Constrained Sampling: Optimum Reconstruction in Subspace with Minimax Regret Constraint

Runyi Yu, *Senior Member, IEEE*, Bashir Sadeghi *Student Member, IEEE*, and Vishnu Naresh Boddeti *Member, IEEE*

Abstract—**This paper considers the problem of optimum reconstruction in generalized sampling-reconstruction processes (GSRPs). We propose constrained GSRP, a novel framework that minimizes the reconstruction error for inputs in a subspace, subject to a constraint on the maximum regret-error for any other signal in the entire signal space. This framework addresses the primary limitation of existing GSRPs (consistent, subspace and minimax regret), namely, the assumption that the *a priori* subspace is either fully known or fully ignored. We formulate constrained GSRP as a constrained optimization problem, the solution to which turns out to be a convex combination of the subspace and the minimax regret samplings. Detailed theoretical analysis on the reconstruction error shows that constrained sampling achieves a reconstruction that is 1) (sub)optimal for signals in the input subspace, 2) robust for signals around the input subspace, and 3) reasonably bounded for any other signals with a simple choice of the constraint parameter. Experimental results on sampling-reconstruction of a Gaussian input and a speech signal demonstrate the effectiveness of the proposed scheme.**

Index Terms—**Consistent sampling, constrained optimization, generalized sampling-reconstruction processes, minimax regret sampling, oblique projection, orthogonal projection, reconstruction error, subspace sampling.**

I. INTRODUCTION

Sampling is the backbone of many applications in digital communications and signal processing; for example, sampling rate conversion for software radio [1], biomedical imaging [2], image super resolution [3], machine learning and signal processing on graph [4], [5], etc. Many of the systems involved in these applications can be modeled as the generalized sampling-reconstruction process (GSRP) as shown in Fig. 1. A typical GSRP consists of a sampling operator S^* associated with a sampling subspace $\mathcal{S}$ in a Hilbert space $\mathcal{H}$, a reconstruction operator W associated with a reconstruction subspace $\mathcal{W}$, and a correction digital filter Q. For a given subspace $\mathcal{W}$, orthogonal projection onto $\mathcal{W}$ minimizes the reconstruction error in $\mathcal{W}$, as measured by the norm of $\mathcal{H}$. As a result, orthogonal projection is considered to be the best possible GSRP. However, the orthogonal projection is not feasible unless the reconstruction space is a subspace of sampling space [6], i.e., $\mathcal{W} \subseteq \mathcal{S}$. Therefore, many solutions have been developed for the GSRP problem under different assumptions on $\mathcal{S}$, $\mathcal{W}$ and the input subspace. These solutions
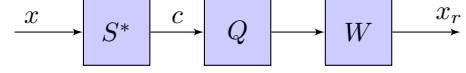
R. Yu is with the Department of Electrical and Electronic Engineering, Eastern Mediterranean University, Gazimagusa, via Merlin 10, Turkey (Tel: +90 392 630 1382; Fax: +90 392 365 0240; E-mail: yu@ieee.org).

B. Sadeghi and V. N. Boddeti are with the Department of Computer Science and Engineering, Michigan State University, USA (E-mails: sadeghib@msu.edu, vishnu@msu.edu).

Fig. 1. A typical GSRP: S^* is a sampling operator, Q is a discrete-time correction filter, and W a reconstruction operator.

can be categorized into *consistent*, *subspace*, and *minimax regret* samplings.

When the inclusion property ($\mathcal{W} \subseteq \mathcal{S}$) does not hold, but one still wants to have the effect of orthogonal projection for any signals in the reconstruction space, Unser *et al* [7], [8] introduced the notion of *consistent* sampling for shiftable spaces. This sampling strategy has later been developed and generalized by Eldar and co-authors [9]–[12]. Common to this body of work is the assumption that the subspace $\mathcal{W}$ and the orthogonal complement of $\mathcal{S}$ (denoted by $\mathcal{S}^\perp$) satisfy the so-called direct-sum condition, i.e., $\mathcal{W} \oplus \mathcal{S}^\perp = \mathcal{H}$. This implies that $\mathcal{W}$ and $\mathcal{S}^\perp$ uniquely decompose $\mathcal{H}$. When the direct-sum condition is relaxed to be a simple sum condition $\mathcal{W} + \mathcal{S}^\perp = \mathcal{H}$, the consistent sampling can still be developed in finite spaces [13], [14]. Further generalization of consistent sampling where even the sum condition is not satisfied can be found in [15], [16].

In many instances, the reconstruction space is usually not the subspace of input signals due to variety of reasons. On one hand, this may be the case due to limitation on physical devices. On the other hand, it can also be advantageous to consider different reconstruction spaces. For example, the sinc function as a generator for the space of band-limited signals suffers from slow convergence in reconstruction; it is more convenient to use a different generator that has finite support. Eldar and Dvorkind in [6] introduced *subspace* sampling and showed that orthogonal projection onto the reconstruction space for signals belonging in a priori subspace is feasible under the direct-sum between $\mathcal{S}^\perp$ and the *a priori* subspace. This subspace can be learned empirically or by a training dataset [17]. Nevertheless, it would still be subject to uncertainties due to, for example, learning imperfection, noise or hardware inability to sample at Nyquist rate. Knyazev *et al* used a convex combination of consistent and subspace GSRP to address the uncertainty of the *a priori* subspace [17]. However, the reconstruction errors of consistent sampling and subspace sampling can be arbitrarily large if the angle between reconstruction (or *a priori*) space and sampling space approaches 90° [6].

Minimax regret GSRP was introduced by Eldar and Dvorkind [6] to address the potential for large errors

associated with consistent (and subspace) sampling, for signals away from the input subspace, by minimizing the maximum regret-error (distance of the reconstructed signal from orthogonal projection). The minimax regret GSRP, however, is found to be conservative as it ignores the *a priori* information on input signals.

In the aforementioned GSRPs the *a priori* subspace is assumed to be either fully known or fully ignored, which is not practically realizable. In addition, the angle between sampling space and input space cannot be controlled (they can get arbitrarily close to $90°$). In this paper, we introduce *constrained* sampling, to address these limitations. We design a robust (in the sense of angle between sampling and input spaces) reconstruction for the signals that approximately lies in the *a priori* subspace. To this end, we introduce a new sampling strategy that exploits the *a priori* subspace information while simultaneously enjoying the reasonably bounded error (for any input) of the minimax regret sampling. This is done by minimizing the maximum possible reconstruction error for the signals lying in the *a priori* subspace while constraining the regret-error to be below certain level for any signal in $\mathcal{H}$. The solution is shown to be a convex combination of minimax regret and consistent sampling. To be specific, given an input x, the reconstruction of the proposed constrained sampling is given as a convex combination below:

$$x_\lambda = \lambda\, x_{\mathrm{sub}} + (1-\lambda) x_{\mathrm{reg}}, \quad \lambda \in [0, 1] \tag{1}$$

where x_{sub} and x_{reg} are the reconstructions of the subspace and minimax sampling, respectively. The result is illustrated in Fig. 2 for a simple case where $\mathcal{H} = \mathbb{R}^2$ and the *a priori* subspace is the same as $\mathcal{W}$ (therefore, subspace sampling would be equivalent to the consistent one). In the figure, x is the input signal; $x_{\mathrm{opt}} = P_\mathcal{W} x$ is the optimal reconstruction, i.e., the orthogonal projection of x onto $\mathcal{W}$; $x_{\mathrm{sub}} = P_{\mathcal{W}\mathcal{S}^\perp} x$ is the oblique projection onto $\mathcal{W}$ along the orthogonal complement of $\mathcal{S}$; and $x_{\mathrm{reg}} = P_\mathcal{W} P_\mathcal{S} x$ is the result of two successive orthogonal projections. The figure shows that as a combination of x_{sub} and x_{reg}, our constrained sampling x_λ can potentially be very close to orthogonal projection. This desirable feature will also be demonstrated in the two examples in Section VI.

The main contributions of this paper can be summarized as follows:

1) We propose and solve a constrained optimization problem which yields reconstruction that is (sub)optimal for signals in input subspace and robust for any other input signals.
2) The relaxed solution to the optimization problem leads to a new sampling strategy (i.e., the constrained sampling) which has consistent (or subspace) and minimax regret samplings as special cases.
3) We provide detailed analysis of reconstruction errors, and obtain reconstruction guarantees in the form of lower and upper bounds of errors.

The organization of the paper is as follows. In Sections II and III, we provide preliminaries and discuss related work,
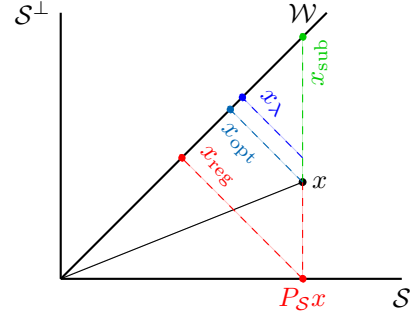

Fig. 2. Comparison of several sampling schemes: $x_{\mathrm{opt}} = P_\mathcal{W} x$ is the orthogonal projection of x onto $\mathcal{W}$; x_{sub} is the oblique projection onto $\mathcal{W}$ along $\mathcal{S}^\perp$; and x_{reg} is orthogonal projection onto $\mathcal{S}$ followed by orthogonal projection onto $\mathcal{W}$. Our constrained reconstruction x_λ is a simple convex combination of x_{sub} and x_{reg} which can potentially get very close to x_{opt}.

respectively. The proposed constrained sampling is described in Section IV. In Section V, we obtain lower and upper bounds on the reconstruction error of the constrained GSRP. We then present two illustrative examples to demonstrate the effectiveness of the new sampling scheme in Section VI. Finally, we conclude the paper in Section VII.

II. PRELIMINARIES

A. Notation

We denote the set of real and integer numbers with $\mathbb{R}$ and $\mathbb{Z}$ respectively. Let $\big(\mathcal{H}, \langle\,\cdot\,\rangle\big)$ be a Hilbert space with the norm $\|\cdot\|$ induced by the inner product $\langle\,\cdot\,\rangle$. We assume throughout the paper that $\mathcal{H}$ is infinite-dimensional unless otherwise stated. Vectors in $\mathcal{H}$ are represented by lowercase letters (e.g., x, v). Capital letters are used to represent operators (e.g., S, W). The (closed) subspaces of $\mathcal{H}$ are denoted by capital calligraphy letters (e.g., $\mathcal{S}$, $\mathcal{W}$). $\mathcal{S}^\perp$ is the orthogonal complement of $\mathcal{S}$ in $\mathcal{H}$. For a linear operator V, its range and nullspace are denoted by $\mathcal{R}(V)$ (or $\mathcal{V}$) and $\mathcal{N}(V)$ respectively. In particular, the Hilbert space of continuous-time square-integrable functions (discrete-time summable sequences, resp) is denoted by L_2 (ℓ_2, resp). At particular time instant $t \in \mathbb{R}$ ($n \in \mathbb{Z}$, resp), the value of signal $x \in L_2$ ($d \in \ell_2$, resp) is denoted by $x(t)$ ($d[n]$, resp).

B. Subspaces and Projections

Given two subspaces $\mathcal{V}_1$, $\mathcal{V}_2$, if they satisfy the direct-sum condition, i.e.,

$$\mathcal{V}_1 \oplus \mathcal{V}_2 = \mathcal{H}$$

we can define an oblique projection onto $\mathcal{V}_1$ along $\mathcal{V}_2$. Let it be denoted as $P_{\mathcal{V}_1\mathcal{V}_2}$. By definition [6], $P_{\mathcal{V}_1\mathcal{V}_2}$ is an unique operator satisfying

$$P_{\mathcal{V}_1\mathcal{V}_2} x = \begin{cases} x, & x \in \mathcal{V}_1 \\ 0, & x \in \mathcal{V}_2. \end{cases}$$

As a result, we have

$$\mathcal{R}(P_{\mathcal{V}_1\mathcal{V}_2}) = \mathcal{V}_1, \quad \mathcal{N}(P_{\mathcal{V}_1\mathcal{V}_2}) = \mathcal{V}_2.$$

Any projection P can be written, in terms of its range and nullspace, as

$$P = P_{\mathcal{R}(P)\mathcal{N}(P)}.$$

By exchanging the role of $\mathcal{V}_1$ and of $\mathcal{V}_2$, we also have the oblique projection $P_{\mathcal{V}_2\mathcal{V}_1}$. And

$$P_{\mathcal{V}_1\mathcal{V}_2} + P_{\mathcal{V}_2\mathcal{V}_1} = I, \tag{2}$$

where $I: \mathcal{H} \to \mathcal{H}$ is the identity operator. In particular, if $\mathcal{V}_1 = \mathcal{V}_2^{\perp} = \mathcal{V}$, then the oblique projections reduce to the orthogonal ones, and (2) specializes to

$$P_{\mathcal{V}} + P_{\mathcal{V}^{\perp}} = I. \tag{3}$$

An important characterization of projection is that a linear operator $P: \mathcal{H} \to \mathcal{H}$ is an oblique projection if and only if $P^2 = P$ [18]. Note that the sum of two projections is generally not a projection. Nevertheless, the following result states that their convex combination remains a projection if both share the same nullspace. This result will be useful in our study of the constrained sampling.

Proposition 1: Let P_1 and P_2 be two projections. If $\mathcal{N}(P_1) = \mathcal{N}(P_2)$, then the following statements hold.

  1) $P_1 P_2 = P_1$ and $P_2 P_1 = P_2$.
  2) $P = \lambda P_1 + (1-\lambda)P_2$ is a projection for any $\lambda \in \mathbb{R}$.

Proof: 1) From (2), it follows

$$P_1 P_2 = P_1(I - P_{\mathcal{N}(P_2)\mathcal{R}(P_2)}) = P_1 - P_1 P_{\mathcal{N}(P_2)\mathcal{R}(P_2)}.$$

If $\mathcal{N}(P_1) = \mathcal{N}(P_2)$, then the last term becomes zero. Hence, $P_1 P_2 = P_1$. Similarly, we have that $P_2 P_1 = P_2$.

2) We can readily verify that $P^2 = P$ in view of the result in 1). $\square$

As consequences of Proposition 1, the following equalities hold (and will be used in Section IV):

$$P_{\mathcal{V}_1} P_{\mathcal{V}_2\mathcal{V}_1^{\perp}} = P_{\mathcal{V}_1} \tag{4}$$

and

$$P_{\mathcal{V}_1\mathcal{V}_2^{\perp}} P_{\mathcal{V}_2} = P_{\mathcal{V}_1\mathcal{V}_2^{\perp}}. \tag{5}$$

C. Angle between Subspaces

The notion of angles between two subspaces indicates how far they are away from each other.

Consider a subspace $\mathcal{V} \subset \mathcal{H}$ and a vector $0 \neq x \in \mathcal{H}$. The angle between x and $\mathcal{V}$, denoted by $(x, \mathcal{V})$, is defined by

$$\cos(x, \mathcal{V}) = \frac{\|P_{\mathcal{V}} x\|}{\|x\|} \tag{6}$$

or equivalently

$$\sin(x, \mathcal{V}) = \frac{\|P_{\mathcal{V}^{\perp}} x\|}{\|x\|}. \tag{7}$$

Let $\mathcal{V}_1, \mathcal{V}_2 \subset \mathcal{H}$ be two subspaces, following [6], the (maximal principal) angle between $\mathcal{V}_1$ and $\mathcal{V}_2$, denoted by $(\mathcal{V}_1, \mathcal{V}_2)$, is defined by

$$\cos(\mathcal{V}_1, \mathcal{V}_2) = \inf_{0 \neq x \in \mathcal{V}_1} \frac{\|P_{\mathcal{V}_2} x\|}{\|x\|} \tag{8}$$

or equivalently

$$\sin(\mathcal{V}_1, \mathcal{V}_2) = \sup_{0 \neq x \in \mathcal{V}_1} \frac{\|P_{\mathcal{V}_2^{\perp}} x\|}{\|x\|}. \tag{9}$$

The angle can also be characterized via any linear operator B whose range is equal to $\mathcal{V}_1$:

$$\cos(\mathcal{V}_1, \mathcal{V}_2) = \inf_{x \notin \mathcal{N}(B)} \frac{\|P_{\mathcal{V}_2} Bx\|}{\|Bx\|} \tag{10}$$

or equivalently

$$\sin(\mathcal{V}_1, \mathcal{V}_2) = \sup_{x \notin \mathcal{N}(B)} \frac{\|P_{\mathcal{V}_2^{\perp}} Bx\|}{\|Bx\|}. \tag{11}$$

Note that $(\mathcal{V}_1, \mathcal{V}_2) \neq (\mathcal{V}_2, \mathcal{V}_1)$ in general. If their orthogonal complements are used instead, the order can be exchanged [6], [7]:

$$(\mathcal{V}_1, \mathcal{V}_2) = (\mathcal{V}_2^{\perp}, V_1^{\perp}). \tag{12}$$

Nevertheless, under a direct-sum condition, the commutativity holds [19]:

$$(\mathcal{V}_1, \mathcal{V}_2) = (\mathcal{V}_2, \mathcal{V}_1) \quad \text{if } \mathcal{V}_1 \oplus \mathcal{V}_2^{\perp} = \mathcal{H}. \tag{13}$$

The angle between subspaces allows descriptions of lower and upper bounds for orthogonal projection of signals in $\mathcal{V}_1$:

$$\cos(\mathcal{V}_1, \mathcal{V}_2)\|x\| \leq \|P_{\mathcal{V}_2} x\| \leq \sin(\mathcal{V}_1, \mathcal{V}_2^{\perp})\|x\|, \quad x \in \mathcal{V}_1 \tag{14}$$

and any signal in $\mathcal{H}$, via a linear operator B with $\mathcal{R}(B) = \mathcal{V}_1$:

$$\cos(\mathcal{V}_1, \mathcal{V}_2)\|Bx\| \leq \|P_{\mathcal{V}_2} Bx\| \leq \sin(\mathcal{V}_1, \mathcal{V}_2^{\perp})\|Bx\|, \quad x \in \mathcal{H}. \tag{15}$$

For oblique projection, the following bounds are proven in [6]

$$\frac{\|P_{\mathcal{V}_2^{\perp}} x\|}{\sin(\mathcal{V}_1, \mathcal{V}_2)} \leq \|P_{\mathcal{V}_1\mathcal{V}_2} x\| \leq \frac{\|P_{\mathcal{V}_2^{\perp}} x\|}{\cos(\mathcal{V}_1, \mathcal{V}_2^{\perp})}. \tag{16}$$

III. RELATED WORK

In this Section, we review four important sampling schemes; namely, orthogonal, consistent, subspace, and minimax regret samplings. For comparison, some properties of these schemes, are summarized in Table I, which also gives the properties of our constrained sampling.

A. Generalized Sampling-Reconstruction Processes

Consider the GSRP in Fig. 1, where $x, x_r \in \mathcal{H}$ are the input and the output signal, respectively; S^* and W are the sampling and reconstruction operators, respectively; and Q is a bounded correction discrete-time filter.

Assume that S^* and W are given in terms of sampling space $\mathcal{S}$ and reconstruction space $\mathcal{W}$, respectively. Let $\mathcal{W}$ be spanned by a set of vectors $\{w_n\}_{n \in \mathcal{I}}$, where $\mathcal{I} \subseteq \mathbb{Z}$ is a set of indexes. Then $W: \ell_2(\mathcal{I}) \to \mathcal{H}$ can be described by the synthesis operator

$$W: c \mapsto Wc = \sum_{n \in \mathcal{I}} c[n]w_n, \quad c \in \ell_2(\mathcal{I}).$$

Note that the range of W is $\mathcal{W}$.

TABLE I
SAMPLING SCHEMES AND THEIR PROPERTIES

Sampling Scheme	GSRP T	Optimal in $\mathcal{A}$?	Error bounded?[a]
Orthogonal[b]	$P_{\mathcal{W}}$	optimal	bounded
Consistent	$P_{\mathcal{WS}^\perp}$	optimal	unbounded
Subspace	$P_{\mathcal{W}}P_{\mathcal{AS}^\perp}$	optimal	unbounded
Regret	$P_{\mathcal{W}}P_{\mathcal{S}}$	non-optimal	bounded
Constrained	$\lambda P_{\mathcal{W}}P_{\mathcal{AS}^\perp}$ $+(1-\lambda)P_{\mathcal{W}}P_{\mathcal{S}}$	sub-optimal	bounded

[a]regardless of $(\mathcal{A}, \mathcal{S})$.
[b]This is the optimal sampling scheme but possible only if $\mathcal{W} \subseteq \mathcal{S}$.

Similarly, let $\mathcal{S}$ be spanned by vectors $\{s_n\}_{n\in\mathcal{I}}$. Then $S^*\colon \mathcal{H} \to \ell_2(\mathcal{I})$ can be described by the adjoint (analysis) operator

$$S^*\colon x \mapsto S^*x = c, \quad c[n] = \langle x, s_n\rangle, \quad n \in \mathcal{I}, \, x \in \mathcal{H}$$

since by definition of adjoint operator [20]

$$\langle Sa, x\rangle = \langle a, S^*x\rangle_{\ell_2} \quad \text{for all} \quad x \in \mathcal{H}, \, a \in \ell_2(\mathcal{I}).$$

Note that the nullspace of S^* is the orthogonal complement of $\mathcal{S}$, i.e., $\mathcal{N}(S^*) = \mathcal{S}^\perp$ (see [20]).

We further assume throughout the paper that set $\{w_n\}$ constitutes a frame of $\mathcal{W}$, that is, there exist two constant scalars $0 < \alpha \le \beta < \infty$ such that

$$\alpha\|x\|^2 \le \sum_{n\in\mathcal{I}} |\langle x, w_n\rangle|^2 \le \beta\|x\|^2, \quad x \in \mathcal{W}.$$

Set $\{s_n\}$ is also assumed to be a frame of $\mathcal{S}$.

The overall GSRP can be described as a linear operator $T = WQS^*\colon \mathcal{H} \to \mathcal{H}$:

$$T\colon x \mapsto x_r = WQS^*x, \quad x \in \mathcal{H}. \tag{17}$$

The reconstruction quality of the GSRP can be studied via the error system

$$E = I - T = I - WQS^*. \tag{18}$$

For any input $x \in \mathcal{H}$, the reconstruction error signal is given as

$$Ex = x - x_r.$$

B. Orthogonal Projection

Consider the optimal reconstruction of signal x by the GSRP. Since $x_r \in \mathcal{W}$, the (norm of) error Ex is minimized by its orthogonal projection on $\mathcal{W}$:

$$x_r = P_{\mathcal{W}}x$$

and the optimal error system is

$$E_{\text{opt}} = I - P_{\mathcal{W}} = P_{\mathcal{W}^\perp}.$$

For each $x \in \mathcal{H}$, the optimal error signal is

$$E_{\text{opt}}x = P_{\mathcal{W}^\perp}x.$$

Under the frame assumption of $\{w_n\}$, $P_{\mathcal{W}}$ can be represented in terms of analysis and synthesis operators as [6]

$$P_{\mathcal{W}} = W(W^*W)^\dagger W^* \tag{19}$$

where "$\dagger$" denotes the Moore-Penrose pseudoinverse.

According to [6], the orthogonal projection $P_{\mathcal{W}}$ is subject to a fundamental limitation on the GSRP: Unless the reconstruction subspace is a subset of the sampling subspace, i.e.,

$$\mathcal{W} \subseteq \mathcal{S} \tag{20}$$

there exists no correction filter Q that renders the GSRP T to be the orthogonal projection $P_{\mathcal{W}}$.

Acknowledging the optimality as well as the limitation of the orthogonal projection, we now introduce the difference between the GSRP T and $P_{\mathcal{W}}$, which is, in the spirit of [6], referred to as the regret-error system:

$$R = P_{\mathcal{W}} - T = P_{\mathcal{W}} - WQS^*. \tag{21}$$

And the regret-error signal is given as

$$Rx = P_{\mathcal{W}}x - x_r = (P_{\mathcal{W}} - WQS^*)x.$$

It is important to note that the two error systems are related as

$$E = R + P_{\mathcal{W}^\perp}. \tag{22}$$

As the optimal sampling, orthogonal projection $P_{\mathcal{W}}$ enjoys two desirable properties:
  1) Error-free in $\mathcal{W}$: i.e., $Ex = 0$ for any $x \in \mathcal{W}$; and
  2) Least-error for $x \in \mathcal{H}$: i.e., $Ex = E_{\text{opt}}x$ for any $x \in H$.
As a result, $\|Ex\| \le \|x\|$ for any $x \in \mathcal{H}$.

C. Consistent Sampling

Consistent sampling achieves the error-free (in $\mathcal{W}$) property of the orthogonal projection without requiring the inclusion condition (20).

Assume that the following direct-sum condition holds

$$\mathcal{W} \oplus \mathcal{S}^\perp = \mathcal{H}. \tag{23}$$

Then, the correction filter

$$Q_{\text{con}} = (S^*W)^\dagger \tag{24}$$

provides an error-free reconstruction for input signals in $\mathcal{W}$ [9].

The resulted GSRP is found to be an oblique projection

$$T_{\text{con}} = W(S^*W)^\dagger S^* = P_{\mathcal{WS}^\perp}. \tag{25}$$

As a result, it is sample consistent, i.e.,

$$S^*(T_{\text{con}}x) = S^*(x - P_{\mathcal{S}^\perp\mathcal{W}}) = S^*x, \; x \in \mathcal{H},$$

where we used (2) and the fact that $\mathcal{N}(S^*) = \mathcal{S}^\perp$.

The error system is

$$E_{\text{con}} = I - P_{\mathcal{WS}^\perp} = P_{\mathcal{S}^\perp\mathcal{W}} \tag{26}$$

and the regret-error system also has a simple form:

$$R_{\text{con}} = P_{\mathcal{W}}P_{\mathcal{S}^\perp\mathcal{W}} \tag{27}$$

since, from (21), (3), and (5), we have

$$
\begin{aligned}
R_{\mathrm{con}} &= P_{\mathcal{W}} - T_{\mathrm{con}} \\
&= P_{\mathcal{W}} - P_{\mathcal{W}\mathcal{S}^\perp} \\
&= P_{\mathcal{W}} - P_{\mathcal{W}} P_{\mathcal{W}\mathcal{S}^\perp} \\
&= P_{\mathcal{W}}(I - P_{\mathcal{W}\mathcal{S}^\perp}) \\
&= P_{\mathcal{W}} P_{\mathcal{S}^\perp \mathcal{W}}.
\end{aligned}
$$

Then, $E_{\mathrm{con}}x = R_{\mathrm{con}}x = 0$ for any $x \in \mathcal{W}$.

The absolute error for each input can be derived as follows:

$$
\|E_{\mathrm{con}}x\|^2 = \|P_{\mathcal{W}^\perp}x\|^2 + \|P_{\mathcal{W}}P_{\mathcal{S}^\perp\mathcal{W}}x\|^2, \quad x \in \mathcal{H}.
$$

And the regret-error is

$$
\|R_{\mathrm{con}}x\| = \|P_{\mathcal{W}}P_{\mathcal{S}^\perp\mathcal{W}}x\|, \quad x \in \mathcal{H}.
$$

From [6], the absolute error can be bounded in terms of the subspace angles as

$$
\frac{E_{\mathrm{opt}}x}{\sin(\mathcal{W}^\perp,\mathcal{S})} \le \|E_{\mathrm{con}}(x)\| \le \frac{E_{\mathrm{opt}}x}{\cos(\mathcal{W},\mathcal{S})}. \tag{28}
$$

The regret-error is shown in Section IV to be bounded as

$$
\frac{\cos(\mathcal{W}^\perp,\mathcal{S})}{\sin(\mathcal{W}^\perp,\mathcal{S})}\|P_{\mathcal{W}^\perp}x\| \le \|R_{\mathrm{con}}x\| \le \frac{\sin(\mathcal{W},\mathcal{S})}{\cos(\mathcal{W},\mathcal{S})}\|P_{\mathcal{W}^\perp}x\|. \tag{29}
$$

It is clear from the left-hand sides of (28) and (29) that absolute error and regret-error for $x \in \mathcal{W}^\perp$ can be arbitrarily large if $(\mathcal{W}^\perp,\mathcal{S})$ is arbitrarily close to zero.

D. Subspace Sampling

The result on consistent sampling in the preceding section can be extended to any reconstruction subspace $\mathcal{A} \subset \mathcal{H}$ that satisfies the direct-sum condition with $\mathcal{S}^\perp$, i.e., $\mathcal{A} \oplus \mathcal{S}^\perp = \mathcal{H}$.

Let $\{a_n\}$ be a frame of subspace $\mathcal{A}$. Denote the corresponding synthesis operator by A. Then the correction filter

$$
Q_{\mathrm{sub}} = (W^*W)^\dagger W^* A (S^*A)^\dagger. \tag{30}
$$

renders the GSRP to be the product of two projection operators:

$$
T_{\mathrm{sub}} = W(W^*W)^\dagger W^* A (S^*A)^\dagger S^* = P_{\mathcal{W}}P_{\mathcal{A}\mathcal{S}^\perp}. \tag{31}
$$

The regret-error system now is

$$
R_{\mathrm{sub}} = P_{\mathcal{W}} - T_{\mathrm{sub}} = P_{\mathcal{W}} - P_{\mathcal{W}}P_{\mathcal{A}\mathcal{S}^\perp} = P_{\mathcal{W}}P_{\mathcal{S}^\perp\mathcal{A}}. \tag{32}
$$

and the error system is

$$
E_{\mathrm{sub}} = P_{\mathcal{W}^\perp} + P_{\mathcal{W}}P_{\mathcal{S}^\perp\mathcal{A}}. \tag{33}
$$

Accordingly, the absolute error and the regret-error are given, respectively, by

$$
\|E_{\mathrm{sub}}x\|^2 = \|P_{\mathcal{W}^\perp}x\|^2 + \|P_{\mathcal{W}}P_{\mathcal{S}^\perp\mathcal{A}}x\|^2, \quad x \in \mathcal{H}
$$

and

$$
\|R_{\mathrm{sub}}x\| = \|P_{\mathcal{W}}P_{\mathcal{S}^\perp\mathcal{A}}x\|, \quad x \in \mathcal{H}.
$$

And regret-error verifies the following error bounds:

$$
\frac{\cos(\mathcal{W}^\perp,\mathcal{S})}{\sin(\mathcal{A}^\perp,\mathcal{S})}\|P_A^\perp x\| \le \|R_{\mathrm{sub}}x\| \le \frac{\sin(\mathcal{W},\mathcal{S})}{\cos(\mathcal{A},\mathcal{S})}\|P_A^\perp x\| \tag{34}
$$

which will be shown in Section IV.

For any $x \in \mathcal{A}$, it holds $P_{\mathcal{S}^\perp\mathcal{A}}x = 0$, thus $E_{\mathrm{sub}}x = E_{\mathrm{opt}}x$ and $R_{\mathrm{sub}}x = 0$. This implies that the optimum reconstruction is achieved for any $x \in \mathcal{A}$. However, the reconstruction error of $E_{\mathrm{sub}}x$ for $x \in \mathcal{A}^\perp$ can still be very large, which can be seen from (34) when angle $(\mathcal{A}^\perp,\mathcal{S})$ is very small.

Recall that, filter Q_{sub} is the minimizer of the reconstruction error for input $x \in \mathcal{A}$, since it is the solution to the following optimization problem [6]:

$$
\min_Q \max_{x \in \mathcal{D}_{\mathcal{A}}} \|Ex\| \tag{35}
$$

where

$$
\mathcal{D}_{\mathcal{A}} = \{x \in \mathcal{A} \colon \|x\| \le L, c = S^*x\} \tag{36}
$$

with $c \in \ell^2$ representing the given sample sequence of input signal x, and scalar $L > 0$ being used as a bound of x so that the objective function in problem (35) is bounded.

E. Minimax Regret Sampling

Introduced in [6], the minimax regret sampling alleviates the drawback of large error associated with the consistent and subspace samplings. This is achieved by minimizing the maximum regret-error rather than the absolute error.

Consider the optimization problem:

$$
\min_Q \max_{x \in \mathcal{D}} \|Rx\| \tag{37}
$$

where

$$
\mathcal{D} = \{x \in \mathcal{H} \colon \|x\| \le L, c = S^*x\}. \tag{38}
$$

Solution to (37) is found to be

$$
Q_{\mathrm{reg}} = (W^*W)^\dagger W^* S (S^*S)^\dagger. \tag{39}
$$

Consequently, the GSRP becomes the product of two orthogonal projections

$$
T_{\mathrm{reg}} = W Q_{\mathrm{reg}} S^* = P_{\mathcal{W}} P_{\mathcal{S}}. \tag{40}
$$

Hence, the regret-error system is

$$
R_{\mathrm{reg}} = P_{\mathcal{W}} - T_{\mathrm{reg}} = P_{\mathcal{W}} P_{\mathcal{S}^\perp}. \tag{41}
$$

And the error system is

$$
E_{\mathrm{reg}} = P_{\mathcal{W}^\perp} + P_{\mathcal{W}} P_{\mathcal{S}^\perp}. \tag{42}
$$

Moreover, the regret-error is shown in [6] to be bounded as

$$
\cos(\mathcal{W}^\perp,\mathcal{S})\|P_{\mathcal{S}^\perp}x\| \le \|R_{\mathrm{reg}}x\| \le \sin(\mathcal{W},\mathcal{S})\|P_{\mathcal{S}^\perp}x\|. \tag{43}
$$

Clearly,

$$
\|R_{\mathrm{reg}}x\| \le \|x\|, \quad x \in \mathcal{H}. \tag{44}
$$

And

$$
\|E_{\mathrm{reg}}x\| \le \sqrt{2}\|x\|, \quad x \in \mathcal{H} \tag{45}
$$

since

$$
\|E_{\mathrm{reg}}x\|^2 \le \left(1 + \sin^2(\mathcal{W},\mathcal{S})\right)\|P_{\mathcal{S}^\perp}x\|^2.
$$

The above error estimates imply that T_{reg} results in good reconstruction for $x \in \mathcal{H}$, at the cost of introducing error for $x \in \mathcal{W}$ (or $\mathcal{A}$). Since it does not differentiate any input signals, it could be very conservative for signals in the input subspace.

IV. Constrained Reconstruction

Suppose that we know *a prior* that input signal x is close to $\mathcal{A}$ (i.e., (x, A) is small), and not necessarily lies in $\mathcal{A}$. This is relevant since in many practical scenarios, input signals cannot be exactly modeled as elements in $\mathcal{A}$. For example when $\mathcal{A}$ is learned via training set and only approximately described as an input subspace. It is also technically necessary when, for example, the sampling hardware is unable to sample at Nyquist rate or the input signal is only approximately bandlimited. We can seek a correction filter to improve the conservativeness of the regret sampling, and in the meantime to achieving minimum error for $x \in \mathcal{A}$ as in the case of subspace sampling. In other words, we wish to reach a trade-off between the two properties of orthorgonal projection $P_\mathcal{W}$.

For this end, we propose the following optimization problem

$$\min_Q \max_{x \in \mathcal{D}_\mathcal{A}} \|Ex\| \tag{46}$$
$$\text{s.t.} \max_{x \in \mathcal{D}} \|Rx\| \leq \beta_0(c)$$

where $\mathcal{D}_\mathcal{A}$ and $\mathcal{D}$ are given in (36) and (38) respectively. By assuming that x belongs to $\mathcal{D}$ or $\mathcal{D}_\mathcal{A}$, we imply that the sample sequence is given (see [6]). The optimization problem (46) makes very good sense because the objective part of it takes care of optimum reconstruction in $\mathcal{A}$ and the constraint part reflects minimax recovery for inputs in the entire signal space.

The regret-error in the constraint above can be relaxed with the error between the GSRP itself and the minimax regret reconstruction rather than the orthogonal projection, i.e.,

$$\|P_\mathcal{W} P_\mathcal{S} x - WQS^*x\| = \|P_\mathcal{W} S(S^*S)^\dagger c - WQc\|.$$

Not only would this realization allow a simple and elegant solution to our search for an alternative sampling scheme, it is also supported by the following arguments. On one hand, from triangular inequality, we have

$$\begin{aligned}
\max_{x \in \mathcal{D}} \|Rx\| &= \max_{x \in \mathcal{D}} \|P_\mathcal{W} x - WQS^*x\| \\
&\leq \max_{x \in \mathcal{D}} \Big\{ \|P_\mathcal{W} P_\mathcal{S} x - WQS^*x\| \\
&\quad + \|P_\mathcal{W} x - P_\mathcal{W} P_\mathcal{S} x\| \Big\} \\
&= \|P_\mathcal{W} S(S^*S)^\dagger c - WQc\| \\
&\quad + \max_{x \in \mathcal{D}} \|P_\mathcal{W} x - P_\mathcal{W} P_\mathcal{S} x\|. \tag{47}
\end{aligned}$$

On the other hand, it is shown in Appendix A that

$$\begin{aligned}
\max_{x \in \mathcal{D}} \|Rx\| &\geq \frac{1}{\sqrt{2}} \Big(\|P_\mathcal{W} S(S^*S)^\dagger c - WQc\| \\
&\quad + \max_{x \in \mathcal{D}} \|P_\mathcal{W} x - P_\mathcal{W} P_\mathcal{S} x\| \Big). \tag{48}
\end{aligned}$$

We complete the argument by noting that the last terms in (47) and (48) are independent of correction filter Q.

In view of the above discussions, we now present the constrained optimization problem:

$$\min_Q \max_{x \in \mathcal{D}_\mathcal{A}} \|x - WQS^*x\| \tag{49}$$
$$\text{s.t.} \|P_\mathcal{W} S(S^*S)^\dagger c - WQc\| \leq \beta_1(c)$$

which would lead to an adequate approximation of the constraint in (46).

The upper bound $\beta_1(c)$ in (49) needs to be properly chosen. Let us consider two extreme cases: $\beta_1(c) = 0$ and $\beta_1(c) = \infty$. If $\beta_1(c) = 0$, the strict constraint implies that the solution to (49) is the standard minimax regret filter in (39). On the other hand, if $\beta_1(c) = \infty$, implying that the constraint is removed, then the optimization problem (49) reduces to the subspace sampling problem (35). Hence, the solution is simply the subspace filter Q_{sub} in (30) and the left-hand side of the constraint part in (49) becomes

$$\beta(c) = \|P_\mathcal{W} S(S^*S)^\dagger c - P_\mathcal{W} A(S^*A)^\dagger c\|. \tag{50}$$

From the above discussions, we conclude that the upper bound in (49) can be set to be $\beta_1(c) = \lambda\beta(c)$ for some parameter $\lambda \in [0, 1]$. Consequently, the constrained optimization problem (49) reduces to

$$\min_Q \max_{x \in \mathcal{D}_\mathcal{A}} \|x - WQS^*x\| \tag{51}$$
$$\text{s.t.} \|P_\mathcal{W} S(S^*S)^\dagger c - WQc\| \leq \lambda\beta(c).$$

In Appendix B, we show that a solution to the above problem is

$$Q_\lambda = \lambda Q_{\text{sub}} + (1 - \lambda)Q_{\text{reg}}. \tag{52}$$

Accordingly, the GSRP can be expressed as

$$T_\lambda = \lambda T_{\text{sub}} + (1 - \lambda)T_{\text{reg}}. \tag{53}$$

Now let us examine the constrained GSRP T_λ. Let B be the convex combination of two projections:

$$B = \lambda P_{\mathcal{A}\mathcal{S}^\perp} + (1 - \lambda)P_\mathcal{S}. \tag{54}$$

In view of (31) and (40), we obtain a compact expression for the GSRP:

$$T_\lambda = P_\mathcal{W} B. \tag{55}$$

The next result, which is proved in Appendix C, states that B is in fact also an oblique projection with the nullspace being $\mathcal{S}^\perp$.

Proposition 2: The linear operator B defined in (53) is given as

$$B = P_{\mathcal{B}\mathcal{S}^\perp}. \tag{56}$$

where $\mathcal{B} = \mathcal{R}(B)$.

Following Proposition 2, the resulting constrained GSRP can be described as the product of two projections:

$$T_\lambda = P_\mathcal{W} P_{\mathcal{B}\mathcal{S}^\perp}. \tag{57}$$

Then, the regret-error system is

$$R_\lambda = P_\mathcal{W} P_{\mathcal{S}^\perp \mathcal{B}}. \tag{58}$$

and the error system is given as

$$E_\lambda = P_{\mathcal{W}^\perp} + P_\mathcal{W} P_{\mathcal{S}^\perp \mathcal{B}}. \tag{59}$$

In view of (22), Similar to the case of subspace sampling, for any $x \in \mathcal{H}$, the error is given by

$$\|E_\lambda x\|^2 = \|P_{\mathcal{W}^\perp} x\|^2 + \|P_\mathcal{W} P_{\mathcal{S}^\perp \mathcal{B}} x\|^2 \tag{60}$$
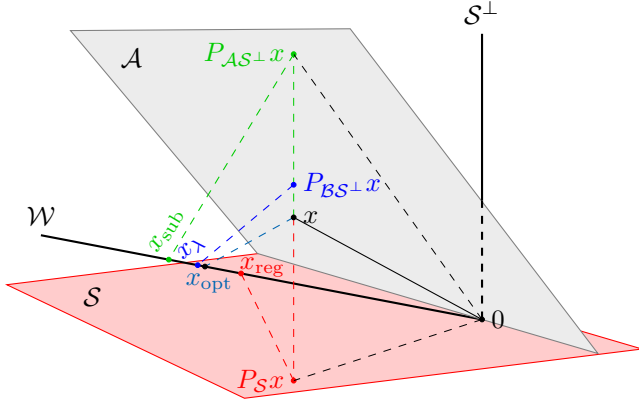
Fig. 3. An illustration of sampling schemes: $\mathcal{S}$ is sampling space, $\mathcal{W}$ is reconstruction space and $\mathcal{A}$ is input space. $x_{\text{opt}} = P_{\mathcal{W}}x$, $x_{\text{sub}} = P_{\mathcal{W}}P_{\mathcal{A}\mathcal{S}^\perp}x$, $x_{\text{reg}} = P_{\mathcal{W}}P_{\mathcal{S}}x$, and $x_\lambda = P_{\mathcal{W}}P_{\mathcal{B}\mathcal{S}^\perp}x$, where $P_{\mathcal{B}\mathcal{S}^\perp} = \lambda P_{\mathcal{A}\mathcal{S}^\perp} + (1-\lambda)P_{\mathcal{S}}$. Note that the constrained reconstruction x_λ has the potential to approach optimum reconstruction x_{opt}.

and the regret-error is

$$\|R_\lambda x\| = \|P_{\mathcal{W}}P_{\mathcal{S}^\perp\mathcal{B}}x\|. \tag{61}$$

It is interesting to see that all the GSRPs discussed have the same expression as in (57). When $\lambda = 0$, then $\mathcal{B} = \mathcal{S}$ and $T_\lambda = T_{\text{reg}}$; and when $\lambda = 1$, then $\mathcal{B} = \mathcal{A}$ and $T_\lambda = T_{\text{sub}}$, which becomes T_{con} if additionally $\mathcal{A} = \mathcal{W}$. This shows that our constrained sampling generalizes all the other three samplings.

Furthermore, (57) shows that the constrained sampling is essentially a subspace sampling with a new modified subspace $\mathcal{B}$, which is comprised of all convex combinations of vectors of $\mathcal{A}$ and $\mathcal{S}$ from (54). Thus $\mathcal{B}$ is closer to $\mathcal{S}$ than $\mathcal{A}$ is, i.e., $(\mathcal{B}, \mathcal{S}) < (\mathcal{A}, \mathcal{S})$, leading to more robust sampling strategy (i.e., better reconstruction for signals not in $\mathcal{A}$). Further explanations on this observation will be given in Section V following the error analysis. A geometrical illustration of all the sampling schemes is provided in Fig. 3.

It should be noted that since $P_{\mathcal{S}^\perp\mathcal{B}}$ is still an *oblique* projection, the error Ex can still be very large in general. However, we shall show in the next section that this concern can be removed by properly choosing the value of parameter λ, one such choice is $\lambda = \cos(\mathcal{A}, \mathcal{S})$.

V. ANALYSIS ON RECONSTRUCTION ERROR

We now derive bounds on regret-error of the constrained sampling T_λ by examining the regret-error system R_λ. These bounds specialize the bounds for the other sampling schemes if $\lambda = 0, 1$. Bounds on absolute error for signal in $\mathcal{A}$ are also provided.

First of all, since $\mathcal{R}(P_{\mathcal{S}^\perp\mathcal{B}}) = \mathcal{S}^\perp$, it follows from (61) and (15) that

$$\cos(\mathcal{W}^\perp, \mathcal{S})\|P_{\mathcal{S}^\perp\mathcal{B}}x\| \leq \|R_\lambda x\| \leq \sin(\mathcal{W}, \mathcal{S})\|P_{\mathcal{S}^\perp\mathcal{B}}x\|, \; x \in \mathcal{H}. \tag{62}$$

Moreover, from (16) it follows that

$$\frac{\|P_{\mathcal{B}^\perp}x\|}{\sin(\mathcal{B}^\perp, \mathcal{S})} \leq \|P_{\mathcal{S}^\perp\mathcal{B}}x\| \leq \frac{\|P_{\mathcal{B}^\perp}x\|}{\cos(\mathcal{B}, \mathcal{S})} \tag{63}$$

where we replace $\sin(\mathcal{S}^\perp, \mathcal{B})$ by $\sin(\mathcal{B}^\perp, \mathcal{S})$ in view of (12). Consequently, the regret-error enjoys the following estimates

$$\frac{\cos(\mathcal{W}^\perp, \mathcal{S})}{\sin(\mathcal{B}^\perp, \mathcal{S})}\|P_{\mathcal{B}^\perp}x\| \leq \|R_\lambda x\| \leq \frac{\sin(\mathcal{W}, \mathcal{S})}{\cos(\mathcal{B}, \mathcal{S})}\|P_{\mathcal{B}^\perp}x\|. \tag{64}$$

The above bounds can be further simplified by applying the following estimates of the trigonometrical functions involving subspace $\mathcal{B}$:

$$\frac{1}{1 + \lambda^2 \frac{\sin^2(\mathcal{A}, \mathcal{S})}{\cos^2(\mathcal{A}, \mathcal{S})}} \leq \cos^2(\mathcal{B}, \mathcal{S}) \leq \frac{1}{1 + \lambda^2 \frac{\cos^2(\mathcal{S}^\perp, \mathcal{A})}{\sin^2(\mathcal{S}^\perp, \mathcal{A})}} \tag{65}$$

and

$$\frac{1}{1 + \lambda^2 \frac{\sin^2(\mathcal{A}, \mathcal{S})}{\cos^2(\mathcal{A}, \mathcal{S})}} \leq \sin^2(\mathcal{B}^\perp, \mathcal{S}) \leq \frac{1}{1 + \lambda^2 \frac{\cos^2(\mathcal{S}^\perp, \mathcal{A})}{\sin^2(\mathcal{S}^\perp, \mathcal{A})}} \tag{66}$$

which are proved in Appendices D and E, respectively.

It is important to point out that $(\mathcal{B}, \mathcal{S}) \leq (\mathcal{A}, \mathcal{S})$ for any $\lambda \in [0, 1]$ since $\cos(\mathcal{B}, \mathcal{S}) \geq \cos(\mathcal{A}, \mathcal{S})$ in view of (65). In other words, the modified subspace $\mathcal{B}$ is closer to $\mathcal{S}$ than input subspace $\mathcal{A}$ is. This explains from another perspective why constrained sampling would generally lead smaller maximum possible error than subspace sampling.

We finally obtain lower and upper bounds on the regret-error:

$$\left(1 + \lambda^2 \frac{\cos^2(\mathcal{S}^\perp, \mathcal{A})}{\sin^2(\mathcal{S}^\perp, \mathcal{A})}\right)^{\frac{1}{2}} \cos(\mathcal{W}^\perp, \mathcal{S})\|P_{\mathcal{B}^\perp}x\|$$
$$\leq \|R_\lambda x\|$$
$$\leq \left(1 + \lambda^2 \frac{\sin^2(\mathcal{A}, \mathcal{S})}{\cos^2(\mathcal{A}, \mathcal{S})}\right)^{\frac{1}{2}} \sin(\mathcal{W}, \mathcal{S})\|P_{\mathcal{B}^\perp}x\|. \tag{67}$$

It is pointed out that with a simple choice of parameter

$$0 \leq \lambda \leq \cos(\mathcal{A}, \mathcal{S}) \tag{68}$$

the reconstruction error in (67) is seen to be bounded as below:

$$\|R_\lambda x\| \leq \sqrt{2}\|x\|, \quad x \in \mathcal{H}. \tag{69}$$

Then, the absolute error is bounded as

$$\|E_\lambda x\| \leq \sqrt{3}\|x\|, \quad x \in \mathcal{H}. \tag{70}$$

We now turn to bounds on errors for signal in subspace $\mathcal{A}$. Let $x \in \mathcal{A}$. Then

$$\begin{aligned}
\|R_\lambda x\| &= \|P_{\mathcal{W}}P_{\mathcal{S}^\perp\mathcal{B}}x\| \\
&= \|P_{\mathcal{W}}[\lambda P_{\mathcal{S}^\perp\mathcal{A}} + (1-\lambda)P_{\mathcal{S}^\perp}]x\| \\
&= (1-\lambda)\|P_{\mathcal{W}}P_{\mathcal{S}^\perp}x\|n \qquad (71) \\
&\leq (1-\lambda)\sin(\mathcal{S}^\perp, \mathcal{W}^\perp)\|P_{\mathcal{S}^\perp}x\|
\end{aligned}$$

where the first step is from (61) and the second step is from (54). Thus, using (12) and (14), we obtain an upper bound on regret-error

$$\|R_\lambda x\| \leq (1-\lambda)\sin(\mathcal{W}, \mathcal{S})\sin(\mathcal{A}, \mathcal{S})\|x\|, \; x \in \mathcal{A}. \tag{72}$$

Similarly, we can also obtain a lower bound on regret-error

$$\|R_\lambda x\| \geq (1-\lambda)\cos(\mathcal{S}^\perp, \mathcal{W})\cos(\mathcal{A}, \mathcal{S}^\perp)\|x\|, \; x \in \mathcal{A} \tag{73}$$

TABLE II
SAMPLING STRATEGIES AND THEIR REGRET-ERRORS

Sampling Scheme	GSRP T	Correction Filter Q	Regret Error $\|Rx\| = \|P_{\mathcal{W}}x - Tx\|$[a]		
			Expression	Lower Bound	Upper Bound
Orthogonal[b]	$P_{\mathcal{W}}$	$(W^*W)^\dagger W^* S(S^*S)^\dagger$	0	0	0
Consistent	$P_{\mathcal{W}\mathcal{S}^\perp}$	$(S^*W)^\dagger$	$\|P_{\mathcal{W}}P_{\mathcal{S}^\perp\mathcal{W}}x\|$	$\frac{\cos(\mathcal{S}^\perp,\mathcal{W})}{\sin(\mathcal{S}^\perp,\mathcal{W})}\|P_{\mathcal{W}^\perp}x\|$	$\frac{\sin(\mathcal{W},\mathcal{S})}{\cos(\mathcal{W},\mathcal{S})}\|P_{\mathcal{W}^\perp}x\|$
Subspace	$P_{\mathcal{W}}P_{\mathcal{A}\mathcal{S}^\perp}$	$(W^*W)^\dagger W^* A(S^*A)^\dagger$	$\|P_{\mathcal{W}}P_{\mathcal{S}^\perp\mathcal{A}}x\|$	$\frac{\cos(\mathcal{S}^\perp,\mathcal{A})}{\sin(\mathcal{S}^\perp,\mathcal{A})}\|P_{\mathcal{A}^\perp}x\|$	$\frac{\sin(\mathcal{W},\mathcal{S})}{\cos(\mathcal{A},\mathcal{S})}\|P_{\mathcal{A}^\perp}x\|$
Regret	$P_{\mathcal{W}}P_{\mathcal{S}}$	$(W^*W)^\dagger W^* S(S^*S)^\dagger$	$\|P_{\mathcal{W}}P_{\mathcal{S}^\perp}x\|$	$\cos(\mathcal{S}^\perp,\mathcal{W})\|P_{\mathcal{S}^\perp}x\|$	$\sin(\mathcal{W},\mathcal{S})\|P_{\mathcal{S}^\perp}x\|$
Constrained[c]	$P_{\mathcal{W}}P_{\mathcal{B}\mathcal{S}^\perp}, \lambda \in [0,1].$	$\lambda(W^*W)^\dagger W^* A(S^*A)^\dagger$ $+(1-\lambda)(W^*W)^\dagger W^* S(S^*S)^\dagger$	$\|P_{\mathcal{W}}P_{\mathcal{S}^\perp\mathcal{B}}x\|$	$\left(1+\lambda^2\frac{\cos^2(\mathcal{S}^\perp,\mathcal{A})}{\sin^2(\mathcal{S}^\perp,\mathcal{A})}\right)^{\frac{1}{2}}$ $\times\cos(\mathcal{S}^\perp,\mathcal{W})\|P_{\mathcal{B}^\perp}x\|$	$\left(1+\lambda^2\frac{\sin^2(\mathcal{A},\mathcal{S})}{\cos^2(\mathcal{A},\mathcal{S})}\right)^{\frac{1}{2}}$ $\times\sin(\mathcal{W},\mathcal{S})\|P_{\mathcal{B}^\perp}x\|$
Constrained $x \in \mathcal{A}$	$P_{\mathcal{W}}-(1-\lambda)P_{\mathcal{W}}P_{\mathcal{S}^\perp}$	$\lambda(W^*W)^\dagger W^* A(S^*A)^\dagger$ $+(1-\lambda)(W^*W)^\dagger W^* S(S^*S)^\dagger$	$(1-\lambda)$ $\times\|P_{\mathcal{W}}P_{\mathcal{S}^\perp}x\|$	$(1-\lambda)\cos(\mathcal{S},\mathcal{A}^\perp)$ $\times\cos(\mathcal{S}^\perp,\mathcal{W})\|x\|$	$(1-\lambda)\sin(\mathcal{A},\mathcal{S})$ $\times\sin(\mathcal{W},\mathcal{S})\|x\|$

[a]Note that the absolute error is given by $\|Ex\|^2 = \|x - Tx\|^2 = \|P_{\mathcal{W}^\perp}x\|^2 + \|Rx\|^2$.

[b]This is the optimal sampling scheme but possible only if $\mathcal{W} \subseteq \mathcal{S}$. The reconstruction error is $\|Ex\| = \|P_{\mathcal{W}}^\perp x\|$.

[c]The modified subspace is $\mathcal{B} = \mathcal{R}\left(\lambda P_{\mathcal{A}\mathcal{S}^\perp} + (1-\lambda)P_{\mathcal{S}}\right)$.

It then follows, from (22), (60), and (14), that absolute error are bounded as

$$\left(\cos^2(\mathcal{A},\mathcal{W}^\perp) + (1-\lambda)^2\cos^2(\mathcal{S}^\perp,\mathcal{W})\cos^2(\mathcal{A},\mathcal{S}^\perp)\right)^{\frac{1}{2}}\|x\|$$
$$\leq \|E_\lambda x\| \leq$$
$$\left(\sin^2(\mathcal{A},\mathcal{W}) + (1-\lambda)^2\sin(\mathcal{W},\mathcal{S})\sin(\mathcal{A},\mathcal{S})\right)^{\frac{1}{2}}\|x\|, x \in \mathcal{A}$$

Table II summaries key results on all the sampling schemes considered in this paper.

VI. EXAMPLES

We now provide two illustrative examples in which reconstruction of a typical Gaussian signal and a speech signal are studied. These examples demonstrate the effectiveness of the proposed constrained sampling.

A. Gaussian Signal

Consider reconstruction of a Gaussian signal of unit energy:

$$x = \left(\frac{1}{\pi\sigma}\right)^{1/4}\exp(\frac{-t^2}{2\sigma}), \tag{74}$$

where $\sigma = 0.09$.

Assume that sampling period T is one (i.e., the Nyquist radian frequency is π) and the sampling space $\mathcal{S}$ is the shiftable subspace generated by the B-spline of order zero:

$$s(t) = \beta^0(t) = \begin{cases} 1, & t \in [-0.5, 0.5) \\ 0, & \text{otherwise.} \end{cases} \tag{75}$$

In other words, $\mathcal{S}$ is spanned by frame vectors $\{\beta^0(t-n)\}_{n\in\mathbb{Z}}$. Since x has its 94% energy in the content of frequencies up to π, it is reasonable to assume that $\mathcal{A}$ to be the subspace of π-bandlimited signals. In this situation, we have $\cos(\mathcal{A},\mathcal{S}) = 0.64$, which can be calculated [7] by

$$\cos^2(\mathcal{A},\mathcal{S}) = \inf_{\omega\in[0,2\pi)}\frac{\left|\sum_{n\in\mathbb{Z}}\widehat{s}^*(\omega+2\pi n)\widehat{a}(\omega+2\pi n)\right|^2}{\sum_{n\in\mathbb{Z}}|\widehat{s}(\omega+2\pi n)|^2\sum_{n\in\mathbb{Z}}|\widehat{a}(\omega+2\pi n)|^2}$$

where "$\widehat{\cdot}$" represents the Fourier transform, and $a(t) = \text{sinc}(t)$).

We let reconstruction space $\mathcal{W}$ be the shiftable subspace generated by the cubic B-splines [21]

$$w(t) = \beta^3(t) = [\beta^0 * \beta^0 * \beta^0 * \beta^0](t) \tag{76}$$

where $*$ is the convolution operator.

Fig. 4 presents the signal-to-noise ratio (SNR) in dB[1] of the reconstruction error Ex for the three sampling schemes. We can observe from Fig. 4 that 1) the performance of the constrained sampling is never below that of the minimax regret regardless of the value of λ, demonstrating the conservativeness of the regret sampling for inputs close to $\mathcal{A}$; 2) the constrained sampling achieves better reconstruction than the subspace sampling for any $\lambda \in (0.20, 1)$; 3) with the simple choice of $\lambda = \cos(\mathcal{A},\mathcal{S}) = 0.64$, the improvement of the constrained sampling over the subspace and minimax regret GSRPs are 1.26dB and 2.40dB, respectively.

The improvements are somehow surprising in view of the fact that only 6% of energy of x falls out of $\mathcal{A}$. It is more worth pointing out the existence of the optimal value (i.e., $\lambda = 0.60 \approx \cos(\mathcal{A},\mathcal{S})$) such that $\|E_\lambda x\|$ is very close to the optimal error $\|E_{\text{opt}}x\|$, demonstrating high potential of constrained sampling in approaching the orthogonal projection.
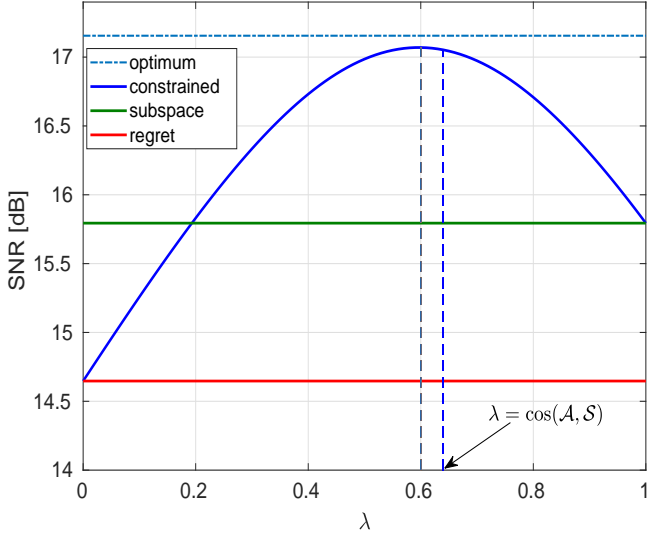
---

[1] SNR$= 20\log\left(\|x\|/\|Ex\|\right)$ dB

Fig. 4. Reconstruction error of a Gaussian signal for all sampling schemes ($\mathcal{S}$ and $\mathcal{W}$ are generated by β^0 and β^3, respectively, and $\mathcal{A}$ is the π-bandlimited subspace).

Fig. 5. Reconstruction error of a speech signal for all sampling schemes ($\mathcal{S}$ is generated by $\beta^0(t)$, $\mathcal{W}$ is generated by a non-ideal low-pass filter associated with a time-support of $[-4T, 4T]$ with $T = 4000^{-1}$s, and $\mathcal{A}$ is the 8kHz-bandlimited subspace).

B. Speech Signal

In this example, input signal is chosen to be a speech signal[2] which is sampled at the rate of 16kHz. Since the sampling rate is sufficiently high, the discrete-time speech signal $x[n]$ can approximate accurately the continuous-time signal $x(t)$ on the fine grid. We assume that the sampling process is an integration over one sampling duration T:

$$c[n] = \frac{1}{T} \int_{nT-T/2}^{nT+T/2} x(t)\mathrm{d}t,$$

where $T = 4000^{-1}$sec. This is equivalent to assuming $s(t) = \frac{1}{T}\beta^0(\frac{t}{T})$ or discrete-time filtering on the fine grid with filter whose impulse response is

$$s[k] = \begin{cases} \frac{1}{3}, & k = -1,\, 0,\, 1 \\ 0, & \text{otherwise}. \end{cases}$$

Since the original continuous-time signal is sampled at 16kHz, we assume that subspace $\mathcal{A}$ is the space of 8kHz-bandlimited signals. For calculation, we use a zero-phase discrete-time FIR low pass filter with cutoff frequency at $1/2$ and of order 100 to simulate $\mathcal{A}$ on the fine grid. The selected $\mathcal{A}$ is equivalent to continuous-time low-pass filter with support $t \in [-25T,\ 25T]$ which approximates $\operatorname{sinc}(4t/T)$. As a result of this approximation, the performance of subspace GSRP is lower than that of minimax GSRP. We obtain $\cos(\mathcal{A}, \mathcal{S}) = 0.55$. For the synthesis, we let $w_n(t) = w(t - nT)$, where $w(t)$ is chosen to have a time-support of $t \in [-4T,\ 4T]$ and to render a low pass filter with cutoff frequency (i.e., Nyquist frequency) $1/(2T)$. On the fine grid, this synthesis process is implemented via a discrete-time low-pass FIR filter of order 16 and with cutoff frequency $1/8$.

In the experiment, following [6], we randomly chose 5000 segments (each with 400 consecutive samples) of the speech
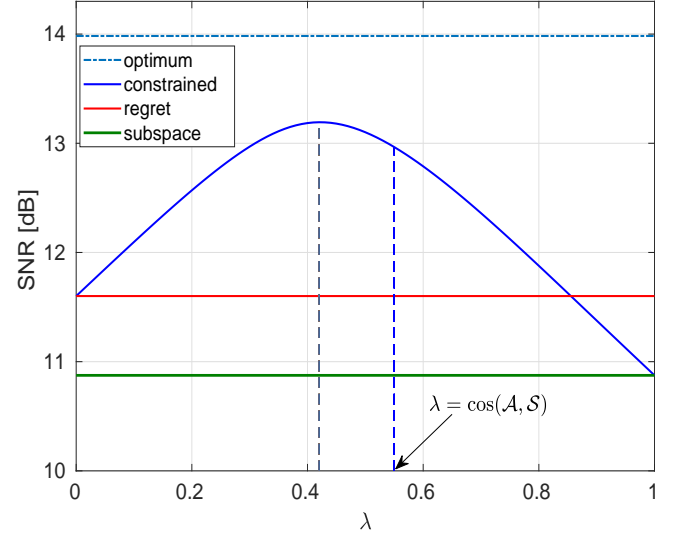
[2]downloaded from https://catalog.ldc.upenn.edu/

signal. The reconstruction errors of all the sampling schemes are shown in Fig. 5. We can observe that 1) the performance of constrained sampling is always better than that of the subspace sampling; 2) it is better than that of regret sampling when $\lambda \in [0, 0.85]$; 3) If $\lambda = \cos(\mathcal{A}, S) = 0.55$, the improvement over the subspace and minimax regret GSRPs are 2.00dB and 1.37dB, respectively. Also note at the optimum value of $\lambda = 0.42$, the performance of the constrained GSRP is only 0.81dB away from that of the orthogonal projection; which again shows the potential of the former in approaching the latter.

VII. Conclusions

This paper re-examined the sampling schemes for generalized sampling-reconstruction processes (GSRPs). Existing GSRP, namely, consistent, subspace, and minimax regret GSRPs, either assume that the *a priori* subspace is fully known or fully ignored. To address this limitation, we proposed, *constrained sampling*, a new sampling scheme that is designed to minimize the reconstruction error for inputs that lie within a known subspace while simultaneously bounding the maximum regret error for all other signals. The constrained sampling formulation leads to a convex combination of the subspace and the minimax regret samplings. It also yields an equivalent subspace sampling process with a modified input space. The constrained sampling is shown to be 1) (sub)optimal for signals in the input subspace, 2) robust for signals around the input subspace, 3) reasonably bounded for any signal in the entire space, and 4) flexible and easy to be implemented as combination of the subspace and regret samplings. We also presented a detailed theoretical analysis of reconstruction error of the proposed scheme. Additionally, we demonstrated the efficiency of constrained sampling through two illustrative examples. Our results suggest that the proposed scheme could potentially approach the optimum

reconstruction (i.e., the orthogonal projection). It would be intriguing to study the optimal selection of the parameter in the convex combination when more *a priori* information about input signals become available.

APPENDIX A
PROOF OF INEQUALITY (48)

As in the proof in [6, theorem 3], we represent any x in $\mathcal{D} = \{x : \|x\| \leq L, c = S^*x\}$ as

$$
\begin{aligned}
x &= P_{\mathcal{S}}x + P_{\mathcal{S}}^{\perp}x \\
&= S(S^*S)^{\dagger}c + v
\end{aligned}
$$

for some v in $\mathcal{G} = \{v \in \mathcal{S}^{\perp} : \|v\|^2 \leq L^2 - \|S(S^*S)^{\dagger}c\|^2\}$.

Let $a_c = WQc - P_{\mathcal{W}}S(S^*S)^{\dagger}c$. Then

$$
\begin{aligned}
\|Rx\|^2 &= \|P_{\mathcal{W}}x - WQS^*x\|^2 \\
&= \|P_{\mathcal{W}}S(S^*S)^{\dagger}c + P_{\mathcal{W}}v - WQc\|^2 \\
&= \|P_{\mathcal{W}}v - a_c\|^2 \\
&= \{\|P_{\mathcal{W}}v\|^2 - 2\mathrm{Re}\{\langle P_{\mathcal{W}}v, a_c\rangle\} + \|a_c\|^2\}.
\end{aligned}
$$

Let

$$
v_1 = -\frac{\langle P_{\mathcal{W}}v, a_c\rangle}{|\langle P_{\mathcal{W}}v, a_c\rangle|}v.
$$

Clearly, $\|v_1\| = \|v\|$ and $v_1 \in \mathcal{G}$ if and only if $v \in \mathcal{G}$. Consequently

$$
\begin{aligned}
&\max_{x \in \mathcal{D}} \|Rx\|^2 \\
=\ & \max_{v \in \mathcal{G}} \{\|P_{\mathcal{W}}v\|^2 + 2|\langle P_{\mathcal{W}}v, a_c\rangle| + \|a_c\|^2\} \\
\geq\ & \|a_c\|^2 + \max_{v \in \mathcal{G}} \|P_{\mathcal{W}}v\|^2 \\
=\ & \|a_c\|^2 + \max_{x \in \mathcal{D}} \|P_{\mathcal{W}}(x - P_{\mathcal{S}}x)\|^2 \\
=\ & \|WQc - P_{\mathcal{W}}S(S^*S)^{\dagger}c\|^2 + \max_{x \in \mathcal{D}} \|P_{\mathcal{W}}x - P_{\mathcal{W}}P_{\mathcal{S}}x\|^2.
\end{aligned}
$$

On the other hand, since for any complex numbers z_1 and z_2,

$$
|z_1|^2 + |z_2|^2 \geq \frac{1}{2}(|z_1| + |z_2|)^2,
$$

we get

$$
\begin{aligned}
\max_{x \in \mathcal{D}} \|Rx\| \geq\ & \frac{1}{\sqrt{2}}(\|WQc - P_{\mathcal{W}}S(S^*S)^{\dagger}c\| \\
& + \max_{x \in \mathcal{D}} \|P_{\mathcal{W}}x - P_{\mathcal{W}}P_{\mathcal{S}}x\|).
\end{aligned}
$$

The proof is complete. $\qquad\square$

APPENDIX B
PROOF OF SOLUTION OF (51)

Due to the direct-sum property ($\mathcal{A} + \mathcal{S}^{\perp} = \mathcal{H}$), $\mathcal{D}_{\mathcal{A}}$ contains only one element, that is $x = A(S^*A)^{\dagger}c$. Define

$$
\mathcal{D}_Q = \{Q : \|P_{\mathcal{W}}S(S^*S)^{\dagger}c - WQc\| \leq \lambda\beta(c)\}
$$

where $\beta(c)$ is given in (50) and $\lambda \in [0, 1]$. The optimization problem given in (51) becomes

$$
\min_{Q \in \mathcal{D}_Q} \|A(S^*A)^{\dagger}c - WQc\|^2. \tag{77}
$$

Invoking orthogonal decomposition of $A(S^*A)^{\dagger}c - WQc$ onto $\mathcal{W}$ and $\mathcal{W}^{\perp}$ and using the triangular inequality, we have for any $Q \in \mathcal{D}_Q$,

$$
\begin{aligned}
&\min_{Q \in \mathcal{D}_Q} \|A(S^*A)^{\dagger}c - WQc\|^2 \\
=\ & \min_{Q \in \mathcal{D}_Q} \|P_{\mathcal{W}}A(S^*A)^{\dagger}c - WQc\|^2 + \|P_{\mathcal{W}^{\perp}}A(S^*A)^{\dagger}c\|^2 \\
\geq\ & \|P_{\mathcal{W}^{\perp}}A(S^*A)^{\dagger}c\|^2 + \min_{Q \in \mathcal{D}_Q} \Big| \|P_{\mathcal{W}}S(S^*S)^{\dagger}c - WQc\| \\
& - \|P_{\mathcal{W}}S(S^*S)^{\dagger}c - P_{\mathcal{W}}A(S^*A)^{\dagger}c\| \Big|^2 \\
=\ & \min_{Q \in \mathcal{D}_Q} \big| \|P_{\mathcal{W}}S(S^*S)^{\dagger}c - WQc\| - \beta(c) \big|^2 \\
& + \|P_{\mathcal{W}^{\perp}}A(S^*A)^{\dagger}c\|^2 \\
\geq\ & (1 - \lambda)^2 B^2(c) + \|P_{\mathcal{W}^{\perp}}A(S^*A)^{\dagger}c\|^2. \tag{78}
\end{aligned}
$$

Substituting

$$
Q = \lambda Q_{\mathrm{sub}} + (1 - \lambda)Q_{\mathrm{reg}}
$$

into (77), we see that the lower bound in (78) is reached. That completes the proof. $\qquad\square$

APPENDIX C
PROOF OF PROPOSITION 2

Since $P_{\mathcal{AS}^{\perp}}$ and $P_{\mathcal{S}}$ have the same nullspace $\mathcal{S}^{\perp}$, applying Proposition 1 on B in (54) concludes B is also an projection.

It remains to be shown that $\mathcal{N}(B) = \mathcal{S}^{\perp}$. It suffices if we show that $Bx = 0$ if and only if $P_{\mathcal{S}}x = 0$, which can be proved by an alternative expression of B (in terms of $P_{\mathcal{S}}$ and $P_{\mathcal{S}^{\perp}\mathcal{A}}$):

$$
\begin{aligned}
B &= \lambda P_{\mathcal{AS}^{\perp}} + (1 - \lambda)P_{\mathcal{S}} \\
&= \lambda P_{\mathcal{AS}^{\perp}} + (1 - \lambda)P_{\mathcal{S}}P_{\mathcal{AS}^{\perp}} \\
&= [\lambda I + (1 - \lambda)P_{\mathcal{S}}]P_{\mathcal{AS}^{\perp}} \\
&= [P_{\mathcal{S}} + \lambda(I - P_{\mathcal{S}})]P_{\mathcal{AS}^{\perp}} \\
&= [P_{\mathcal{S}} + \lambda P_{\mathcal{S}^{\perp}}]P_{\mathcal{AS}^{\perp}} \\
&= P_{\mathcal{S}} + \lambda P_{\mathcal{S}^{\perp}}P_{\mathcal{AS}^{\perp}} \tag{79}
\end{aligned}
$$

where the second step is from (4), the second to the last step is due to (2), and the last step is from (4). For any $x \in \mathcal{H}$, since $P_{\mathcal{S}}x$ and $P_{\mathcal{S}^{\perp}}P_{\mathcal{AS}^{\perp}}x$ are perpendicular to each other, hence, the statement follows immediately. The proof is complete. $\square$

APPENDIX D
PROOF OF BOUNDS OF $\cos(\mathcal{B}, \mathcal{S})$ IN (65)

From (8),

$$
\cos^2(\mathcal{B}, \mathcal{S}) = \inf_{x \notin \mathcal{S}^{\perp}} f(x) \tag{80}
$$

where

$$
f(x) = \frac{\|P_{\mathcal{S}}Bx\|^2}{\|Bx\|^2}. \tag{81}
$$

Since $B = P_{\mathcal{S}} + \lambda P_{\mathcal{S}^{\perp}}P_{\mathcal{AS}^{\perp}}$ (see (79)),

$$
\begin{aligned}
f(x) &= \frac{\|P_{\mathcal{S}}(P_{\mathcal{S}} + \lambda P_{\mathcal{S}^{\perp}}P_{\mathcal{AS}^{\perp}})\|^2}{\|P_{\mathcal{S}}x + \lambda P_{\mathcal{S}^{\perp}}P_{\mathcal{AS}^{\perp}}x\|^2} \\
&= \frac{\|P_{\mathcal{S}}x\|^2}{\|P_{\mathcal{S}}x\|^2 + \lambda^2\|P_{\mathcal{S}^{\perp}}P_{\mathcal{AS}^{\perp}}x\|^2} \\
&= \frac{1}{1 + \lambda^2\frac{\|P_{\mathcal{S}^{\perp}}P_{\mathcal{AS}^{\perp}}x\|^2}{\|P_{\mathcal{S}}x\|^2}} \tag{82}
\end{aligned}
$$

where the second step for the denominator is due to the orthogonality of $P_{\mathcal{S}^\perp} P_{\mathcal{AS}^\perp} x$ to $P_{\mathcal{S}} x$. From (14), it holds

$$\cos\left(\mathcal{S}^\perp, \mathcal{A}\right) \|P_{\mathcal{AS}^\perp} x\| \le \|P_{\mathcal{S}^\perp} P_{\mathcal{AS}^\perp} x\| \le \sin\left(\mathcal{A}, \mathcal{S}\right) \|P_{\mathcal{AS}^\perp} x\|. \tag{83}$$

Then, from (16), it follows that $P_{\mathcal{AS}^\perp} x$ satisfies

$$\frac{\|P_{\mathcal{S}} x\|}{\sin\left(\mathcal{S}^\perp, \mathcal{A}\right)} \le \|P_{\mathcal{AS}^\perp} x\| \le \frac{\|P_{\mathcal{S}} x\|}{\cos\left(\mathcal{A}, \mathcal{S}\right)}. \tag{84}$$

Combining (83) and (84) yields

$$\frac{\cos(\mathcal{S}^\perp, \mathcal{A})}{\sin\left(\mathcal{S}^\perp, \mathcal{A}\right)} \|P_{\mathcal{S}} x\| \le \|P_{\mathcal{S}^\perp} P_{\mathcal{AS}^\perp} x\| \le \frac{\sin(\mathcal{A}, \mathcal{S})}{\cos\left(\mathcal{A}, \mathcal{S}\right)} \|P_{\mathcal{S}} x\|. \tag{85}$$

As a result, we have from (82) that

$$\frac{1}{1 + \lambda^2 \frac{\sin^2(\mathcal{A}, \mathcal{S})}{\cos^2(\mathcal{A}, \mathcal{S})}} \le f(x) \le \frac{1}{1 + \lambda^2 \frac{\cos^2(\mathcal{S}^\perp, \mathcal{A})}{\sin^2(\mathcal{S}^\perp, \mathcal{A})}} \tag{86}$$

and (65) follows immediately from (80) and (86). □

APPENDIX E
PROOF BOUNDS OF $\sin(\mathcal{B}^\perp, \mathcal{S})$ IN (66)

From (9),

$$\sin^2\left(\mathcal{B}^\perp, \mathcal{S}\right) = \sup_{x \notin \mathcal{S}} g(x) \tag{87}$$

where

$$g(x) = \frac{\|P_{\mathcal{S}^\perp} P_{\mathcal{B}^\perp \mathcal{S}} x\|^2}{\|P_{\mathcal{B}^\perp \mathcal{S}} x\|^2}.$$

According to [18], the adjoint operator of a projection $P_{\mathcal{AS}}$ is also a projection:

$$P_{\mathcal{AS}}^* = P_{\mathcal{S}^\perp \mathcal{A}^\perp}. \tag{88}$$

We can then show that

$$\begin{aligned}
P_{\mathcal{B}^\perp \mathcal{S}} &= I - P_{\mathcal{S}\mathcal{B}^\perp} \\
&= I - B^* \\
&= I - \left(\lambda P_{\mathcal{AS}^\perp} + (1 - \lambda) P_{\mathcal{S}}\right)^* \\
&= I - \left(\lambda P_{\mathcal{S}\mathcal{A}^\perp} + (1 - \lambda) P_{\mathcal{S}^\perp}\right) \\
&= \lambda(I - P_{\mathcal{S}\mathcal{A}^\perp}) + (1 - \lambda)(I - P_{\mathcal{S}}) \\
&= \lambda P_{\mathcal{A}^\perp \mathcal{S}} + (1 - \lambda) P_{\mathcal{S}^\perp}.
\end{aligned}$$

Note that $g(x)$ has the same form as $f(x)$, except that all the subspaces involved are replaced by their orthogonal complements. Using (86) and noting $(\mathcal{A}^\perp, \mathcal{S}^\perp) = (\mathcal{S}, \mathcal{A})$, $(\mathcal{A}, \mathcal{S}) = (\mathcal{S}, \mathcal{A})$, we have

$$\frac{1}{1 + \lambda^2 \frac{\sin^2(\mathcal{A}, \mathcal{S})}{\cos^2(\mathcal{A}, \mathcal{S})}} \le g(x) \le \frac{1}{1 + \lambda^2 \frac{\cos^2(\mathcal{S}^\perp, \mathcal{A})}{\sin^2(\mathcal{S}^\perp, \mathcal{A})}}.$$

Then, inequality (66) follows immediately. □

REFERENCES

[1] T. Hentschel and G. Fettweis, "Sample rate conversion for software radio," *IEEE Commun. Mag.*, vol. 38, no. 8, pp. 142–150, 2000.

[2] T. M. Lehmann, C. Gonner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE Trans. Med. Imag.*, vol. 18, no. 11, pp. 1049–1075, 1999.

[3] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *Int. J. Imag. Syst. Technol.*, vol. 14, no. 2, pp. 47–57, 2004.

[4] A. Ortega, P. Frossard, J. Kovačević, J. M. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[5] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: Sampling theory," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6510–6523, 2015.

[6] Y. C. Eldar and T. G. Dvorkind, "A minimum squared-error framework for generalized sampling," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2155–2167, 2006.

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

[8] M. Unser and J. Zerubia, "A generalized sampling theory without band-limiting constraints," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 45, no. 8, pp. 959–969, 1998.

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

[10] Y. C. Eldar, "Sampling without input constraints: Consistent reconstruction in arbitrary spaces," in *Sampling, Wavelets, and Tomography*, pp. 33–60, Springer, 2004.

[11] T. G. Dvorkind, Y. C. Eldar, and E. Matusiak, "Nonlinear and nonideal sampling: Theory and methods," *IEEE Trans. Signal Process.*, vol. 56, no. 12, pp. 5874–5890, 2008.

[12] Y. C. Eldar and M. Unser, "Nonideal sampling and interpolation from noisy observations in shift-invariant spaces," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2636–2651, 2006.

[13] A. Hirabayashi and M. Unser, "Consistent sampling and signal recovery," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4104–4115, 2007.

[14] T. G. Dvorkind and Y. C. Eldar, "Robust and consistent sampling," *IEEE Signal Proc. Lett.*, vol. 16, no. 9, pp. 739–742, 2009.

[15] M. L. Arias and C. Conde, "Generalized inverses and sampling problems," *J. Math. Anal. Appl.*, vol. 398, no. 2, pp. 744–751, 2013.

[16] K. H. Kwon and D. G. Lee, "Generalized consistent sampling in abstract Hilbert spaces," *J. Math. Anal. Appl.*, vol. 433, no. 1, pp. 375–391, 2016.

[17] A. Knyazev, A. Gadde, H. Mansour, and D. Tian, "Guided signal reconstruction theory," *arXiv preprint arXiv:1702.00852*, 2017.

[18] O. Christensen and Y. C. Eldar, "Oblique dual frames and shift-invariant spaces," *Appl. Comput. Harmon. Anal.*, vol. 17, no. 1, pp. 48–68, 2004.

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

[20] E. Kreyszig, *Introductory Functional Analysis with Applications*. Wiley New York, 1978.

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

[22] B. Adcock, A. C. Hansen, and C. Poon, "Beyond consistent reconstructions: Optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem," *SIAM J. Math. Anal.*, vol. 45, no. 5, pp. 3132–3167, 2013.

[23] A. Anis, A. Gadde, and A. Ortega, "Efficient sampling set selection for bandlimited graph signals using graph spectral proxies," *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3775–3789, 2016.

[24] T. Blu, P. Thévenaz, and M. Unser, "Linear interpolation revitalized," *IEEE Trans. Image Process.*, vol. 13, no. 5, pp. 710–719, 2004.

[25] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques: Part I— Interpolators and projectors," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2783–2795, 1999.

[26] T. G. Dvorkind, H. Kirshner, Y. C. Eldar, and M. Porat, "Minimax approximation of representation coefficients from generalized samples," *IEEE Trans. Signal Process.*, vol. 55, no. 9, pp. 4430–4443, 2007.

[27] Y. C. Eldar, "Mean-squared error sampling and reconstruction in the presence of noise," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4619–4633, 2006.

[28] Y. C. Eldar and T. Michaeli, "Beyond bandlimited sampling," *IEEE Signal Process. Mag.*, vol. 26, no. 3, pp. 48–68, 2009.

[29] B. B. Haro and M. Vetterli, "Sampling continuous-time sparse signals: A frequency-domain perspective," *IEEE Trans. Signal Process.*, vol. 66, no. 6, pp. 1410–1424, 2018.

[30] A. Hirabayashi, "Consistent sampling and efficient signal reconstruction," *IEEE Signal Process. Lett.*, vol. 16, no. 12, pp. 1023–1026, 2009.

[31] A. Knyazev, A. Jujunashvili, and M. Argentati, "Angles between infinite dimensional subspaces with applications to the Rayleigh-Ritz and alternating projectors methods," *arXiv preprint arXiv:0705.1023*, 2007.

[32] T. Košir and M. Omladič, "Normalized tight vs. general frames in sampling problems," *Adv. Oper. Theory*, vol. 2, no. 2, pp. 114–125, 2017.

[33] B. Sadeghi and R. Yu, "Shift-variance and cyclostationarity of linear periodically shift-variant systems," in *10th Int. Conf. Sampling Process. Theory Appl., Bremen*, 2013.

[34] B. Sadeghi and R. Yu, "Shift-variance and nonstationarity of linear periodically shift-variant systems and applications to generalized sampling-reconstruction processes," *IEEE Trans. Signal Process.*, vol. 64, no. 6, pp. 1493–1506, 2016.

[35] B. Sadeghi, R. Yu, and R. Wang, "Shifting interpolation kernel toward orthogonal projection," *IEEE Trans. Signal Process.*, vol. 66, no. 1, pp. 101–112, 2018.

[36] J. Shi, X. Liu, L. He, M. Han, Q. Li, and N. Zhang, "Sampling and reconstruction in arbitrary measurement and approximation spaces associated with linear canonical transform," *IEEE Trans. Signal Process.*, vol. 64, no. 24, pp. 6379–6391, 2016.

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

[38] M. Vetterli, J. Kovačević, and V. K. Goyal, *Foundations of Signal Processing*. Cambridge University Press, 2014.

[39] L. Xu, R. Tao, and F. Zhang, "Multichannel consistent sampling and reconstruction associated with linear canonical transform," *IEEE Signal Process. Lett.*, vol. 24, no. 5, pp. 658–662, 2017.