Consistent Sampling and Signal Recovery

Akira Hirabayashi, *Member, IEEE*, and Michael Unser, *Fellow, IEEE*

***Abstract*—An attractive formulation of the sampling problem is based on the principle of a consistent signal reconstruction. The requirement is that the reconstructed signal is indistinguishable from the input in the sense that it yields the exact same measurements. Such a system can be interpreted as an oblique projection onto a given reconstruction space. The standard formulation requires a one-to-one relationship between the input measurements and the reconstructed model. Unfortunately, this condition fails when the cross-correlation matrix between the analysis and reconstruction basis functions is not invertible; in particular, when there are less measurements than the number of reconstruction functions. In this paper, we propose an extension of consistent sampling that is applicable to those singular cases as well, and that yields a unique and well-defined solution. This solution also makes use of projection operators and has a geometric interpretation. The key idea is to exclude the null space of the sampling operator from the reconstruction space and to enforce consistency on its complement. We specify a class of consistent reconstruction algorithms corresponding to different choices of complementary reconstruction spaces. The formulation includes the Moore-Penrose generalized inverse, as well as other potentially more interesting reconstructions that preserve certain preferential signals. In particular, we display solutions that preserve polynomials or sinusoids, and therefore perform well in practical applications.**

***Index Terms*—Consistency, preferential components, sampling method, signal reconstruction, underdetermined scenario.**

I. Introduction

A sampling theorem usually refers to a mathematical formula or a process for reconstructing a continuously defined signal from its discrete measurements. The classical example is Shannon's sampling theorem which has been extended in a variety of ways (see, e.g. [1]–[4]). Sampling theorems typically guarantee a *perfect reconstruction* under suitable mathematical conditions which often represent an idealized situation. For example, Shannon's sampling theorem assumes that the underlying signal is bandlimited and requires an infinite number of samples. Real world signals or images, however, are never exactly bandlimited and the number of sampled values available is usually finite.

A more recent trend has been to approach sampling from the perspective of *approximation* [5]–[10]. Here, the goal is no longer to obtain an exact reconstruction, but rather a solution that is as close as possible to the original signal according to some criterion. This gives rise to a more realistic formulation of the sampling problem because there are no restrictions of bandlimitation on the class of input signals. In this work, we will assume that the measurements (generalized samples) are obtained by taking a series of inner products between the input f and a set of analysis functions $\{\psi_n\}_{n=1,\cdots,N}$. Moreover, we will concentrate on linear algorithms that reconstruct the signal as a linear combination of reconstruction functions $\{\varphi_k\}_{k=1,\cdots,K}$. An attractive solution for determining the optimal signal coefficients is to attempt to minimize the squared norm of the difference between a reconstructed signal and the target function [5]–[7]. It turns out that this can be achieved only when the reconstruction functions are included in the subspace spanned by the sampling functions. In the sequel, we will refer to the subspaces spanned by the two sets of basis functions as the *sampling space* and the *reconstruction space*, respectively.

When both the sampling functions and the reconstruction functions are given *a priori*, the sampling and reconstruction spaces are usually distinct, and the minimum error reconstruction can, in general, no longer be recovered from the measurements. Unser and Aldroubi investigated such a case, and proposed the use of a consistency criterion [8], [9]. This means that, if the reconstruction is measured using the same sampling functions, then the measurements will be equal to the original measurements. The sampling theorem in [8] reconstructs signals in an infinite dimensional subspace that is T-shift-invariant. The authors also provide a bound that guarantees that the consistent reconstruction is a good approximation of the minimum error solution, which is typically not accessible from the measurements alone [11]. An extended framework for general—but finite dimensional—subspace reconstruction (not necessarily shift-invariant) is discussed in [10].

The above formulations all require a one-to-one relation between the measurements and the reconstructed function. However, one can easily conceive of situations where this condition is not satisfied. This happens, for instance, when the dimensions of the sampling and reconstruction spaces are not matched. The case $N > K$ where the number of measurements (N) is greater than the number of degrees of freedom in the reconstruction (K) can be dealt with easily using least squares techniques. Of greater interest to us are the situations where the sampling problem is ill-posed. This happens, for example, when there is missing data: $N < K$. There are also instances with $N = K$ for which consistent signal recovery is not feasible. A well-known example is the problem of signal interpolation using polynomial splines of even degree with knots at the integers, which gives rise to a singular system matrix [12]. This latter example is meant to show that we are not completely free to select the reconstruction space independently of the analysis functions.

In this paper, we propose a formulation that can also deal with the cases where the consistent sampling problem is ill-posed.

Manuscript received February 28, 2006; revised November 5, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Karim Drouiche. This work was supported in part by a Grant from the Venture Business Laboratory, Yamaguchi University, Yamaguchi, Japan.

A. Hirabayashi is with the Faculty of Engineering, Yamaguchi University, Ube, 755-8611 Japan (e-mail: a-hira@yamaguchi-u.ac.jp).

M. Unser is with the Biomedical Imaging Group, STI-IOA, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland (e-mail: michael.unser@epfl.ch).

Digital Object Identifier 10.1109/TSP.2007.895996

1053-587X/$25.00 © 2007 IEEE

The key idea is to essentially exclude the null space of the sampling operator from the reconstruction space, and to enforce the consistency constraint on a reduced subspace. An important aspect is that the choice of the complement of the null space is not unique, a fact that can be used to our advantage for introducing special reconstruction constraints. For instance, we will show that enforcing the preservation of certain preferential signals (such as polynomials or sinusoids) can facilitate signal recovery and improve reconstruction.

This paper is organized as follows. In Section II, we formulate the sampling problem within a Hilbert space framework, and recall the principle of a consistency signal recovery. We also display some examples where the one-to-one condition fails, implying that the standard approach is not applicable. In Section III, we clarify the meaning of the one-to-one condition from the viewpoints of over-sampling and under-sampling. In Section IV, we propose an extended sampling theorem that yields a well-defined consistent solution in all cases ($N \leq K$), including the singular ones. The key point is that this sampling theorem depends on the specification on an appropriate reconstruction subspace L, the choice of which is not necessarily unique. In Section V, we look at special strategies for specifying L and propose a reconstruction process that preserves certain preferential signals. We also present explicit computational formulas. In Section VI, we discuss the connection with traditional consistency sampling and clarify the type of extension that is provided here. One special case corresponds to the generalized inverse version of the classical algorithm. However, the present formulation also covers alternative reconstruction strategies (e.g., preservation of polynomials) that are less standard and more interesting from an approximation theoretic point of view. This latter point is demonstrated in Section VII where we present some examples of signal recovery and compare the two types of approaches.

A. Notations and Mathematical Preliminaries

We will make use of the following notations. The measurements of a signal are represented as a vector in the N-dimensional unitary space $\mathbf{C}^N$. The reconstructed signal, on the other hand, will be parameterized by a vector in $\mathbf{C}^K$ with $K \geq N$. The standard bases for $\mathbf{C}^N$ and $\mathbf{C}^K$, are $\{\boldsymbol{e}_n^{(N)}\}_{n=1,\cdots,N}$ and $\{\boldsymbol{e}_k^{(K)}\}_{k=1,\cdots,K}$, respectively. That is, $\boldsymbol{e}_n^{(N)}$ (respectively, $\boldsymbol{e}_k^{(K)}$) is the N-dimensional (respectively, K-dimensional) vector consisting of zero elements except for the nth (respectively, kth) element which is equal to 1. I is the identity operator.

The orthogonal complement of a subspace S is denoted by $S^{\perp}$. $\mathcal{R}(A)$ and $\mathcal{N}(A)$ stand for the range and the null space of an operator A, respectively. A^* is the adjoint operator of A.

Let $A^{\dagger}$ and P_S denote the Moore-Penrose generalized inverse of an operator A and the orthogonal projection operator onto S, respectively. It holds that

$$AA^{\dagger} = P_{\mathcal{R}(A)}, \quad A^{\dagger}A = P_{\mathcal{R}(A^*)} \tag{1}$$

$$A^{\dagger} = A^*(AA^*)^{\dagger} = (A^*A)^{\dagger}A^* \tag{2}$$

$$P_S(AP_S)^{\dagger} = (AP_S)^{\dagger}, \text{ and } (P_SA)^{\dagger}P_S = (P_SA)^{\dagger}. \tag{3}$$

The following lemma will be used to prove our main theorem.

Lemma 1: [13] The operator equation

$$AXB = C$$

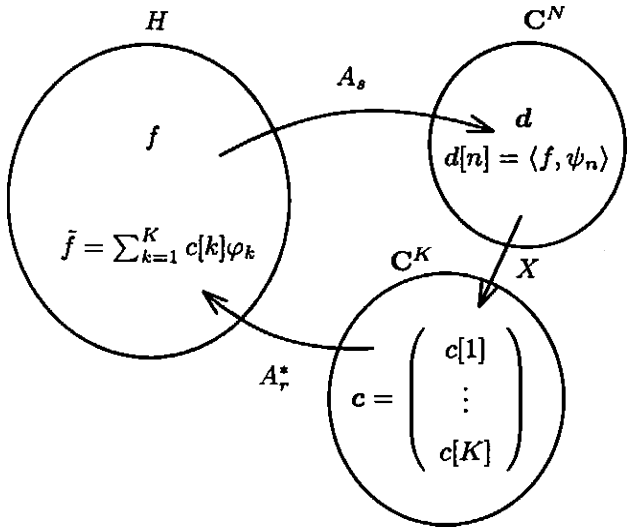


Fig. 1. Sampling and reconstruction formulation.

has a solution X if and only if

$$\mathcal{R}(A) \supset \mathcal{R}(C)$$

and

$$\mathcal{N}(B) \subset \mathcal{N}(C).$$

In this case, the general solution X is

$$X = A^{\dagger}CB^{\dagger} + Y - A^{\dagger}AYBB^{\dagger}$$

where Y is an arbitrary operator.

Let α and β be elements of two Hilbert spaces H_1 and H_2, respectively. Let $(\alpha \otimes \overline{\beta})$ be an operator from H_2 to H_1 defined by

$$(\alpha \otimes \overline{\beta})\gamma = \langle \gamma, \beta \rangle \alpha \quad \text{for any} \quad \gamma \in H_2 \tag{4}$$

where $\langle \cdot, \cdot \rangle$ is the inner product in H_2. This operator is called the Neumann-Schatten product [14] and it satisfies the relation

$$(\alpha \otimes \overline{\beta})^* = \beta \otimes \overline{\alpha}. \tag{5}$$

II. Consistency

A. Formulation of the Sampling Problem

We start with the formulation of the sampling problem, which is illustrated in Fig. 1. The original input signal f is defined over a continuous domain $\mathcal{D}$ and is assumed to belong to a Hilbert space $H = H(\mathcal{D})$. The measurements of f, denoted by $d[n]$ $(n = 1, 2, \ldots, N)$, are given by the inner product in H of f with the sampling functions $\{\psi_n\}_{n=1,\cdots,N}$

$$d[n] = \langle f, \psi_n \rangle.$$

The N-dimensional vector consisting of $d[n]$ is denoted by $\boldsymbol{d}$. Let A_s be the operator that maps f into $\boldsymbol{d}$

$$A_s f = \boldsymbol{d}. \tag{6}$$

By using the Neumann-Schatten product, the operator A_s is expressed without f as

$$A_s = \sum_{n=1}^{N} \boldsymbol{e}_n^{(N)} \otimes \overline{\psi_n}. \tag{7}$$

The reconstructed signal $\tilde{f} \in H$ is given by a linear combination of reconstruction functions $\{\varphi_k\}_{k=1,\cdots,K}$

$$\tilde{f} = \sum_{k=1}^{K} c[k]\varphi_k. \tag{8}$$

The K-dimensional vector of signal coefficients $c[k]$ is denoted by $\boldsymbol{c}$. We now introduce the (adjoint) reconstruction operator

$$A_r = \sum_{k=1}^{K} \boldsymbol{e}_k^{(K)} \otimes \varphi_k. \tag{9}$$

It follows from (8), (4), (5), and (9) that

$$\tilde{f} = A_r^* \boldsymbol{c}. \tag{10}$$

Let X be the $K \times N$ matrix that maps $\boldsymbol{d}$ to $\boldsymbol{c}$

$$X\boldsymbol{d} = \boldsymbol{c}. \tag{11}$$

Then, (10), (11), and (6) yield

$$\tilde{f} = A_r^* X A_s f. \tag{12}$$

With this formulation, the sampling problem becomes equivalent to finding a suitable matrix X so that $\tilde{f}$ satisfies some optimality criterion.

Let V_s and V_r be subspaces in H spanned by $\{\psi_n\}_{n=1,\cdots,N}$ and $\{\varphi_k\}_{k=1,\cdots,K}$, respectively. They are called the *sampling space* and the *reconstruction space*, respectively. They play important roles in this paper. It holds that

$$V_s = \mathcal{R}(A_s^*)$$
$$V_r = \mathcal{R}(A_r^*).$$

There are no particular assumptions on the functions $\{\psi_n\}_{n=1,\cdots,N}$ and $\{\varphi_k\}_{k=1,\cdots,K}$. However, it makes good sense to select the functions in each set to be linearly independent so that they constitute bases for V_s and V_r, respectively.

B. Conventional Consistency Sampling Theorems

Consistency is the concept proposed in [8]. The requirement is that the reconstructed signal $\tilde{f}$ should yield the same measurements as the underlying original signal f if it was reinjected into the system. By using the formulation mentioned above, consistency is expressed as

$$A_s\tilde{f} = A_s f.$$

The consistency sampling theorems that have been discussed so far in the literature [8]–[10] all make the assumption that

$$V_r \cap V_s^{\perp} = \{0\}. \tag{13}$$

This condition guarantees a one-to-one relation between the measurements $\{d[n]\}_{n=1,\cdots,N}$ and the reconstructed function $\tilde{f}(x)$. It ensures both the existence and the uniqueness of the consistent reconstruction. In that case, it has been shown that the consistent reconstruction of any f in H is given by its oblique projection onto V_r along $V_s^{\perp}$.

However, one can easily think of situations where (13) does not hold. The most obvious case is when the number of sampling functions is less than that of the reconstruction functions. This corresponds to the missing data scenario mentioned in the introduction. The standard setting $N = K$ is not completely safe either, as illustrated by the following example.

Example 1: Let $\beta^0(x)$ and $\beta^1(x)$ be the B-splines of degree 0 and 1 defined by

$$\beta^0(x) = \begin{cases} 1 & (0 \le x < 1) \\ 0 & (x < 0, x \ge 1), \end{cases}$$

and

$$\beta^1(x) = (\beta^0 * \beta^0)(x)$$

respectively, where $*$ is the convolution operator. The domain $\mathcal{D}$ of the target signal is $[0, N]$. Let ψ_n and φ_k be

$$\psi_n(x) = \beta^0(x - n + 1)$$
$$\varphi_k(x) = \beta_K^1(x - k + 1)$$

respectively, where

$$\beta_K^1(x) = \sum_{k \in \mathbb{Z}} \beta^1(x - kK)$$

This corresponds to the periodized version of a system where the sampling is performed by integrating the signal over the sampling period ($T = 1$) and where the reconstruction is performed using piecewise linear splines.

Now, in the case where $N = K$ is even, one can verify that $\tilde{f}$ in (10) with

$$\boldsymbol{c} = (1, -1, \ldots, 1, -1) \tag{14}$$

belongs to $V_s^{\perp}$, which contradicts hypothesis (13).

This means that there are situations, such as the example here, where the consistency sampling theorems that have been proposed in the literature are not directly applicable (concretely, this means that the underlying system matrix is not invertible in the conventional sense). This motivates us to investigate viable methods for achieving consistency in the underdetermined scenario that is specified by

$$V_r \cap V_s^{\perp} \neq \{0\} \tag{15}$$
$$V_r + V_s^{\perp} = H. \tag{16}$$

Note that (16) may hold if the dimension of V_s is lesser or equal to that of V_r but not otherwise; this excludes all overdetermined scenarios where there are more independent measurements than degrees of freedom in the reconstruction. Let us also emphasize that the symbol $+$ in the left-hand side (LHS) of (16) is not the direct sum because of (15).

III. Over-Sampling and Under-Sampling

Before proceeding to the main result, we clarify the meaning of (15) from the viewpoints of over-sampling and under-sampling. Usually, these concepts are defined with reference to Shannon's sampling theorem for band-limited signals. In that context, over-sampling means sampling at a rate that is above the critical Nyquist rate, while under-sampling means sampling at a lesser rate. It is obviously possible to reconstruct band-limited signals perfectly in the former case but generally not in the latter.

With the formulation in Section II-A, these concepts are translated as follows. First, the band-limited property is generalized to the fact that f belongs to V_r. Second, the over-sampling scenario corresponds to the case where the sampled measurements $\{d[n]\}_{n=1,\cdots,N}$ are linearly dependent for any f in V_r. Third, the perfect reconstruction property means that there exists X that satisfies

$$A_r^* X A_s f = f \tag{17}$$

for any f in V_r. We can also distinguish between different sampling scenarios.

Definition 1: With the formulation in Section II-A, if there exists some X that satisfies (17) for any f in V_r, then we have an oversampling (respectively, normal-sampling) scenario over V_r depending on whether the sampled measurements $\{d[n]\}_{n=1,\cdots,N}$ are linearly dependent for any f in V_r or not. If, on the other hand, there is no X that satisfies (17) for any f in V_r, then we have an undersampling scenario.

The next proposition reemphasizes the crucial role of (13) in the standard formulations of the sampling theorem.

Proposition 1: There exists an operator X that satisfies (17) for any f in V_r, if and only if (13) holds.

Proof: Assume that some X satisfies (17) for any f in V_r. Then, it holds for any f in $V_r \cap V_s^\perp$ that

$$f = A_r^* X A_s f = 0$$

because $A_s f = 0$. Hence, (13) holds.

Conversely, assume that (13) holds. Then, Proposition 1 in [10] says that $A_r^*(A_s A_r^*)^\dagger A_s$ is the oblique projection operator onto V_r along $V_s^\perp$. Hence, by letting $X = (A_s A_r^*)^\dagger$, it holds that

$$A_r^* X A_s f = A_r^* \left(A_s A_r^*\right)^\dagger A_s f = f$$

which implies (17). ∎

Here, we will depart from the traditional normal and over-sampling scenarios covered by Proposition 1 and concentrate on the under-sampled case instead.

IV. Generalized Consistency Sampling Theorem

We now discuss the consistency sampling problem under (15) and (16). The first equation implies that a consistent reconstruction is generally not unique. Indeed, let f_1 be a consistent reconstruction, and g be a non-zero element of $V_r \cap V_s^\perp$. Then, $f_1 + g$ is also a consistent reconstruction that is obviously different from f_1.

In order to recover uniqueness, we consider some complementary subspace of $V_r \cap V_s^\perp$ in V_r, denoted by L, which plays a central role in our formulation. This subspace is such that

$$L \oplus \left(V_r \cap V_s^\perp\right) = V_r \tag{18}$$

where $\oplus$ denotes the direct sum, and

$$L \cap V_s^\perp = \{0\}. \tag{19}$$

Proposition 2: Let L be a given complementary subspace of $V_r \cap V_s^\perp$ in V_r. Then, any $f \in H$ has a corresponding consistent reconstruction in L that is uniquely defined.

Proof: Let both f_1 and f_2 be consistent reconstructions in L. Then, $A_s f_1 = A_s f_2$. Hence, it follows that

$$A_s(f_1 - f_2) = 0$$

which implies $f_1 - f_2$ belongs to $\mathcal{N}(A_s) = V_s^\perp$. Since $f_1 - f_2$ also belongs to L, it follows from (19) that

$$f_1 - f_2 \in L \cap V_s^\perp = \{0\}$$

which implies $f_1 = f_2$. ∎

The uniqueness of the consistent reconstruction is guaranteed by forcing the solution to lie in the complementary subspace L. We can characterize the reconstruction by an oblique projection as follows. Equations (16) and (18) yield

$$L \oplus V_s^\perp = H.$$

That is, H is decomposed into the direct sum of L and $V_s^\perp$. Then, we can define the oblique projection operator onto L along $V_s^\perp$ denoted by $P_{L,V_s^\perp}$.

Theorem 1: Let L be a fixed complementary subspace of $V_r \cap V_s^\perp$ in V_r. For any f in H, the function $\tilde{f}$ in (12) is a unique consistent reconstruction in L if and only if $\tilde{f}$ is the oblique projection of f onto L along $V_s^\perp$.

Proof: Let $U = A_r^* X A_s$. Assume that $\tilde{f}$ is a unique consistent reconstruction for any f in H. Then, since $\tilde{f} = A_r^* X A_s f = Uf$, it holds that

$$A_s U = A_s. \tag{20}$$

Hence, we have

$$U^2 = A_r^* X A_s U = A_r^* X A_s = U$$

which implies

$$U^2 = U. \tag{21}$$

Equation (20) also implies that

$$\mathcal{N}(U) \subset \mathcal{N}(A_s U) = \mathcal{N}(A_s) \subset \mathcal{N}\left(A_r^* X A_s\right) = \mathcal{N}(U)$$

which means

$$\mathcal{N}(U) = \mathcal{N}(A_s) = V_s^\perp. \tag{22}$$

Since $\tilde{f}$ for any f in H lies in L, we have

$$\mathcal{R}(U) = L. \tag{23}$$

Equations (21), (22), and (23) imply that $\tilde{f}$ is the oblique projection of f onto L along $V_s^{\perp}$.

Conversely, assume that $\tilde{f} = A_r^* X A_s f$ for any f in H is the oblique projection of f onto L along $V_s^{\perp}$. Then, (21), (22), and (23) hold. Since (21) yields $U(I-U)=0$, it follows from (22) that

$$\mathcal{R}(I-U) \subset \mathcal{N}(U) = \mathcal{N}(A_s)$$

which implies

$$A_s(I-U) = 0.$$

Hence, we have (20), which implies that $\tilde{f}$ is a consistent reconstruction. ■

Theorem 1 means that the unique consistent reconstruction in L is obtained by reconstructing the oblique projection:

$$\tilde{f} = P_{L,V_s^{\perp}} f. \tag{24}$$

This leads to a consistency sampling theorem for the case of (15).

Theorem 2: Let L be a fixed complementary subspace of $V_r \cap V_s^{\perp}$ in V_r. The unique consistent reconstruction in L is obtained by

$$X = (A_r^*)^{\dagger} P_{L,V_s^{\perp}} A_s^{\dagger} + Y - A_r A_r^{\dagger} Y A_s A_s^{\dagger} \tag{25}$$

where Y is an arbitrary linear operator (i.e., rectangular matrix) from $\mathbf{C}^N$ to $\mathbf{C}^K$.

Proof: Because of Theorem 1, we prove that the oblique projection in (24) is obtained by (12) if and only if X is given by (25). Assume that $\tilde{f}$ is the oblique projection in (24) for any f in H. Then, it follows from (12) that

$$A_r^* X A_s = P_{L,V_s^{\perp}}. \tag{26}$$

Since L is a subspace of V_r, it holds that

$$\mathcal{R}(A_r^*) = V_r \supset L = \mathcal{R}\left(P_{L,V_s^{\perp}}\right).$$

Further, we have that

$$\mathcal{N}(A_s) = V_s^{\perp} = \mathcal{N}\left(P_{L,V_s^{\perp}}\right).$$

These two equations and Lemma 1 imply that (26) has a solution, and that its general form is given by (25) since $(A_r^*)^{\dagger} A_r^* = A_r A_r^{\dagger}$.

Conversely, assume that X is given by (25). Since $\mathcal{R}(A_r^*) \supset L = \mathcal{R}(P_{L,V_s^{\perp}})$, it holds that

$$A_r^* (A_r^*)^{\dagger} P_{L,V_s^{\perp}} = P_{L,V_s^{\perp}}.$$

Further, since $\mathcal{N}(P_{L,V_s^{\perp}}) = V_s^{\perp} = \mathcal{N}(A_s)$, we have that

$$P_{L,V_s^{\perp}} A_s^{\dagger} A_s = P_{L,V_s^{\perp}}.$$

Hence, it follows from (12) and (25) that

$$\tilde{f} = A_r^* X A_s f = A_r^* (A_r^*)^{\dagger} P_{L,V_s^{\perp}} A_s^{\dagger} A_s f = P_{L,V_s^{\perp}} f$$

which implies (24). ■

If $V_r \cap V_s^{\perp} = \{0\}$, the only possible choice of subspace L is V_r itself, and the oblique projection in (24) yields that onto V_r along $V_s^{\perp}$. Hence, Theorem 2 reduces to the conventional consistency sampling theorems in [8] and [10].

Even though the solution specified by (24) is unique, there is a whole family of equivalent operators X in (25) that depend on the choice of Y. The canonical choice is $Y = 0$ but there may also be alternative ones that are more attractive computationally.

V. Sampling Theorem With Preservation of Preferential Signals

In Theorem 2, we are free to choose the subspace L as long as it is a complementary subspace of $V_r \cap V_s^{\perp}$ in V_r. From the viewpoint of consistency, there is no difference among all possible choices of L. However, $\tilde{f}$ itself changes depending on L. This suggests that there may be ways of selecting L that are better than others. Here, we propose a selection strategy that favors a particular subclass of signals.

Even though we have enough degrees of freedom to reconstruct all signals in V_r, the implication of (15) is that we do not have enough information (measurements) to distinguish between all of them. Suppose now that we have identified a subclass of signals of *special* interest in V_r that do not belong to $V_r \cap V_s^{\perp}$. Then, we can at least make sure that these will be reconstructed perfectly by selecting a subspace L that contains them. A typical example is the constant that often accounts for a large portion of the signal energy; especially in image processing applications.

We now assume that we are given a collection, $\{\phi_i\}_{i=1}^{I}$, of such preferential signals in V_r; these span the subspace $V_i = \operatorname{span}\{\phi_i\}_{i=1}^{I}$ where the subscript i means "interest."

If the direct sum of $V_r \cap V_s^{\perp}$ and V_i is equal to V_r, that is, if

$$V_i \oplus \left(V_r \cap V_s^{\perp}\right) = V_r$$

then L is uniquely determined as V_i. Otherwise, we need a second condition. To this end, we impose that the remainder of $V_i \oplus (V_r \cap V_s^{\perp})$ is its orthogonal complement in V_r, which is denoted by V_c. The subscript c means "complement." This allows us to specify the reconstruction subspace L as

$$L = V_i \oplus V_c. \tag{27}$$

The application of Theorem 2 then yields to our next result.

Theorem 3: The unique consistent reconstruction in $L = V_i \oplus V_c$ of a signal $f \in H$ is obtained by (12) if and only if X is given by

$$X = (A_r^*)^{\dagger} (A_s W)^{\dagger} + Y - A_r A_r^{\dagger} Y A_s A_s^{\dagger} \tag{28}$$

where Y is arbitrary and where W is an operator defined by

$$W = P_{V_i} + P_{V_c}. \tag{29}$$

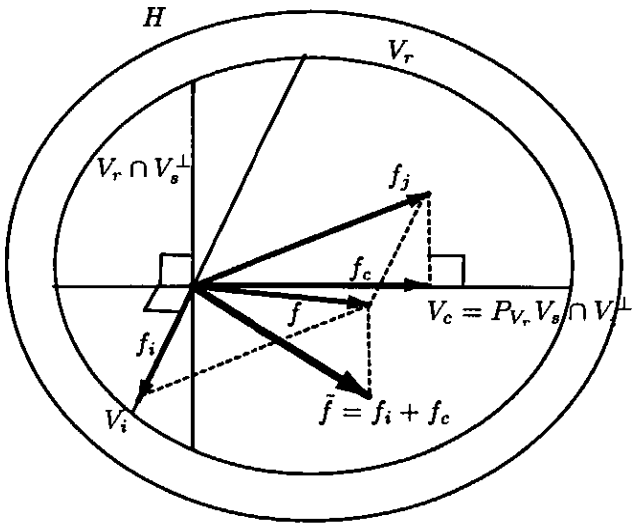


Fig. 2. Reconstruction $\tilde{N}$ by Theorem 3. Note that V_i and $V_r \cap V_s^\perp$ are not perpendicular in general.

Proof: Since V_c is perpendicular to V_i, the sum of the two orthogonal projection operators P_{V_i} and P_{V_c} is that onto L in (27). Then, Proposition 1 in [10] implies that the oblique projection operator $P_{L,V_s^\perp}$ onto L in (27) along $V_s^\perp$ is given by $W(A_sW)^\dagger A_s$. Hence, it follows from the first equation of (3) that

$$P_{L,V_s^\perp} = W(A_sW)^\dagger A_s = (A_sW)^\dagger A_s.$$

Moreover, we have that

$$\begin{aligned}(A_r^*)^\dagger P_{L,V_s^\perp} A_s^\dagger &= (A_r^*)^\dagger (A_sW)^\dagger A_s A_s^\dagger \\ &= (A_r^*)^\dagger \left(P_{\mathcal{R}(A_s)} A_sW\right)^\dagger P_{\mathcal{R}(A_s)} \\ &= (A_r^*)^\dagger \left(P_{\mathcal{R}(A_s)} A_sW\right)^\dagger \\ &= (A_r^*)^\dagger (A_sW)^\dagger,\end{aligned}$$

which implies that (25) takes the special form given by (28). ■

In order to verify our second condition, let us decompose f in V_r into two components:

$$f = f_i + f_j$$

where f_i and f_j lie in V_i and $\{V_c \oplus (V_r \cap V_s^\perp)\}$, respectively, as shown in Fig. 2. The first component f_i is perfectly reconstructed because f_i lies in V_i. On the other hand, the other component f_j is not perfectly reproduced in general because it does not belong to L. It is projected onto V_c and its projection is denoted by f_c. Then, the reconstruction for f is given by

$$\tilde{f} = f_i + f_c.$$

Note that f_c is the minimum error approximation of f_j in V_c because V_c is perpendicular to $V_r \cap V_s^\perp$. This guarantees that the proposed sampling theorem provides a good approximation of the input signal (i.e., better than for any other possible choice of V_c such that $L = V_i + V_c$).

The important term of X in (28) is the first one. It involves the factors A_sW and A_r, which are abstract operators from the function space H to the vector spaces $\mathbf{C}^N$ and $\mathbf{C}^K$, respectively, and which are not suitable for computer calculations.

In order to specify X numerically, we now provide its matrix expression. First, let G be the Gram matrix of $\{\varphi_k\}_{k=1,\cdots,K}$ given by

$$G = \begin{pmatrix} \langle\varphi_1,\varphi_1\rangle & \langle\varphi_2,\varphi_1\rangle & \cdots & \langle\varphi_K,\varphi_1\rangle \\ \langle\varphi_1,\varphi_2\rangle & \langle\varphi_2,\varphi_2\rangle & \cdots & \langle\varphi_K,\varphi_2\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\varphi_1,\varphi_K\rangle & \langle\varphi_2,\varphi_K\rangle & \cdots & \langle\varphi_K,\varphi_K\rangle \end{pmatrix}.$$

Next, we introduce the cross-correlation matrix between $\{\varphi_k\}_{k=1,\cdots,K}$ and $\{\psi_n\}_{n=1,\cdots,N}$:

$$B = \begin{pmatrix} \langle\varphi_1,\psi_1\rangle & \langle\varphi_2,\psi_1\rangle & \cdots & \langle\varphi_K,\psi_1\rangle \\ \langle\varphi_1,\psi_2\rangle & \langle\varphi_2,\psi_2\rangle & \cdots & \langle\varphi_K,\psi_2\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\varphi_1,\psi_N\rangle & \langle\varphi_2,\psi_N\rangle & \cdots & \langle\varphi_K,\psi_N\rangle \end{pmatrix}$$

and define three more matrices

$$C = \begin{pmatrix} \langle\phi_1,\varphi_1\rangle & \langle\phi_2,\varphi_1\rangle & \cdots & \langle\phi_I,\varphi_1\rangle \\ \langle\phi_1,\varphi_2\rangle & \langle\phi_2,\varphi_2\rangle & \cdots & \langle\phi_I,\varphi_2\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\phi_1,\varphi_K\rangle & \langle\phi_2,\varphi_K\rangle & \cdots & \langle\phi_I,\varphi_K\rangle \end{pmatrix}$$

$$D = \begin{pmatrix} \langle\phi_1,\psi_1\rangle & \langle\phi_2,\psi_1\rangle & \cdots & \langle\phi_I,\psi_1\rangle \\ \langle\phi_1,\psi_2\rangle & \langle\phi_2,\psi_2\rangle & \cdots & \langle\phi_I,\psi_2\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\phi_1,\psi_N\rangle & \langle\phi_2,\psi_N\rangle & \cdots & \langle\phi_I,\psi_N\rangle \end{pmatrix}$$

$$E = \begin{pmatrix} \langle\phi_1,\phi_1\rangle & \langle\phi_2,\phi_1\rangle & \cdots & \langle\phi_I,\phi_1\rangle \\ \langle\phi_1,\phi_2\rangle & \langle\phi_2,\phi_2\rangle & \cdots & \langle\phi_I,\phi_2\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\phi_1,\phi_I\rangle & \langle\phi_2,\phi_I\rangle & \cdots & \langle\phi_I,\phi_I\rangle \end{pmatrix}.$$

Theorem 4: The predominant term of X in (28) admits the following explicit matrix form

$$(A_r^*)^\dagger (A_sW)^\dagger = G^\dagger U V^\dagger \tag{30}$$

where

$$U = CE^\dagger D^* + GG^\dagger B^* \Phi^\dagger Z \tag{31}$$

$$V = DE^\dagger D^* + \Phi\Phi^\dagger Z \tag{32}$$

with G, B, C, D and E as defined above, and where Φ, Ψ, and Z are auxiliary matrices given by

$$\Phi = BG^\dagger B^* \tag{33}$$

$$\Psi = BG^\dagger C \tag{34}$$

$$Z = \Phi - \Psi(\Psi^*\Phi^\dagger\Psi)^\dagger \Psi^*\Phi^\dagger\Phi. \tag{35}$$

Proof: Let A_i be the operator defined by

$$A_i = \sum_{i=1}^{I} \boldsymbol{e}_i^{(I)} \otimes \overline{\phi_i}. \tag{36}$$

It follows from (7), (9), and (36) that

$$G = A_r A_r^* \tag{37}$$

$$B = A_s A_r^* \tag{38}$$

$$C = A_r A_i^* \tag{39}$$

$$D = A_s A_i^* \tag{40}$$
$$E = A_i A_i^*. \tag{41}$$

Equation (1) implies that the orthogonal projection operator P_{V_i} in the right-hand side (RHS) of (29) is given by $A_i^\dagger A_i$. Hence, it follows from (2) and (39)–(41) that

$$A_r P_{V_i} A_s^* = C E^\dagger D^* \tag{42}$$
$$A_s P_{V_i} A_s^* = D E^\dagger D^*. \tag{43}$$

The other projection operator P_{V_c} in (29) is expressed as follows. The subspace V_c is the intersection of the orthogonal complements of V_i and $V_r \cap V_s^\perp$. Since Theorem 4 in [15] implies that V_r can be decomposed into the orthogonal direct sum of

$$V_r = P_{V_r} V_s \oplus \left(V_r \cap V_s^\perp\right) \tag{44}$$

the orthogonal complement of $V_r \cap V_s^\perp$ in V_r is expressed as $P_{V_r} V_s$. Therefore, V_c is determined by

$$V_c = P_{V_r} V_s \cap V_i^\perp. \tag{45}$$

By a decomposition similar to (44), $P_{V_r} V_s$ can be expressed as the orthogonal direct sum of

$$P_{V_r} V_s = P_{P_{V_r} V_s} V_i \oplus \left(P_{V_r} V_s \cap V_i^\perp\right). \tag{46}$$

Equations (45) and (46) yield

$$P_{V_r} V_s = P_{P_{V_r} V_s} V_i \oplus V_c.$$

Hence, the orthogonal projection operator onto V_c is given by the difference of those onto $P_{V_r} V_s$ and $P_{P_{V_r} V_s} V_i$. That is

$$P_{V_c} = \left(P_{V_r} A_s^*\right)\left(P_{V_r} A_s^*\right)^\dagger - \left\{\left(P_{V_r} A_s^*\right)\left(P_{V_r} A_s^*\right)^\dagger A_i^*\right\} \times \left\{\left(P_{V_r} A_s^*\right)\left(P_{V_r} A_s^*\right)^\dagger A_i^*\right\}^\dagger. \tag{47}$$

It follows from (2), (37), (38), and (33) that

$$\left(P_{V_r} A_s^*\right)\left(P_{V_r} A_s^*\right)^\dagger = A_r^* G^\dagger B^* \Phi^\dagger B G^\dagger A_r. \tag{48}$$

Further, (48), (39), (34), (2), (37), and (33) imply that

$$\left\{\left(P_{V_r} A_s^*\right)\left(P_{V_r} A_s^*\right)^\dagger A_i^*\right\}\left\{\left(P_{V_r} A_s^*\right)\left(P_{V_r} A_s^*\right)^\dagger A_i^*\right\}^\dagger = A_r^* G^\dagger B^* \Phi^\dagger \Psi (\Psi^* \Phi^\dagger \Psi)^\dagger \Psi^* \Phi^\dagger B G^\dagger A_r. \tag{49}$$

Hence, from (47)–(49), (37), (38), and (33), we have

$$A_r P_{V_c} A_s^* = G G^\dagger B^* \Phi^\dagger Z \tag{50}$$
$$A_s P_{V_c} A_s^* = \Phi \Phi^\dagger Z. \tag{51}$$

Then, (29), (42), (50), and (31) yield

$$A_r W A_s^* = C E^\dagger D^* + G G^\dagger B^* \Phi^\dagger Z = U. \tag{52}$$

Further, (29), (43), (51), and (32) give

$$A_s W A_s^* = D E^\dagger D^* + \Phi \Phi^\dagger Z = V. \tag{53}$$

Finally, it follows from (2), (52), and (53) that

$$\left(A_r^*\right)^\dagger \left(A_s W\right)^\dagger = \left(A_r A_r^*\right)^\dagger A_r^* W A_s^* \left(A_s W A_s^*\right)^\dagger = G^\dagger U V^\dagger$$

which implies (30) ■

When the reconstruction basis is orthogonal, the matrix G reduces to an identity. In that case, the formulas in Theorem 4 get simplified as follows.

Corollary 1: If G is the identity matrix, then Z, U, and V in (35), (31), and (32) are expressed as

$$Z = BB^* - BC(C^* B^\dagger BC)^\dagger C^* B^*$$
$$U = C E^\dagger D^* + B^\dagger Z$$
$$V = D E^\dagger D^* + Z$$

respectively, and the first term of X in (29) is given by

$$\left(A_r^*\right)^\dagger \left(A_s W\right)^\dagger = U V^\dagger.$$

The proof is straightforward.

VI. Moore-Penrose Generalized Inverse Reconstruction

The consistent sampling theorems that have been described in the literature all require (13) that guarantees that the solution is unique. The formulation that is the closest to ours is that of Eldar [10]. In our notation, her main result reads.

Proposition 3: [10] Let $V_r \cap V_s^\perp = \{0\}$. Then, any $f \in H$ can be consistently reconstructed from the measurements $\boldsymbol{d}$ by (8) if X in (11) is given by

$$X = \left(A_s A_r^*\right)^\dagger. \tag{54}$$

This reconstruction is uniquely specified from the measurements and corresponds to the oblique projection of f onto V_r along the direction specified by $V_s^\perp$.

Interestingly, Eldar did choose to specify her solution using a generalized inverse operator even though this precaution was not really necessary in the above context. Since the Moore-Penrose generalized inverse in (54) is always well defined, it is possible to apply her reconstruction in the singular case as well, with the caveat that one uses unicity for the signals belonging to V_r. We will now derive this generalized inverse reconstruction algorithm as a special case of our formulation. This will also clarify the behavior of this type of operator for the case of (15).

Theorem 5: The unique consistent reconstruction in $L = \mathcal{R}(A_r^* A_r A_s^*)$ is obtained by (12) if and only if X is given by

$$X = \left(A_s A_r^*\right)^\dagger + Y - A_r A_r^\dagger Y A_s A_s^\dagger$$

where Y is arbitrary.

In order to prove Theorem 5, we use the following two lemmas. The proof of these is deferred to the Appendix.

Lemma 2: The function space H can be decomposed into the direct sum of

$$H = \mathcal{R}\left(A_r^* A_r A_s^*\right) \oplus V_s^{\perp}. \tag{55}$$

Lemma 3: The operator $A_r^*(A_s A_r^*)^{\dagger} A_s$ is the oblique projection operator onto $\mathcal{R}(A_r^* A_r A_s^*)$ along $V_s^{\perp}$.

(Proof of Theorem 5): Lemma 3 implies that $P_{L,V_s^{\perp}}$ for $L = \mathcal{R}(A_r^* A_r A_s^*)$ is given by

$$P_{L,V_s^{\perp}} = A_r^* \left(A_s A_r^*\right)^{\dagger} A_s.$$

Hence, it follows from (1) and (3) that

$$\begin{aligned}
\left(A_r^*\right)^{\dagger} P_{L,V_s^{\perp}} A_s^{\dagger} &= \left(A_r^*\right)^{\dagger} A_r^* \left(A_s A_r^*\right)^{\dagger} A_s A_s^{\dagger} \\
&= P_{\mathcal{R}(A_r)} \left(P_{\mathcal{R}(A_s)} A_s A_r^* P_{\mathcal{R}(A_r)}\right)^{\dagger} P_{\mathcal{R}(A_s)} \\
&= \left(P_{\mathcal{R}(A_s)} A_s A_r^* P_{\mathcal{R}(A_r)}\right)^{\dagger} \\
&= \left(A_s A_r^*\right)^{\dagger}
\end{aligned}$$

so that

$$\left(A_r^*\right)^{\dagger} P_{L,V_s^{\perp}} A_s^{\dagger} = \left(A_s A_r^*\right)^{\dagger}.$$

Hence, Theorem 2 reduces to the theorem. ■

Theorem 5 means that the reconstruction $\tilde{f}$ by $(A_s A_r^*)^{\dagger}$ (g-inverse) lies in $\mathcal{R}(A_r^* A_r A_s^*)$. While unicity and perfect reconstruction are guaranteed for this particular subspace, it is not so for the whole reconstruction space V_r when the condition in Proposition 3 fails. The Moore-Penrose g-inverse solution has the property that the norm of the coefficients of the reconstructed signal is minimum, which may or may not be relevant here. However, its disadvantage over the more general solution described in Theorem 3 is that it arbitrarily limits our freedom in the choice of the reconstruction subspace. In the sequel, we will show that the perfect reconstruction of certain classes of preferential signals is an interesting option and that the choice of an appropriate subspace L can improve the results substantially.

VII. Application Examples

A. Signal Recovery Example

First, we consider a simple example to show the effectiveness of the proposed sampling theorem. Let H be $L^2[0, K]$, where K is the number of the reconstruction functions. Functions f in H satisfy

$$\int_0^K |f(x)|^2 \, dx < \infty$$

and the corresponding inner product is

$$\langle f, g \rangle = \frac{1}{K} \int_0^K f(x) \overline{g(x)} dx$$

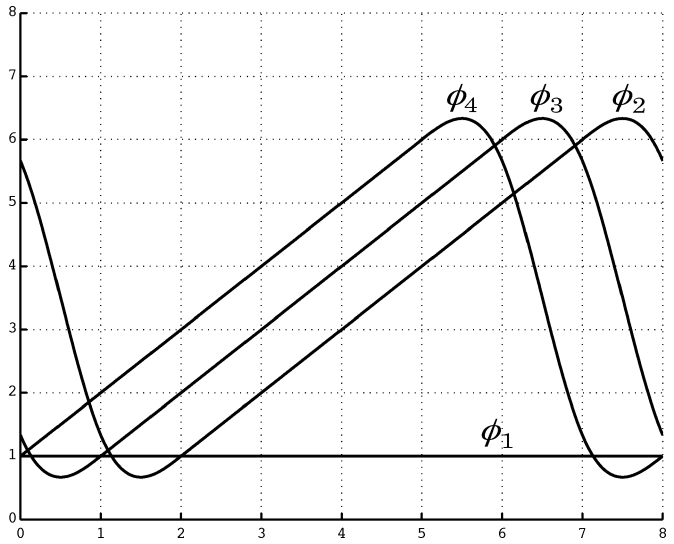


Fig. 3. Preferential signals $\{\phi_i\}_{i=1}^{I}$ in the simulation that are preserved by the proposed sampling theorem.

The sampling functions are integer-shifts of the B-spline of degree 0

$$\psi_n(x) = \beta^0(x - n + 1). \tag{56}$$

The reconstruction functions are integer-shifts of the cyclic B-spline of degree 3. That is, by letting

$$\beta_K^3(x) = \sum_{k \in \mathbb{Z}} \beta^3(x - kK)$$

the reconstruction function is given by

$$\varphi_k(x) = \beta_K^3(x - k + 1).$$

We assume that K is even, and $N = K$. In this case, similar to Example 1, the one-to-one condition does not hold. Indeed, $\tilde{f}$ in (10) with $\boldsymbol{c}$ in (14) belongs to $V_s^{\perp}$.

In order to use the proposed sampling theorem, we need to select a class of preferential signals $\{\phi_i\}_{i=1}^{I}$ in V_r. In this example, we choose four reference signals $\{\phi_i\}_{i=1,2,...,4}$ given by

$$\phi_i = A_r^* \boldsymbol{c}_i$$

with

$$\begin{aligned}
\boldsymbol{c}_1 &= (1, 1, \ldots, 1) \\
\boldsymbol{c}_2 &= (1, 2, \ldots, K-1, 0) \\
\boldsymbol{c}_3 &= (2, 3, \ldots, K-1, 0, 1) \\
\boldsymbol{c}_4 &= (3, 4, \ldots, K-1, 0, 1, 2).
\end{aligned}$$

These signals for $K = 8$ are shown in Fig. 3. In particular, ϕ_1 corresponds to the constant signal, and is included in V_r because the B-splines satisfy the partition of unity. Indeed, the classical property $\sum_{k \in \mathbb{Z}} \beta^n(x - k) = 1$, which holds for any B-spline of degree n [16], implies that $\sum_{k=1,\ldots,K} \beta_K^n(x - k) = 1$.

The best way to emphasize the difference between the proposed sampling theorem (Theorem 3) and the g-inverse solution in Theorem 5 is to consider a target signal f that belongs

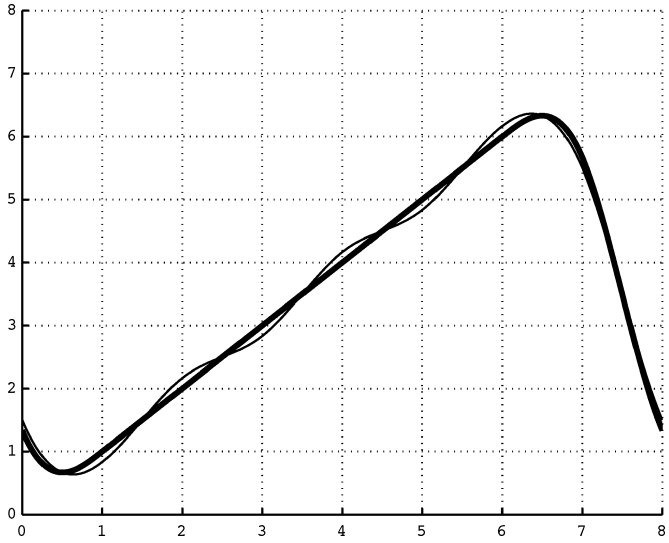


Fig. 4. Reconstructed signals by the proposed sampling theorem (the thick line) and the g-inverse solution in Theorem 5 (the thinner line) when the target is ϕ_3, which is perfectly reconstructed by the proposed one.

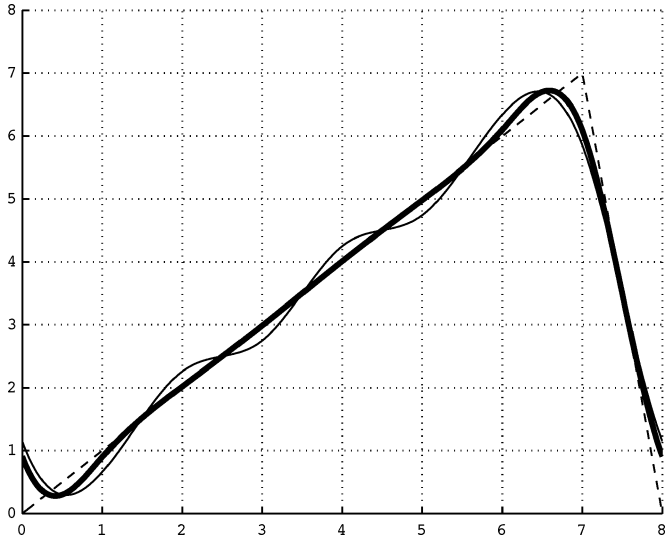


Fig. 5. Reconstructed signals by the proposed sampling theorem (the thick line) and the g-inverse solution (the thin line) when the target (the dashed line) is the signal given in (57).

to the subspace V_i. In the present case, the constant signal incidentally belongs to $\mathcal{R}(A_r^* A_r A_s^*)$, and is perfectly reconstructed by g-inverse solution as well as the proposed one. We therefore concentrate on the target signal ϕ_3, which is not included in $\mathcal{R}(A_r^* A_r A_s^*)$. The reconstructed results are shown in Fig. 4. The thick line shows the reconstructed signal given by the proposed sampling theorem. As expected, the target signal is perfectly reconstructed since it is included in V_i. The thinner line shows the reconstructed signal obtained by the g-inverse solution in Theorem 5, which is less satisfactory and oscillating.

Fig. 5 shows the reconstructed results when the target is

$$f(x) = \begin{cases} x & (0 \le x < K-1) \\ -(K-1)(x-K) & (K-1 \le x \le K), \end{cases} \quad (57)$$

which is displayed with a dashed line. Note that $f(x)$ does not belong to V_r, and hence is not perfectly reconstructed in either cases. The thick and the thin lines show signals obtained by the proposed sampling theorem and the g-inverse solution, respectively. We can see that the proposed sampling theorem gives a better result—essentially perfect for $2 \le x \le 6$—than the g-inverse solution. Obviously, the result depends on the choice of the preferential signals which plays a key role in our formulation.

B. Application to Image Magnification

We now present an application of the proposed sampling theorem to an image magnification problem. Fig. 6 shows three images (256×256 pixels) taken from the SIDBA standard image library. These were downsized by a factor of two using 2×2 averaging. The reduced 128×128 images were further divided into 8×8 pixels patches. We set the task to magnify these patches back to the original size (16×16) using a full DCT expansion, and to recover the original 256×256 image by juxtaposition. This was achieved by reconstructing a continuous image from the 8×8 subimages (measurements) by the proposed sampling theorem and the g-inverse solution in Theorem 5. The 16×16 magnified subimages were obtained by resampling the continuous reconstruction. For comparison, we also include the results of a zooming by two obtained by bicubic spline interpolation which is considered to be one of the better methods for image resampling [17], [18]. This latter, more standard approach to resizing has the advantage of simplicity but it should be noted that it does not ensure consistency with the measurements (local averages).

In order to specify our consistent reconstruction algorithm, we select H to be the space of square-integrable functions on the area $[0, N_x] \times [0, N_y]$, where $N_x = N_y = 8$. The inner product $\langle f, g \rangle$ in H is defined by

$$\langle f, g \rangle = \frac{1}{N_x N_y} \int_0^{N_x} \int_0^{N_y} f(x,y) g(x,y) dx dy.$$

The sampling functions $\psi_{n_x,n_y}(x,y)$ corresponding to our set-up (downsizing by averaging) are

$$\psi_{n_x,n_y}(x,y) = \psi_{n_x}(x) \psi_{n_y}(y)$$

where ψ_{n_x} and ψ_{n_y} are functions defined by (56) with $n_x = 1, 2, \ldots, N_x$ and $n_y = 1, 2, \ldots, N_y$.

The reconstruction is performed in a local DCT basis, which is defined by

$$\varphi_{k_x,k_y}(x,y) = \varphi_{k_x}(x) \varphi_{k_y}(y)$$

for $k_x = 1, 2, \ldots, K_x$ and $k_y = 1, 2, \ldots, K_y$, where

$$\varphi_{k_x}(x) = \begin{cases} 1 & (k_x = 1) \\ \sqrt{2} \cos \frac{(k_x - 1)\pi x}{N_x} & (k_x = 2, 3, \ldots, K_x), \end{cases}$$

$$\varphi_{k_y}(y) = \begin{cases} 1 & (k_y = 1) \\ \sqrt{2} \cos \frac{(k_y - 1)\pi y}{N_y} & (k_y = 2, 3, \ldots, K_y). \end{cases}$$

Here, $K_x = K_y = 16$, which is the same as the pixel number in each axis in a magnified image, meaning that we are in an under-sampling scenario.

(a) (b) (c)

Fig. 6. Standard images used in the simulation. (a) Airplane. (b) Girl. (c) Lenna.

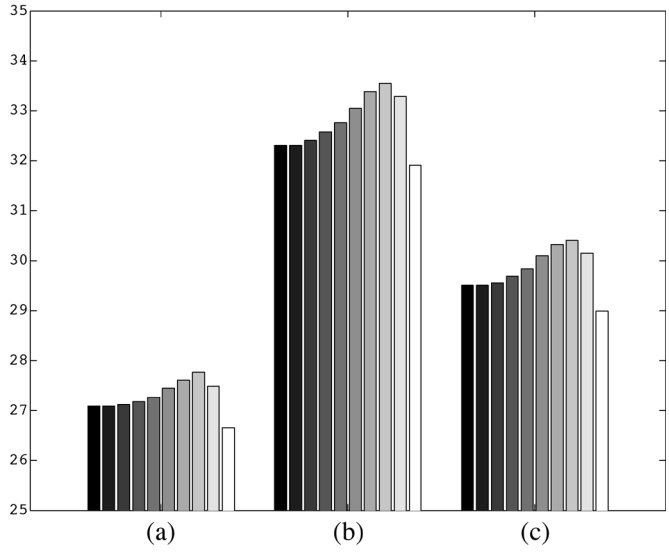


Fig. 7. PSNR values in dB for the three standard images shown in Fig. 6. The 10 bars for each image quantify the PSNR values obtained by the g-inverse solution, the proposed method with the subset (i), $\cdots$, the proposed method with the subset (viii), and the bicubic interpolation from left to right, respectively.

For the preferential signals in the proposed sampling theorem, we used some smooth functions (lower frequency cosinusoids) included in our reconstruction set. Since we have 64 measurements, we can specify up to 64 functions. Taking this into account, we used eight subsets of the reconstruction functions: (i) $\{\varphi_{k_x,k_y}\}_{k_x=1}^{1}{}_{k_y=1}^{1}$ $(I = 1^2)$; (ii) $\{\varphi_{k_x,k_y}\}_{k_x=1}^{2}{}_{k_y=1}^{2}$ $(I = 2^2), \cdots$; (viii) $\{\varphi_{k_x,k_y}\}_{k_x=1}^{8}{}_{k_y=1}^{8}$ $(I = 8^2)$. Note that the proposed method without any specification of functions reduces to the g-inverse solution.

Fig. 7 shows PSNR values in dB calculated by

$$\mathrm{PSNR} = 20\log_{10} \times \frac{255}{\sqrt{\sum_{n_x=1}^{256}\sum_{n_y=1}^{256}\left(\mathrm{RstImg}[n_x,n_y] - \mathrm{OrgImg}[n_x,n_y]\right)^2}}$$

where $\mathrm{RstImg}[n_x,n_y]$ and $\mathrm{OrgImg}[n_x,n_y]$ are the n_xth, n_yth pixel values of the reconstructed and the target images, respectively. The horizontal axis indicates the input images in Fig. 6. The 10 bars for each image quantify the PSNR values obtained by the g-inverse solution, the proposed method with the subset (i), $\cdots$, the proposed method with the subset (viii), and the bicubic interpolation from left to right, respectively. We see that, in each case, the best result is obtained by the proposed method with the subset (vii). The same tendency was also observed with other images of the SIDBA image library, except for the fact that the optimal subset depends on the target image (results not shown). For the present data set, the best reconstruction was obtained by the proposed method with the subset (vii), shown in Fig. 8(a). Fig. 8(b) and (c) shows the result by the g-inverse solution and the bicubic interpolation. In order to facilitate the comparison, one of the 16×16 areas indicated by the white box in Fig. 9(a) is displayed in (b), and the downsized image is shown in (c). Fig. 9(d) and (e) shows the resampled images from the reconstructed continuous images by the proposed sampling theorem and the g-inverse solution, respectively. We can see blocking artifacts in (e), while there is no such effect in (d). Fig. 9(f) shows the reconstructed image by the bicubic interpolation, which lacks sharpness comparatively. We believe that this is partly due to the lack of consistency of this type of reconstruction (interpolation) which does not take into account the fact that the image was reduced by 2×2 averaging.

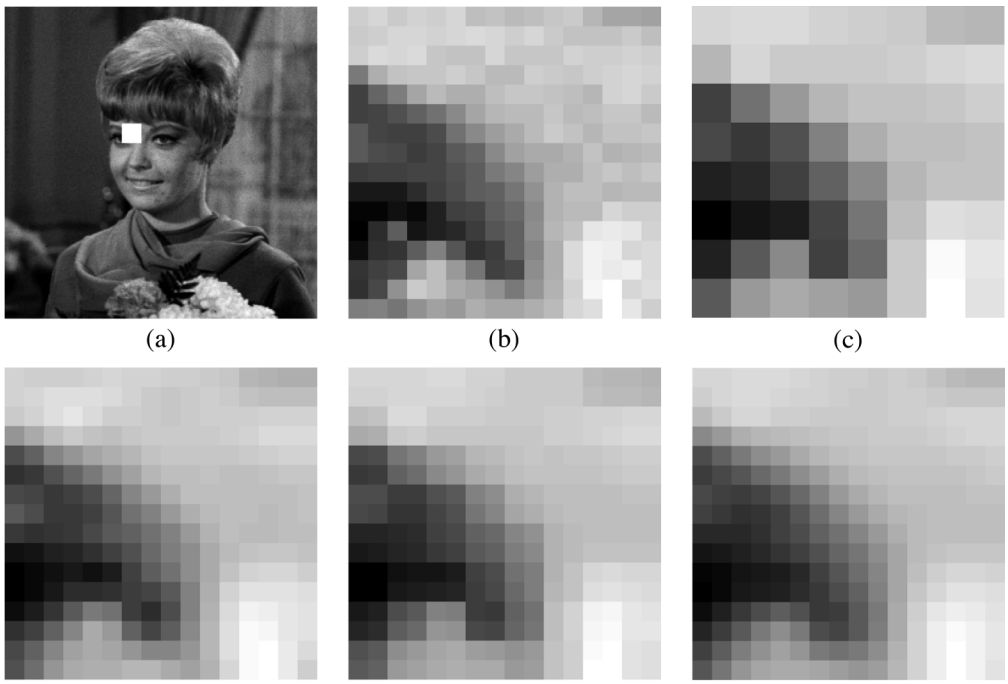
While the present magnification experiment is somewhat contrived (zooming factor of two applied to 8×8 image patches), it shows the potential of the proposed approach and the superiority of consistent sampling over more traditional interpolation approaches which do not take into account the measurement process. An important factor for the success of the proposed reconstruction procedure is the appropriate selection of the class of preferential signals. In our experiments, this has been done in an ad hoc fashion and it is probable that one may be able to further improve the results by optimizing this selection process as well.

VIII. Conclusion

We proposed a consistent sampling theorem that yields a unique and well-defined solution even in the cases where there is no one-to-one mapping between the measurements and the reconstructed signal. The key idea was to exclude the null space of the sampling operator from the reconstruction space, and to impose consistency to its complement. This has a simple geometric interpretation and has the advantage of removing all restrictions on the choice of sampling and reconstruction functions. By taking advantage of the fact that the choice of the complement is not unique, we proposed the

(a) (b) (c)

Fig. 8. Reconstructed images. (a) Proposed method. (b) g-inverse solution. (c) Bicubic interpolation.



(a) (b) (c)

(d) (e) (f)

Fig. 9. Details of the reconstruction shown in Fig. 8. (a) Input image with the white box indicating the area magnified in (b). (c) Downsized image. (d) Proposed method. (e) g-inverse solution. (f) Bicubic interpolation.

main result: a consistency sampling theorem that preserves certain preferential signals. We have also shown that a particular solution to the consistency sampling problem is provided by the Moore-Penrose generalized inverse. However, we did also propose other solutions (in particular, polynomial and sinusoid-preserving ones) that can give better results in practice.

APPENDIX

A. *Proof of Lemma 2*

Let u be an element in $\mathcal{R}(A_r^* A_r A_s^*)^\perp \cap V_s$. Then, it holds that

$$\|A_r u\|^2 = \langle A_r^* A_r u, u\rangle = 0$$

which implies

$$\mathcal{N}(A_r) \supset \mathcal{R}\left(A_r^* A_r A_s^*\right)^\perp \cap V_s.$$

Taking orthogonal complement of the equation yields

$$V_r \subset \mathcal{R}\left(A_r^* A_r A_s^*\right) + V_s^\perp.$$

Hence, it follows from (16) that

$$H = V_r + V_s^\perp \subset \mathcal{R}\left(A_r^* A_r A_s^*\right) + V_s^\perp$$

which implies

$$H = \mathcal{R}\left(A_r^* A_r A_s^*\right) + V_s^\perp. \tag{58}$$

Let u be an element in $\mathcal{R}(A_r^* A_r A_s^*) \cap V_s^\perp$. Then, u satisfies $A_s u = 0$, and there exist some v such that $u = A_r^* A_r A_s^* v$. It holds that

$$0 = \langle A_s u, v\rangle = \langle A_s A_r^* A_r A_s^* v, v\rangle = \|A_r A_s^* v\|^2.$$

Hence, $A_r A_s^* v = 0$ which implies $u = 0$. Therefore, the RHS of (58) is a direct sum, and (55) holds. ■

B. Proof of Lemma 3

According to the interpretation given in [10], the operator is a projection along $V_s^{\perp}$. Hence, we show that

$$\mathcal{R}\left(A_r^*\left(A_sA_r^*\right)^{\dagger}A_s\right)=\mathcal{R}\left(A_r^*A_rA_s^*\right). \tag{59}$$

It holds that

$$\begin{aligned}\mathcal{R}\left(A_r^*\left(A_sA_r^*\right)^{\dagger}A_s\right)&\subset\mathcal{R}\left(A_r^*\left(A_sA_r^*\right)^{\dagger}\right)\\&=\mathcal{R}\left(A_r^*\left(A_sA_r^*\right)^{*}\right)\\&=\mathcal{R}\left(A_r^*A_rA_s^*\right)\\&=\mathcal{R}\left(A_r^*\left(A_sA_r^*\right)^{\dagger}A_sA_r^*\left(A_sA_r^*\right)^{*}\right)\\&\subset\mathcal{R}\left(A_r^*\left(A_sA_r^*\right)^{\dagger}A_s\right),\end{aligned}$$

which implies (59). ■

ACKNOWLEDGMENT

The authors wish to thank the anonymous reviewers for their valuable comments.

REFERENCES

[1] A. Jerri, "The Shannon sampling theorem—Its various extensions and applications: A tutorial review," *Proc. IEEE*, vol. 65, no. 11, pp. 1565–1596, 1977.

[2] R. Marks, II, *Introduction to Shannon Sampling and Interpolation Theory*. New York: Springer-Verlag, 1991.

[3] A. Zayed, *Advances in Shannon's Sampling Theory*. New York: CRC Press, 1993.

[4] J. Higgins, *Sampling Theory in Fourier and Signal Analysis: Foundations*. Oxford, U.K.: Oxford Univ. Press, 1996.

[5] H. Ogawa, "A unified approach to generalized sampling theorems," in *Proc. IEEE-IECEJ-ASJ Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, , Tokyo, 1986, vol. 3, pp. 1657–1660.

[6] H. Ogawa, "A generalized sampling theorem," *Electron. Commun. Jpn., Part 3*, vol. 72, no. 3, pp. 97–105, 1989.

[7] A. Aldroubi and M. Unser, "Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory," *Numer. Function. Anal. Optimiz.*, vol. 15, no. 1–2, pp. 1–21, 1994.

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

[9] M. Unser, "Sampling-50 years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, Apr. 2000.

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

[11] M. Unser and J. Zerubia, "Generalized sampling: Stability and performance analysis," *IEEE Trans. Signal Process.*, vol. 45, no. 12, pp. 2941–2950, Dec. 1997.

[12] I. Shoenberg, *Cardinal Spline Interpolation*. Philadelphia, PA: SIAM, 1973.

[13] A. Albert, *Regression and the Moore–Penrose Pseudoinverse*. New York: Academic, 1972.

[14] R. Schatten, *Norm Ideals of Completely Continuous Operators*. Berlin, Germany: Springer-Verlag, 1960.

[15] H. Ogawa and S. Hara, "Properties of partial projection filter," (in Japanese) *Trans. Inst. Electron., Inf., Commun. Eng. A*, vol. J71-A, no. 2, pp. 527–534, 1988.

[16] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Process. Mag.*, vol. 16, no. 6, pp. 22–38, Nov. 1999.

[17] E. H. W. Meijering, W. J. Niessen, and M. A. Viergever, "Quantitative evaluation of convolution-based methods for medical image interpolation," *Med. Image Anal.* , vol. 5, no. 2, pp. 111–126, Jun. 2001.

[18] P. Thévenaz, T. Blu, and M. Unser, "Interpolation revisited," *IEEE Trans. Med. Imag.*, vol. 19, no. 7, pp. 739–758, Jul. 2000.

Akira Hirabayashi (S'93–M'95) was born in Nagoya, Japan, on February 2, 1971. He received the B.E., M.E., and D.E. degrees in computer science in 1993, 1995, and 1999, respectively, from the Tokyo Institute of Technology, Tokyo, Japan.

From 1995 to 2000, he was a research associate with the Graduate School of Information Science and Engineering, Tokyo Institute of Technology. He is now an Associate Professor with the Faculty of Engineering, Yamaguchi University, Yamaguchi, Japan. From 2004 to 2005, he was an Invited Professor with the Biomedical Imaging Group, Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne. His research interests include sampling theory, signal and image processing, learning theory, and pattern recognition.

Dr. Hirabayashi is a member of the Institute of Electronics, Information, and Communication Engineers (IEICE), Japan, and the Society of Instrument and Control Engineeres (SICE), Japan.

Michael Unser (M'88–SM'94–F'99) received the M.S. (*summa cum laude*) and Ph.D. degrees in electrical engineering in 1981 and 1984, respectively, from the Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne. From 1985 to 1997, he was a scientist with the National Institutes of Health, Bethesda, MD. He is now a Professor and Director of the Biomedical Imaging Group, EPFL. His main research areas are biomedical image processing, splines, and wavelets.

Dr. Unser is the Associate Editor-in-Chief of the IEEE TRANSACTIONS ON MEDICAL IMAGING and the Editor-in-Chief of the *Wavelet Digest*. He organized the first IEEE International Symposium on Biomedical Imaging (ISBI'2002) and currently chairs the Technical Committee of the IEEE-SP Society on Bio Imaging and Signal Processing (BISP). He is the recipient of three Best Paper Awards from the IEEE Signal Processing Society.