1}$
Regret	$\left\{b^{2n_w+1}\right\}^{-1} * b^{n_w+n_s+1} * \left\{b^{2n_s+1}\right\}^{-1}$
Subspace	$\left\{b^{2n_w+1}\right\}^{-1} * b^{n_w+n_a+1} * \left\{b^{n_a+n_s+1}\right\}^{-1}$

7   Error Analysis

The reconstruction algorithms we discussed in previous sections are based on minimizing an appropriate cost function, and differ in their assumptions on the input signal. The various approaches are summarized in Table 2. In the table, the notation $H(S^*x)$ means that H is a function of S^*x, $\mathcal{L} = \{x \mid \|x\| \leq L\}$, and $\mathcal{G} = \{x \mid S^*x = c, \|x\| \leq L\}$.

Table 2: Comparison of reconstruction methods

	Assumptions	Criterion	Optimal H	Reconstruction
Consistent	$\mathcal{H} = \mathcal{W} \oplus \mathcal{S}^\perp$	$\min_H \|S^*WH(S^*x) - S^*x\|^2$	$(S^*W)^\dagger$	$\hat{x} = E_{\mathcal{W}\mathcal{S}^\perp} x$
Subspace $\mathcal{A}$	$\mathcal{H} = \mathcal{A} \oplus \mathcal{S}^\perp$ $x \in \mathcal{A}$	$\min_H \|WH(S^*x) - x\|^2$	$(W^*W)^\dagger W^*A(S^*A)^\dagger$ where $\mathcal{R}(A) = \mathcal{A}$	$\hat{x} = P_\mathcal{W} x$
Minimax regret	$\|x\| \leq L$	$\min_H \max_{x \in \mathcal{L}} \|WHS^*x - P_\mathcal{W}x\|^2$ $\min_d \max_{x \in \mathcal{G}} \|Wd - P_\mathcal{W}x\|^2$	$(W^*W)^\dagger W^*S(S^*S)^\dagger$	$\hat{x} = P_\mathcal{W} P_\mathcal{S} x$

In Sections 7.1 and 7.2 we derive tight bounds on the norm of the error when using the regret and consistent reconstruction methods, respectively. Based on these bounds, in Section 7.3, we compare the performance of both approaches. Before proceeding to the detailed development, we note that if we know a priori that x lies in a subspace $\mathcal{A}$ such that $\mathcal{H} = \mathcal{A} \oplus \mathcal{S}^\perp$, then the subspace technique will yield the minimal error approximation of x and therefore is optimal in the squared-norm sense. When $\mathcal{A} = \mathcal{S}$ this strategy reduces to the minimax regret method, while if $\mathcal{A} = \mathcal{W}$, then we obtain the consistent reconstruction.

Unfortunately, in many cases we do not have prior knowledge on the subspace in which x is contained. Therefore, we must resort to the minimax regret or the consistent reconstruction techniques. Our analysis below shows that if the spaces $\mathcal{S}$ and $\mathcal{W}$ are sufficiently far apart, or if x has enough energy in $\mathcal{S}$, then the minimax regret method is preferable in a squared-norm error sense to the consistent reconstruction approach. These analytical results are also demonstrated through simulation in Section 8.

7.1   Error Bounds Using the Minimax Regret Method

Theorem 4 provides tight bounds on the error resulting from the minimax regret design strategy.

Theorem 4. *Let $e_{\mathrm{REG}}(x) = x - P_{\mathcal{W}}P_{\mathcal{S}}x$ denote the error resulting from the minimax regret reconstruction of Theorem 3 and let $e_{\mathrm{OPT}}(x) = P_{\mathcal{W}^\perp}x$ be the optimal error in the squared-norm sense. Then*

$$\|e_{\mathrm{OPT}}(x)\|^2 + \cos^2\left(\mathcal{S}^\perp, \mathcal{W}\right)\|P_{\mathcal{S}^\perp}x\|^2 \leq \|e_{\mathrm{REG}}(x)\|^2 \leq \|e_{\mathrm{OPT}}(x)\|^2 + \sin^2\left(\mathcal{W}, \mathcal{S}\right)\|P_{\mathcal{S}^\perp}x\|^2, \qquad (60)$$

where $\cos(\cdot)$ and $\sin(\cdot)$ are defined in (11).

Before proving the theorem, we note that if we know the norm bound $\|x\| \leq L$, then $\|P_{\mathcal{S}^\perp}x\|^2 \leq L^2 - \|S(S^*S)^\dagger c\|^2$ with strict equality if $\|x\| = L$.

Proof. Writing $e_{\mathrm{REG}}(x) = P_{\mathcal{W}^\perp}e_{\mathrm{REG}}(x) + P_{\mathcal{W}}e_{\mathrm{REG}}(x)$, we have that

$$\|e_{\mathrm{REG}}(x)\|^2 = \|e_{\mathrm{OPT}}(x)\|^2 + \|P_{\mathcal{W}}P_{\mathcal{S}^\perp}x\|^2. \qquad (61)$$

Note that for $x \in \mathcal{S}$, $\|e_{\mathrm{REG}}(x)\|^2 = \|e_{\mathrm{OPT}}(x)\|^2$, so that the minimax regret reconstruction is optimal. If $x \notin \mathcal{S}$, then $\|P_{\mathcal{S}^\perp}x\| \neq 0$ and we can rewrite (61) as

$$\|e_{\mathrm{REG}}(x)\|^2 = \|e_{\mathrm{OPT}}(x)\|^2 + \|P_{\mathcal{W}}v\|^2 \|P_{\mathcal{S}^\perp}x\|^2, \qquad (62)$$

where we defined $v = P_{\mathcal{S}^\perp}x / \|P_{\mathcal{S}^\perp}x\|$. Since v is a normalized vector in $\mathcal{S}^\perp$ which is orthogonally projected onto $\mathcal{W}$, we can use the definitions in (11) to obtain

$$\cos^2\left(\mathcal{S}^\perp, \mathcal{W}\right) \leq \|P_{\mathcal{W}}v\|^2 \leq \sin^2\left(\mathcal{S}^\perp, \mathcal{W}^\perp\right). \qquad (63)$$

Combining (63) with (62) and (12), results in (60).

If $v \in \mathcal{S}^\perp$ achieves the maximum (minimum)[3] angle with $\mathcal{W}$, then with $x = v + s$, where $s = P_{\mathcal{S}}x = S(S^*S)^\dagger c$, we achieve the upper (lower) bound of (60). Therefore, the bounds of (60) are tight. □

7.2 Error Bounds Using the Consistent Method

An upper bound on the norm of the error resulting from the consistent reconstruction $e_{\mathrm{CON}}(x) = x - E_{\mathcal{W}\mathcal{S}^\perp}x = E_{\mathcal{S}^\perp\mathcal{W}}x$ was developed in [1] using the relation

$$e_{\mathrm{OPT}}(x) = P_{\mathcal{W}^\perp}e_{\mathrm{CON}}(x). \qquad (64)$$

Specifically, it was shown that

$$\|e_{\mathrm{CON}}(x)\|^2 \leq \frac{\|e_{\mathrm{OPT}}(x)\|^2}{\cos^2\left(\mathcal{W}, \mathcal{S}\right)}. \qquad (65)$$

---

[3]Here we assume that the inf and the sup in (11) can be replaced by min and max, respectively. We refer the reader to [1, Theorem 2] for sufficient conditions for the above to hold, in the case of SI spaces.

Note that although $\mathcal{H} = \mathcal{S} \oplus \mathcal{W}^\perp$ implies $\cos(\mathcal{W}, \mathcal{S}) > 0$ [21], with proper choice of $\mathcal{S}$ and $\mathcal{W}$, the reconstruction error can be made arbitrarily large.

We now show that using (64) we can also obtain a tight lower bound on the error (this bound is usually higher than the trivial bound $\|e_{\mathrm{CON}}(x)\| \geq \|e_{\mathrm{OPT}}(x)\|$ stated in [1]). If $e_{\mathrm{CON}}(x) = e_{\mathrm{OPT}}(x) = 0$, then the consistent reconstruction is optimal. Next suppose that $e_{\mathrm{CON}}(x) \neq 0$, which occurs if and only if[4] $e_{\mathrm{OPT}}(x) \neq 0$. In this case we derive from (64):

$$0 < \frac{\|e_{\mathrm{OPT}}(x)\|^2}{\|e_{\mathrm{CON}}(x)\|^2} = \frac{\|P_{\mathcal{W}^\perp} e_{\mathrm{CON}}(x)\|^2}{\|e_{\mathrm{CON}}(x)\|^2} \leq \sin^2\left(\mathcal{S}^\perp, \mathcal{W}\right), \tag{66}$$

obtaining the tight lower bound

$$\frac{\|e_{\mathrm{OPT}}(x)\|^2}{\sin^2\left(\mathcal{S}^\perp, \mathcal{W}\right)} \leq \|e_{\mathrm{CON}}(x)\|^2. \tag{67}$$

Note that $\sin^2\left(\mathcal{S}^\perp, \mathcal{W}\right) \leq 1$ with equality only if $\mathcal{W} = \mathcal{S}$ (in which case $e_{\mathrm{CON}}(x) = e_{\mathrm{REG}}(x) = e_{\mathrm{OPT}}(x)$). Combining (65) and (67),

$$\frac{\|e_{\mathrm{OPT}}(x)\|^2}{\sin^2\left(\mathcal{S}^\perp, \mathcal{W}\right)} \leq \|e_{\mathrm{CON}}(x)\|^2 \leq \frac{\|e_{\mathrm{OPT}}(x)\|^2}{\cos^2\left(\mathcal{W}, \mathcal{S}\right)}. \tag{68}$$

As in the case of the bounds (60), it can be shown that the bounds of (68) are tight, by taking $v \in \mathcal{S}^\perp$ which achieves the maximum (minimum) angle with respect to $\mathcal{W}^\perp$ and constructing $x = v + w$ where $w \in \mathcal{W}$ satisfies $P_{\mathcal{S}} w = S(S^*S)^\dagger c$ (so that $S^*x = c$).

7.3 Bound Comparison

Using the bounds in the previous subsections we can identify regions of $\|e_{\mathrm{OPT}}(x)\|$ for which the regret approach is preferable to the consistent method, for all values of x, and vice versa. Specifically, if the upper bound in (60) is smaller than the lower bound in (68), then the norm of the error resulting from the consistent reconstruction scheme will be larger than that resulting from the regret approach. Manipulating the equations, it can be shown that this occurs when

$$\|e_{\mathrm{OPT}}(x)\|^2 \geq \gamma_1 \|P_{\mathcal{S}^\perp} x\|^2, \tag{69}$$

where the constant γ_1 is given by

$$\gamma_1 = \frac{\sin^2\left(\mathcal{S}^\perp, \mathcal{W}\right) \sin^2\left(\mathcal{W}, \mathcal{S}\right)}{\cos^2\left(\mathcal{S}^\perp, \mathcal{W}\right)}. \tag{70}$$

Since the numerator of (70) is no larger than 1, a sufficient condition to ensure a lower error using the regret reconstruction is

$$\|e_{\mathrm{OPT}}(x)\|^2 \geq \frac{1}{\cos^2\left(\mathcal{S}^\perp, \mathcal{W}\right)} \|P_{\mathcal{S}^\perp} x\|^2. \tag{71}$$

---

[4]An equivalent claim is $e_{\mathrm{CON}}(x) = 0$ if and only if $e_{\mathrm{OPT}}(x) = 0$. Indeed, when $e_{\mathrm{CON}}(x) = 0$, trivially $e_{\mathrm{OPT}}(x) = 0$. On the other hand, assuming $e_{\mathrm{OPT}}(x) = 0$, (64) implies $e_{\mathrm{CON}}(x) \in \mathcal{W}$. Since $e_{\mathrm{CON}}(x) \in \mathcal{S}^\perp$, and $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$, we must have $e_{\mathrm{CON}}(x) = 0$.

Evidently, if $\mathcal{W}$ is close to $\mathcal{S}^\perp$ and most of the signal energy is within the sampling space, then the minimax regret method will result in a lower error than the consistent reconstruction approach.

Similarly, by comparing the worst-case bound on the consistent reconstruction error with the best-case bound on the regret error, we can show that if

$$\|e_{\mathrm{OPT}}(x)\|^2 \leq \gamma_2 \|P_{\mathcal{S}^\perp} x\|^2, \tag{72}$$

where

$$\gamma_2 = \frac{\cos^2\left(\mathcal{W}, \mathcal{S}\right) \cos^2\left(\mathcal{S}^\perp, \mathcal{W}\right)}{\sin^2\left(\mathcal{W}, \mathcal{S}\right)}, \tag{73}$$

then the consistent reconstruction scheme will result in a lower error. A sufficient condition is

$$\|e_{\mathrm{OPT}}(x)\|^2 \leq \cos^2\left(\mathcal{W}, \mathcal{S}\right) \cos^2\left(\mathcal{S}^\perp, \mathcal{W}\right) \|P_{\mathcal{S}^\perp} x\|^2. \tag{74}$$

Using (11) and (12), we can readily see that $\gamma_2 \leq \gamma_1$ by comparing the numerators and denominators of the two terms. These results are illustrated in Fig. 5.

$$\|e_{\mathrm{CON}}(x)\|^2 \leq \|e_{\mathrm{REG}}(x)\|^2 \qquad \qquad \|e_{\mathrm{CON}}(x)\|^2 \geq \|e_{\mathrm{REG}}(x)\|^2$$

$$\gamma_2 \|P_{\mathcal{S}^\perp} x\|^2 \quad \gamma_1 \|P_{\mathcal{S}^\perp} x\|^2 \qquad\qquad \|e_{\mathrm{OPT}}(x)\|^2$$

Figure 5: Regions of $\|e_{\mathrm{OPT}}(x)\|^2$ in which the regret reconstruction leads to a smaller error than the consistent reconstruction, and vice versa.

As evident from the figure, when $\|e_{\mathrm{OPT}}(x)\|^2$ is large (*i.e.,* most of the signal energy is not within the reconstruction space), or the bound $\gamma_1 \|P_{\mathcal{S}^\perp} x\|^2$ is small (*i.e.,* most of the signal energy is within the sampling space and $\mathcal{W}$ is 'close' to $\mathcal{S}^\perp$) the regret scheme will outperform the consistent reconstruction method. Conversely, for small values of $\|e_{\mathrm{OPT}}(x)\|^2$, the consistent reconstruction approach is preferred. These results are intuitive as illustrated geometrically in Fig. 6. In Fig. 6(a) we depict the consistent and regret reconstruction when $\mathcal{W}$ is far from $\mathcal{S}$. As can be seen in the figure, in this case the error resulting from the consistent reconstruction is large with respect to the regret reconstruction error. In Fig. 6(b), $\mathcal{W}$ and $\mathcal{S}$ are close, and the errors have roughly the same magnitude.

8 Examples

We now present several examples illustrating the minimax regret reconstruction and compare it with the consistent reconstruction method. In Section 8.1 we consider a speech processing example, and in Section 8.2 we provide an image processing example, which also demonstrates the error analysis of Section 7.

(a)   (b)
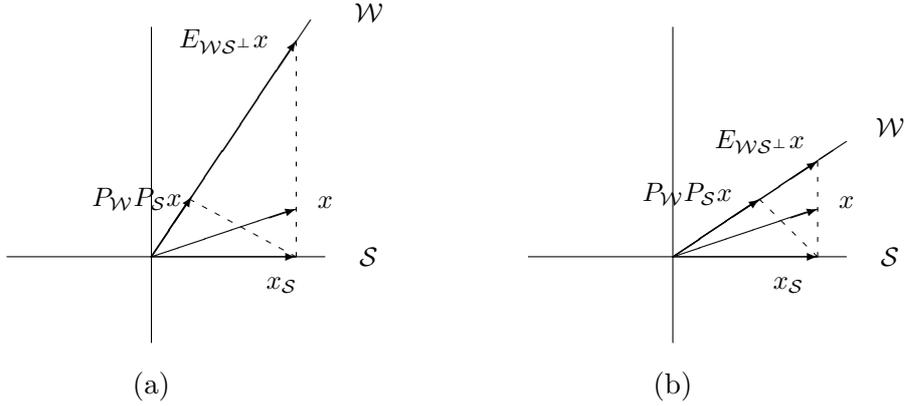
Figure 6: Comparison of the minimax regret reconstruction ($P_{\mathcal{W}}P_{\mathcal{S}}x$) and consistent reconstruction ($E_{\mathcal{W}\mathcal{S}^\perp}x$) for two different choices of $\mathcal{W}$ (a) $\mathcal{W}$ 'far' from $\mathcal{S}$ (b) $\mathcal{W}$ 'close' to $\mathcal{S}$.

8.1   Speech Processing Example

Suppose we sample a continuous-time speech signal $x(t)$ using a non-ideal sampler, so that the samples $c[n]$ are equal to the average of the signal over intervals of length Δ:

$$c[n] = \frac{1}{\Delta} \int_{nT-\Delta}^{nT} x(t)dt. \tag{75}$$

The samples $c[n]$ can be obtained by filtering $x(t)$ with a filter whose impulse response $s(t)$ is given by

$$s(t) = \begin{cases} \frac{1}{\Delta}, & 0 \le t \le \Delta; \\ 0, & \text{otherwise}, \end{cases} \tag{76}$$

and then sampling the output at times $t = nT$. The filter $s(t)$ can be viewed as a (non-ideal) low-pass filter (LPF). In the simulations below, we use $T = 4000^{-1}$s and $\Delta = 1.125$ms. The reconstructed output $\hat{x}(t)$ is obtained from the transformed sequence $d = Hc$ using an interpolation kernel $w(t)$, i.e., $\hat{x}(t) = \sum_n d[n]w(t - nT)$, where we choose $w(t)$ as a non-ideal LPF with support on $t \in [0, 1.75]$ms which approximates an ideal LPF with cutoff frequency 2kHz.

For the purpose of simulation we approximate the continuous-time signal $x(t)$ with a discrete sequence $x[n]$ on a fine grid. The signal was chosen as a speech fragment, taken from the Timit database [26], at a sample rate of 8kHz. The continuous time integration kernel $s(t)$ is approximated by the discrete filter

$$s[n] = \begin{cases} \frac{1}{N}, & 0 \le n \le N - 1; \\ 0, & \text{otherwise}, \end{cases} \tag{77}$$

with $N = 10$ samples. The ideal sampling is implemented by down-sampling the filter output with a
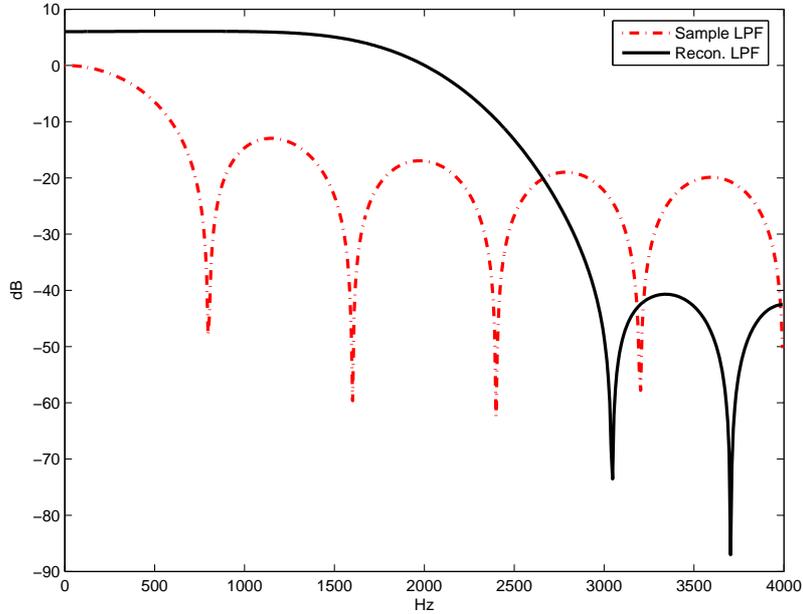
Figure 7: Frequency responses of the non ideal low-pass filters.

decimation factor of 2. The (non-ideal) LPF followed by decimation can be described by proper construction of the sampling matrix S^*.

To implement the reconstruction we use a linear-phase FIR filter of order 14 (with cutoff frequency 2kHz) as the interpolation kernel. Here as well, the discrete-time interpolation kernel simulates the continuous-time interpolation kernel, by constructing it over the 8kHz fine grid and up-sampling the input sequence d by a factor of 2, prior to filtering. The up-sampling followed by the filtering operation can be described by proper construction of the matrix W. The frequency responses of the non-ideal low-pass filters are presented in Fig. 7. Figure 8 shows an example of an input sequence $x[n]$ and 3 different reconstructed signals, corresponding to $H = I$ (that is, direct reconstruction without a correcting transformation), $H = (S^*W)^\dagger$ (consistent reconstruction) and $H = H_{\text{REG}}$ of Theorems 2 and 3. As can be seen from the figure, the results of direct reconstruction are poor. This is despite the fact that delay and gain compensations were applied for this method.[5] In particular, the reconstruction filter was multiplied by a factor of 2 as custom in a down-up sampling scheme with a factor of 2. As can be seen from the figure, the consistent and minimax regret reconstruction methods perform much better. However, it can be seen that the minimax regret reconstruction leads to better results than the consistent reconstruction method. For statistical significance, the experiment was repeated with 1000 different sections of speech of length 200 samples each, and we evaluated the normalized error norm $\|x - \hat{x}\| / \|x\|$ for each fragment. The average normalized errors obtained where $0.83, 0.744, 0.413$, for the direct, consistent and minimax regret reconstruction methods, respectively.

We also note that subjective listening tests performed for this setup, confirmed that the minimax regret

---

[5]Such compensations are not required in the consistent and regret methods, as they are taken care of automatically by H.

24
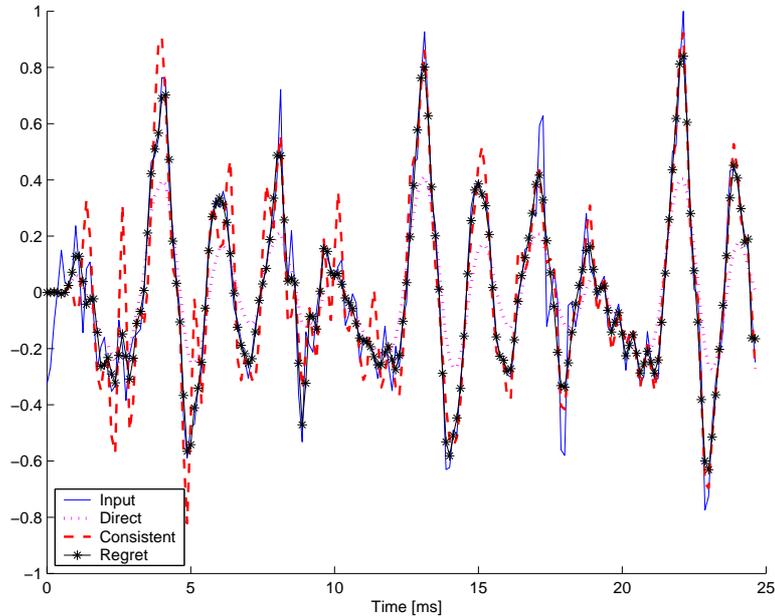
Figure 8: The original speech fragment $x[n]$ and the reconstructed signal using three different methods: direct reconstruction, consistent reconstruction and the minimax regret reconstruction.

estimator outperforms the other methods.

8.2   Image Processing Example

We next consider an image processing example, with the goal of partially demonstrating the error analysis of Section 7, in the special case of spline subspaces. For that, we assume that the sampling space $\mathcal{S}$ and the reconstruction space $\mathcal{W}$ are generated by splines of order $n_s = 0$ and $n_w = 2$, respectively. In this example, we take the input signal x to be the 512×512 gray-scale Lena image. To obtain a continuous-time representation for the image, we assume the following continuous-time model within each axis:

$$x(t) = \alpha \sum_n y_1[n]\beta^{n_s}(t-n) + (1-\alpha) \sum_n y_2[n]\beta^{n_w}(t-n), \tag{78}$$

where β^{n_a} is a B-spline generator of order n_a, as defined in Section 6.1, and y_1, y_2 are the spline coefficients. Thus, within each axis we model the image as present both in the sampling space and in the reconstruction space, where the parameter $0 < \alpha < 1$ controls the amount of energy of our signal within each of these spaces. We note that the suggested convex representation preserves the values of the image on the original grid.

As the given image is known only at a discrete grid, we first calculate (within each axis) the spline coefficients $y_1[n], y_2[n]$ of (78) to obtain a continuous-time representation for the image. These coefficients can be computed by direct B-spline filtering [13] with the symmetric IIR filter $\{b^{n_s}\}^{-1}[n]$ (to obtain $y_1[n]$) or $\{b^{n_w}\}^{-1}[n]$ (to obtain $y_2[n]$). For example, $y_1[n] = \left\{\{b^{n_s}\}^{-1} * x\right\}[n]$, where x are the given pixel values.

The samples $c = S^*x(t)$ can then be obtained directly from the sequences $y_1[n], y_2[n]$ as

$$c[n] = \langle \beta^{n_s}(t-n), x(t) \rangle = \alpha \left\{ y_1 * b^{2n_s+1} \right\} [n] + (1-\alpha) \left\{ y_2 * b^{n_s+n_w+1} \right\} [n], \qquad (79)$$

where we used our model (78) for x and evaluated spline inner-products based on the results in [13, 14]. Given the samples $c[n]$, we compute $d = Hc$ using three different choices of H: No correction at all *i.e.,* the direct method $H = I$, consistent reconstruction $H = H_{\text{CON}}$ and regret reconstruction $H = H_{\text{REG}}$. The reconstructed signal is then given by $\hat{x}(t) = \sum_n d[n]\beta^{n_w}(t-n)$. Under this model, we can compute the exact value of the reconstruction error in the continuous-time domain, as it only involves the computation of B-spline inner products.

In Fig. 9 we plot the normalized error $\|x - \hat{x}\| / \|x\|$ using the different methods (direct (*i.e., $H = I$*), consistent, regret, and the optimal (but usually unattainable) least-squares reconstruction $P_{\mathcal{W}}x$).
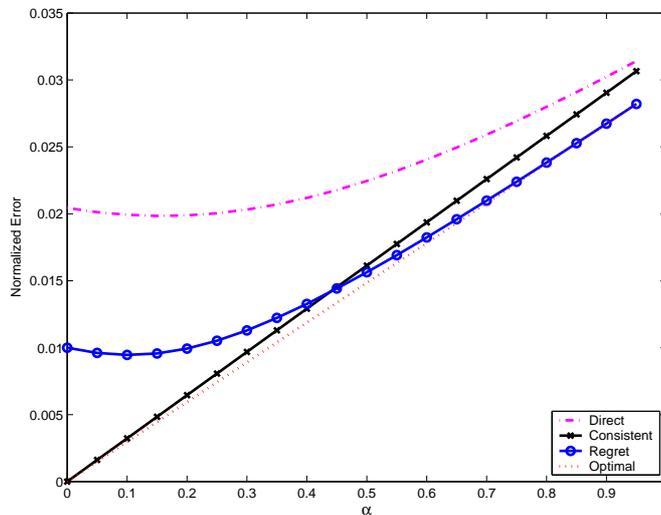
Figure 9: Evaluation of the normalized error $\|x - \hat{x}\| / \|x\|$ for the different reconstruction methods. The relative energy of the input signal within the sampling space (spline of order zero) and the reconstruction space (spline of order 2) is determined by the parameter α.

As can be seen from the figure, when α is small, most of the energy of the signal lies in the reconstruction space and the consistent reconstruction outperforms the regret method. On the other hand, when α is close to one, most of the signal energy is in the sampling space and the regret reconstruction is superior. This example also demonstrates that in the current setup, the direct reconstruction leads to the poorest results. Finally, we note that this simulation suggests a way to obtain bounds on the constants γ_1 and γ_2 of (70) and (73), respectively, when direct calculations of these quantities is difficult. Specifically, since at $\alpha = 0.45$ (there $\|P_{\mathcal{W}^\perp}x\|^2$ was calculated to be $\|P_{\mathcal{W}^\perp}x\|^2 \approx 0.75 \|P_{\mathcal{S}^\perp}x\|^2$) the regret scheme outperforms the consistent reconstruction, we conclude that for spline subspaces with $n_s = 0, n_w = 2$, γ_2 must be smaller than 0.75. Using similar arguments, γ_1 must be larger than 0.5.

9 Conclusion

In this paper, we treated the problem of sampling and reconstruction in general vector spaces, using the squared-norm error as the performance measure. If the input signal x lies in an appropriate subspace $\mathcal{A}$ of $\mathcal{H}$, then we showed that a linear reconstruction can be obtained that minimizes the squared-norm error. However, if x is an arbitrary input signal, then the squared-error cannot be minimized. Instead, we proposed a minimax regret approach that minimizes the worst-case difference between the squared-error and the smallest possible error. We showed that the resulting reconstruction can be interpreted geometrically in terms of the orthogonal projections onto the sampling space $\mathcal{S}$ and the reconstruction space $\mathcal{W}$. We also considered efficient implementations of the proposed schemes in the case of spline subspaces.

Finally, we compared the performance of the minimax regret reconstruction with that of the previously proposed consistent reconstruction, and demonstrated both analytically and through simulation that the minimax regret method can often outperform the consistent reconstruction strategy. We then identified the regions in which each of the approaches should be used.

References

[1] M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices," *IEEE Trans. Signal Processing*, vol. 42, no. 11, pp. 2915–2925, Nov. 1994.

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

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

[4] M. Unser, "Sampling—50 years after Shannon," *IEEE Proc.*, vol. 88, pp. 569–587, Apr. 2000.

[5] Y. C. Eldar, "Sampling and reconstruction in arbitrary spaces and oblique dual frame vectors," *J. Fourier Analys. Appl.*, vol. 1, no. 9, pp. 77–96, Jan. 2003.

[6] P. P. Vaidyanathan, "Generalizations of the sampling theorem: Seven decades after Nyquist," *IEEE Trans. Circuit Syst. I*, vol. 48, no. 9, pp. 1094–1109, Sep. 2001.

[7] Y. C. Eldar, "Sampling without input constraints: Consistent reconstruction in arbitrary spaces," in *Sampling, Wavelets and Tomography*, A. I. Zayed and J. J. Benedetto, Eds., pp. 33–60. Boston, MA: Birkhauser, 2004.

[8] I. Djokovic and P. P. Vaidyanathan, "Generalized sampling theorems in multiresolution subspaces," *IEEE Trans. Signal Processing*, vol. 45, pp. 583–599, Mar. 1997.

[9] A. Aldroubi, "Non-uniform weighted average sampling and exact reconstruction in shift-invariant and wavelet spaces," *Appl. Comp. Harmonic. Anal.*, vol. 13, pp. 151–161, 2002.

[10] T. Werther and Y. C. Eldar, "General framework for consistent sampling in Hilbert spaces," to appear in *International Journal of Wavelets, Multiresolution and Information Processing*.

[11] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Linear minimax regret estimation of deterministic parameters with bounded data uncertainties," *IEEE Trans. Signal Processing*, vol. 52, pp. 2177–2188, Aug. 2004.

[12] Y. C. Eldar and N. Merhav, "A competitive minimax approach to robust estimation of random parameters," *IEEE Trans. Signal Processing*, vol. 52, pp. 1931–1946, July 2004.

[13] M. Unser, A. Aldroubi, and M. Eden, "B-Spline signal processing: Part I - Theory," *IEEE Trans. Signal Processing*, vol. 41, no. 2, pp. 821–833, Feb 1993.

[14] M. Unser, A. Aldroubi, and M. Eden, "B-Spline signal processing: Part II - Efficient design and applications," *IEEE Trans. Signal Processing*, vol. 41, no. 2, pp. 834–848, Feb 1993.

[15] A. Muñoz, T. Blu, and M. Unser, "Least-squares image resizing using finite differences," *IEEE Trans. Image Processing*, vol. 10, no. 9, pp. 1365–1378, Sep 2001.

[16] A. Ben-Israel and T. N.E.Greville, *Generalized Inverses: Theory and Applications*, John Wiley and Sons, Inc., 1974.

[17] S. Kayalar and H. L. Weinert, "Oblique projections: Formulas, algorithms, and error bounds," *Math. Contr. Signals Syst.*, vol. 2, no. 1, pp. 33–45, 1989.

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

[19] R. M. Young, *An Introduction to Nonharmonic Fourier Series*, New York: Academic Press, 1980.

[20] Y. C. Eldar and O. Christansen, "Characterization of frame duals," to appear in J. Applied Signal Processing.

[21] W. S. Tang, "Oblique projections, biorthogonal Riesz bases and multiwavelets in Hilbert space," *Proc. Amer. Math. Soc.*, vol. 128, no. 2, pp. 463–473, 2000.

[22] A. Aldroubi and M. Unser, "Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory," vol. 15, no. 1-2, pp. 1–21, February 1994.

[23] A. Aldroubi and K. Grochenig, "Nonuniform sampling and reconstruction in shift invariant spaces," *SIAM Review*, vol. 43, pp. 585–620, 2001.

[24] O. Christensen, *An Introduction to Frames and Riesz Bases*, Boston, MA: Birkhauser, 2002.

[25] O. Christansen and Y. C. Eldar, "Oblique dual frames and shift-invariant spaces," *Appl. Comp. Harm. Anal.*, vol. 17, no. 1, 2004.

[26] National Institute of Standards and Technology, "The DARPA TIMIT acoustic-phonetic continuous speech corpus," CD-ROM NIST Speech Disc 1-1.1, Oct. 1991.