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.

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 \( \{\psi_n\}_{n=1}^{N} \). Moreover, we will concentrate on linear algorithms that reconstruct the signal as a linear combination of reconstruction functions \( \{\varphi_k\}_{k=1}^{K} \). 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 \([5] \sim [7]\). 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 \textit{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. 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 \( \mathbb{C}^N \). The reconstructed signal, on the other hand, will be parameterized by a vector in \( \mathbb{C}^K \) with \( K \geq N \). The standard bases for \( \mathbb{C}^N \) and \( \mathbb{C}^K \) are \( \{ e_n^{(N)} \}_{n=1}^N \) and \( \{ e_k^{(K)} \}_{k=1}^K \), respectively. That is, \( e_n^{(N)} \) (resp., \( e_k^{(K)} \)) is the \( N \)-dimensional (resp., \( K \)-dimensional) vector consisting of zero elements except for the \( n \)-th (resp., \( k \)-th) 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

\[
AA^\dagger = P_{\mathcal{R}(A)^\perp}, \quad A^\dagger A = P_{\mathcal{R}(A^*)^\perp},
\]

\[
A^\dagger = A^*(AA^* \dagger) = (A^*A)^\dagger A^*,
\]

\[
P_S(A P_S) \dagger = (A P_S)^\dagger, \quad \text{and} \quad (P_S A)^\dagger P_S = (P_S A)^\dagger.
\]

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

**Lemma 1:** [13] The operator equation

\[
AXB = C
\]

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

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

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 \bar{\beta})^* = \beta \otimes \overline{\alpha}.

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 $\{\psi_n\}_{n=1}^N$:

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

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

$$A_s f = \mathbf{d}.

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

$$A_s = \sum_{n=1}^N e_n^{(N)} \otimes \overline{\psi_n}.

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

$$\hat{f} = \sum_{k=1}^K e[k] \varphi_k.

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

$$A_r = \sum_{k=1}^K e[k]^{(K)} \otimes \overline{\varphi_k}.

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

$$\hat{f} = A_r^* \mathbf{e}.

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

$$X \mathbf{d} = \mathbf{c}.

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

$$\hat{f} = A_r^* X A_s f.

With this formulation, the sampling problem becomes equivalent to finding a suitable matrix $X$ so that $\hat{f}$ satisfies some optimality criterion.
Let $V_s$ and $V_r$ be subspaces in $H$ spanned by $\{\psi_n\}_{n=1,...,N}$ and $\{\varphi_k\}_{k=1,...,K}$, 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 $\{\psi_n\}_{n=1,...,N}$ and $\{\varphi_k\}_{k=1,...,K}$. 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 re-injected 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], [9], [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 $\{d[n]\}_{n=1,...,N}$ 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 $V_s^\perp$.  

---

Fig. 1. Sampling and Reconstruction Formulation.
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
\[
\beta^0(x) = \begin{cases} 
1 & (0 \leq x < 1), \\
0 & (x < 0, x \geq 1), 
\end{cases}
\]
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 $\psi_n$ and $\varphi_k$ be
\[
\psi_n(x) = \beta^0(x - n + 1),
\]
\[
\varphi_k(x) = \beta^1_k(x - k + 1),
\]
respectively, where
\[
\beta^1_k(x) = \sum_{k \in \mathbb{Z}} \beta^1(x - kK),
\]
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
\[
c = (1, -1, \ldots, 1, -1)
\]
belongs to $V_s^\perp$, which contradicts hypothesis (13).

This means that there are situations, such as the example above, 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 the two conditions
\[
V_r \cap V_s^\perp \neq \{0\}, 
\]
\[
V_r + V_s^\perp = H. 
\]
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 over-determined 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 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

\[
A_s^* X A_s f = f
\]  

(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 over-sampling (resp. 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 under-sampling scenario.

The next proposition re-emphasizes the crucial role of Condition (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

\[
f = A_s^* X A_s f = 0
\]

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

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

\[
A_s^* X A_s f = A_s^* (A_s A_s^*)^\dagger A_s f = f,
\]

which implies (17).  

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

$$L \oplus (V_r \cap V_s^\perp) = V_r,$$  \hfill (18)

where $\oplus$ denotes the direct sum, and

$$L \cap V_s^\perp = \{0\}. \hfill (19)$$

**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 $f_1 = f_2$. \hfill \blacksquare

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

$$A_s U = A_s. \hfill (20)$$

Hence, we have

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

which implies

$$U^2 = U. \hfill (21)$$

Equation (20) also implies that

$$\mathcal{N}(U) \subset \mathcal{N}(A_s U) = \mathcal{N}(A_s) \subset \mathcal{N}(A_s^* X A_s) = \mathcal{N}(U),$$

which means

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

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

$$\mathcal{R}(U) = L. \hfill (23)$$
Equations (21), (22), and (23) imply that \( \tilde{f} \) is the oblique projection of \( f \) onto \( L \) along \( V_r^\perp \).

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 \( V_r^\perp \). Then, (21), (22), and (23) hold. Since (21) yields \( U(I - U) = 0 \), it follows from (22) that
\[
\mathcal{R}(I - U) \subseteq \mathcal{N}(U) = \mathcal{N}(A_s),
\]
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:
\[
\tilde{f} = P_{L,V_r^\perp} f.
\]
(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 (12) if and only if \( X \) is given by
\[
X = (A_r^*)^\dagger P_{L,V_r^\perp} A_s^\dagger + Y - A_r A_r^\dagger Y A_s A_s^\dagger,
\]
where \( Y \) is an arbitrary linear operator (i.e., rectangular matrix) from \( \mathbb{C}^N \) to \( \mathbb{C}^K \).

(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
\[
A_r^* X A_s = P_{L,V_r^\perp}.
\]
(26)
Since \( L \) is a subspace of \( V_r \), it holds that
\[
\mathcal{R}(A_r^*) = V_r \supset L = \mathcal{R}(P_{L,V_r^\perp}).
\]
Further, we have that
\[
\mathcal{N}(A_s) = V_s^\perp = \mathcal{N}(P_{L,V_r^\perp}).
\]
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 \( \mathcal{R}(A_r^*) \supset L = \mathcal{R}(P_{L,V_r^\perp}) \), it holds that
\[
A_r^* (A_r^*)^\dagger P_{L,V_r^\perp} = P_{L,V_r^\perp}.
\]
Further, since \( \mathcal{N}(P_{L,V_r^\perp}) = V_s^\perp = \mathcal{N}(A_s) \), we have that
\[
P_{L,V_r^\perp} A_s^\dagger A_s = P_{L,V_r^\perp}.
\]
Hence, it follows from (12) and (25) that
\[
\tilde{f} = A_r^* X A_s f = A_r^* (A_r^*)^\dagger P_{L,V_r^\perp} A_s^\dagger A_s f = P_{L,V_r^\perp} f,
\]
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, \( \hat{f} \) 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 \( V_i = \text{span} \{\phi_i\}_{i=1}^I \) 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

\[
V_i \oplus (V_r \cap V_s^\perp) = V_r,
\]

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

\[
X = (A_s^\dagger)^\dagger (A_s W)^\dagger + Y - A_s A_s^\dagger Y A_s A_s^\dagger,
\]

where \( Y \) is arbitrary and where \( W \) is an operator defined by

\[
W = P_{V_i} + P_{V_s^\perp}. \tag{29}
\]

(Proof) Since \( V_c \) is perpendicular to \( V_i \), the sum of the two orthogonal projection operators \( P_{V_i} \) and \( P_{V_s^\perp} \) 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 \( W (A_s W)^\dagger A_s \). Hence, it follows from the first equation of (3) that

\[
P_{L,V_s^\perp} = W (A_s W)^\dagger A_s = (A_s W)^\dagger A_s.
\]
Moreover, we have that
\[
(A_x^*)^\dagger P_{L,V_r} A_s^\dagger = (A_x^*)^\dagger (A_s W)^\dagger A_s A_s^\dagger \\
= (A_x^*)^\dagger (P_{\mathcal{R}(A_s)} A_s W)^\dagger P_{\mathcal{R}(A_s)} \\
= (A_x^*)^\dagger (P_{\mathcal{R}(A_s)} A_s W)^\dagger \\
= (A_x^*)^\dagger (A_s W)^\dagger,
\]
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
\[
\hat{f} = f_i + f_c.
\]
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 \( A_s W \) and \( A_r \), which are abstract operators from the function space \( H \) to the vector spaces \( \mathbb{C}^N \) and \( \mathbb{C}^K \), 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
\( \{ \varphi_k \}_{k=1, \ldots, K} \) 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 \( \{ \varphi_k \}_{k=1, \ldots, K} \) and \( \{ \psi_n \}_{n=1, \ldots, N} \):
\[
B = \begin{pmatrix}
\langle \varphi_1, \psi_1 \rangle & \langle \varphi_2, \psi_1 \rangle & \cdots & \langle \varphi_K, \psi_1 \rangle \\
\langle \varphi_1, \psi_2 \rangle & \langle \varphi_2, \psi_2 \rangle & \cdots & \langle \varphi_K, \psi_2 \rangle \\
\vdots & \vdots & \ddots & \vdots \\
\langle \varphi_1, \psi_N \rangle & \langle \varphi_2, \psi_N \rangle & \cdots & \langle \varphi_K, \psi_N \rangle
\end{pmatrix}
\]

and define three more matrices
\[
C = \begin{pmatrix}
\langle \phi_1, \varphi_1 \rangle & \langle \phi_2, \varphi_1 \rangle & \cdots & \langle \phi_J, \varphi_1 \rangle \\
\langle \phi_1, \varphi_2 \rangle & \langle \phi_2, \varphi_2 \rangle & \cdots & \langle \phi_J, \varphi_2 \rangle \\
\vdots & \vdots & \ddots & \vdots \\
\langle \phi_1, \varphi_K \rangle & \langle \phi_2, \varphi_K \rangle & \cdots & \langle \phi_J, \varphi_K \rangle
\end{pmatrix}
\]
\[
D = \begin{pmatrix}
\langle \phi_1, \psi_1 \rangle & \langle \phi_2, \psi_1 \rangle & \cdots & \langle \phi_J, \psi_1 \rangle \\
\langle \phi_1, \psi_2 \rangle & \langle \phi_2, \psi_2 \rangle & \cdots & \langle \phi_J, \psi_2 \rangle \\
\vdots & \vdots & \ddots & \vdots \\
\langle \phi_1, \psi_N \rangle & \langle \phi_2, \psi_N \rangle & \cdots & \langle \phi_J, \psi_N \rangle
\end{pmatrix}
\]
\[
E = \begin{pmatrix}
\langle \phi_1, \phi_1 \rangle & \langle \phi_2, \phi_1 \rangle & \cdots & \langle \phi_I, \phi_1 \rangle \\
\langle \phi_1, \phi_2 \rangle & \langle \phi_2, \phi_2 \rangle & \cdots & \langle \phi_I, \phi_2 \rangle \\
\vdots & \vdots & \ddots & \vdots \\
\langle \phi_1, \phi_I \rangle & \langle \phi_2, \phi_I \rangle & \cdots & \langle \phi_I, \phi_I \rangle
\end{pmatrix}
\]

**Theorem 4:** The predominant term of \( X \) in (28) admits the following explicit matrix form
\[
(A_s^*)^\dagger (A_s W)^\dagger = G^\dagger UV^\dagger.
\] (30)

where
\[
U = CE^\dagger D^* + GG^\dagger B^* \Phi^\dagger Z,
\] (31)
\[
V = DE^\dagger D^* + \Phi \Phi^\dagger Z,
\] (32)

with \( G, B, C, D \) and \( E \) as defined above, and where \( \Phi, \Psi, \) and \( Z \) are auxiliary matrices given by
\[
\Phi = BG^\dagger B^*,
\] (33)
\[
\Psi = BG^\dagger C,
\] (34)
\[
Z = \Phi - \Psi (\Psi^\dagger \Psi)^\dagger \Psi^\dagger \Phi.
\] (35)
(Proof) Let $A_i$ be the operator defined by

$$A_i = \sum_{l=1}^I e_l^{(j)} \otimes \delta_l.$$  \hfill (36)

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

$$G = A_r A_r^*,$$  \hfill (37)
$$B = A_s A_s^*,$$  \hfill (38)
$$C = A_r A_r^*,$$  \hfill (39)
$$D = A_s A_s^*,$$  \hfill (40)
$$E = A_i A_i^*.$$  \hfill (41)

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

$$A_r P_{V_i} A_r^* = CE^\dagger D^*,$$  \hfill (42)
$$A_i P_{V_i} A_i^* = DE^\dagger D^*.$$  \hfill (43)

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

$$V_e = P_{V_e} V_s \oplus (V_r \cap V_s^\perp),$$  \hfill (44)

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

$$V_e = P_{V_r} V_s.$$  \hfill (45)

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

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

Equations (45) and (46) yield

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

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

$$P_{V_r} = (P_{V_r} A_s^*) (P_{V_r} A_s^*)^\dagger - \{(P_{V_r} A_s^*) (P_{V_r} A_s^*)^\dagger A_s^*\} \{(P_{V_r} A_s^*) (P_{V_r} A_s^*)^\dagger A_s^*\}^\dagger.$$  \hfill (47)

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

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

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

$$\{(P_{V_r} A_s^*) (P_{V_r} A_s^*)^\dagger A_s^*\} \{(P_{V_r} A_s^*) (P_{V_r} A_s^*)^\dagger A_s^*\}^\dagger = A_s^* G^\dagger B^* \Phi^\dagger \Psi (\Phi^\dagger \Psi)^\dagger B G^\dagger A_r.$$  \hfill (49)
Hence, from (47) ~ (49), (37), (38), and (33), we have
\[ A_r P V_1 A_r^* = G G^\dagger B^* \Phi^\dagger Z, \]  
(50)
and
\[ A_r P V_1 A_r^* = \Phi \Phi^\dagger Z. \]  
(51)

Then, (29), (42), (50), and (31) yield
\[ A_s W A_s^* = C E^\dagger D^* + G G^\dagger B^* \Phi^\dagger Z = U. \]  
(52)

Further, (29), (43), (51), and (32) give
\[ A_s W A_s^* = D E^\dagger D^* + \Phi \Phi^\dagger Z = V. \]  
(53)

Finally, it follows from (2), (52), and (53) that
\[ (A_s^*)^\dagger (A_s W)^\dagger = (A_r A_r^*)^\dagger A_s^* W A_s^* (A_s W A_s^*)^\dagger = G^\dagger U V^\dagger, \]
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
\[ Z = B B^* = B C (C^* B^\dagger B C)^\dagger C^* B^*, \]
\[ U = C E^\dagger D^* + B^\dagger Z, \]
\[ V = D E^\dagger D^* + Z, \]
respectively, and the first term of $X$ in (29) is given by
\[ (A_s^*)^\dagger (A_s W)^\dagger = U V^\dagger. \]

The proof is straightforward.

**VI. Moore-Penrose Generalized Inverse Reconstruction**

The consistent sampling theorems that have been described in the literature all require the condition (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 $d$ by (8) if $X$ in (11) is given by
\[ X = (A_s A_r^*)^\dagger. \]
(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 \( L = \mathcal{R}(A_r^*A_r^*) \) is obtained by (12) if and only if \( X \) is given by

\[
X = (A_xA_x^*)^\dagger + Y - A_xA_x^*[YA_xA_x^*],
\]

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

\[
H = \mathcal{R}(A_r^*A_r^*) \oplus V_s^\perp.
\]

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

(Proof of Theorem 5) Lemma 3 implies that \( (A_xA_x^*)^\dagger \) lies in \( \mathcal{R}(A_r^*A_r^*) \) is given by

\[
P_{L,V_s^\perp} = A_r^*(A_xA_x^*)^\dagger A_s.
\]

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

\[
(A_x^*)^\dagger P_{L,V_s^\perp}A_x^* = (A_x^*)^\dagger (A_xA_x^*)^\dagger A_s A_x^*
\]

\[
= P_{\mathcal{R}(A_r)}(P_{\mathcal{R}(A_r)}A_xA_x^*P_{\mathcal{R}(A_r)})^\dagger P_{\mathcal{R}(A_r)}
\]

\[
= (P_{\mathcal{R}(A_r)}A_xA_x^*P_{\mathcal{R}(A_r)})^\dagger
\]

\[
= (A_xA_x^*)^\dagger,
\]

so that

\[
(A_x^*)^\dagger P_{L,V_s^\perp}A_x^* = (A_xA_x^*)^\dagger.
\]

Hence, Theorem 2 reduces to the theorem.

Theorem 5 means that the reconstruction \( \hat{f} \) by \( (A_xA_x^*)^\dagger \) (g-inverse) lies in \( \mathcal{R}(A_r^*A_r^*) \). 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

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

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

$$\psi_n(x) = \beta^0(x - n + 1).$$

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

$$\varphi_k(x) = \beta_k^3(x - k + 1),$$

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, $\hat{f}$ in (10) with $c$ in (14) belongs to $V_r^\perp$.

In order to use the proposed sampling theorem, we need to select a class of preferential signals $\phi_i$ in $V_r$. In this example, we choose four reference signals $\phi_i$ given by $\phi_i = A_e c_i$, with

$$c_1 = (1, 1, \ldots, 1),$$
$$c_2 = (1, 2, \ldots, K - 1, 0),$$
$$c_3 = (2, 3, \ldots, K - 1, 0, 1),$$
$$c_4 = (3, 4, \ldots, K - 1, 0, 1, 2).$$

These signals for $K = 8$ are shown in Fig. 3. In particular, $\phi_1$ corresponds to the constant signal, and is included in $V_r$ because the B-splines satisfy the partition of unity. Indeed, the classical property $\sum_{k \in \mathbb{Z}} \beta^m(x - k) = 1$, which holds for any B-spline of degree $n$ [16], implies that $\sum_{k=1,\ldots,K} \beta_k^0(x - k) = 1$.

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 $\hat{f}$ that belongs to the subspace $V_r$. In the present case, the constant signal incidentally belongs to $R(A_e^* A_r A_e^*)$, and is perfectly reconstructed by g-inverse solution as well as the proposed one. We therefore concentrate on the target signal $\phi_2$, which is not included in $R(A_e^* A_r A_e^*)$.

The reconstructed results are shown in Fig. 4. The thick line shows the reconstructed signal given by the proposed
Fig. 3. Preferential signals $\{\phi_k\}_{k=1}^T$ in the simulation that are preserved by the proposed sampling theorem.

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 $\phi_0$, 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).
Fig. 6. Standard images used in the simulation. (a) Airplane. (b) Girl. (c) Lenna.

Sampling theorem. As expected, the target signal is perfectly reconstructed since it is included in $V_r$. 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

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

(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 $2 \leq x \leq 6$ — 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 3 images ($256 \times 256$ pixels) taken from the SIDBA standard image library. These were downsized by a factor of two using $2 \times 2$ averaging. The reduced $128 \times 128$ images were further divided into $8 \times 8$ pixels patches. We set the task to magnify these patches back to the original size ($16 \times 16$) using a full DCT expansion, and to recover the original $256 \times 256$ image by juxtaposition. This was achieved by reconstructing a continuous image from the $8 \times 8$ subimages (measurements) by the proposed sampling theorem and the g-inverse solution in Theorem 5. The $16 \times 16$ 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).
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), ..., the proposed method with the subset (viii), and the bicubic interpolation from left to right, respectively.

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

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

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

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 $\psi_{n_x}$ and $\psi_{n_y}$ 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.

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 8 subsets of the reconstruction functions: (i) $\{ \varphi_{k_x, k_y} \}_{k_x=1}^{16} \cdots_{k_y=1}^{16}$ ($I = 1^2$), (ii) $\{ \varphi_{k_x, k_y} \}_{k_x=1}^{16} \cdots_{k_y=1}^{16}$ ($I = 2^2$), $\cdots$, (viii) $\{ \varphi_{k_x, k_y} \}_{k_x=1}^{16} \cdots_{k_y=1}^{16}$ ($I = 8^2$). 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
\[
\text{PSNR} = 20 \log_{10} \frac{255}{\sqrt{\sum_{n_x=1}^{256} \sum_{n_y=1}^{256} (\text{RstImg}[n_x, n_y] - \text{OrgImg}[n_x, n_y])^2}}
\]
where RstImg$[n_x, n_y]$ and OrgImg$[n_x, n_y]$ are the $n_x$-th, $n_y$-th pixel value of the reconstructed and the target image, 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). Figs. 8 (b) and (c) show the result by the g-inverse solution and the bicubic interpolation. In order to facilitate the comparison, one of the $16 \times 16$ areas indicated by the white box in Fig. 9 (a) is displayed in (b), and the downsized image is shown in (c). Figs. 9 (d) and (e) show 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 $2 \times 2$ averaging.

While the present magnification experiment is somewhat contrived (zooming factor of two applied to 8 x 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 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 \( \mathcal{R}(A_s^* A_r A_s^*)^\perp \cap V_s \). Then, it holds that

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

which implies

\[
\mathcal{N}(A_r) \supset \mathcal{R}(A_s^* A_r A_s^*)^\perp \cap V_s.
\]

Taking orthogonal complement of the equation yields

\[
V_r \subset \mathcal{R}(A_s^* A_r A_s^*) + V_s^\perp.
\]

Hence, it follows from (16) that

\[
H = V_r + V_s^\perp \subset \mathcal{R}(A_s^* A_r A_s^*) + V_s^\perp,
\]

which implies

\[
H = \mathcal{R}(A_s^* A_r A_s^*) + V_s^\perp.
\]

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

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

Hence, \( A_r A_s^* v = 0 \) which implies \( u = 0 \). Therefore, the right-hand side 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

\[
\mathcal{R}(A_s^* (A_r A_s^*)^\dagger A_r) = \mathcal{R}(A_s^* A_r A_s^*).
\]

It holds that

\[
\mathcal{R}(A_s^* (A_s A_r^*)^\dagger A_r) \subset \mathcal{R}(A_s^* (A_r A_s^*)^\dagger) = \mathcal{R}(A_s^* A_r A_s^*) = \mathcal{R}(A_s^* (A_s A_r^*)^\dagger A_r A_s^* (A_s A_r^*)^\dagger) \subset \mathcal{R}(A_s^* (A_r A_s^*)^\dagger A_r),
\]

which implies (59).
ACKNOWLEDGMENT

The authors wish to thank the anonymous reviewers for their valuable comments. This work is partially supported by the grant from the Venture Business Laboratory, Yamaguchi University, Yamaguchi, Japan.

REFERENCES

Michael Unser received the M.S. (summa cum laude) and Ph.D. degrees in Electrical Engineering in 1981 and 1984, respectively, from the Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland. From 1985 to 1997, he worked as a scientist with the National Institutes of Health, Bethesda USA. He is now professor and Director of the Biomedical Imaging Group at the 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).

Dr. Unser is a fellow of the IEEE and the recipient of three Best Paper Awards from the IEEE Signal Processing Society.